

CV Date	07/12/2025
---------	------------

Part A. PERSONAL INFORMATION

First Name	Fermín		
Family Name	Cruz Mata		
Sex	Male	Date of Birth	
ID number Social Security, Passport			
URL Web			
Email Address			
Open Researcher and Contributor ID (ORCID)	0000-0001-9029-9963		

A.1. Current position

Job Title	Profesor Titular de Universidad		
Starting date	2019		
Institution	Universidad de Sevilla		
Department / Centre	Lenguajes y Sistemas Informáticos		
Country	Spain	Phone Number	
Keywords			

Part B. CV SUMMARY

My research career began in 2004 as a doctoral fellow at the IBM Speech Technology Center in Seville, continuing as a researcher in 2005 at Indisys S.L., a company dedicated to the development of natural language dialogue systems. At the end of 2005, I began to work as an Assistant Professor in the Department of Computer Languages and Systems, continuing in this occupation until today (I am currently an Associate Professor). My area of research has remained from the beginning in the area of Natural Language Processing, always from a perspective of approaches based on machine learning. My Ph.D. thesis, which focused on the extraction of opinions from product reviews, obtained in 2011 a national award granted by the Spanish Society for Natural Language Processing for the best research work in the area in that year. After the doctoral thesis, my lines of research have focused not only on the analysis of opinions but also on the processing of publications on social networks and the analysis of communities on the Internet.

In all these years, I have maintained a high research activity, which has been reflected in the publication of 18 papers in journals indexed in JCR, multiple participations in national and international congresses, two research stays (at the Universidad Politécnica de Cataluña, together with Dr. Horacio Rodríguez, and at the Institute of Information Science and Technology in Pisa, Italy, collaborating with Dr. Fabrizio Sebastiani), and the supervision of a Ph.D. thesis. In addition, the various contributions made to the state of the art in the area have been used for technology transfer, through various contracts 68/83 with the company Itelligent S.L.; the transferred technology has been used as a central part of the GEOM Index (<http://www.geomindex.com>) and netOpinion (<https://itelligent.es/es/netopinion/>) platforms. GEOM Index has appeared in numerous press articles, and netOpinion has important clients such as Chanel, DKV Seguros, Conde Nast, Clínicas Baviera, Cortefiel and GSM Management Consulting.

Regarding participation in publicly funded research projects, in the last 10 years, I have actively participated in four projects, three national and two regional. Most recently, I have been the principal investigator of a research project focused on the automatic identification and classification of argumentative fallacies in political discourse, with a special emphasis on the development of high-quality annotated corpora and reproducible evaluation benchmarks.

Within my work as a teacher, I have been teaching for 6 years a subject in the Master in Software Engineering: Cloud, Data and IT Management of the University of Seville, focused

on natural language processing, where every year I update the contents to transmit the latest advances in the field to my students. In addition, in terms of management tasks, I have been the Secretary of the Department of Computer Languages and Systems from 2017 to 2021, an occupation that, although it has caused a period of lower productivity in terms of scientific publications, has provided me with other skills related to management that I hope can be applied in the coordination of the project presented here.

Part C. RELEVANT ACCOMPLISHMENTS

C.1. Most important publications in national or international peer-reviewed journals, books and conferences

AC: corresponding author. (n° x / n° y): position / total authors. If applicable, indicate the number of citations

- 1 Scientific paper.** Noelia Lopez-Duran; David Romero-Organvdez; (3/4) Fermín L. Cruz; David Benavides. 2025. Software bug report dataset from Eclipse projects. DATA IN BRIEF. ELSEVIER SCIENCE BV; ELSEVIER; Elsevier BV. 62, pp.112016-112016. ISSN 2352-3409. SCOPUS (0), WOS (0) <https://doi.org/10.1016/j.dib.2025.112016>
- 2 Scientific paper.** (1/4) Cruz, Fermín L.; Enríquez, Fernando; Ortega, F. Javier; Troyano, José A.2025. FallacyES-Political: A Multiclass Dataset of Fallacies in Spanish Political Debates. PROCESAMIENTO DEL LENGUAJE NATURAL. SOC ESPANOLA PROCESAMIENTO LENGUAJE NATURAL-SEPLN. 74-74, pp.127-138. ISSN 1135-5948, ISSN 1989-7553. SCOPUS (1), WOS (1), Dialnet (0) <https://doi.org/10.26342/2025-74-9>
- 3 Scientific paper.** Alvarez-Galvez, Javier; (2/3) Cruz, Fermin L.; Troyano, Jose A.2023. Discovery and characterisation of socially polarised communities on social media. Scientific reports. NATURE PUBLISHING GROUP; NATURE PORTFOLIO. 13-1. ISSN 2045-2322, ISSN 2045-2322. SCOPUS (1), WOS (1) <https://doi.org/10.1038/s41598-023-42592-2>
- 4 Scientific paper.** (1/4) Cruz, Fermín L. (AC); Troyano, José A.; Enríquez, Fernando; Ortega, F. Javier. 2023. Detección y clasificación de falacias prototípicas y espontáneas en español. PROCESAMIENTO DEL LENGUAJE NATURAL. SOC ESPANOLA PROCESAMIENTO LENGUAJE NATURAL-SEPLN. 71-71, pp.53-62. ISSN 1135-5948, ISSN 1989-7553. SCOPUS (3), WOS (0), Dialnet (2) <https://doi.org/10.26342/2023-71-4>
- 5 Scientific paper.** Vallejo, Carlos G.; Troyano, José A.; Enríquez, Fernando; Ortega, F. Javier; (5/5) Cruz, Fermín L.2020. MCFS: Min-cut-based feature-selection. KNOWLEDGE-BASED SYSTEMS. ELSEVIER SCIENCE BV; ELSEVIER. 195. ISSN 0950-7051, ISSN 1872-7409. SCOPUS (3), WOS (3) <https://doi.org/10.1016/j.knosys.2020.105604>
- 6 Scientific paper.** Cotelo, Juan M.; Ortega, F. Javier; Troyano, José A.; Enríquez, Fernando; (5/5) Cruz, Fermín L.2020. Known by who we follow: a biclustering application to community detection. IEEE ACCESS. IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC; Institute of Electrical and Electronics Engineers (IEEE). 8, pp.192218-192228. ISSN 2169-3536, ISSN 2169-3536. SCOPUS (2), WOS (1) <https://doi.org/10.1109/ACCESS.2020.3032015>
- 7 Scientific paper.** Cotelo, J. M.; (2/4) Cruz, F. L.; Enríquez, F.; Troyano, J. A.2016. Tweet categorization by combining content and structural knowledge. INFORMATION FUSION. ELSEVIER SCIENCE BV; ELSEVIER. 31, pp.54-64. ISSN 1566-2535, ISSN 1872-6305. SCOPUS (29), WOS (24) <https://doi.org/10.1016/j.inffus.2016.01.002>
- 8 Scientific paper.** Cotelo, J. M.; (2/4) Cruz, F. L.; Troyano, J. A.; Ortega, F. J.2015. A modular approach for lexical normalization applied to Spanish tweets. EXPERT SYSTEMS WITH APPLICATIONS. PERGAMON-ELSEVIER SCIENCE LTD; Elsevier BV. 42-10, pp.4743-4754. ISSN 0957-4174, ISSN 1873-6793. SCOPUS (13), WOS (10) <https://doi.org/10.1016/j.eswa.2015.02.003>

- 9 **Scientific paper.** Franco-Salvador, Marc; (2/4) Cruz, Fermín L.; Troyano, José A.; Rosso, Paolo. 2015. Cross-domain polarity classification using a knowledge-enhanced meta-classifier. KNOWLEDGE-BASED SYSTEMS. ELSEVIER SCIENCE BV; ELSEVIER. 86, pp.46-56. ISSN 0950-7051, ISSN 1872-7409. SCOPUS (33), WOS (23) <https://doi.org/10.1016/j.knosys.2015.05.020>
- 10 **Data Paper.** (1/4) Cruz, Fermín L. (AC); Enríquez, Fernando; Ortega, F. Javier; Troyano, José A. 2025. Annotated Spanish general election debate transcriptions 1993-2023. SCIENTIFIC DATA. NATURE PUBLISHING GROUP; NATURE PORTFOLIO. 12-1. ISSN 2052-4463. SCOPUS (0), WOS (0) <https://doi.org/10.1038/s41597-025-05818-8>

C.3. Research projects and contracts

- 1 **Project.** SOL2024-31823, Nuevo Impulso a la Investigación en Ingeniería Informática (NI-FOUR). Universidad de Sevilla. Riquelme Santos, José Cristóbal. 01/01/2025-31/12/2027. 48.643,98 €.
- 2 **Project.** PID2021-123005NB-I00, Detección y generación automática de falacias mediante el uso de modelos de lenguaje masivos basados en aprendizaje profundo. Ministerio de Ciencia e Innovación. Cruz Mata, Fermín. 01/09/2022-28/02/2026. 37.389 €.
- 3 **Project.** US-1380960, Explorando la generación de contenido online por el usuario y su influencia predictiva en la calidad relacional. Aplicación al sector hotelero de Andalucía. Consejería de Economía, Conocimiento, Empresas y Universidad. Sánchez Franco, Manuel Jesús. 01/01/2022-31/05/2023. 40.243,7 €.
- 4 **Project.** P11-TIC-7684, Aorescu: Análisis de Opinión en Redes Sociales y Contenidos Generados por Usuarios. Junta de Andalucía - Consejería de Innovación, Ciencia y Empresas. Troyano Jiménez, José Antonio. 26/03/2013-25/03/2016. 151.685 €.
- 5 **Project.** TIN2012-38536-C03-02, Análisis de Contenidos Generados por Usuarios. Ministerio de Economía y Competitividad. Troyano Jiménez, José Antonio. 01/01/2013-31/12/2015. 26.863,2 €.
- 6 **Contract.** Minería de Opiniones para Textos sobre Automóviles Intelligent Information Technologies, S.L.. Troyano Jiménez, José Antonio. 28/04/2017-06/04/2019. 5.088 €.