

Curriculum vitae

Quochi, Valeria

RESEARCH INTEREST AND TRACK RECORD

In the last decade Valeria Quochi essentially focused on two main activities: scientific research in computational lexicography, automatic acquisition and standard representation of lexical (semantic) information; and institutional research on infrastructural issues and community building initiatives.

She actively worked on the acquisition of lexical information from corpora and of their digital representation using standard formats, or close adaptations of them. She participated to EU R&D projects such as OPENER (FP7-ICT-2011-SME-DCL-296451), on the creation of language resources and technology for multilingual sentiment and opinion mining, and PANACEA (FP7 ICT-2009--248064), on the creation of a platform for the automatic acquisition of language resources online, in which she led WP7 on the validation and evaluation of the platform and acted as deputy team leader. Her interests leads her to get involved in projects that require the study, creation, improvement, maintenance and update of lexical resources such as the PAROLE-SIMPLE-CLIPS and ItalWordNet lexicons, the ImagACT ontology of action types; as well as in the design and testing of lexicon manipulation and editing software that would help in such tasks. She is currently the team leader in the national project “Languages and Culture of Ancient Italy. Historical Linguistics and Digital Models”, for which the CNR-ILC team is developing the digital platform that will help create dictionaries for ancient fragmentary languages and integrate/link them to their related epigraphic texts and relevant bibliography.

Regarding institutional research, certain of their potential for effectively improving cross-disciplinary and cross-boundary collaboration in all fields of research also in the SSH, she strongly supports digital Research Infrastructures and has actively taken part in the pan-European Common Language and technology Research Infrastructure (CLARIN) since its preparatory phase started in 2008. Dr. Quochi has contributed to the definition and adoption of standard data and metadata representation formats and has participated to the set-up and construction of the Italian CLARIN-IT national infrastructural network. She currently cares for the data curation of the national repository ILC4CLARIN, takes part in the executive board of the national consortium (<http://www.clarin-it.it/en/content/governance>) and represents Italy within the CLARIN-ERIC User Involvement Committee (<https://www.clarin.eu/governance/user-involvement-committee>).

Since 2018 she has been involved in the European LEXicography InfraStructure - ELEXIS (H2020-INFRAIA-02-2017-731015), with a special interest in the creation of the Italian sense inventory and annotation of a test corpus for multilingual WSD.

Finally, she is involved in the recent ParlaMint project, a CLARIN-sponsored project on the creation of comparable, linguistically annotated, multilingual corpus of transcriptions of parliamentary debates of various European national parliaments.

EDUCATION

2007	PhD in Linguistics Dipartimento di Linguistica “T. Bolelli”, Università di Pisa, Italy
2002	Laurea Degree Facoltà di Lettere, Università di Pisa, Italy

CURRENT POSITION(S)

2014 –	Permanent Researcher, Liv. III, Istituto di Linguistica Computazionale “A. Zampolli”,
--------	---

PREVIOUS POSITIONS

- 2010 – 2013 Fixed term Researcher, Liv. III, Istituto di Linguistica Computazionale “A. Zampolli”, Consiglio Nazionale delle Ricerche, Pisa, Italy
- 1998 – 2005 Occasional Collaborator, Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche, Pisa, Italy
- 2003 Language Consultant, Google Inc, remotely.
- 2002 Research Collaborator, Consorzio Pisa Ricerche, Pisa, Italy

FELLOWSHIPS

- 2009 – 2010 Post-doc Research Grant, Istituto di Linguistica Computazionale “A. Zampolli”, Consiglio Nazionale delle Ricerche, Pisa, Italy
- 2007- 2008 Research Grant, Istituto di Linguistica Computazionale “A. Zampolli”, Consiglio Nazionale delle Ricerche, Pisa, Italy
- 2006 – 2007 CNR Research Scholarship, Istituto di Linguistica Computazionale “A. Zampolli”, Consiglio Nazionale delle Ricerche, Pisa, Italy
- 2003-2005 PhD Scholarship, Dipartimento di Linguistica “T. Bolelli”, Università di Pisa.

RESEARCH PROJECTS

- 2020- Languages and Cultures of Ancient Italy. Historical Linguistics and Digital Models, PRIN 2017 SH5 Linea A. Role: Leader of the CNR research unit.
- 2020 - ParlaMint-IT (a CLARIN-ERIC funded project)
- Since 2018: European LEXicography InfraStructure (ELEXIS), 2020-INFRAIA-02-2017-731015.
- Since 2015: CLARIN-IT (now funded nationally as CLARIN FOE). Roles: User Involvement referent.
- 2015-2017: The Digital Language Diversity Project (DLDP), Erasmus+ Programme, 2015-1-IT02-KA204-015090.
- 2013-2014: Open Polarity Enhanced Name Entity Recognition (OPENER), FP7-ICT-2011-SME-DCL-296451.
- 2013: Ontologia per immagini dell'azione per l'acquisizione di L2 in ambiente multilingue (IMAGACT), PAR FAS REGIONE TOSCANA, B11J1000055002.
- 2010-2012: Platform for Automatic Normalized Annotation and Cost-Effective Acquisition of Language Resources for Human Language Technologies (PANACEA), FP7-ICT-2009-248064. Roles: Workpackage Leader (WP7 Evaluation).
- 2008-2011: Fostering Language Resources Network (FLaReNet), ECP-2007-LANG-617001. Roles: coordination staff member, editorial group member.
- 2008-2011: Common Language Resources and Technology Infrastructure (CLARIN), ESFRI Preparatory Phase, FP7-INFRASTRUCTURES-2007-212230. Role: Deputy of the CNR group leader, scientific participant.
- 2006-2009: Bootstrapping Of Ontologies and Terminologies Strategic REsearch Project (BootSTREP), FP6-IST-2004-028099.
- 2002-2003: International Standards for Language Engineering (ISLE), FP5-IST-1999-10647.

TEACHING ACTIVITIES / INVITED SPEAKER

- 20/05/2021 Invited talk at Università di Firenze, Progetto Dipartimenti Eccellenti, “Lessicologia e lessicografia computazionale nell’era del web semantico e dell’intelligenza artificiale.
- 21/06/2011 Instructor at the CLARA training course on Multilingual Lexical Resources and Tools. Bergen, Norway
- 2006 Instructor, seminar series on "Generative Lexicon" and "Construction Grammar" for the BA Lingua 1 course, LEU, Univ. Pisa

VISITING PERIODS

- 2005 2005 visiting scholar at the Linguistics Programme of the Princeton University (Princeton, NJ, USA), guest of Prof. Adele E. Goldberg.

ORGANISATION OF SCIENTIFIC MEETINGS (if applicable)

- 2012 Member of the Organising Committee, international workshop: LREC 2012 Workshop on Language Resource Merging.
- 2011 Member of the Organising Committee, international conference: “3rd European Language Resources and Technologies Forum: Language Resources in the Sharing Age - the Strategic Agenda” 26-27 May 2011, Venezia, Italy.
- 2010 Member of the Organising Committee, international conference: “2 nd European Language Resources and Technologies Forum: Language Resources of the future - the future of Language Resources”. 11-12 February 2010, Barcelona, Spain.
- 2010 Member of the Organising Committee, international workshop “Semeval-2” Evaluation Campaign and workshop, Task 7 Argument Selection and Coercion. 17 July 2010, Uppsala, Finland.
- 2009 Member of the Organising Committee, international conference: “5th International Conference on Generative Approaches to the Lexicon, Pisa, Italia, 17-19 September 2009.
- 2009 Member of the Organising Committee, international conference: “1st European Language Resources and Technologies Forum: Shaping the Future of the Multilingual Digital Europe.” 12-13 February 2009, Vienna, Austria.
- 2008 Member of the Organising Committee, logistics and PRs, international conference: “Sixth International Language Resource European Conference LREC 2008”, Marrakech, Morocco.
- 2006 Member of the Organising Committee, national conference “Tecnologie per l’e-participation TALeP 2006: Freno o opportunità per la partecipazione democratica?”. 27-28 November 2006, Pisa.

INSTITUTIONAL RESPONSIBILITIES

- 2017 – Member of the User Involvement Committee, Italian Representative, CLARIN-ERIC research infrastructure.
- 2015 – Data Manager and curator for the ILC4CLARIN center of the CLARIN-IT infrastructure.
- 2015 – Member of the CLARIN-IT executive committee and User Involvement referent
- 2014 – Member of the Committee for international relations for the Area di Ricerca di Pisa, CNR.

REVIEWING ACTIVITIES

- 2018 – 2021 Scientific Committee member, Peer-reviewer, CLARIN annual conference
- 2017 Evaluator for the *"Rita Levi Montalcini" Italian Junior research programme*
- 2016 Scientific Committee member, Peer-reviewer, International Conference on Computational Linguistics (COLING)
- 2014 – 2020 Scientific Committee member, Peer-reviewer, CLIC-IT Italian Conference on Computational Linguistics

- 2012, 2014 Scientific Committee member, Peer-reviewer, annual conference of the Association for Computational Linguistics.
- 2011, 2012 Scientific Committee member, Peer-reviewer, European Intelligence and Security Informatics Conference (EISIC).
- 2011 Scientific Committee member, Peer-reviewer, International Symposium on Open Source Intelligence & Web Mining (OSINT-WM 2011)
- 2010 – Reviewer, *Language Resources and Evaluation Journal*, Springer.
- 2009, 2013 Scientific Committee member, Peer-reviewer, program committee member, International Conference on Generative Approaches to the Lexicon (GL2009).
- 2008 – 2020 Scientific Committee member, Peer-reviewer, International Language Resources and Evaluation Conference (LREC)

MAJOR COLLABORATIONS

- Prof. Nuria Bel, acquisition of language resources, Universitat Pompeu Fabra, Barcellona, Spain
- Prof. Francesca Murano, representation of lexical information for ancient fragmentary languages, DILEF, Università di Firenze.

SELECTED PUBLICATIONS

Elisa Dominutti, Lucia Pifferi, Felice Dell'Orletta, Simonetta Montemagni, Valeria Quochi (2019). Building an Italian Written-Spoken Parallel Corpus: a Pilot Study. In: Raffaella Bernardi, Roberto Navigli e Giovanni Semeraro (a cura di) *Proceedings of the Sixth Italian Conference on Computational Linguistics, (CLiC-it 2019)*, Bari, Italy, November 13-15 2019. CEUR Workshop Proceedings, vol. 2481. ISSN: 91613-0073.

Nicolas Lionel, Alexander König, Monica Monachini, Riccardo Del Gratta, Silvia Calamai, Andrea Abel, Alessandro Enea, Francesca Biliotti, Valeria Quochi, Francesco Vincenzo Stella (2018). CLARIN-IT: State of Affairs, Challenges and Opportunities. In: *Post-conference proceedings of the CLARIN Annual Conference 2017*. Linköping: Linköping University Electronic Press., Budapest, 18-20 September 2017

Claudia Soria, Valeria Quochi, Irene Russo (2018). The DLDP Survey on Digital Use and Usability of EU Regional and Minority Languages. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. ISBN: 979-10-95546-00-9. European Language Resources Association (ELRA), Paris (Francia), pp. 4155-4160.

Quochi Valeria (2016). Development and representation of Italian light-fare constructions. In: (a cura di): Jiyoung Yoon and Th. Stefan Gries, *Corpus-based Approaches to Construction Grammar*. vol. 19, p. 39-64, Amsterdam/Philadelphia: John Benjamins Publishing Company, ISBN: 9789027204417, doi: 10.1075/cal.19.03quo

Bartolini Roberto, Valeria Quochi, Irene De Felice, Irene Russo, Monica Monachini (2014). From Synsets to Videos: Enriching ItalWordNet Multimodally. In: (a cura di): N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odiijk, S. Piperidis, *Proceedings of the 9th Language Resources and Evaluation Conference (LREC 2014)*. p. 3110-3117, PARIS: European Language Resources Association (ELRA), ISBN: 978-2-9517408-8-4, Reykjavik, Iceland, 26-31 May 2014

Calzolari Nicoletta, Monachini Monica, Valeria Quochi, Soria Claudia, Toral Antonio (2014). Lexicons, Terminologies, Ontologies: Reflections from Experiences in Resource Construction. In: (a cura di) Nachum Dershowitz and Ephraim Nissan, *Language, Culture, Computation. Computational Linguistics and Linguistics. Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday*, Part III. LECTURE NOTES IN COMPUTER SCIENCE, vol. 8003, p. 103-121, Berlin - Heidelberg: Springer-Verlag, ISBN: 978-3-642-45326-7, ISSN: 0302-9743, doi: 10.1007/978-3-642-45327-4_7

Frontini Francesca, Valeria Quochi, Sebastian Padó, Jason Utt, Monica Monachini (2014). Polysemy Index for Nouns: an Experiment on Italian using the PAROLE SIMPLE CLIPS Lexical Database. In: (a cura di): N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odiijk, S. Piperidis, *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014)*. p.

2955-2963, PARIS: European Language Resources Association (ELRA), ISBN: 978-2-9517408-8-4, 26-31 May 2014

Soria Claudia, Nicoletta Calzolari, Monica Monachini, Valeria Quochi, Nuria Bel, Khalid Choukri, Joseph Mariani, Jan Odiijk, Stelios Piperidis (2014). The language resource Strategic Agenda: the FLaReNet synthesis of community recommendations. *Language Resources and Evaluation Journal*, vol. 48, p. 753-775, ISSN: 1574-020X, doi: 10.1007/s10579-014-9279-y

Quochi Valeria, Frontini, Francesca, Rubino, Francesco (2012). A MWE Acquisition and Lexicon Builder Web Service. In: (a cura di) Martin Kay and Christian Boitet, *Proceedings of the 24th International Conference on Computational Linguistics, COLING 2012*. p. 2291-2306, Association for Computational Linguistics (ACL), ISBN: 9781627483896, Mumbai, India, 8-15 December 2012

Thompson Paul, John McNaught, Simonetta Montemagni, Nicoletta Calzolari, Riccardo Del Gratta, Vivian Lee, Simone Marchi, Monica Monachini, Piotr Pezik, Valeria Quochi, C. J. Rupp, Yutaka Sasaki, Giulia Venturi, Dietrich Rebholz-Schuhmann, Sophia Ananiadou (2011). The BioLexicon: a large-scale terminological resource for biomedical text mining. *BMC Bioinformatics*, vol. 12, ISSN: 1471-2105, doi: 10.1186/1471-2105-12-397