

# Curriculum Vitae

---

**Alfio Massimiliano Gliozzo <gliozzo@itc.it>**

(+39)0461314573  
via Sommarive, 18

<http://tcc.itc.it/people/gliozzo/>  
Povo-Trento, Italy, 38059

---

## SHORT BIOGRAPHY

My research field is Artificial Intelligence, focusing on corpus based approaches for knowledge acquisition and retrieval. I got my Ph.D. in Information and Communication Technologies at the University of Trento in December 2005. I'm author of more than thirty scientific publications, published in top level conference proceedings and international journals. Currently, I'm working as a research consultant for both the Cognitive and Communication Technologies (TCC) division at FBK-irst, Trento, and the Laboratory for Applied Ontology (LOA) at the Institute for Cognitive Sciences and Technology of the National Research Council (ISTC-CNR), Rome. My research work includes the participation at many UE founded projects, program committee membership at international conferences, students supervision, teaching in advanced courses and the development of state of the art NLP tools, some of them reporting the best performances in the context of international evaluation challenges. Apart from my academic activities, I'm also collaborating with some ITC company as a consultant for technology transfer.

---

## RESEARCH INTERESTS

**Computational Linguistics** Lexical Semantics, Word Sense Disambiguation, Distributional Similarity, Multilinguality, Textual Entailment.

**Knowledge Engineering** Knowledge Representation, Conceptual Modelling, Ontology Learning from Texts, Semantic WEB.

**Information Retrieval** Multilingual Text Categorization, Topic Detection, Semi supervised text categorization.

**Machine Learning** kernel methods, unsupervised learning, bootstrapping, Support Vector Machines, one class learning.

---

## EDUCATION

**Ph.D. in Information and Communication Technology, University of Trento, Italy** 2002-2005

**"Laurea" degree summa cum laude in Philosophy (logic and epistemology) at the University of Bologna** 1996-2002

**Qualified Programmer's Diploma I.T.I.S. "G. Marconi" Catania** 1990-1995

---

## RESEARCH PROJECTS

**Research Consultant for the FIRB-israel project** 2007-

- Institution: Fondazione Bruno Kessler (FBK) Istituto per la Ricerca Scientifica e Tecnologica (IRST), Trento
- Duties: research on relationship between textual Entailment and Ontology Learning

**Research Consultant for the NEON UE project** 2007-

- Institution: Laboratory for Applied Ontology (LOA) at the Institute of Cognitive Sciences and Technologies (ISTC) of the Italian National Research Council (CNR), Rome
- Duties: (i) coordination of task 7.3 - Ontology learning technology transfer at the Food and Agriculture Organization (FAO) of the United Nations, (ii) research and coordination of technology transfer of Semantic WEB technologies to enhance the Organization Workflow at CNR.

**Research Consultant for the ONTOTEXT project** 2005-2006

- Institution: Istituto Trentino di Cultura (ITC) - Istituto per la Ricerca Scientifica e Tecnologica (IRST), Trento
- Duties: (i) Development of Topic Detection systems (ii) Research on Ontology learning from texts

- Research Consultant for the Presto Space UE project** 2005-2006
- Institution: University of Rome Tor Vergata
  - Duties: (i) research on Cross Language Information Retrieval (ii) development of a module for Latent Semantic Analysis
- Research Consultant for the project MEANING UE project** 2001-2004
- Institution: Istituto Trentino di Cultura (ITC) - Istituto per la Ricerca Scientifica e Tecnologica (IRST), Trento
  - research and development of Word Sense Disambiguation systems

## **INDUSTRIAL EMPLOYMENT**

---

- Consultant for P.M.F. S.r.l.** 2007 -
- Catania, Italy.
  - Responsibilities: coordination of ITC technology transfer to the e-Learning area.
- Consultant for Viteco S.r.l.** 2007 -
- Catania, Italy.
  - Responsibilities: writing project proposals
- Consultant for Shopping.com** 2006
- an eBay limited company, London, U.K.
  - Responsibilities: research on optimizing text categorization systems for commercial advertising.
- Consultant for BC Software S.r.l.** 2006
- Catania, Italy.
  - Responsibilities: Design of Natural Language Interfaces to databases.

## **STAGES**

---

- Visiting Researcher at CNTS** *March 2005 - September 2005*
- Institution: CNTS, Language Technology Group, Department of Linguistics of the University of Antwerp
  - Supervisor: Walter Daelemans
  - Duties: Research on Ontology Learning from Texts
- Visiting Researcher at UPC** *September 2002*
- Institution: TALP - Research Center for Technologies and Applications in Language and Speech at the Universitat Politècnica de Catalunya, Barcelona, Spain
  - Supervisor: Lluís Marqués
  - Duties: Research on Machine Learning methods for Word Sense Disambiguation
- Visiting Researcher at the University of the Basque Country** *October 2002*
- Institution: IXA -Natural Language Processing group at the University of the Basque Country, Computer Science Faculty, San Sebastian, Spain
  - Supervisor: German Rigau
  - Duties: Research on Semantic Domains and Topic Signatures
- Master Thesis stage at ITC-irst** *2000 - 2001*
- Institution: Istituto Trentino di Cultura (ITC) - Istituto per la Ricerca Scientifica e Tecnologica (IRST), Trento
  - Supervisor: Bernardo Magnini
  - Duties: Research on Semantic Domains and Word Sense Disambiguation

## **ORGANIZATION OF INTERNATIONAL CONFERENCES**

---

- Reviewer for the the 10th Congress of the Italian Association for Artificial Intelligence, 2007.
- Program committee member the 45th Annual Meeting of the Association for Computational Linguistics, ACL-07, Prague, Czech Republic, Main conference, Semantics track

- Program committee member of the ACL 2007 Student Research Workshop
- Scientific committee member of LREC 2006: International Conference on Language, Resources and Evaluation

## **TEACHING**

---

- Semantic Domains in Natural Language Processing. Advanced course at ESSLLI 2006. 31 July, 11 August, Malaga, Spain.
- An Introduction to Kernel methods for Natural Language Processing. Tutorial, University of Antwerp, October 2005.

## **EVALUATION CHALLENGES**

---

- Participated at the English Lexical Sample task at the 4th International Workshop on Semantic Evaluation (SemEval-2007): best performing system on 13 participants.
- Participated at the English Lexical Substitution task at the 4th International Workshop on Semantic Evaluation (SemEval-2007): best performing system on 9 participants.
- Participated at the Pascal Challenge on the Evaluation of Machine Learning for Information Extraction achieving competitive results.
- Participated at the SENSEVAL-3 WSD competition developing the best performing systems for the following tasks: English lexical sample (coarse-grained evaluation), Italian lexical sample, Spanish lexical sample and Catalan lexical sample, 2004
- Participated at the SENSEVAL-2 WSD competition, developing the IRST Systems for the following tasks: English all words, English Lexical Sample, Italian lexical sample, 2001

## **INVITED TALKS**

---

- October 6, 2006: Invited talk at the Dipartimento di Ingegneria Informatica, University of Palermo, Italy. Talk Title "Domain Driven Technologies for Natural Language Processing"
- July 28, 2006: Invited talk at Yahoo! Research, Barcelona, Spain. Talk Title "Domain Driven Technologies for Natural Language Processing"
- July 20, 2005: Invited talk at the Institut für Informatik of the University of Leipzig, Germany. Talk title: "Domain Models for Lexical Semantics and their applications in Natural Language Processing"
- March 18, 2005: Talk at the Department of Linguistics of the University of Antwerp, Belgium. Talk title: "Domain Models for Lexical Semantics and their applications in Natural Language Processing"; Speaker at the International Colloquium on "Word structure and lexical systems: models and applications", December 16-17, 2004, Pavia, Italy. "Domain models for lexical semantics"
- December 2, 2004: Invited talk at the Institute for Knowledge and Language Engineering of the University of Magdeburg, Germany. "Kernel Methods for Word Sense Disambiguation"
- October 6, 2004: Invited talk at the Department of Computer Science, Systems, and Production of the University of Rome Tor Vergata, Artificial Intelligence Research Group, Rome, Italy. Talk title: "Kernel Methods for Word Sense Disambiguation"
- October 5, 2004: Invited talk at the Institute of Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy. Talk title: "Domain Models for Lexical Semantics and their applications in Natural Language Processing"

## **HONORS AND AWARDS**

---

- Honored with biography selected for inclusion in the 25th Silver Anniversary Edition of Who's Who in the World.

## **COMPUTER SCIENCE BACKGROUND**

---

- Programming languages: Java, Lisp, perl, C++
- Experience in developing robust systems for Natural Language Processing, Information Retrieval and Machine Learning
- Webdesign

## **SPOKEN LANGUAGES**

---

- Italian (native language)
- English (fluent)

## **LIST OF PUBLICATIONS**

---

### **Books:**

- Alfio Gliozzo, Francesco Bianchini, Maurizio Matteuzzi(Eds.)." Instrumentum Vocale: intelligenza artificiale e linguaggio", Bononia University Press, to be published.
- Alfio Gliozzo, Carlo Strapparava."Semantic Domains in Natural Language Processing", Springer, to be published.

### **Journal Articles:**

- Roberto Basili, Alfio Massimiliano Gliozzo, Paolo Marocco. "Teorie geometriche del significato", Res Cogitans, ISSN 1828-7778, 2006.
- D. Pérez, A. Gliozzo, E. Alfonseca, C. Strapparava, B. Magnini, and P. Rodriguez. Sobre los efectos de combinar análisis de semántica oculta con otras técnicas de procesamiento de lenguaje natural para la evaluación de preguntas abiertas, Revista Signos, 38(59), November 2005.
- Alfio Massimiliano Gliozzo, Claudio Giuliano and Raffaella Rinaldi. "Instance Filtering for Entity Recognition", in ACM SIGKDD Explorations, special Issue on Natural Language Processing and Text Mining, 7(1), pp. 11-18, June 2005.
- A. Gliozzo, C. Strapparava, I. Dagan. "Unsupervised and Supervised Exploitation of Semantic Domains in Lexical Disambiguation", in Computer Speech and Language,18(3), pp. 275-299 , 2004.
- B. Magnini, C. Strapparava, G. Pezzulo, A. Gliozzo. "The Role of Domain Information in Word Sense Disambiguation", Natural Language Engineering, special issue on Word Sense Disambiguation, 8(4), pp. 359-373, Cambridge University Press, 2002.

### **Papers at Refereed Conferences and Workshops:**

- Alfio Gliozzo, Aldo Gangemi, Valentina Presutti, Elena Cardillo, Enrico Daga, Alberto Salvati and Gianluca Troiani. "A Semantic Web Layer to Enhance Legacy Systems, to appear in Proceedings of 6th International Semantic Web Conference, Busan, Korea, 2007
- Chris Biemann, Claudio Giuliano and Alfio Gliozzo. "Unsupervised Part-Of-Speech Tagging Supporting Supervised Methods". Proceedings of the International Conference Recent Advantages in Natural Language Processing (RANLP), September 2007, Bulgaria.
- Roberto Basili, Alfio Gliozzo and Marco Pennacchiotti. "Harvesting Ontologies from Open Domain Corpora: a Dynamic Approach". Proceedings of the International Conference Recent Advantages in Natural Language Processing (RANLP), September 2007, Bulgaria.
- Claudio Giuliano, Alfio Gliozzo and Carlo Strapparava. "Lexical Substitution Task Exploiting Domain and Syntagmatic coherence. Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), Prague, 28-30 June 2007.
- Claudio Giuliano and Alfio Gliozzo. "Instance Based Lexical Entailment for Ontology Population". Proceeding of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), pp. 248–256, Prague, 2007.
- Davide Picca, Alfio Massimiliano Gliozzo and Massimiliano Ciaramita. 2007. "Semantic Domains and Supersense Tagging for Domain-Specific Ontology Learning", in Proceedings of RIAO'2007, Pittsburgh.
- Alfio Massimiliano Gliozzo, Marco Pennacchiotti and Patrick Pantel. 2007. "The Domain Restriction Hypothesis: Relating Term Similarity and Semantic Consistency". In Proceedings of North American Association for Computational Linguistics / Human Language Technology (NAACL HLT 07). pp. 131-138. Rochester, NY.
- Roberto Basili, Marco Cammisa and Alfio Gliozzo. "Integrating Domain and Paradigmatic Similarity for Unsupervised Sense Tagging." in Proceedings of ECAI 2006, Riva del Garda, Italy
- Ido Dagan, Oren Glickman, Alfio Gliozzo, Efrat Marmorshtein and Carlo Strapparava. "Direct Word Sense Matching for Lexical Substitution". Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL'06), Sydney, Australia, 2006.
- Alfio Massimiliano Gliozzo and Carlo Strapparava. "Exploiting Comparable Corpora and Bilingual Dictionaries for Cross-Language Text Categorization". Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL'06), Sydney, Australia, 2006.
- Alfio Massimiliano Gliozzo. "The GOD model", EACL-06 poster session, Trento, Italy.
- Alfio Massimiliano Gliozzo. "Semantic Domains and Linguistic Theory" in Proceedings of the LREC 2006 workshop "Toward Computational Models of Literary Analysis", Genova, Italy, 2006.
- Claudio Giuliano, Alfio Gliozzo and Carlo Strapparava. "Syntagmatic Kernels: a Word Sense Disambiguation case study", in proceedings of the EACL-06 workshop on Learning Structured Information in Natural Language Applications. Trento, Italy.
- Alfio Gliozzo, Carlo Strapparava and Ido Dagan. "Investigating Unsupervised Learning for Text Categorization Bootstrapping". Proceedings of the joint Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP), 2005, Vancouver, B.C., Canada

- Alfio Massimiliano Gliozzo, Claudio Giuliano and Carlo Strapparava. "Domain Kernels for Word Sense Disambiguation". Proceedings of the 43th Annual Meeting of the Association for Computational Linguistics (ACL'05), pages 403-410, Ann Arbor, Michigan, June, 2005.
- Alfio Massimiliano Gliozzo and Carlo Strapparava. "Domain Kernels for Text Categorization". Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005), pages 56-63, Ann Arbor, Michigan, June, 2005.
- Alfio Massimiliano Gliozzo and Carlo Strapparava. "Cross Language Text Categorization by acquiring Multilingual Domain Models from Comparable Corpora". Proceedings of the ACL Workshop on Building and Using Parallel Texts, pages 9-16, Ann Arbor, Michigan, June, 2005.
- D. Perez, A. Gliozzo, C. Strapparava, Enrique Alfonseca, Pilar Rodriguez and Bernardo Magnini. "Automatic Assessment of Students' free-text Answers underpinned by the Combination of a Bleu-inspired algorithm and LSA". Proceedings of FLAIRS-2005, Clearwater Beach, Florida, May 16-18, 2005.
- Claudio Giuliano, Alfio Massimiliano Gliozzo, Alberto Lavelli, Lorenza Romano, "Filtering Uninformative Words to Speed up IE: ITC-irst Participation in the PASCAL Challenge". Proceedings of the PASCAL Challenges Workshop, Southampton, U.K., pages 11-13, April 2005.
- Ernesto D'Avanzo, Alfio Massimiliano Gliozzo, Carlo Strapparava "Automatic Acquisition of Domain Information for Lexical Concepts". Proceedings of the MEANING-2005 workshop, Trento, Italy, 2005.
- Alfio Gliozzo, Bernardo Magnini and Carlo Strapparava, "Unsupervised Domain Relevance Estimation for Word Sense Disambiguation", in Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-04), Barcelona, Spain, 25-26 July 2004.
- Carlo Strapparava, Claudio Giuliano, Alfio Gliozzo. "Pattern Abstraction and Term Similarity for Word Sense Disambiguation: IRST at Senseval-3". Proceedings of the Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (Senseval-3), Barcelona, Spain, 25-26 July 2004.
- Bernardo Magnini, Carlo Strapparava, Giovanni Pezzulo and Alfio Gliozzo. "Using Domain Information for Word Sense Disambiguation". Proceedings of SENSEVAL-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems, Toulouse(France), 2001.
- Alfio Massimiliano Gliozzo, Claudio Giuliano and Raffaella Rinaldi. "Instance Pruning by filtering Uninformative Words: an Information Extraction Case Study", Proceedings of CICLING 2005, Mexico, 2005.
- Alfio Massimiliano Gliozzo, Carlo Strapparava and Marcello Ranieri. "Crossing Parallel Corpora and Multilingual Lexical Databases for WSD", in Proceedings of CICLING 2005, Mexico, 2005
- Bernardo Magnini, Carlo Strapparava, Giovanni Pezzulo, Alfio Gliozzo. "Comparing ontology-Based and Corpus-based Domain Annotations in Wordnet". Proceedings of the First Global WordNet Conference, Mysore, India, 2002.
- Bernardo Magnini, Carlo Strapparava, Giovanni Pezzulo and Alfio Gliozzo. "Using Domain Information for Word Sense Disambiguation". in Proceedings of SENSEVAL-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems, Toulouse(France), 2001.

#### **Miscellaneous:**

- Alfio Massimiliano Gliozzo. "Semantic Domains in Computational Linguistics", Ph.D. Thesis, University of Trento, 2005
- Alfio Massimiliano Gliozzo. "Ruolo dei campi semantici nella struttura di un lessico computazionale: utilizzo per la disambiguazione automatica di senso", Master Thesis, University of Bologna, March, 2001.