



Name: José Ángel González Barba

Birth date: 16/08/94

Address: Dos de Mayo nº 10 Moncada, Valencia, Spain

Phone number: +34685479514

E-mail: jgonba2elirf@gmail.com

Languages: Spanish (native), Catalan, English

Social networks:     

Bio

Hi! 🙋😊. I am José, an NLP researcher deeply passionate about exploring the limitless possibilities of NLP and Deep learning. I earned a degree in Informatics from the Polytechnic University of Valencia, where I also pursued a Master's in Artificial Intelligence and Pattern Recognition, followed by a PhD in Computer Science. During my academic journey, I worked in the ELiRF research team at the same institution.

My PhD research centered on summarization and attention-based models, and it received the best NLP thesis by the Spanish Society for Natural Language Processing. But my expertise extends across a wide range of NLP topics, including zero- and few-shot text classification, sentiment and emotion analysis, figurative language understanding, dialogue systems, information extraction, and machine-generated text detection.

In 2022, I joined as Senior Research Scientist at the Research Team of Symanto, a company combining NLP with psychology to understand people's opinions. My work spanned diverse areas, including topic discovery, detecting machine-generated text, parameter-efficient fine-tuning, and few-shot approaches for multimodal classification and text generation.

You can know more about me and my skills on [Github](#) and find my publications in [Google Scholar](#) 😊

Current situation

Company: Symanto Research GmbH & Co. KG

Position: Principal Research Scientist

Dedication: Full time

Education

Title	University	Link	Date
PhD in Computer Science (FPI-UPV grant)	Polytechnic University of Valencia	Link	07/07/2021
Master's Degree in Artificial Intelligence, Pattern Recognition and Digital Imaging	Polytechnic University of Valencia	Link	25/07/2017
Grado en Ingeniería Informática	Polytechnic University of Valencia	Link	14/07/2016

Experience

Position	Company	Date
Principal Research Scientist	Symanto	2024 – present
Professor	European University	2023 – present
Professor	University of Barcelona (IL3)	June 2023 – present
Senior Research Scientist	Symanto	2022-2024
Senior Research Technician	Valencian Research Institute for Artificial Intelligence	2020-2022
Senior Research Technician	Department of Computer Systems and Computation (UPV)	2018-2020
Junior Research Technician	Department of Computer Systems and Computation (UPV)	2016-2017

R&D projects

Title	University	Funded by	Date
GUAITA: Monitoring and analyzing social networks to aid decision making INNVA1/2020/61	Valencian Research Institute for Artificial Intelligence (9 researchers)	Valencian Innovation Agency + Apunt TV 148,764.78€	2020-2022
AI models for oil spill detection in aperture radar synthetic images	Collaboration between Valencian Research Institute for Artificial Intelligence y ORBITAL EOS	ORBITAL EOS 10,000€	2022-2022
AMIC-UPV: Affective analysis of multimedia information with inclusive and natural communication TIN2017-85854-C4-1-R	Department of Computer Systems and Computation (UPV) (8 researchers)	Ministry of Economy and Competitiveness + FEDER funds 108,174.00€	2018-2020
MULAN: Audio, speech and language processing for analyzing multimedia information TIN2014-54288-C4	Department of Computer Systems and Computation (UPV) (8 researchers)	Ministry of Economy and Competitiveness + FEDER funds 87,604.00€	2016-2017
IberBench: An open-source benchmark for evaluating LLMs on Iberian languages	Symanto Spain SLU	IVACE 127,962€	2023-2024

Teching experience

Accredited as “Assistant Professor Doctor” by the National Agency for Quality Assessment and Accreditation of Spain

Subject	University	Date
Object Oriented Programming	European University	2025-present
Advanced Machine Learning Techniques	University of Barcelona (IL3)	2024-present
Intelligent Systems and Knowledge Representation	European University	2023-present
Computational Linguistics	Polytechnic University of Valencia	2022-present
Algorithmics	Polytechnic University of Valencia	2019/2020 y 2020/2021
Information Retrieval	Polytechnic University of Valencia	2020/2021
Introduction to computing and programming	Polytechnic University of Valencia	2019/2020
Programming	Polytechnic University of Valencia	2018/2019

Tutor of 6 Final Degree Projects:

Title	Student	Date	Qualification
Comparative study of tools for Natural Language Processing tasks	DIAZ ROUSSEL, NICOLAS	14/07/2020	8.0 / 10
Contextual and incontextual word embeddings for Twitter	ARIAS MONCHO, JOSE	15/09/2020	10 / 10
Semantic search of profiles on social networks	ROS PAGAN, DIEGO	29/03/2021	9.0 / 10
Development of a platform for Twitter analytics	PITARCH MIRAVETE, JOSE LUIS	21/07/2021	7.5 / 10
Question Answering using Neural Networks	SALYUK KULINICH, MYKOLA	19/09/2022	9.5 / 10
Search assistant for the European University	ZHAN BEI, CHEN	07/07/2025	TBD

Co-director of one PhD student:

Title	Student	Date	Qualification
Few-shot models as annotators	BASILE, ANGELO	TBD	TBD

Journal publications

1. Areg Mikael Sarvazyan, José Ángel González, Marc Franco-Salvador, Francisco Rangel, Berta Chulvi, Paolo Rosso. Overview of IberAuTeXtification at IberLEF 2024: Detection and attribution of machine-generated text on languages of the iberian peninsula. *Procesamiento del Lenguaje Natural*, 72(0), SJR Q4(CS)
2. Ahuir, V., González, J. Á., Hurtado, L. F., & Segarra, E. (2024). Abstractive Summarizers Become Emotional on News Summarization. *Applied Sciences*, 14(2), 713. Q2 (CS)
3. Areg Mikael Sarvazyan, José Ángel González, Marc Franco-Salvador, Francisco Rangel, Berta Chulvi, Paolo Rosso. Overview of AuTeXtification at IberLEF 2023: Detection and Attribution of Machine-Generated Text in Multiple Domains. *Procesamiento del Lenguaje Natural*, 71(0), 275–288. SJR Q4(CS)
4. José Ángel González, Encarna Segarra, Fernando García-Granada, Emilio Sanchis, Lluís-F. Hurtado. Attentional Extractive Summarization. *Applied Sciences* (ISSN 2076-3417). Volumen 13, open access. Q2 (CS)
5. Sanchis-Font, Rosario; Castro-Bleda, Maria Jose; González-Barba, José Ángel; Pla Santamaría, Ferran; Hurtado Oliver, Lluís Felip. Cross-Domain Polarity Models to Evaluate User eXperience in E-learning. *Neural Processing Letters* (ISSN 1370-4621). 2021. Volumen 53, páginas 3199-3215. Q2 (CS)
6. Ahuir-Esteve, Vicent; Hurtado Oliver, Lluís Felip; González-Barba, José Ángel; Segarra Soriano, Encarnación. NASca and NASes: Two Monolingual Pre-Trained Models for Abstractive Summarization in Catalan and Spanish. *Applied Sciences* (ISSN 2076-3417). 2021. Volumen 11, páginas 1-16. Q2 (CS)
7. González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. TWiLBert: Pre-trained deep bidirectional transformers for Spanish Twitter. *Neurocomputing* (ISSN 0925-2312). 2021. Volumen 426, páginas 58-69. Q1 (CS)
8. González-Barba, José Ángel; Segarra Soriano, Encarnación; García-Granada, Fernando; Sanchis Arnal, Emilio; Hurtado Oliver, Lluís Felip. Extractive summarization using siamese hierarchical transformer encoders. *Journal of Intelligent & Fuzzy Systems* (ISSN 1064-1246). 2020. Volumen 39, páginas 2049-2419. Q3 (CS)
9. González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. Self-attention for Twitter sentiment analysis in Spanish. *Journal of Intelligent & Fuzzy Systems* (ISSN 1064-1246). 2020. Volumen 39, páginas 2165-2175. Q3 (CS)
10. González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. Transformer based contextualization of pre-trained word embeddings for irony detection in Twitter. *Information Processing & Management* (ISSN 0306-4573). 2020. Volumen 57, páginas 1-15. Q1 (CS)
11. González-Barba, José Ángel; Julien Delonca; Sanchis Arnal, Emilio; García-Granada, Fernando; Segarra Soriano, Encarnación. Applying Siamese Hierarchical Attention Neural Networks for multi-document summarization. *Sociedad Española para el Procesamiento del Lenguaje Natural*. 2019. Páginas 111-118.

- 12.** Hurtado Oliver, Lluís Felip; González-Barba, José Ángel; Pla Santamaría, Ferran. Choosing the right loss function for multi-label Emotion Classification. *Journal of Intelligent & Fuzzy Systems* (ISSN 1064-1246). 2019. Volumen 36, páginas 4697 – 4708. Q3 (CS)
- 13.** González-Barba, José Ángel; Segarra Soriano, Encarnación; García-Granada, Fernando; Sanchís Arnal, Emilio; Hurtado Oliver, Lluís Felip. Siamese hierarchical attention networks for extractive summarization. *Journal of Intelligent & Fuzzy Systems* (ISSN 1064-1246). 2019. Volumen 36, páginas 4599 – 4607. Q3 (CS)
- 14.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Segarra Soriano, Encarnación; García-Granada, Fernando; Sanchís Arnal, Emilio. Summarization of Spanish Talk Shows with Siamese Hierarchical Attention Networks. *Applied Sciences* (ISSN 2076-3417). Volumen 9, páginas 1-13. Q2 (CS)

Conference publications

- 1.** Sarvazyan, A., Gonzalez, J., & Franco-salvador, M. (2024). 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). Association for Computational Linguistics.
- 2.** Sarvazyan, A. M., González, J. Á., & Franco-Salvador, M. (2024). TextMachina: Seamless Generation of Machine-Generated Text Datasets. 28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems). Core B.
- 3.** Sarvazyan, A.M., González, J.Á., Rosso, P., Franco-Salvador, M. (2023). Supervised Machine-Generated Text Detectors: Family and Scale Matters. In: *Experimental IR Meets Multilinguality, Multimodality, and Interaction. CLEF 2023. Lecture Notes in Computer Science*, vol 14163. Springer, Cham.
- 4.** José Ángel González, Annie Louis, and Jackie Chi Kit Cheung. 2022. Source-summary Entity Aggregation in Abstractive Summarization. *Proceedings of the 29th International Conference on Computational Linguistics*. Páginas 6019–6034, Gyeongju, Republic of Korea. International Committee on Computational Linguistics. Core A.
- 5.** Encarnación Segarra Soriano, Vicent Ahuir, Lluís-F. Hurtado, and José González. 2022. DACSA: A large-scale Dataset for Automatic summarization of Catalan and Spanish newspaper Articles. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5931–5943, Seattle, United States. Association for Computational Linguistics. Core A.
- 6.** González, J.Á., Buscaldi, D., Sanchis, E., Hurtado, LF. (2022). Transformer-Based Models for the Automatic Indexing of Scientific Documents in French. In *Natural Language Processing and Information Systems. NLDB 2022. Lecture Notes in Computer Science*, vol 13286. Springer, Cham
- 7.** Vicent Ahuir, José-Ángel González, & Lluís-Felip Hurtado (2022). Enhancing Categorization in Low-data Situations. In *Proceedings of the Iberian Languages Evaluation Forum. Spain, September 20, 2022. CEUR-WS.org*.González-Barba, José Ángel; Arias-Moncho, José; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at TASS 2020: TWiLBERT for Sentiment Analysis and Emotion Detection in Spanish Tweets. *Iberian Languages Evaluation Forum (IberLEF 2020)*. 2020. Páginas 179 – 186. Málaga, Spain.
- 8.** Nina-Alcocer, Victor Manuel; González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. Aggressiveness Detection through Deep Learning Approaches. *Iberian Languages Evaluation Forum (IberLEF 2019)*. 2019. Páginas 544-549. Bilbao, Spain.
- 9.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at IroSvA: Transformer Encoders for Spanish Irony Detection. *Iberian Languages Evaluation Forum (IberLEF 2019)*. 2019. Páginas 278-284. Bilbao, Spain.
- 10.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at TASS 2019: Transformer Encoders for Twitter Sentiment Analysis in Spanish. *Iberian Languages Evaluation Forum (IberLEF 2019)*. Páginas 571-578. Bilbao, Spain.
- 11.** García-Granada, Fernando; Sanchís Arnal, Emilio; Castro-Bleda, Maria Jose; González-Barba, José Ángel; Hurtado Oliver, Lluís Felip. Word Discovering in Low-Resources Languages Through Cross-Lingual Phonemes. 21st International Conference on Speech and Computer (SPECOM 2019). 2019. Páginas 133 – 141. Istanbul, Turkey.
- 12.** Sanchis-Font, Rosario; Castro-Bleda, Maria Jose; González-Barba, José Ángel. Applying sentiment analysis with cross-domain models to evaluate user experience in virtual learning environments. . 15th International Work-Conference on Artificial Neural Networks, IWANN 2019. 2019. Páginas 609-620. Gran Canaria, Spain.
- 13.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at SemEval-2019 Task 3: Snapshot Ensemble of Hierarchical Convolutional Neural Networks for Contextual Emotion Detection. 13th International Workshop on Semantic Evaluation (SemEval 2019). 2019. Páginas 195-199. Minneapolis, USA.
- 14.** García-Granada, Fernando; Sanchís Arnal, Emilio; Castro-Bleda, Maria Jose; González-Barba, José Ángel; Hurtado Oliver, Lluís Felip. Using out of domain units for word discovering in raw speech. 23rd
- 15.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at MultiStanceCat 2018. Third Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2018). 2018. Páginas 173-179. Sevilla, Spain.

- 16.** González-Barba, José Ángel; Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran. ELiRF-UPV at TASS 2018: Sentiment Analysis in Twitter based on Deep Learning. Workshop on Sentiment Analysis at SEPLN (TASS 2018). 2018. Páginas 37-44. Sevilla, Spain.
- 17.** José-Ángel González; Lluís-F. Hurtado; Encarna Segarra; Ferran Pla. ELiRF-UPV at SemEval-2018 Task 10: Capturing Discriminative Attributes with Knowledge Graphs and Wikipedia. 12th International Workshop on Semantic Evaluation (SemEval 2018). 2018. Páginas 968-971. New Orleans, USA.
- 18.** José-Ángel González; Lluís-F. Hurtado; Encarna Segarra; Ferran Pla. ELiRF-UPV at SemEval-2018 Task 11: Machine Comprehension using Commonsense Knowledge. 12th International Workshop on Semantic Evaluation (SemEval 2018). 2018. Páginas 1034-1037. New Orleans, USA.
- 19.** José-Ángel González; Lluís-F. Hurtado; Ferran Pla. ELiRF-UPV at SemEval-2018 Tasks 1 and 3: Affect and Irony Detection in Tweets. 12th International Workshop on Semantic Evaluation (SemEval 2018). 2018. Páginas 565-569. New Orleans, USA.
- 20.** González-Barba, José Ángel; Pla Santamaría, Ferran; Hurtado Oliver, Lluís Felip. ELiRF-UPV at IberEval 2017: Classification Of Spanish Election Tweets (COSET). 2nd Workshop on Human Language Technologies for Iberian languages (IberEval 2017). 2017. Páginas 55-60. Murcia, Spain.
- 21.** González-Barba, José Ángel; Pla Santamaría, Ferran; Hurtado Oliver, Lluís Felip. ELiRF-UPV at IberEval 2017: Stance and Gender Detection in Tweets. 2nd Workshop on Human Language Technologies for Iberian languages (IberEval 2017). 2017. Páginas 193-198. Murcia, Spain.
- 22.** Hurtado Oliver, Lluís Felip; Pla Santamaría, Ferran; González-Barba, José Ángel. ELiRF-UPV en TASS 2017: Análisis de Sentimientos en Twitter basado en Aprendizaje Profundo. Workshop on Sentiment Analysis at SEPLN (TASS 2017). 2017. Páginas 29-34. Murcia, Spain.
- 23.** González-Barba, José Ángel; Pla Santamaría, Ferran; Hurtado Oliver, Lluís Felip. ELiRF-UPV at SemEval-2017 Task 4: Sentiment Analysis using Deep Learning. 11th International Workshop on Semantic Evaluation (SemEval 2017). 2017. Páginas 723-727. Vancouver, Canada.
- 24.** Hurtado Oliver, Lluís Felip; Segarra Soriano, Encarnación; Pla Santamaría, Ferran; Carrasco-Gómez, Pascual Andrés; González-Barba, José Ángel. ELiRF-UPV at SemEval-2017 Task 7: Pun Detection and Interpretation. 11th International Workshop on Semantic Evaluation (SemEval 2017). 2017. Páginas 440-443. Vancouver, Canada.

Organized shared tasks

Title	Forum	Year
AuTexTification: Detection and Attribution of Machine-Generated Text in Multiple Domains	Iberian Language and Evaluation Forum	2023
IberAuTexTification: Detection and Attribution of Machine-Generated Text on Languages of the Iberian Peninsula	Iberian Language and Evaluation Forum	2024
MIMIC: Multi-Modal AI Content Moderation	Iberian Language and Evaluation Forum	2025

Research internships

University	Location	Supervisor	Funded by	Purpose
- McGill University - MILA Research Institute	Montreal , Canada	Jackie Chi Kit Cheung	Movility grant (FPI-UPV, 4000€)	Semantic entity aggregation for automatic summarization.

Awards

- Award for the [best doctoral thesis](#) in Natural Language Processing, by the Spanish Society of Natural Language Processing.
- Doctoral thesis evaluated as cum laude with international mention.
- Award for the best record in the 2015-2016 academic year of the Computer Engineering degree (UPV).
- Best academic record of the university master's degree in Artificial Intelligence, Pattern Recognition and Digital Imaging (2017).
- 19 honors in bachelor's and master's degrees, including bachelor's and master's thesis.