

Javier García Gilabert

☎ (+34) 618 29 22 09 | ✉ javier.garcia1.bsc@gmail.com | 🏠 Scholar | 🐙 Github | 🔗 LinkedIn | 🏠 Homepage

Education



Master's Exchange Student (Erasmus)

AALTO UNIVERSITY

Helsinki, FI

Sep. 2023 - Jan. 2024

- Got the best Exchange student award in my master's program.



Master in Artificial Intelligence

POLYTECHNIC UNIVERSITY OF CATALONIA (UPC) & UNIVERSITY OF BARCELONA (UB) & UNIVERSITY OF ROVIRA I VIRGILI (URV)

Barcelona, ES

2022 - 2024



- Interuniversity master (taught in English).
- Achieved a mean degree mark of 8.85/10.
- Master Thesis: *Improving Multilingual Neural Machine Translation by projecting language representations.*



Bachelor's degree in Data Science

POLYTECHNIC UNIVERSITY OF VALENCIA (UPV)

Valencia, ES

2018 - 2022

- Got the best student award in Data Science. Honor granted for achieving the highest marks in my university degree.
- Achieved a mean degree mark of 9.3/10.
- Subjects passed with honors: 13.
- Received a Special Mention for my Bachelor Thesis: *Image Captioning using pre-trained GPT-2 models.*

Experience



Research Engineer in NLP

INSTITUTO DE TELECOMUNICAÇÕES - PI: ANDRÉ MARTINS

Lisboa, PT

Feb. 2026 - Jun. 2026

Working as a research engineer in the SMURF4EU project to develop a sovereign European suite of Multimodal Reasoning Foundation Models, leveraging 1.9M GPU-hours via EuroHPC Extreme Scale Access to build open-source, long-context systems across 24 EU languages.



Research Engineer in NLP

BARCELONA SUPERCOMPUTING CENTER - PI: MAITE MELERO

Barcelona, ES

Jan. 2024 - Jan. 2026

Working as a research engineer in the Language Technologies Group at the Barcelona Supercomputing Center.

- NLP Research Engineer contributing to key large-scale initiatives, including the European **Eloquence** project, as well as major Spanish and Catalan national projects such as **ALIA**, **ILENIA**, and **AINA**.
- Developed Multilingual Neural Machine Translation models LLM-based (SalamandraTA-2b-instruct, SalamandraTA-7b-instruct, Plume).
- Worked in low-resource machine translation.
- Experience in training LLMs in multi-node distributed environments using DeepSpeed, Nemo and Megatron frameworks.



Research Engineer in NLP

MT@UPC (MACHINE TRANSLATION GROUP AT UPC) - PI: JOSÉ ADRIÁN RODRÍGUEZ FONOLLOSA

Barcelona, ES

Oct. 2022 - Sep. 2023

National Spanish Project ROB-IN. The primary goal of this project was to create a robot designed for continuous personalized assistance with self-explanatory capabilities. My role in this project involved contributing to the development of the Natural Language Processing (NLP) component.



Undergraduate Research Assistant

DEPARTMENT OF APPLIED STATISTICS AT UPV - PI: JOSÉ MIGUEL CAROT SIERRA

Valencia, ES

Oct. 2019 - Aug. 2022

My role was mainly as a data scientist, dealing with data collection, processing, analysis, and the application of machine learning techniques.

Awards and Grants

2024	Best Exchange Student award in Master in Artificial Intelligence Honor granted for achieving the highest marks in my master's exchange program.	UPC-URV-UB
2022	Best Student award in Data Science Honor granted for achieving the highest marks in my university degree.	UPV
2022	Research Collaboration Fellowship Collaboration fellowship with the PRHLT group at UPV.	Spanish Ministry of Education
2022	Beca Santander Estudios Progreso 2021-2022 The "Becas Santander Progreso" program is a grant given by Banco Santander S.A., whose objective is to facilitate access to education for students with the best academic records at a national level who are pursuing official Bachelor's or Master's degree studies.	Banco Santander
2021	Awards to Graduates of the UPV with Best Academic Records Honor granted for achieving the highest marks in the second academic year of my University degree.	UPV

Publications



Sara Papi, **Javier Garcia Gilabert**, Zachary Hopton, Vilém Zouhar, Carlos Escolano, Gerard I. Gállego, Jorge Iranzo-Sánchez, Ahrii Kim, Dominik Macháček, Patricia Schmidtova, Maike Züfle. 2025. Hearing to Translate: The Effectiveness of Speech Modality Integration into LLMs. *Transactions of the Association for Computational Linguistics*



Iñaki Lacunza, **Javier Garcia Gilabert**, Francesca De Luca Fornaciari, Javier Aula-Blasco, Aitor Gonzalez-Agirre, Maite Melero, Marta Villegas. 2025. ACADATA: Parallel Dataset of Academic Data for Machine Translation. *In Proceedings of the International Conference on Language Resources and Evaluation (LREC'26), Palma de Mallorca, Spain. European Language Resources Association (ELRA).*



Javier Garcia Gilabert, Carlos Escolano, Xixian Liao and Maite Melero. 2025. Terminology-Constrained Translation from Monolingual Data using GRPO. *In Proceedings of the Tenth Conference on Machine Translation.*



Javier Garcia Gilabert, Xixian Liao, Severino Da Dalt, Ella Bohman, Audrey Mash, Francesca De Luca Fornaciari, Irene Baucells, Joan Llop, Miguel Claramunt Argote, Carlos Escolano and Maite Melero. 2025. From Salamandra to SalamandraTA: BSC Submission for WMT25 General Machine Translation Shared Task. *In Proceedings of the Tenth Conference on Machine Translation.*



Javier Garcia Gilabert, Carlos Escolano, Aleix Sant, Francesca De Luca Fornaciari, Audrey Mash, Xixian Liao, and Maite Melero. 2024. Investigating the translation capabilities of Large Language Models trained on parallel data only. *In Proceedings of Machine Translation Summit XX: Volume 1, pages 24–53, Geneva, Switzerland. European Association for Machine Translation.*



Javier Garcia Gilabert, Carlos Escolano, Audrey Mash, Xixian Liao, and Maite Melero. 2024. MT-LENS: An all-in-one Toolkit for Better Machine Translation Evaluation. *In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (System Demonstrations), pages 51–60, Albuquerque, New Mexico. Association for Computational Linguistics.*



Javier Garcia Gilabert, Aleix Sant, Carlos Escolano, Francesca De Luca Fornaciari, Audrey Mash, and Maite Melero. 2024. BSC Submission to the AmericasNLP 2024 Shared Task. *In Proceedings of the 4th Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP 2024), pages 143–149, Mexico City, Mexico. Association for Computational Linguistics.*



Javier Garcia Gilabert, Carlos Escolano, and Marta R. Costa-Jussà. 2024. ReSeTOX: Re-learning attention weights for toxicity mitigation in machine translation. *In Proceedings of the 25th Annual Conference of the European Association for Machine Translation, pages 37–58, Sheffield, England. European Association for Machine Translation.*

Peer collaborations:



Aleix Sant, Daniel Bardanca, José Ramon Pichel Campos, Francesca De Luca Fornaciari, Carlos Escolano, **Javier Garcia Gilabert**, Pablo Gamallo, Audrey Mash, Xixian Liao and Maite Melero. 2024. Training and Fine-Tuning NMT Models for Low-Resource Languages using Apertium-Based Synthetic Corpora. *In Proceedings of the Ninth Conference on Machine Translation, pages 925–933, Miami, Florida, USA. Association for Computational Linguistics.*



Xixian Liao, Carlos Escolano, Audrey Mash, Francesca De Luca Fornaciari, **Javier García Gilabert**, Miguel Claramunt Argote, Ella Bohman and Maite Melero. 2025. Culture-aware machine translation: the case study of low-resource language pair Catalan-Chinese. *In Proceedings of Machine Translation Summit XX: Volume 1, pages 150–161, Geneva, Switzerland. European Association for Machine Translation.*

Service in the field

- 2025 **Reviewer** EMNLP Demo Track 2025
- 2025 **Reviewer** WMT 2025
- 2025 **Reviewer** MT Summit 2025
- 2025 **Reviewer** NAACL Demo track 2025
- 2025 **Reviewer** IJCNLP-AAACL Demo track 2025
- 2025 **Reviewer** EAACL Demo track 2025

Certificates

- 2023 **DeepLearn 2023 summer: 10th International Gran Canaria School on Deep Learning** *DeepLearn*
- 2017 **First Certificate of English (FCE)** *Cambridge*

Skills

- Programming** Python, R, Bash.
- Frameworks** PyTorch, numpy, pandas, nemo, megatron, deepspeed, slurm, gradio, vllm, fastapi, lm-evaluation-harness, accelerate, transformers.
- Languages** Spanish (Native), Catalan (Native), English (Fluent), Portuguese (Basic).

Volunteering Activities

- Mentor for international exchange students** *UPV*
- MENTOR *2022*

Assisting international exchange students in their adaptation to the university and the city of Valencia. As part of the university's MENTOR support program: <https://www.upv.es/entidades/OPII/infoweb/pi/info/820230normali.html>

- Mentor for new university students** *UPV*
- PATU *2019-2021*

Advising new Data Science students in their transition to the university, as part of PATU (Programa de Acción Tutorial): <https://www.upv.es/perfiles/futuro-alumno/integra-patu-es.html>