Franco Salvador, Marc

Valencia, Valencian Community, Spain

Personal information

Nationality	Spanish
Websites	LinkedIn · Google Scholar

Vocational area

R&D in Natural Language Processing · Deep Learning · Artificial Intelligence · Information Retrieval · Generative AI · Data, Image and Speech Processing

Professional Experience

References available upon request

Nov. 2024 – present	**Head of the Artificial Intelligence and Machine Learning Unit** at the United Nations International Computing Centre (UNICC), Valencia (Spain)
Apr. 2022 – Oct. 2024	**Chief Scientific Officer (NLP & AI)** at Symanto, Valencia (Spain)
	National and international R&D projects:
	ANDHI - Anomalous Diffusion of Harmful Information. 2-year R&D project. Funded by the Ministry of Science and Innovation, Spain (MICIN CPP2021-008994).
	OBULEX - Observatory on the use of sexist language on the Internet. 15-month R&D project. Funded by the Valencian Institute of Business Competitiveness (IVACE), Spain (IVACE IMINOD/2022/106).
Aug. 2023 – Oct. 2024	**Scientific Advisor** at Genaios, Valencia (Spain)
Apr. 2022 – Oct. 2024	**Scientific Advisor** at AI Zwei, Valencia (Spain)
Sep. 2021 – Oct. 2024	**Scientific Advisor** at Coinradar.ai
Oct. 2023 – Sep. 2024	**Adjunt Professor** at Universidad Europea, Remote (Spain)
	Subjects: Artificial Intelligence (AI) projects; Project management; Intelligence and reasoning in the context of AI.

May 2022 - Jan. 2024	**Deep Learning lecturer - Big data & data science master** at Institut de Formació Contínua IL3 - Universitat de Barcelona, Remote (Spain)
Feb. 2021 – Mar. 2022	**Director of Research** at Symanto, Valencia (Spain) National and international R&D projects: ProHaters - Proactive Profiling of Hate Speech Spreaders. 3-year R&D project. Funded by the CDTI - Ministry of Science and Innovation, Spain (IDI-20210776). XAI-DisInfodemics: eXplainable AI for disinformation and conspiracy detection during infodemics. 3-year R&D project. Funded by the Ministry of Science and Innovation, Spain (PLEC2021-007681). DETEMP - Early detection of depression symptoms in social media. 15-month R&D project. Funded by the Valencian Institute of Business Competitiveness (IVACE), Spain (IMINOD/2021/72).
June. 2021 – July 2021	**Collaborator - Big data & data science master** at Institut de Formació Contínua IL3 - Universitat de Barcelona, Barcelona (Spain)
May 2018 – Jan. 2021	**Principal Research Scientist** at Symanto, Valencia (Spain)
Apr. 2017 – Apr. 2018	**Innovation Lead** at Symanto, Valencia (Spain)
Dec. 2016 – Apr. 2017	**Data scientist** at Symanto, Valencia (Spain)
Sep. 2016 – Nov. 2016	**Research fellow** at Data and Web Science Group, Universität Mannheim, Mannheim (Germany) Funded by DAAD (German Academic Exchange Service) in accordance with the DAAD scholarship program "STIBET-Doktoranden".
June 2016 – Aug. 2016	**Data scientist** at Symanto, Nuremberg (Germany)
July 2014 – May 2016	**Senior technician** at Pattern Recognition and Human Language Technology (PRHLT) Research Center - Universitat Politècnica de València, Valencia (Spain) Funded by: The Spanish Ministry of Economy and Competitiveness 3-year R&D project (TIN2015-71147-C2-1-P) SOMEMBED - SOcial Media language understanding - EMBEDing contexts. The Spanish Ministry of Economy and Competitiveness 3-year R&D project (MEC TIN2012-38603-C02-01) DIANA-APPLICATIONS - Finding Hidden Knowledge in Text.
Apr. 2014 – June 2014	**Technical specialist** at Itálica Research Group - Universidad de Sevilla, Seville (Spain) Funded by the Spanish Ministry of Economy and Competitiveness 3.5-year R&D project (TIN2011-14726-E) *Destilado de opiniones desde contenidos generados por usuarios.*
Apr. 2013 – Mar. 2014	**Research fellow** at Linguistic Computing Laboratory, Computer Science Department, Sapienza University of Rome, Rome (Italy) Funded by the 5-year ERC Starting Grant No. 259234 MultiJEDI: Multilingual Joint word sEnse DIsambiguation.
Aug. 2012 – Dec. 2012	**Senior technician** at Natural Language Engineering Lab, Department of Information Systems and Computing, Universitat Politècnica de València, Valencia (Spain) Funded by the Spanish Ministry of Education and Science 4-year R&D project (CICYT TIN2009-13391-C04-03) TEXT-ENTERPRISE 2.0: Text comprehension techniques applied to the needs of the Enterprise 2.0.

Education

2013–2017	**Doctor of Philosophy (PhD), Computer Science, Universitat Politècnica de València, Spain**
	Final score: Excellent with **international** and **cum laude** distinction
	PhD thesis: A Cross-domain and Cross-language Knowledge-based Representation of Text and its Meaning
	Advisor: Paolo Rosso
2011–2013	**Master of Science (MSc), Artificial Intelligence, Pattern Recognition and Digital Image, Universitat Politècnica de València, Spain**
	Avg. score: 8.7/10
	MSc thesis: Cross-Language Plagiarism Detection using a Multilingual Semantic Network
	Score: 10/10 - Cum laude
	Advisor: Paolo Rosso
2006–2011	**Bachelor of Science (BSc), Computer Science Engineering (5-year degree), specialization in Artificial Intelligence, Universitat Politècnica de València, Spain**
	Avg. score: 7/10
	BSc final project: Voice-based web surfing
	Score: 10/10
	Advisor: Carlos D. Martínez Hinarejos

Volunteer experience

April. 2021 – Sep. 2024	**City lead** at Valencia AI@Spain AI, Valencia (Spain)

Languages

Spanish	Mother tongue
Valencian	Mother tongue
Catalan	Mother tongue
English	**Professional working proficiency**
Italian	Basic competence

Certifications

Apr. 2025	Associate Professor Accreditation *Acreditación de Profesor Contratado Doctor (PCD)* ANECA
Oct. 2024	Google Cloud Certified: Professional Machine Learning Engineer Google
June 2024	Professional Scrum™ with Kanban I (PSK I) Scrum.org
Jan. 2024	Data Protection for Employees heyData

April 2023	Microsoft Certified: Azure AI Engineer Associate
	Microsoft
July 2022	Professional Scrum Master™ I (PSM I)
	Scrum.org
Apr. 2022	Strategic Leadership and Management Specialization
	University of Illinois Urbana-Champaign
Dec. 2021	Digitization applied to the productive sector
	Confederación Española de Organizaciones Empresariales (CEOE)
Jan. 2016	Cambridge First Certificate in English (FCE) - B2 level of the Common European Framework of Reference for Languages (CEFR)
	University of Cambridge
Jan. 2016	Elemental Grade in Valencian - B1 level of the Common European Framework of Reference for Languages (CEFR)
	Generalitat Valenciana
Dec. 2015	Occupational safety: laboratory researcher profile (register number: 15/45480)
	Universitat Politècnica de València

Courses

Feb. 2022	First aid in babies and children.
	Cruz Roja Española
Oct. 2017	Bitcoin and Cryptocurrency Technologies.
	Offered by the Princeton University at Coursera
Feb. 2013	PROMISE Winter School 2013 - Bridging between Information Retrieval and Databases.
	Bressanone, Italy.
Jul. 2009	UIMP English Language Immersion Course - Intensive English Course with Native Teachers and Monitors.
	Barcelona, Spain.

Achievements and Awards

2024	First place at the SemEval-2024 Task 8: Multigenerator, Multidomain, and Multilingual Black-Box Machine-Generated Text Detection - Monolingual subtask A) Binary Human-Written vs. Machine-Generated Text Classification.
	Awarded by the 18th International Workshop on Semantic Evaluation (SemEval-2024).
2021	First place in several evaluation metrics at the eRisk 2021: Early risk prediction on the Internet shared task@CLEF - subtasks on "Early detection of signs of self-harm" and "Measuring the severity of the signs of depression".
2018	PhD thesis award for the best dissertation in the field of Information Retrieval.
	Awarded by the 5th Spanish Conference in Information Retrieval (CERI 2018).
2017	Cum laude PhD thesis: A Cross-domain and Cross-language Knowledge-based Representation of Text and its Meaning.
2016	First place at the SemEval-2016, task 3: community question answering - subtask B: question-question similarity.
	Awarded by the 10th International Workshop on Semantic Evaluation (SemEval-2016).

2013	Winner of the MAVIR 2013 national prize to the best M.Sc. thesis on Language Technologies applied to Intelligent Systems of Access to the Multilingual and Multimedia Information.
	Awarded by MAVIR Consortium. Ssponsored by Córex and DAEDALUS.
2013	Cum laude MSc thesis: Cross-language Plagiarism Detection using a Multilingual Semantic Network.
2013	Third prize winner in the poster session of PROMISE Winter School 2013.

Research Internships

2015	National Institute of Astrophysics, Optics and Electronics (INAOE), Puebla, Mexico
	Mentor: Dr. Manuel Montes-y-Gómez
	Funded by the Web Information Quality Evaluation Initiative (WIQ-EI) IRSES 4-year project (grant no. 269180) within the FP 7 Marie Curie People Framework.
	Date: Feb. - Mar. 2015
2014	International Institute of Information Technology (IIIT), Hyderabad, India
	*During this internship I have been also research intern at Veooz - URL: www.veooz.com
	Mentor: Prof. Dr. Vasudeva Varma
	Funded by the Web Information Quality Evaluation Initiative (WIQ-EI) IRSES 4-year project (grant no. 269180) within the FP 7 Marie Curie People Framework.
	Date: Oct - Dec. 2014
2012	Computational Research Center (CIC), Mexico D.F., Mexico & National Institute of Astrophysics, Optics and Electronics (INAOE), Puebla, Mexico
	Mentor: Prof. Dr. Grigori Sidorov & Dr. Manuel Montes-y-Gómez
	Funded by the Web Information Quality Evaluation Initiative (WIQ-EI) IRSES 4-year project (grant no. 269180) within the FP 7 Marie Curie People Framework.
	Date: Nov - Dec. 2012

Competences

Artificial Intelligence	Deep learning, machine learning, data mining, natural language processing, information retrieval, machine translation
Cloud Computing	Microsoft Azure, Google Cloud
Data Mining & Machine Learning	TensorFlow, Keras, PyTorch, scikit-learn, huggingface-transformers, KNIME, Weka
Programming Languages	Python, Java, C, C++, Bash
Software Development	Kanban, Scrum, Azure DevOps, Eclipse, Maven, Visual Studio, Visual Studio Code, Git

Containers & Orchestration	Docker, Kubernetes
Big Data, DB	BigQuery, Spark, SQL, PostgreSQL, Oracle

Publications

2025	J. Á. González, I. Borrego-Obrador, A. Romo-Herrero, A. Mikael Sarvazyan, M. Chinea-Ríos, A. Basile, and M. Franco-Salvador. **IberBench: LLM evaluation on Iberian languages.** In *Computer Speech & Language*, 96 (101899), 2025, DOI: 10.1016/j.csl.2025.101899 **(Impact Factor: 3.4)**
2025	E. Corbí and M. Franco-Salvador. **ECORBI-UPV at EXIST 2025: Large Language Models and Embedding Strategies for Sexism Detection in Tweets.** In *Guglielmo Faggioli et al., editors, Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum*, September 2025. CEUR-WS.org.
2025	K. Korre, A. Basile, S. Yenikent, B. Spallaccia, M. Franco-Salvador, and A. Barrón-Cedeño. **Examining Inferred Author and Textual Correlates of Harmful Language Annotation.** In *Language Resources & Evaluation* (2025), DOI: 10.1007/s10579-025-09822-7 **(Impact Factor: 1.7)**
2024	A. Mikael Sarvazyan, J. Á. González, and M. Franco-Salvador. **TextMachina: Seamless Generation of Machine-Generated Text Datasets**. In *Proceedings of the 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024)*, Seville, Spain, pp. 566-575, September 11-13, 2024. **(CORE B)**
2024	K. Myśliwiec, M. Chinea-Rios, I. Borrego-Obrador, and M. Franco-Salvador. **AI-Q: a Framework to Measure the Robustness of Text Classification Models**. In *Proceedings of the 28th International Conference on Knowledge Based and Intelligent information and Engineering Systems (KES 2024)*, Seville, Spain, pp. 596-605, September 11-13, 2024. **(CORE B)**
2024	A. Mikael Sarvazyan, J. Á. González, F. Rangel, P. Rosso, and M. Franco-Salvador. **Overview of IberAuTexTification at IberLEF 2024: Detection and Attribution of Machine-Generated Text on Languages of the Iberian Peninsula**. In *Sociedad Española de Procesamiento del Languaje Natural (SEPLN)*, num. 73, pp. 421–434. Sep 1, 2024.
2024	A. Mikael Sarvazyan, J. Á. González, and M. Franco-Salvador. **Genaios at SemEval-2024 Task 8: Detecting Machine-Generated Text by Mixing Language Model Probabilistic Features**. In *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)*, pp. 101–107, Mexico City, Mexico.
2024	A. Basile, P. Rosso. and M. Franco-Salvador. **PyRater: A Python Toolkit for Annotation Analysis**. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pp. 13356–13362, Torino, Italy. **(CORE B)**
2023	A. Basile, M. Franco-Salvador, and P. Rosso. **Zero-Shot Data Maps. Efficient Dataset Cartography Without Model Training**. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 8264–8277, Singapore. **(CORE A)**

2023   A. Mikael Sarvazyan, J. Á. González, M. Franco-Salvador, F. Rangel, B. Chulvi, and P. Rosso. **Overview of AuTexTification at IberLEF 2023: Detection and Attribution of Machine-Generated Text in Multiple Domains**. In *Sociedad Española de Procesamiento del Lenguaje Natural (SEPLN)*, num. 71, pp. 275–288. Sep 1, 2023.

2023   A. Mikael Sarvazyan, J. Á. González, P. Rosso, and M. Franco-Salvador. **Supervised Machine-Generated Text Detectors: Family and Scale Matters**. In *14th International Conference of CLEF on Experimental IR meets Multilinguality, Multimodality, and Interaction (CLEF 2023)*, Springer-Verlag, LNCS(14163), pp. 121–132.

2023   M. Chinea-Rios, I. Borrego-Obrador, M. Franco-Salvador, F. Rangel, and P. Rosso. **Profiling Cryptocurrency Influencers with Few-shot Learning**. In *Avi Arampatzis et al., editors, Working Notes of CLEF 2023 - Conference and Labs of the Evaluation Forum*, September 2023. CEUR-WS.org.

2023   J. Bevendorff, I. Borrego-Obrador, M. Chinea-Ríos, M. Franco-Salvador, M. Fröbe, A. Heini, K. Kredens, E. Körner, M. Mayerl, P. Pezik, M. Potthast, F. Rangel, P. Rosso, E. Stamatatos, B. Stein, M. Wiegmann, M. Wolska, and E. Zangerle. **Overview of PAN 2023: Authorship Verification, Multi-Author Writing Style Analysis, Profiling Cryptocurrency Influencers, and Trigger Detection.** In *In 14th International Conference of CLEF on Experimental IR meets Multilinguality, Multimodality, and Interaction (CLEF 2023)*, Springer-Verlag, LNCS(14163), pp. 459–481.

2023   J. Bevendorff, M. Chinea-Ríos, M. Franco-Salvador, A. Heini, E. Körner, K. Kredens, M. Mayerl, P. Pezik, M. Potthast, F. Rangel, P. Rosso, E. Stamatatos, B. Stein, M. Wiegmann, M. Wolska, and E. Zangerle. **Overview of PAN 2023: Authorship Verification, Multi-Author Writing Style Analysis, Profiling Cryptocurrency Influencers, and Trigger Detection.** In *Proc. of the 45th European Conference on Information Retrieval (ECIR 2023)*, Dublin, Ireland, 2nd-6th April, 2023. **(CORE A)**

2022   E. Whitehouse, W. Gerard, Y. Klimovich, and M. Franco-Salvador. **Programming by Example and Text-to-Code Translation for Conversational Code Generation.** In *arXiv preprint*, 2022.

2022   T. Muller, G. Pérez-Torró, and M. Franco-Salvador. **Few-Shot Learning with Siamese Networks and Label Tuning.** In *Proc. of the 60th Annual Meeting of the Association for Computational Linguistics - Volume 1: Long Papers (ACL 2022)*, Dublin, Ireland, May 22-27, pp. 853–8545. **(CORE A)**

2022   T. Muller, G. Pérez-Torró, and M. Franco-Salvador. **Active Few-Shot Learning with FASL.** In *Proc. of the 27th International Conference on Natural Language & Information Systems (NLDB 2022)*, Valencia, Spain, June 15-17, pp. 98–110. **(CORE C)**

2022   A. Basile, M. Franco-Salvador, and P. Rosso. **Unsupervised Ranking and Aggregation of Label Descriptions for Zero-Shot Classifiers.** In *Proc. of the 27th International Conference on Natural Language & Information Systems (NLDB 2022)*, Valencia, Spain, June 15-17, pp. 119–126. **(CORE C)**

2022   M. Chinea-Rios, T. Muller, G. de la Peña, F. Rangel, and M. Franco-Salvador. **Zero and Few-shot Learning for Author Profiling.** In *Proc. of the 27th International Conference on Natural Language & Information Systems (NLDB 2022)*, Valencia, Spain, June 15-17, pp. 333–344. **(CORE C)**

2021   C. Cohrdes, S. Yenikent, J. Wu, B. Ghanem, M. Franco-Salvador, and F. Vogelgesang. **Indications of Depressive Symptoms During the COVID-19 Pandemic in Germany: Comparison of National Survey and Twitter Data.** In *JMIR Ment Health* 2021;8(6):e27140, DOI: 10.2196/27140 **(Impact Factor: 4.39)**

2021   A. Basile, G. Pérez-Torró, and M. Franco-Salvador. **Probabilistic Ensembles of Zero- and Few-Shot Learning Models for Emotion Classification.** In *Proc. of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, Bulgaria, September 1-4, pp. 128–137.

2021   A. Basile, M. Chinea-Rios, A.S. Uban, T. Muller, L. Rossler, S. Yenikent, B. Chulví, P. Rosso, and M. Franco-Salvador. **UPV-Symanto at eRisk 2021: Mental Health Author Profiling for Early Risk Prediction on the Internet.** In *CLEF 2021 Labs and Workshops, Notebook Papers.* CEUR-WS.org, http://ceur-ws.org/Vol-2936/paper-75.pdf.

2021   S. Stajner, S. Yenikent, B. Ghanem, and M. Franco-Salvador. **What Motivates You? Benchmarking Automatic Detection of Basic Needs From Short Posts.** In *Proc. of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*, August 1th through August 6th, 2021, pp. 803-–81. **(CORE A)**

2021   S. Stajner, S. Yenikent, and M. Franco-Salvador. **Five Psycholinguistic Characteristics for Better Interaction with Users.** In *Proceedings of the 8th IEEE International Conference on Behavioural and Social Computing (BESC)*, November 1, 2021.

2020   M. Chinea-Rios, M. Franco-Salvador, and Y. Benajiba. **Aspect On: an Interactive Solution for Post-Editing the Aspect Extraction based on Online Learning.** In *Proceedings of The 12th Language Resources and Evaluation Conference (LREC 2020)*, Marseille, France, May 13-16, pp. 4974–4981. **(CORE C)**

2019   A. Basile, M. Franco-Salvador, N. Pawar, S. Štajner, M. Chinea Rios, and Y. Benajiba. **SymantoResearch at SemEval-2019 Task 3: Combined Neural Models for Emotion Classification in Human-Chatbot Conversations.** In *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*, Minneapolis, USA, June 6-7.

2018   M. Franco-Salvador, and L. A. Leiva. **Multilingual phrase sampling for text entry evaluations.** In *International Journal of Human-Computer Studies*, vol. 113, pp. 15-31. DOI: 10.1016/j.ijhcs.2018.01.006 **(Impact Factor: 2.13)**

2018   R. M. Giménez-Pérez, M. Franco-Salvador, and P. Rosso. **String Kernels for Polarity Classification: A Study across Different Languages.** In *Proc. of the 23rd International Conference on Natural Language & Information Systems (NLDB 2018)*, Paris, France, 13-15 June, pp. 489–493. **(CORE C)**

2018	M. A. Álvarez Carmona, M. Franco-Salvador, M. Montes-Y-Gómez, P. Rosso, L. Villaseñor-Pineda, and E. Villatoro Tello. **Semantically-informed Distance and Similarity Measures for Paraphrase Plagiarism Identification.** In *Journal of Intelligent and Fuzzy Systems, Special Issue 4th Int. Symposium on Language & Knowledge Engineering*, vol. 34, pp. 2983–2990. DOI: 10.3233/JIFS-169483 **(Impact Factor: 1.26)**
2018	S. Štajner, M. Franco-Salvador, P. Rosso, and S. P. Ponzetto. **CATS: A Tool for Customised Alignment of Text Simplification Corpora.** In *Proc. 11th Int. Conf. on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 7-12. **(CORE C)**
2017	G. Glavaš, M. Franco-Salvador, S. P. Ponzetto, and P. Rosso. **A Resource-light Method for Cross-lingual Semantic Textual Similarity.** In *Knowledge-Based Systems*, vol. 143, pp. 1-9. DOI: 10.1016/j.knosys.2017.11.041 **(Impact Factor: 4.53)**
2017	S. Štajner, M. Franco-Salvador, S. P. Ponzetto, P. Rosso, and H. Stuckenschmidt. **Sentence Alignment Methods for Improving Text Simplification Systems.** In *Proc. of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017)*, Vancouver, Canada, July 30th through August 4th, 2017, pp. 97–102. **(CORE A)**
2017	M. Franco-Salvador. **A Cross-domain and Cross-language Knowledge-based Representation of Text and its Meaning.** In *PhD Thesis, Universitat Politècnica de València.* http://hdl.handle.net/10251/84285
2017	M. Franco-Salvador, N. Plotnikova, N. Pawar, and Y. Benajiba. **Subword-based Deep Averaging Networks for Author Profiling in Social Media - Notebook for PAN at CLEF 2017.** In *CLEF 2017 Labs and Workshops, Notebook Papers.* CEUR-WS.org, http://ceur-ws.org/Vol-1866/
2017	M. Franco-Salvador, Greg Kondrak, and P. Rosso. **Bridging the Native Language and Language Variety Identification Tasks.** In *Proc. of the 21st International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2017)*, Marseille, France, 6-8 September, 2017, pp. 1554–1561. DOI: 10.1016/j.procs.2017.08.06 **(CORE B)**
2017	R. M. Giménez-Pérez, M. Franco-Salvador, and P. Rosso. **Single and Cross-domain Polarity Classification using String Kernels.** In *Proc. of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017)*, Valencia, Spain, 3-7 April, 2017, pp. 558-563. **(CORE A)**
2016	M. Franco-Salvador, P. Gupta, P. Rosso, and R. E. Banchs. **Cross-language Plagiarism Detection Over Continuous-space- and Knowledge Graph-based Representations of Language.** In *Knowledge-Based Systems*, vol. 111, pp. 87–99. DOI: 10.1016/j.knosys.2016.08.004 **(Impact Factor: 4.53)**
2016	M. Franco-Salvador, P. Rosso, and M. Montes-y-Gómez. **A Systematic Study of Knowledge Graph Analysis for Cross-language Plagiarism Detection.** In *Information Processing & Management*, vol. 52(4), pp. 550–570. DOI: 10.1016/j.ipm.2015.12.004 **(Impact Factor: 2.39)**

2016   F. Rangel, M. Franco-Salvador, and P. Rosso. **A Low Dimensionality Representation for Language Variety Identification.** In *Proc. of the 17th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2016)*, Springer, Cham, 2016. pp. 156-169. **(CORE C)**

2016   M. Franco-Salvador, S. Kar, T. Solorio, and P. Rosso. **UH-PRHLT at SemEval-2016 Task 3: Combining Lexical and Semantic-based Features for Community Question Answering.** In *Proc. of the 10th International Workshop on Semantic Evaluation, SemEval '16*, San Diego, California. Association for Computational Linguistics.

2015   M. Franco-Salvador, F. L. Cruz, J. A. Troyano, and P. Rosso. **Cross-domain Polarity Classification using a Knowledge-enhanced Meta-classifier.** In *Knowledge-Based Systems*, vol. 86, pp. 46–56. DOI: 10.1016/j.knosys.2015.05.020 **(Impact Factor: 4.53)**

2015   M. Franco-Salvador, F. Rangel, P. Rosso, M. Taulé, and A. Martí. **Language Variety Identification using Distributed Representations of Words and Documents.** In *6th International Conference of CLEF on Experimental IR meets Multilinguality, Multimodality, and Interaction (CLEF 2015)*, Springer-Verlag, LNCS(9283), pp. 28–40.

2015   M. Franco-Salvador, I. Bensalem, E. Flores, P. Gupta, and P. Rosso. **PAN 2015 Shared Task on Plagiarism Detection: Evaluation of Corpora for Text Alignment.** In *CLEF 2015 Labs and Workshops, Notebook Papers*. CEUR-WS.org, http://ceur-ws.org/Vol-1391/

2015   M. Franco-Salvador, P. Rosso, and F. Rangel. **Distributed Representations of Words and Documents for Discriminating Similar Languages.** In *Proc. of the RANLP Joint Workshop on Language Technology for Closely Related Languages, Varieties and Dialects (LT4VarDial)*.

2014   M. Franco-Salvador, P. Rosso, and R. Navigli. **A Knowledge-based Representation for Cross-Language Document Retrieval and Categorization.** In *Proc. of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014)*, Gothenburg, Sweden, 26-30 April, 2014, pp. 414–423. **(CORE A)**

2014   M. Franco-Salvador, P. Gupta and P. Rosso. **Knowledge Graphs as Context Models: Improving the Detection of Cross-Language Plagiarism with Paraphrasing.** In *Revised Papers PROMISE Winter School 2013*, Springer-Verlag, LNCS(8173), pp. 227–236.

2013   M. Franco-Salvador, P. Gupta and P. Rosso. **Graph-Based Similarity Analysis: A New Approach to Cross-Language Plagiarism Detection.** In *Sociedad Española de Procesamiento del Languaje Natural (SEPLN)*, num. 50, pp. 21–28.

2013   M. Franco-Salvador, P. Gupta and P. Rosso. **Cross-Language Plagiarism Detection using a Multilingual Semantic Network.** In *Proc. of the 35th European Conference on Information Retrieval (ECIR 2013)*, Moscow, Russia, 24-27 March, 2013, Springer-Verlag, LNCS(7814), pp. 710–713. **(CORE B)**

2012   M. Franco-Salvador, P. Gupta and P. Rosso. **Cross-language Plagiarism Detection Using BabelNet's Statistical Dictionary.** In *Computación y Sistemas, Revista Iberoamericana de Computación*, ISSN 1405-5546, vol. 16, num. 4, pp. 383–390.