

## Short CV with selected publications

**Paolo Rosso**, born 7th July 1964 in Genova, Italy.  
Italian citizenship. Spanish Permanent Residence Permit.

Contact work address:  
Dept. Sistemas Informáticos y Computación  
Universidad Politécnica de Valencia  
Camino de Vera, s/n  
46022 Valencia, Spain

Tel. +34-963877007 ext. 73571  
Email: [proso@dsic.upv.es](mailto:proso@dsic.upv.es)  
Website: <http://www.dsic.upv.es/~proso/>

## Position

**November 2008 – present:** Associate Professor, Technical University of Valencia, Spain

**July 1999 – October 2008:** Assistant Professor, Technical University of Valencia, Spain

## Education

PhD in Computer Science: Trinity College University of Dublin, Ireland, May 1999.  
Title thesis: Decision Support in Speech and Language Therapy: Managing Phonologically Disordered Speech in Childhood  
Advisor: Prof. Jane Grimson

## Professional Activities

Deputy Steering Committee Chair for the Conference at the CLEF Initiative (Conference and Labs of the Evaluation Forum).

## Editorial Services to Scholarly Publications

**Editorial Board:** Information Processing and Management, Computer Applications in Engineering Education, Research and Development in Computer Science and Engineering with Applications, Computación y Sistemas.

**Guest editor:** Information Processing and Management (2016: special issue on Emotion and Sentiment in Social and Expressive Media); ACM Transactions on Internet Technology (2016: special issue on Affect and Interaction in Agent-based Systems and Social Media); Language Resources and Evaluation Journal (2013: special issue on Analysis of Short Texts on the Web); ACM Transactions on Intelligent Systems and Technology (2012: special issue on Search and Mining User Generated Contents).

**Reviewer in journals:** Information Processing and Management (27 completed assignments; top reviewer in 2015); Computational Linguistics, Information Sciences; Transactions on Knowledge and Data Engineering; Artificial Intelligence; Knowledge and Information Systems; Language Resources and Evaluation; IEEE Transactions on

Systems, Man, and Cybernetics; ACM Transactions on Intelligent Systems and Technology; Journal of Intelligent & Fuzzy Systems, etc.

PC member / reviewer in **conferences**: Association for Computational Linguistics (ACL), European Chapter of the Association for Computational Linguistics (EACL), Association for Computational Linguistics: Human Language Technologies (ACL-HLT); Conference on Computational Linguistics (COLING), Empirical Methods on Natural Language Processing (EMNLP), World Wide Web (WWW), Research and Development in Information Retrieval (SIGIR), Conference on Information and Knowledge Management (CIKM), Association for the Advancement of Artificial Intelligence (AAAI), European Conference on Information Retrieval (ECIR), Recent Advances in Natural Language Processing (RANLP), Computational Linguistics and Natural Language Processing (CICLing), Language Resource and Evaluation Conference (LREC), etc.

### Organising Committees

**EACL 2017** (European Chapter of the Association for Computational Linguistics): Chair of local organizing team: <http://eacl2017.org/>

**CLEF 2013** (Conference and Labs of the Evaluation Forum): General chair <http://clef2013.clef-initiative.eu/index.php?page=Pages/organizers.html>

**CERI 2012** (Spanish Conference on Information Retrieval): Chair of the organizing committee: [http://users.dsic.upv.es/grupos/nle/ceri/en/comite\\_organizador.html](http://users.dsic.upv.es/grupos/nle/ceri/en/comite_organizador.html)

Chair at PAN (2009-2016) and organiser of digital text forensics shared tasks at CLEF and FIRE on plagiarism detection and author profiling: <http://pan.webis.de/>

Organiser of workshops at several conferences: LREC (ES<sup>3</sup>-2012 ES<sup>3</sup>LOD-2014), CIKM (SMUC-2010, 2011), ECML-PKDD (DyNaK-2010), etc.

### Research Interests

Natural Language Processing and Information Retrieval (plagiarism detection, author profiling, irony and sarcasm detection in social media, etc.)

Member of the Pattern Recognition and Human Language Technology (PRHLT) research center: <https://www.prhlt.upv.es/>

### Awarded Research Projects

The following list includes projects for which I acted at the Technical University of Valencia as Principal Investigator during the last five years:

#### **SOCOCODE : Social Copying Community Detection**

Army Research Office (project n. W911NF-14-1-0254)

Duration: 36 months (from September 2014)

Research center: PRHLT, Technical University of Valencia

#### **SomEMBED: SOcial Media language understanding-EMBEDding contexts**

Spanish Ministry of Economy and Competitiveness (project n. TIN2009-13391-C04)

Duration: 36 months (from January 2016)

Research centers: Technical University of Valencia, University of Barcelona

### **WIQ-EI: Web Information Quality Evaluation Initiative**

IRSES project (grant n. 269180) within the FP 7 Marie Curie People Framework

Duration: for 48 months (from June 2011)

Research centers: Know-Center Graz (Austria), Technical University of Valencia (Spain), Bauhaus-Universität Weimar (Germany), University of the Aegean (Greece), National University of San Luis (Argentina), INAOE (Mexico), Polytechnic National Institute (Mexico), IIIT Hyderabad (India), AU-KBC Research Centre (India)

### **DIANA-APPLICATIONS Finding Hidden Knowledge in Texts: Applications**

Spanish Ministry of Science and Innovation (project n. TIN2012-38603-C02-01)

Duration: 36 months (from January 2013)

Research centers: Technical University of Valencia, University of Barcelona

### **TEXT-ENTERPRISE 2.0: Text comprehension techniques applied to the needs of the Enterprise 2.0**

Spanish Ministry of Science and Innovation (project n. TIN2009-13391-C04)

Duration: 36 months (from January 2010)

Research centers: University of Alicante, Technical University of Valencia, University of Jaén, University of Barcelona

### **Application of Natural Language Processing to the Need of the University**

Spanish Association of International Cooperation, Spanish Ministry of Foreign Affairs and Cooperation (project n. AECID-PCI AP/043848/11)

Duration: 12 months (from December 2011)

Bilateral research project with the Université Mentouri Constantine (Algeria)

### **Supervisions and services**

Advisor of 15 PhD theses: <http://www.dsic.upv.es/~proso/theses.html>

### **List of Selected Publications**

400+ papers published in journals, conferences and workshops.

H-index 36

### **Journal Papers and Chapters in Books (2011-2016)**

Gupta P., Banchs R., Rosso P. (2016) Continuous Space Models for CLIR. In: Information Processing & Management (accepted)

Franco-Salvador M., Rosso P., Montes-y-Gómez M. (2016) A Systematic Study of Knowledge Graph Analysis for Cross-language Plagiarism Detection. In: Information Processing & Management, vol. 52, issue 4, pp. 550–570

Rangel F., Rosso P. (2016) On the Impact of Emotions on Author Profiling. In: Information Processing & Management, vol. 52, issue 1, pp. 73-92

Alonso J., Paredes R., Rosso P. (2016) Feature Representation for Social Circles Detection Using MAC. In: Neural Computing and Applications. DOI:

10.1007/s00521-016-2222-y

Errecalde M., Cagnina L., Rosso P. (2016) Silhouette + Attraction: A Simple and Effective Method for Text Clustering. In: *Natural Language Engineering*, vol. 22, issue 5, pp. 687-726

Sulis E., Hernández I., Rosso P., Patti V., Ruffo G. (2016) Figurative Messages and Affect in Twitter: Differences Between #irony, #sarcasm and #not. In: *Knowledge-Based Systems*, vol. 108, pp. 132–143

Hernández I., Patti V., Rosso P. (2016) Irony Detection in Twitter: The Role of Affective Content. In: *ACM Transactions on Internet Technology*, vol. 16, issue 3, pp. 1-24

Gupta P., Banchs R., Rosso P. (2016) Squeezing Bottlenecks: Exploring the Limits of Autoencoder Semantic Representation Capabilities. In: *Neurocomputing*, vol. 175. part B, pp. 1001-1008.

Gupta P., Costa-jussà Ruiz M., Rosso P., Banchs R. (2016) A Deep Source-Context Feature for Lexical Selection in Statistical Machine Translation. In: *Pattern Recognition Letters*, vol. 75, issue C, pp.24-29

Franco-Salvador M., Gupta P., Rosso P., Banchs R. (2016) Cross-language Plagiarism Detection Over Continuous-space- and Knowledge Graph-based Representations of Language. In: *Knowledge-Based Systems*, vol. 111, pp. 87–99

Pérez F., Cardiff J., Rosso P., Pinto D. (2016) Prototype/topic based Clustering Method for Weblogs. In: *Intelligent Data Analysis*, vol. 20, no. 1, pp. 47-65

Baca-Gomez, Y., Martínez A., Rosso P., Estrada H., Hernández I. (2016) Web Service SWePT: A Hybrid Opinion Mining Approach. In: *Journal of Universal Computer Science*, vol. 22, no. 5, pp. 671-690

Hernández I., Rosso P. (2016) Irony, Sarcasm, and Sentiment Analysis. Chapter 7 In: *Sentiment Analysis in Social Networks*, F.A. Pozzi, E. Fersini, E. Messina, and B. Liu (Eds.), Elsevier Science and Technology

Rosso P., Hernández I., Rangel F. (2016) Universality and Creativity: The usage of Language in Gender and Irony. In: *Creativity and Universality in Language, Lecture Notes in Morphogenesis*, E.Altman, M.Degli Esposti, F.Pachet (Eds.), Springer-Verlag, pp. 177-186

Hernández D., Montes-y-Gómez M., Rosso P., Guzmán R. (2015) Detecting Positive and Negative Deceptive Opinions using PU-learning. In: *Information Processing & Management*, vol. 51, issue 4, pp. 433–443

Pertile S., Moreira V., Rosso P. (2015) Comparing and Combining Content- and Citation-based Approaches for Plagiarism Detection. In: *Journal of the Association for Information Science and Technology*. DOI: 10.1002/asi.23593

Flores E., Barrón-Cedeño A., Moreno L., Rosso P. (2015) Uncovering Source Code Re-Use in Large-Scale Programming Environments. In: *Computer Applications in*

Engineering and Education, vol. 23, issue 3, pp. 383–390.

Flores E., Barrón-Cedeño A., Moreno L., Rosso P. (2015) Cross-Language Source Code Re-Use Detection using Latent Semantic Analysis. In: Journal of Universal Computer Science, vol. 21, num. 13, pp. 1708-1725.

Outahajala M., Benajiba Y., Rosso P., Zenkour L. (2015) Using Confidence and Informativeness Criteria to improve POS-Tagging in Amazigh. In: Journal of Intelligent and Fuzzy Systems, vol. 28, no. 3, pp. 1319-1330.

Pérez F., Cardiff J., Rosso P., Pinto D. (2015) Disambiguating Company Names in Microblog Text Using Clustering for Online Reputation Management. In: Signos, vol. 48 num. 87, pp. 54-77

Franco-Salvador M., Cruz F., Troyano J.A., Rosso P. (2015) Cross-domain Polarity Classification using a Knowledge-enhanced Meta-classifier. In: Knowledge-Based Systems, vol. 86, pp. 46-56

Roshchina A., Cardiff J., Rosso P. (2015) TWIN: Personality-based Intelligent Recommender System. In: Journal of Intelligent and Fuzzy Systems, vol. 28, num. 5, pp. 2059-2071

Rosso P. (2015) Author Profiling and Plagiarism Detection. Chapter in: Information Retrieval. 8th RuSSIR. Braslavski, Karpov, Worring, Volkovich, Ignatov (Eds.). Series: Communications in Computer and Information Science, vol. 505, pp. 229-250

Benajiba Y., Rosso P., Abouenour L., Trigui O., Bouzoubaa K., Belguith L.H. (2014) Question Answering. Chapter 11 In: Natural Language Processing of Semitic Languages. Zitouni I. (Ed.). Series: Theory and Applications of Natural Language Processing. Springer

Reyes A., Rosso P. (2014) On the Difficulty of Automatically Detecting Irony: Beyond a Simple Case of Negation. In: Knowledge and Information Systems, vol. 40, issue 3, pp. 595-614

Cagnina L., Errecalde M., Ingaramo D., Rosso P. (2014) An efficient Particle Swarm Optimization approach to cluster short texts. In: Information Sciences, vol. 265, pp. 36-49

Danger R., Pla F., Molina A., Rosso P. (2014) Towards a Protein-Protein Interaction Information Extraction System: Recognizing Named Entities. In: Knowledge-Based Systems, vol. 57, pp. 104-118

Bogdanova D., Rosso P., Solorio T. (2014) Exploring High-Level Features for Detecting Cyberpedophilia. In: Computer Speech & Language, vol. 28, issue 1, pp. 108-120

Abouenour L., Nasri M., Bouzoubaa K., Kabbaj A., Rosso P. (2014) Construction of an Ontology for Intelligent Arabic QA systems Leveraging the Conceptual Graphs Representation. In: Journal of Intelligent and Fuzzy Systems, vol. 27, num. 6, pp. 2869-2881

Pérez F., Cardiff J., Rosso P., Pinto D. (2014) Weblog and Short Text Feature Extraction

and Impact on Categorisation. In: Journal of Intelligent and Fuzzy Systems, vol. 27, num. 5, pp. 2529-2544

Barrón-Cedeño A., Vila M., Martí M.A., Rosso P. (2013) Plagiarism meets Paraphrasing: Insights for the Next Generation in Automatic Plagiarism Detection. In: Computational Linguistics, vol. 39, issue 4, pp. 917-947

Barrón-Cedeño A., Gupta P., Rosso P. (2013) Methods for Cross-Language Plagiarism Detection. In: Knowledge-Based Systems, vol. 50, pp. 11-17

Reyes A., Rosso P., Veale T. (2013) A Multidimensional Approach for Detecting Irony in Twitter. In: Language Resources and Evaluation, vol. 47, issue 1, pp. 239-268

Sánchez F., Villatoro E., Montes-y-Gómez M., Villaseñor L., Rosso P. (2013) Determining and Characterizing the Reused Text for Plagiarism Detection. In: Expert Systems with Applications, vol. 40, issue 5, pp. 1804-1813

Panicheva P., Cardiff J., Rosso P. (2013) Identifying Subjective Statements in News Titles Using a Personal Sense Annotation Framework. In: Journal of the American Society for Information Science and Technology, vol. 47, issue 3, pp. 1411-1422

Abouenour L., Bouzoubaa K., Rosso P. (2013) On the Evaluation and Improvement of Arabic WordNet Coverage and Usability. In: Language Resources and Evaluation, vol. 47, issue 3, pp. 891-917

Reyes A., Rosso P. (2012) Making Objective Decisions from Subjective Data: Detecting Irony in Customers Reviews. In: Journal on Decision Support Systems, vol. 53, issue 4, pp. 754-760

Reyes A., Rosso P., Buscaldi D. (2012) From Humor Recognition to Irony Detection: The Figurative Language of Social Media. In: Data & Knowledge Engineering, vol. 74, pp.1-12.

Rosso P., Hurtado L., Segarra E., Sanchis E. (2012) On the Voice-activated Question Answering. In: IEEE Transactions on Systems, Man, and Cybernetics--Part C, vol. 42, num. 1, pp.75-85

Rosso P., Correa S., Buscaldi D. (2011) Passage retrieval in legal texts. In: The Journal of Logic and Algebraic Programming, vol. 80, pp. 139-153

Potthast M., Barrón-Cedeño A., Stein B., Rosso P. (2011) Cross-Language Plagiarism Detection. In: Languages Resources and Evaluation. Special Issue on Plagiarism and Authorship Analysis, vol. 45, num. 1, pp.45-62

Pinto D., Rosso P., Jiménez H. (2011) A Self-Enriching Methodology for Clustering Narrow Domain Short Texts. In: The Computer Journal, vol. 54, issue 7, pp. 1148-1165