

## Personal information

First name/ Surname **Marco Dinarelli**  
Address B.P. 133, Building S, LIMSI-CNRS, 91403 Orsay Cedex (France)  
Telephone +33 1 69 85 80 96 Mobile +33 672011942  
Web site [www.marcodinarelli.it](http://www.marcodinarelli.it)  
E-mail [marcod@limsi.fr](mailto:marcod@limsi.fr), [marco.dinarelli@gmail.com](mailto:marco.dinarelli@gmail.com)  
Nationality Italian

## Current Position **Post doctoral research associate**

Subject Application of paraphrases to statistical machine translation  
(within the project TRACE, <http://anrtrace.limsi.fr> )

## Research Interests

Spoken Dialog Systems (Automatic Speech Recognition and Understanding)  
Text Categorization  
Data-driven Stochastic Approaches  
Machine Learning  
Kernel Methods  
Natural Language Processing

## Previous Positions

Jun 2010 – November 2011 **Post doctoral research associate**  
Subject Extended named entity detection on transcriptions of French broadcast news and on old documents acquired via optical character recognition  
(within the project Quaero, <http://www.quaero.org>)

## Education

November 2006 – March 2010 **Ph.D. in Information and Communication Technology**  
Ph.D. Thesis "Spoken Language Understanding: from Spoken Utterances to Semantic Structures"  
Subjects covered Automatic Speech Recognition and Understanding, Language Modelling, Machine Learning, Kernel Methods, Text Categorization, Semantic Representations, Information Retrieval.  
Name and type of organisation providing education International Doctoral School in Information and Communication Technology (University)  
14, Via Sommarive, 38100 Trento (Italy)

16/07/2007 - 27/07/2007 (2 weeks) **Summer School Attendant**  
Subjects covered Advanced Dialogue Systems: Affectivity, Adaptability and Multimodality  
Name and type of organisation providing education Computer Science Queen's University (Summer School ELSNET 2007)  
18, Malone Road, BT7 1NN Belfast (United Kingdom)

October 2003 – April 2006 **Master Degree in Computer Science** ("Laurea Specialistica")  
Master Thesis "Data Clusterisation and Compression Using Locality Sensitive Hashing and a Variable Window Burrows-Wheeler Transform Algorithm"  
Name and type of organisation providing education Computer Science Department of University of Pisa  
3, Largo Bruno Pontecorvo, 56100 Pisa (Italy)  
Level international classification M.Sc. Degree

October 1998 – December 2003 **First Level Degree in Computer Science**

Name and type of organisation providing education Department of Computer Science of University of Pisa  
3, Largo Bruno Pontecorvo, 56100 Pisa (Italy)  
Level in international classification B.Sc. Degree

### Teaching

2008 – 2009  
FBK (Bruno Kessler Foundation), University of Trento.  
Assistantship of the course *Spoken Dialog Systems* (20 hours)  
Master in Human Language Technologies for Graduate Students

### Work Experiences

May 2006 – July 2006 Software Developer  
Main activities and responsibilities Development of a tool for Data Clustering and Compression  
Name and address of employer Department of Computer Science  
3, Largo Bruno Pontecorvo, 56100 Pisa (Italy)

September 2002 – January 2003 Software Developer  
Main activities and responsibilities Development of a Network Stack Protocol Layer for MOTES Sensor Networks  
Name and address of employer National Council of Research (CNR) (Research Laboratory)  
1, Via Giuseppe Moruzzi, 56100 Pisa (Italy)

## Publications

### International Journals

Marco Dinarelli, A. Moschitti, G. Riccardi

***Discriminative Reranking for Spoken Language Understanding***

IEEE Journal of Transactions on Audio, Speech and Language Processing (TASLP), volume 20, issue 2, pages 526 - 539, 2011.

S. Hahn, Marco Dinarelli, C. Raymond, F. Lefèvre, P. Lehnen, R. De Mori, A. Moschitti, H. Ney, G. Riccardi

***Comparing Stochastic Approaches to Spoken Language Understanding in Multiple Languages***

IEEE Journal of Transactions on Audio, Speech and Language Processing (TASLP), volume 19, issue 6, pages 1569 - 1583, 2010.

### International Conferences

Marco Dinarelli, S. Rosset

***Models Cascade for Tree-Structured Named Entity Detection***

In Proceedings of International Joint Conference on Natural Language Processing (IJCNLP), Chiang Mai, Thailand, 2011.

A. Garcia-Fernandez, A.L. Ligozat, Marco Dinarelli, D. Bernhard

***When was it written ? Automatically Determining Publication Dates***

In proceedings of String Processing and Information Retrieval (SPIRE), Pisa, Italy, 2011.

Marco Dinarelli, S. Rosset

***Hypotheses Selection Criteria in a Reranking Framework for Spoken Language Understanding***

In Proceedings of Empirical Methods for Natural Language Processing (EMNLP), Edinburgh, U.K., 2011.

Marco Dinarelli, E. Stepanov, S. Varges, G. Riccardi

***The LUNA Spoken Dialog System: Beyond Utterance Classification***

In Proceedings of International Conference of Acoustics, Speech and Signal Processing (ICASSP), Dallas, USA, 2010.

Marco Dinarelli, A. Moschitti, G. Riccardi

***Reranking Models Based On Small Training Data For Spoken Language Understanding***

In Proceedings of Empirical Methods for Natural Language Processing (EMNLP), Singapore, 2009.

S. Quarteroni, G. Riccardi, Marco Dinarelli

***What's In An Ontology For Spoken Language Understanding***

In Proceedings of Interspeech, Brighton, U.K., 2009.

Marco Dinarelli, A. Moschitti, G. Riccardi

***Concept Segmentation And Labeling For Conversational Speech***

In Proceedings of Interspeech, Brighton, U.K., 2009.

Marco Dinarelli, A. Moschitti, G. Riccardi

***Reranking Models for Spoken Language Understanding***

In Proceedings of the European chapter of the Association for Computational Linguistics (EACL), Athens, Greece, 2009.

## International Workshops

C. Grouin, Marco Dinarelli, S. Rosset, G. Wisniewski, P. Zweigenbaum  
**Coreference Resolution in Clinical Reports. The LIMS Participation in the i2b2/VA 2011 Challenge**  
In Proceedings of i2b2/VA 2011 Coreference Resolution Workshop, 2011.

Marco Dinarelli, A. Moschitti, G. Riccardi  
**Hypotheses Selection For Re-ranking Semantic Annotations**  
IEEE Workshop on Spoken Language Technology (SLT), Berkeley, U.S.A., 2010.

S. Quarteroni, Marco Dinarelli, G. Riccardi  
**Ontology-Based Grounding Of Spoken Language Understanding**  
IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Merano, Italy, 2009.

Marco Dinarelli, S. Quarteroni, S. Tonelli, A. Moschitti, G. Riccardi  
**Annotating Spoken Dialogs: from Speech Segments to Dialog Acts and Frame Semantics**  
EACL Workshop on Semantic Representation of Spoken Language, Athens, Greece, 2009.

A. Bisazza, Marco Dinarelli, S. Quarteroni, S. Tonelli, A. Moschitti, G. Riccardi  
**Semantic Annotation for Conversational Speech**  
IEEE Workshop on Spoken Language Technology (SLT), Goa, India, 2008.

Marco Dinarelli, A. Moschitti, G. Riccardi  
**Joint Generative and Discriminative Models for Spoken Language Understanding**  
IEEE Workshop on Spoken Language Technology (SLT), Goa, India, 2008.

## National Conferences

P. Zweigenbaum, C. Grouin, Marco Dinarelli, G. Wisniewski, S. Rosset  
**Résolution des coréférences dans des comptes rendus cliniques. Une expérimentation issue du défi i2b2/VA 2011**  
In proceedings of Reconnaissance des Formes et l'Intelligence Artificielle (RFIA), 2012.

## National Workshops

A. Garcia-Fernandez, A.L. Ligozat, Marco Dinarelli, D. Bernhard  
**Méthodes pour l'archéologie linguistique : datation par combinaison d'indices temporels**  
Workshop DEfi Fouille de Text (DEFT) associé à la conférence francophone TALN, Montpellier, France, 2011.

## Ph.D. Thesis

Marco Dinarelli  
**Spoken Language Understanding: from Spoken Utterances to Semantic Structures**  
Ph.D. Dissertation, University of Trento  
Department of Computer Science and Information Engineering (DISI), Italy, 2010.  
Available at: <http://eprints-phd.biblio.unitn.it/280/>

## Contributed Talks

Spoken Language Understanding: from User Utterances to Semantic Structures  
Talk for the group ILES at LIMS  
LIMS-CNRS, Paris, November 2010

Joint Generative and Discriminative Models  
Talk for visiting Dr. Allen Gorin, director of the Human Language Technology Research, Department of Defense, USA.  
Department of Computer Science and Information Engineering (DISI)  
University of Trento, December 2009.

LUNA project review meeting  
Brussels, Belgium, October 2009.

Reranking Models for Spoken Language Understanding  
Talk for visiting Prof. Nicolò Cesa Bianchi, Full Professor at "Politecnico di Milano"  
Department of Computer Science and Information Engineering (DISI)  
University of Trento, June 2009.

LUNA project review meeting  
Avignon, France, October 2008.

## Contributed Conference Posters&Demos

IEEE International Conference on Acoustics, Speech and Signal Processing  
(ICASSP'09), Dallas, U.S.A., March 2010.

IEEE Workshop of Automatic Speech Recognition and Understanding  
(ASRU'09), Merano, Italy, December 2009.

International Conference of the International Speech Communication Association  
(INTERSPEECH'09), Brighton, UK, September 2009.

Empirical Methods for Natural Language Processing  
(EMNLP'09), Singapore, August 2009.

IEEE Workshop of Spoken Language Technologies  
(SLT'08), Goa, India, December 2008.

## Personal skills and competences

Mother tongue **Italian**

Other languages

Self-assessment

*European level*

**English**

**French**

### Understanding

Listening

Reading

### Speaking

Spoken interaction

Spoken production

### Writing

B2 Independent user C1 Proficient user B2 Independent user B2 Independent user B2 Independent user

B1 Independent user B2 Independent user B1 Independent user B1 Independent user A2 Basic User

(\*) [Common European Framework of Reference \(CEF\) level](#)

Organisational skills and competences

During my Ph.D. and my pos-doc, I was working in European funded projects where I have been responsible for writing and reviewing several shared documents. Many deadlines must be respected: Work Package Deliverables, Quarterly and Annual Reports etc.  
I have attended most of the plenary meetings of the projects giving presentations of my research activity progress.

I have worked and published papers with several persons of my research group but also located in different countries. I had the opportunity to share data and software and so to adapt my work to other needs.

Computer skills and competences	Programming Languages: C, C++, Java, C#, Perl, PHP, Shell scripting (Linux), HTML, XML, Latex Databases: MySQL, PostgreSQL Software Libraries/Technologies: STL, OpenGL, QT, Zlib, BerkeleyDB, JSP, Servlet, JDBC, RMI, RPC, CORBA, IPC System V, ASSIST, STXXL Tools: Apache and Tomcat, Matlab, Octave, HTK, Microsoft Office Programming Environment: Microsoft Visual Studio, Borland C++ Builder, Dev-C++, Emacs
Other skills and competences	Martial Arts, Skating, Swimming, Music, Cinema, Comics, Science Fiction Books
Driving licence	B

**Additional information** My Ph.D. at University of Trento has been funded by the following european projects:  
“LUNA: Spoken Language Understanding in Multi-Lingual Communication Systems”  
<http://www.ist-luna.eu/>  
“LiveMemories: Active Digital Memories of Collective Life”  
<http://www.livememories.org/Home.aspx>  
(Only the last 5 months)

### **For Italy Only**

Le informazioni contenute nel presente documento vengono rese ai sensi e per gli effetti degli artt. 46 e 47 del DPR 455-200 e corrispondono a verità  
Dicembre 2011