

EUROPEAN  
CURRICULUM VITAE  
FORMAT



PERSONAL INFORMATION

Name **MARCO TURCHI**

Address **ITALY: VIA SAN GIMIGNANO 46, 53036 POGGIBONSI (SI)**  
**ENGLAND: 6 THE CLOISTERS 22 COLLEGE ROAD, BS8 3HZ, BRISTOL**

Telephone **Italian: +39 3382809829**  
**English: +44 7804458481**

E-mail **marco.turchi@gmail.com**  
**marco.turchi@bris.ac.uk**

Personal Website **www.marcoturchi.com**

Nationality **Italian**

Date of birth **21/07/1975**

WORK EXPERIENCE

- Dates (from – to) *January 2007 – July 2009*
- Name and address of employer Department of Engineering Mathematics, University of Bristol
- Type of business or sector Statistical Multilingual Analysis for Retrieval and Translation (EU IST Project SMART (FP6-033917)), Web Mining, Machine Learning
- Occupation or position held Research Assistant Advisor: Prof. Nello Cristianini.
- Main activities and responsibilities *Researching different aspects of Statistical Machine Translation (SMT) using Machine Learning approaches. Includes a statistical and computational analysis of a SMT system, pointing out its learning capability under different conditions and performance limitations, and Confidence Estimation analysis of a SMT system.*
  
- Dates (from – to) *September 2006 – January 2007*
- Name and address of employer Department of Engineering Mathematics, University of Bristol
- Type of business or sector Machine Learning, Natural Language Processing and Web Mining.
- Occupation or position held Visiting Fellow Advisor: Prof. Nello Cristianini.
- Main activities and responsibilities Extraction and analysis of multimedia content from the Web focused on pattern analysis and named entity recognition from multilanguage news contents
  
- Dates (from – to) *March 2006 – September 06*
- Name and address of employer Xerox Research Centre Europe, Grenoble France
- Type of business or sector Automatic Hyperlink Assistant for Wikipedia Writer
- Occupation or position held Research Internship Advisors: Dr. Eric Gaussier and Dr. Cyril Goutte
- Main activities and responsibilities Analyzing how to automatically hyperlink “interesting” word or set of words in a document with external web pages that could help the reader to understand the content of the document. Includes using of probabilistic and string methods.
  
- Dates (from – to) *September 2005 – March 2006*
- Name and address of employer Statistics Department, University of California, Davis
- Type of business or sector Machine Learning, Natural Language Processing and Web Mining.

- Occupation or position held
- Main activities and responsibilities

Visiting Fellow Advisor: Prof. Nello Cristianini

Researching and studying the introduction of the variable “time” in document representation. This has produced two different analysis:

- given a sequence of written documents in different languages, find temporal relations among languages (language evolution);
- given a sequence of documents in time, find patterns which merge textual information and temporal relations.

- Dates (from – to)

*June 2005 - September 2005*

- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities

Yahoo Research Lab, Pasadena California U.S.

Feature Selection using unlabelled documents

Research Internship Advisor: Dr. Dennis De Coste

Researching alternative feature selection methods for unlabelled data using SVM, LDA, PCA and clustering algorithms.

- Dates (from – to)

*December 1995 - current*

- Main activities and responsibilities

Several experiences as:

- Lecturer and Network Administrator of Windows 2000
- Programmer and System Administrator
- Student Co-advisor for Master and Degree thesis on Text Analysis
- Reviewer of international conferences and journals

## EDUCATION AND TRAINING

- Dates (from – to)

October 2002 - March 2006

- Name and type of organization providing education and training
- Principal subjects/occupational skills covered
- Title of qualification awarded

Department of Information Engineering, University of Siena

Doctor in “Computer Engineering, adaptive information processing”

Title of PhD Thesis: “Time Series Analysis of Textual Data” Supervisor: Prof. M. Maggini  
PhD

- Dates (from – to)

2001/2002

- Name and type of organization providing education and training
- Principal subjects/occupational skills covered

Department of Information Engineering, University of Siena

Master degree in Information Engineering, University of Siena.

Title of the Dissertation: “Detection of Web page components: an approach based on visual layout” Supervisor: Prof. Ing. Marco Maggini, Ing. Michelangelo Diligenti

- Title of qualification awarded
- Level in national classification (if appropriate)

MS

Marks: 102/110

## PERSONAL SKILLS AND COMPETENCES

## TECHNICAL SKILLS AND COMPETENCES

- EXCELLENT IN JAVA, PYTHON, MATLAB, MYSQL AND LINUX.
- BASIC KNOWLEDGE OF ANSI C, C++, HTML, PERL.
- EXCELLENT USER’S EXPERIENCE OF JAVA SEARCH ENGINE: APACHE LUCENE, GENERAL ARCHITECTURE FOR TEXT ENGINEERING: GATE, WEKA, APACHE CAYENNE, STATISTICAL MACHINE TRANSLATION TOOLKIT MOSES, LINGPIPE, SCIPY.
- EXCELLENT EXPERIENCE OF HIGH PERFORMANCE COMPUTING (HPC), LARGE-SCALE DATASET

<b>Italian</b>	Mother Tongue
<b>English</b>	Proficient
<b>French</b>	Basic

**ORGANIZATIONAL SKILLS  
AND COMPETENCES**

**ORGANIZER OF:**

1. INTELLIGENT ANALYSIS AND PROCESSING OF WEB NEWS CONTENT WORKSHOP AT WI-IAT - MILAN 15 SEPTEMBER 2009
2. STATISTICAL MULTILINGUAL ANALYSIS FOR RETRIEVAL AND TRANSLATION ASSOCIATED WORKSHOP AT EAMT - BARCELONA 13 MAY 2009
3. EUROPEAN PROJECT SMART MEETING IN BRISTOL, MAY, 2008.

COORDINATOR AND HEAD COACH OF BASKETBALL TEAMS FROM SEPTEMBER 1993.

STUDENT CO-ADVISOR FOR MASTER AND DEGREE THESIS ON TEXT ANALYSIS.

**ORAL PRESENTATION**

SEVERAL PLATFORM PRESENTATIONS TO INTERNATIONAL CONFERENCES AND WORKSHOPS.

**SELECTED CONFERENCES:**

- THIRD WORKSHOP ON STATISTICAL MACHINE TRANSLATION, ACL'08
- IEEE/WIC/ACM INTERNATIONAL JOINT CONFERENCE ON WEB INTELLIGENCE WI'04

**ADDITIONAL  
INFORMATION**

**PUBLICATIONS**

**JOURNAL PAPERS:**

- M. BRESCO, M. TURCHI, T. DE BIE, N. CRISTIANINI "MODELLING SEQUENCE EVOLUTION WITH KERNEL", COMPUTATIONAL OPTIMIZATION AND APPLICATIONS JOURNAL, 2007.

**REVIEWED BOOK CHAPTERS:**

- M. TURCHI, A. MAMMONE, N. CRISTIANINI "ANALYSIS OF TEXT PATTERNS USING KERNEL METHODS". TEXT MINING: THEORY AND APPLICATIONS. EDITORS: ASHOK N. SRIVASTAVA AND MEHRAN SAHAMI IN PRESS: TAYLOR AND FRANCIS PUBLISHER, 2008.
- A. MAMMONE, M. TURCHI, N. CRISTIANINI "SUPPORT VECTOR MACHINES". WIRES: WILEY'S INTERDISCIPLINARY REVIEWS IN COMPUTATIONAL STATISTICS. EDITORS: EDWARD J. WEGMAN, YASMIN H. SAID AND DAVID W. SCOTT LAUNCHING ONLINE: JUNE 2009. TO APPEAR.

**PEER-REVIEWED CONFERENCE PAPERS:**

- L.SPECIA \*, M. TURCHI \*, N. CANCEDDA, M. DYMETMAN AND N. CRISTIANINI "ESTIMATING THE SENTENCE-LEVEL QUALITY OF MACHINE TRANSLATION SYSTEMS", EAMT 2009, BARCELONA, SPAIN, MAY 2009 (\* EQUAL CONTRIBUTORS.)
- M. TURCHI, T. DE BIE, N. CRISTIANINI "LEARNING PERFORMANCE OF A MACHINE TRANSLATION SYSTEM: A STATISTICAL AND COMPUTATIONAL ANALYSIS", IN PROCEEDINGS OF THE ACL 2008 THIRD WORKSHOP ON STATISTICAL MACHINE TRANSLATION, JUNE 2008.
- M. TURCHI, N. CRISTIANINI "A STATISTICAL ANALYSIS OF LANGUAGE EVOLUTION" IN PROCEEDING OF EVOLUTION OF LANGUAGE SIXTH INTERNATIONAL CONFERENCE ROME, 12-15 APRIL 2006.
- M. TURCHI, L. RIGUTINI, M. MAGGINI "PSEUDO-SUPERVISED CLUSTERING FOR TEXT DOCUMENTS", IN PROCEEDING OF IEEE/WIC/ACM INTERNATIONAL JOINT CONFERENCE ON WEB INTELLIGENCE, WI'04, BEIJING, SEPTEMBER 2004

**PROJECTS**

SMART, Pascal, Pascal 2, Pascal Pump Priming Project (Semantic Information Extraction using Integrated Probabilistic Techniques).

## SELECTED COURSES

- SUMMER SCHOOL ON PROBABILISTIC TECHNIQUES IN COMPUTER SCIENCE, BRISTOL, JULY 6 - 11 2008.
- THE 2008 UK SYMPOSIUM ON KNOWLEDGE DISCOVERY AND DATA MINING , BRISTOL UK, 16 APRIL 2008.
- THE ANALYSIS OF PATTERNS, UNIVERSITY RESIDENTIAL CENTER BERTINORO (FORLÌ-CESENA), ITALY, OCTOBER, 2007.
- MLSS, MACHINE LEARNING SUMMER SCHOOL, MAX PLANCK CAMPUS, TÜBINGEN, GERMANY, AUGUST , 2007.
- FIRST MACHINE TRANSLATION MARATHON, EDINBURGH UNIVERSITY, EDINBURGH, SCOTLAND, APRIL, 2007.
- THE ANALYSIS OF PATTERNS INTERNATIONAL SCHOOL "EDUARDO R. CAIANIELLO", ERICE, ITALY, 2005.

## COLLABORATION

Xerox Research Centre Europe, Grenoble, France  
National Research Council Canada, Gatineau, QC, Canada

More information about me and my research are available on my website.

References: available upon request.