

Curriculum Vitae

Nicoletta Calzolari Zamorani

Research Associate, former *Director of Research* and *Director* of the Istituto di Linguistica Computazionale “Antonio Zampolli” of the National Research Council (ILC – CNR), Pisa, Italy.

Awards

- Awarded the title *ACL Fellow* for “significant contributions to computational lexicography, and for the creation and dissemination of language resources” in the founding group of the ACL Fellows Program (http://aclweb.org/aclwiki/index.php?title=ACL_Fellows) of the Association for Computational Linguistics
- *Honorary Doctorate* in Philosophy received from the University of Copenhagen, November 2007, “for her significant contribution to the field of Computational Linguistics”.

Professional Experience

- *Director of the Istituto di Linguistica Computazionale* (ILC-CNR) (2003-2008)
- Member of the “*Comitato Ordinatore*” of CNR for setting up the ICT (Information and Communication Technology) Department (2004-06)
- *Director of Research* at ILC-CNR, since 1992
- Researcher at Pisa University, Department of Linguistics (1981-1992)
- “*Contratto quadriennale*” (4-year grant) at Pisa University, first at the Institute of Philosophy, then at the Institute of Glottology – Mathematical linguistics, (1974-1981)
- Ministerial fellowship (Borsa di Studio Ministeriale di addestramento didattico e scientifico) at CNUCE (Centro Nazionale Universitario di Calcolo Elettronico), Group of Computational Linguistics (1972-1974)

Academic Experience

- *Visiting Professor* at Kyoto University (December 2014 - February 2015)
- Member of the *PhD Board for Linguistics* (Consiglio di Dottorato in Linguistica) at Pisa University
- Member of many PhD juries in Italy and abroad
- Invited to teach many courses on many aspects of Computational Linguistics at many International Schools, PhD courses, in Italy and abroad
- Co-Director, with B. Quemada, of the *3rd Escuela Interlatina de Altos Estudios en Linguística Aplicada*, organised by the Real Academia Espanola de la Lengua, 2003

Research Interests and Research Activities

Computational Linguistics; Language Technology; Language Resources and Knowledge Resources; Computational Lexicography and Lexicology: Lexical Knowledge Bases, Multilingual lexicons, Machine-readable dictionaries, relations between Lexicons, Ontologies and Terminologies; Standardisation and reusability of lexical resources; Corpus Linguistics and statistical analysis of textual corpora; Knowledge Acquisition from multiple (lexical and textual) sources, integration and representation; Morphology and word-formation; Lexical Semantics and Semantic Annotation;

Collocations and Multi-words; Natural Language Processing and Language Engineering Applications: of various types (filtering, summarisation, information extraction, etc.) and in different areas (e-government, humanities and social sciences, medicine, biology, etc.); Validation of Language Resources; Infrastructural issues related to Language Resources.

Professional Activities

Coordination/Management of Projects

– International and European Projects:

Coordinating and managing International and European projects/initiatives:

- *Chair of ISO/TC 37/SC 4 on Language Resource Management and committee member of ISO/TC 37/AG 0 (since 2015)*
- *EU Coordinator of the European eContentplus Thematic Network FLaReNet – Fostering Language Resources Network (2008-11), whose mission is to identify short, medium, and long-term strategic objectives and provide consensual recommendations in the form of a plan of action for EC, national organisations and industry*
- *Convener of the ISO (International Organization for Standardization) Working Group: ISO/TC 37/SC 4/WG 4 – Lexical Databases (2003-14), for the definition of international standards for lexical resources*
- *Chair of the Scientific Board of the ESFRI Research Infrastructure CLARIN – Common Language Resources and Technology Infrastructure (2008-10), whose mission is offering a stable, persistent, accessible and extendable infrastructure and enabling eHumanities*
- *EU Coordinator of the European Safe Internet Action Plan POESIA Project POESIA – Public Opensource Environment for a Safer Internet Access (IAP 2117/27572) (2002-04)*
- *EU Coordinator of the European IST Thematic Network ENABLER – European National Activities for Basic Language Resources (European Network of National Projects) (IST-2000-31069) (2001-03). This was the first Network of National projects in Europe in the field of Language Resources*
- *EU Coordinator of the European MLIS MUSI Project – Multilingual Summarization for the Internet (MLIS-5015) (2000-02)*
- *EU Coordinator of the European LE SPARKLE Project – Shallow Parsing and Knowledge Extraction for Language Engineering (LE1-2111) (1995-97)*
- *Responsible of the European ISLE Computational Lexicon Working Group of the International project between Europe and US: IST-HLT US-NSF ISLE – International Standards for Language Engineering (IST-1999-10647) (2000-03)*
- *Chief Co-Editor and Responsible of the EAGLES Computational Lexicon Working Group of the European Projects:*
 - *LRE EAGLES I – Expert Advisory Group on Language Engineering Standards ([LRE-61100](#)) (1993-96)*
 - *LE EAGLES II – Expert Advisory Group on Language Engineering Standards (LE3-4244 10484/0) (1996-99)*

Coordinating and managing the Pisa Unit in International and European projects:

- *ESPRIT BRA (Basic Research Action) ACQUILEX I – Acquisition of Lexical Knowledge for Natural Language Processing Systems (1989-92)*

- EUROTRA-7 Feasibility Study – *Reusability of lexical and terminological resources*. (1990-91)
- ESPRIT MULTILEX – [*A Multi-Functional Standardised Lexicon for European Community Languages*](#) (ESPRIT II-5304) (1990-93)
- EUREKA GENELEX – Generic Lexicon (1990-94), an industrial European project to define and establish a generic model for computational lexicons.
- MLAP NERC – *Network of European Reference Corpora* (1991-93)
- *Multilingual Corpora* – Feasibility project
- EUROTRA ET-10/51 – *Semantic Analysis Using a Natural Language Dictionary* (1992-1993)
- EUROTRA ET-10/63 – *Probabilistic and Corpus--based Methods in EUROTRA: Terminology, Lexicon and Preference* (1992-1993)
- EUROTRA ET-10/75 – *Collocations and the Lexicalisation of Semantic Operations* (1992-1993)
- ESPRIT BRA (Basic Research Action) ACQUILEX II – *Acquisition of Lexical Knowledge for Natural Language Processing Systems* (1992-95)
- ESPRIT DARPA/NSF Co-operation on *Lexical Databases* (ECUS-04) (1992-93)
- LRE DELIS – *Descriptive Lexical Specifications and Tools for corpus-based Lexicon building* (LRE 61034) (1993-95)
- LRE RENOS – *Reduction of Noise and Silence in Full Text Retrieval Systems for Legal Texts* (LRE62074) (1993-95)
- LRE RELATOR (LRE62056) (1993-95) – Feasibility study for a European Infrastructure for Language Resources
- LRE MEMORIA (LRE63257) (1994-95)
- LRE PAROLE – *Preparatory Action for Linguistic Resources Organization for Language Engineering* (LRE63368) (1994-95)
- LRE MULTEXT – *Multilingual Text Tools and Corpora* (LRE62050) (1994-96)
- LE PAROLE – *Permanent Action for Linguistic Resources Organization for Language Engineering* (LE2-4017) (1996-97)
- LE EuroWordNet I – *Building a multilingual WordNet database with semantic relations between words* (LE2-4003) (1996-99)
- LE EuroWordNet II – *Building a multilingual WordNet database with semantic relations between words* (1997-99) (LE4-8328)
- UNL – *Universal Networking Language* - United Nations University Project
- LE ELSE – *Evaluation of Language and Speech Engineering* (LE4-8340) (1998-99)
- LE SIMPLE – *Semantic Information for Multifunctional Plurilingual LEXicons* (LE4-8346) (1998-2000)
- US NSF XMELLT – *Cross-lingual Multi-word Expression Lexicons for Language Technology* (1999-2001)
- ELSNET-4 IST Thematic Network – *European Network of Excellence in Human Language Technologies* (IST-2001-34571) (2002-04)
- e-Content INTERA – *Integrated European language data Repository Area* (2003-05)
- e-Content LIRICS – *Linguistic Infrastructure for Interoperable Resources and Systems* (2005-07)
- Japanese NEDO – *Developing International Standards of Language Resources for Semantic Web Applications*, International Joint Research Program of the New Energy and Industrial Technology Development Organization (2005-08)
- IST Network of Excellence DEMO-net: *The e-Participation Network* (FP6-2004-27219) (2006-10)

- IST BOOTStrep – *Bootstrapping Of Ontologies and Terminologies STRategic REsearch Project* (2006-09), in the biological domain.
- Japanese NICT *Infrastructure for Intercultural Collaboration* (2007-09)
- ESFRI CLARIN – *Common Language Resources and Technology Infrastructure* (2008-10), for Social Sciences and the Humanities
- 7thFP ICT PANACEA – *Platform for Automatic, Normalized Annotation and Cost-Effective Acquisition of Language Resources for Human Language Technologies* (2010-13)
- 7thFP ICT Network of Excellence METANET – T4ME – *Technologies for the Multilingual European Information Society* (2010-13), whose mission is to forge a strategic alliance for creating the technologies and applications needed for making the linguistic diversity of our multicultural society sustainable
- CIP ICT-PSP *Multilingual Web – Standards and best practices for the Multilingual Web Thematic Network* (2010-13)

– **National Projects:**

Coordinating and managing National projects or the Pisa Unit in National projects:

- Coordinator of the *Project on Lexical Acquisition* in the Italian National Project – Cluster 18 “Linguistica” – Law 488 (2000-03)
- Responsible for the global Technical Coherence in two Italian National Projects:
 - *TAL – Infrastruttura nazionale per le Risorse Linguistiche nel settore del trattamento automatico della lingua naturale parlata e scritta* – Legge 46/82 art. 10 (1999-2001)
 - *LCRMM – Linguistica Computazionale: ricerche monolingui e multilingui* - Cluster 18 “Linguistica” – Law 488 (2000-03)
- Responsible for the Computational Lexicon in the *CLIPS* National Project– Cluster 18 “Linguistica” – Law 488 (2000-03)
- Responsible for ILC of the project *ADAPT TECLA* of the Ministero del Lavoro with the Istituto Universitario Orientale di Napoli (1999-2000)
- Responsible for ILC of the CNR Project “*ITAL*”
- Responsible of the Unit “Models of Lexical information” in the Italian National Strategic Project on “Natural Language Processing in the area Information Technology”, 19867-88
- Responsible of the Unit “Representation of linguistic properties in an Italian Lexical Database” in the Italian National Strategic Project on “Methods and Tools for Language Industry in the area Information Technology”, 1986

Honorary/Scientific positions

– **International and National Committees (also past member):**

- *President of ELRA (European Language Resources Association)* (since 2012), and previously Vice-President from 2010 and 2004-08) and Board Member (since 2010, and from 2002-08), association with both academic and industrial institutional members, for the coordination of the field of language resources
- Permanent Member (since 1973) of the *ICCL (International Committee of Computational Linguistics)*
- Member of the *Board of UNDL Foundation (Universal Networking Digital Language Foundation)* (since 2014) (<http://www.undlfoundation.org/undlfoundation/>)

- Member of the *Advisory Board* (since 2014) of *LIDER (Linked Data as an enabler of cross-media and multilingual content analytics for enterprises across Europe)*, FP7 project (<http://www.lider-project.eu/>) on Content analytics and language technologies
- *Vice President* of the Legal organisation *META-TRUST* (<http://www.meta-net.eu/meta-trust>), “the legal person of the network of excellence META-NET and the Multilingual Europe Technology Alliance (META)”
- Member of the *Advisory Board of ESICT* (Experience-oriented Sharing of health knowledge via Information and Communication Technology) Danish Project (2010-13)
- *President* of the *International Association PAROLE* (since 2005), with institutional members involved in language resource creation
- *Standing Committee Member* (since 2010) of the *META Technology Council* of the META-NET Network of Excellence, for the preparation of a strategic research agenda for language technology
- Member (since 2010) of the *Vision Group on Media and Information Services* of the META-Net Network of Excellence
- Member (since 2009) of the *ESFRI Social Sciences and Humanities Thematic Working Group*, with the mandate of developing the European landscape for Research Infrastructures
- Member of the *ACL Executive Committee* (2007-09)
- *Chair* (since 2007) of the Scientific Board of the ESFRI Research Infrastructure *CLARIN (Common Language Resources and Technology Infrastructure)*
- *Founding Member* of the *Italian Forum for Language Technology (Forum TAL)*, nominated by the Ministry of Communication (since 2002)
- Member of *IULA’s Advisory Committee*, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra, Barcelona (since 2005)
- *Chair* of the *ELRA Production Committee*
- Member of the *ELRA Validation Committee*
- Member of the *Advisory Committee for the 21st Century COE (Center of Excellence) Program "Framework for Systematization and Application of Large-scale Knowledge Resources" (COE 21 LKR)*, Tokyo Institute of Technology, Japan (2004-09)
- Member of the *Comitato Ordinatore of CNR for ICT Department* (2004-06)
- *Convenor* of the *WRITE (Written Resources Infrastructure Technology and Evaluation) International Committee*, international federation for language resources (2003-08)
- Member of *ELSNET (European Network of Excellence of Language and Speech) Board* (since 2003)
- Member of the *International Scientific Advisory Board of the Austrian Academy Corpus (AAC)* (since 2003)
- Member of *SENSEVAL Advisory Committee*
- Member of *SIGLEX-ACL (Special Interest Group on the Lexicon) Committee*, and member of the *Nominating Committee* (since 1999)
- Research Partner of *ECOR (European Centre for Ontological Research)*
- Member of the *Committee DIAM (Documentazione, informazione automatica, micrografia) of UNI (Ente Nazionale Italiano di Unificazione, Italian Body for Standardisation)* (since 2004)
- Member of the *METALEX Technical Working Group of OLAC (Open Language Archives Community)* (since 2001)
- Member of the *Technical Committee of the European project ECHO*
- Member of the *Steering Board of the ISLE Metadata Initiative (IMDI)*
- Member of *ACO-HUM (Advanced Computing in Humanities)*, EC Socrates/Erasmus Thematic Network (1996-2000)
- Member of the *Coordination Technical Committee of the EU-USA EC-NSF ISLE project*

- Member of the *Project Technical Committee of EC LE-SIMPLE*
- Member of the *Steering Committee of the European Corpus Initiative (ECI)* (1992-94)
- Member of the *ELSNET Task Force on Reusability of Linguistic Resources*
- Member of the *ELSNET Task Force on Training*
- Member of the *Nominating Committee of ACL European Chapter* (1993-96)
- *Co-chair of the TEI (Text Encoding Initiative) Work Group AI5 Print Dictionaries* (1989-1993)
- Member of the *TEI Working Committee on Text Analysis and Interpretation* (1990-93)
- Member of the *TEI Work Group AI1 Linguistic description* (1990-93)
- Member of the *TEI Work Group AI6 Machine Lexica* (1990-93)
- Board member of the *MMA (Multimedia Annotation) Consortium* and member of the *Specification Working Group*
- Member of the *EURALEX Executive Board*
- Member of the *Scientific Council of ASSITERM* (Italian Association for Terminology)
- Chairman of the *Specialist Group on Structured Lexical Data Bases of ALLC* (since 1983)
- Member of the *Scientific Council of ILC* and of the “*Comitato di Istituto*” of *ILC*
- Member of the Scientific Council of the Convention between ILC and Istituto Universitario Orientale di Napoli
- Member of the Scientific Council of the Convention between ILC and Università di Pavia

– **Other international and national scientific and/or organisational activities:**

- *Conference Chair of LREC (International Conference on Language Resources and Evaluation): LREC 2016 Portoroz, LREC 2014 Reykjavík, LREC 2012 Istanbul, LREC 2010 Malta, LREC 2008 Marrakech, LREC 2006 Genova, LREC 2004 Lisbon*
- *General Conference Chair of the International Conference COLING 2016, Osaka*
- *General Conference Chair of the International Conference COLING/ACL 2006, Sydney (Australia), the 3rd Joint Conference of the two oldest international conferences of the field, ACL and COLING*
- *Chair of the 1st, 2nd, 3rd FLAReNet Forum, 2009, 2010, 2011, in Venice, Barcelona, Vienna*
- *President of the Conference TAL2006, Rome, with the Ministry of Communications*
- *President of the Conference TALEP2006, Pisa*
- *Co-Chair of the International Advisory Committee of the 2007 IEEE International Conference on Natural Language Processing and Knowledge Engineering, Beijing, China, Aug.30-Sep.1, 2007*
- *Co-chair of LREC 2006 Workshop Roadmap for Language Resources and Evaluation In a Multilingual Environment, with K. Chouhri, C. Cieri, D. Gibbon, C-R. Huang, organised for COCOSDA and WRITE, Genova.*
- Member of the *International Advisory Committee del First International Joint Conference on Natural Language Processing (IJCNLP-04)*, organized by the Asia Federation of Natural Language Processing (AFNLP), Sanya City, Hainan Island (China), 22-24 March 2004
- *Co-Chair of the International Advisory Committee of IEEE NLP-KE-2003 International Conference, Beijing, October 2003*
- *General Workshop Chair of the International Conference ACL-2000, Hong Kong.*
- *Co-organiser of the SENSEVAL and ROMANSEVAL Word Sense Disambiguation international evaluation initiatives (1998 – 2001 – 2004)*
- *Invited at many Experts Panels at the EC, e.g. for the preparation of the 7th FP*

- Invited as “*Research Fellowship*” at *IBM T.J.Watson Scientific Center*, Yorktown (USA), 1986
- Proposer of the “*Action Plan on Lexicons*” within NERC for the EC, 1992
- *Invited Speaker* at numerous (*about 150*) of the most important international scientific Conferences, Workshops, International Committees, Panels, etc. in Italy, Europe, America, Asia, Australia
- *Member of Programme Committees* for hundreds of the most important International Conferences (COLING, ACL, EACL, LREC, etc.), National Conferences, International and National Workshops, etc.
- Member of the major international and national Associations of the field
- Organiser of several International Conferences and Workshops

Editorial Activities

- *Co-editor-in-chief* with Nancy Ide of the International Journal *Language Resources and Evaluation*, Springer (since 2005)
- *Director* of the Editorial Board of the Journal *Linguistica Computazionale*, IEPI, Pisa – Roma (2003-08)

Member of the Editorial Board, Advisory Board or Program Committee of International Scientific Journals, Series of Books, Publications in the fields of Computational Linguistics and Terminology, among which:

- International Editorial Board member of the *Series Studies in Computational and Theoretical Linguistics*, complementary monograph series to the Prague Bulletin of Mathematical Linguistics
- Advisory Board member of the *International Journal on Terminology*, John Benjamins Publishing Company, Amsterdam
- Editorial Board member of the *Prague Bulletin of Mathematical Linguistics*
- Associate Editor of the International Journal *ACM Transactions on Asian Language Information Processing (TALIP)*, Association for Computing Machinery
- Member of the Editorial Board of the series *Studies in Machine Translation and Natural Language Processing*, Official Publications of the European Commission, Luxembourg
- Editorial Board member of *FCCL – Web Journal of Formal, Computational & Cognitive Linguistics*, Russian Association of Artificial Intelligence, Kazan University, ed. Valery Solovyev
- Advisory Board member of the journal *Pragmalingüística*, University of Cadiz
- Member of the *Advisory Board of the Humanities Computing Yearbook*, Toronto University (Canada) (1988)
- Guest Editor, with Don Walker and Antonio Zampolli, of two Special Issues of the International Journal of *ACL Computational Linguistics*, 13(1987)3-4
- Chairman of the Editorial Committee for the *NERC Final Report*

Referee for many international journals and books.

Research Evaluation Activities

- Member of the *International Evaluation Panel of CHIST-ERA* (www.chistera.eu) (since 2011), an ERA-Net on ICT of the "Future and Emerging Technologies" Unit of DG INFSO

- Member of the *STEVIN International Assessment Panel for the Dutch STEVIN Programme* (since 2004), for the evaluation of Language Technology National programmes in The Netherlands
- *Evaluator* of many international (American, Canadian, European, Asian) and national (for MIUR) Human Language Technology and Linguistics R&D projects, project proposals, research grants, national projects, national programmes, ...
- Referee for the *ANVUR* and for “*Comitato di Indirizzo per la Valutazione della Ricerca*” (CIVR) (Committee for Research Evaluation) of the MIUR (since 2004)
- Member of the MIUR Expert Group for the collaboration Italy –Israel (2004)
- Member of the CNR Expert Committee for the Evaluation of Short-term Mobility projects (since 2004)
- Reviewer of proposals for ESF (European Science Foundation)
- Member for Italy of the Evaluation Panel for the *first LRE Call for Projects of the European Commission*, 1992
- Referee for hundreds of the most important International Conferences in NLP and HLT
- Referee for applications for Tenure or Chair positions in Italy and abroad (e.g. Universities of Princeton, Brandeis, Vassar, Maryland in the US)
- Member of Commissions for PhD in Italy and abroad

Publications

The total numbers of publications, as author or co-author, is **more than 400**.

H-index provided by Google Scholar:

<http://scholar.google.com/citations?hl=en&user=mtRkGQwAAAAJ>

On March 2015:

Citation indices	All	Since 2010
Citations	4505	1396
h-index	33	15
i10-index	107	40