

RICARDO MUÑOZ SÁNCHEZ

Gothenburg, Sweden | P: +46 73 704 77 49 | ricardo.munoz.sanchez@svenska.gu.se
[linkedin.com/in/ricardo-muñoz-sánchez/](https://www.linkedin.com/in/ricardo-muñoz-sánchez/) | github.com/rimusa

EDUCATION

THE UNIVERSITY OF GOTHENBURG

Doctoral Student in Natural Language Processing

Gothenburg, Sweden

Oct 2021 - Present

- Working on bias and fairness in educational NLP. Supervisors: Elena Volodina, Simon Dobnik.
- Took courses in deep learning for natural language processing, argumentation mining, research ethics, and academic communication.

THE UNIVERSITY OF EDINBURGH

Master of Science in Artificial Intelligence (Merit)

Edinburgh, United Kingdom

Sep 2019 - Aug 2020

- Earned the CONACYT International Scholarship – Central Region 2019 by the CONACYT (Sep 2019 – Sep 2020).
- Dissertation project: Exploring the Relationship Between Intrinsic and Extrinsic Bias Metrics in Spanish Word Embeddings. Advisors: Seraphina Goldfarb-Tarrant, Adam Lopez. ([link](#))
- Other projects: implemented a network for multimodal sentiment analysis; surveyed the current state of computational linguistics regarding Ancient Egyptian.
- Took courses in data science, machine learning, neural networks, Bayesian neural networks, reinforcement learning, deep learning, natural language processing (NLP), natural language understanding (NLU), natural language generation (NLG), game theory, and computational cognitive neuroscience.

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO (UNAM)

Bachelor of Science in Mathematics

Mexico City, Mexico

Aug 2012 - Sep 2018

- Final Grade: 8.93/10.
- Earned the TELMEX Scholarship for Academic Excellence by the Carlos Slim Foundation (Aug 2011 – Jun 2016).
- Dissertation project: Studied the order type of maximal chains of subgraphs of the Rado graph that are isomorphic to it. Advisor: David Meza Alcántara. ([link](#))
- Took courses in calculus, linear algebra, logic, probability, statistics, mathematical analysis, and abstract algebra.

WORK EXPERIENCE

THE UNIVERSITY OF GOTHENBURG

Doctoral Studentship

Gothenburg, Sweden

Oct 2021 - Present

- Working on bias and fairness in educational NLP. Supervisors: Elena Volodina, Simon Dobnik.
- Taught courses on introduction to data science and AI, machine learning for statistical AI, basic skills for language technology, and representations of meaning.

APLI (20+ employee recruitment solutions startup; [website](#))

Data Scientist

Mexico City, Mexico

Apr 2021 – Dec 2021

- Developed ML and NLP models that provide insights about job applicants to the recruiters of the company's clients.
- Created processes and procedures to measure and mitigate biases, as well as to improve the explainability of the models the company offers.
- Provided input on how to improve the natural language understanding (NLU) of the company's Facebook bot, its flagship product.

REDUCE SOLUCIONES (10+ employee renewable energy startup; [website](#))

Data Science Advisor

Mexico City, Mexico

Sep 2020 – Present

Director of IT

Jul 2016 – Aug 2019

Co-Founder

Jul 2016

- Formulated 23 different processes and procedures for the company by coordinating with areas such as marketing, operations, HR, among others.
- Coordinated the setup of the company's computing systems, enabling remote tracking the performance of installed systems for clients, which doubled the amount of sales.
- Streamlined the design process of photovoltaic and solar thermal systems by implementing machine learning solutions, reducing the time needed for these projects by over 70%.

- The objective of the mission is sending a swarm of nine self-assembling robots to the moon and is expected to launch in 2021.
- Provided input on the initial design of the robots through preliminary research and discussions with a multidisciplinary team.
- Advised on the viability of the project by simulating and illustrating that the robots had desirable clustering patterns over 85% of the time.

RESEARCH AND PUBLICATIONS

FIRST AUTHOR PAPERS

- **Ricardo Muñoz Sánchez**, Simon Dobnik, Elena Volodina. “*Harnessing GPT to Study Second Language Learner Essays: Can We Use Perplexity to Determine Linguistic Competence?*”. To appear at the BEA 2024 Workshop, co-located with NAACL 2024.
- **Ricardo Muñoz Sánchez**. “*When Hieroglyphs Meet Technology: A Linguistic Journey through Ancient Egypt Using Natural Language Processing*”. To appear at the LT4HALA 2024 Workshop, co-located with LREC-COLING 2024.
- **Ricardo Muñoz Sánchez**, Simon Dobnik, Maria Irena Szawerna, Therese Lindström Tiedemann, Elena Volodina. “*Did the Names I Used within My Essay Affect My Score? Diagnosing Name Biases in Automated Essay Scoring*”. CALD-Pseudo Workshop, co-located with EACL 2024. ([link](#))
- **Ricardo Muñoz Sánchez***, Eric Johansson*, Shakila Tayefeh*, Shreyash Kad*. “*A First Attempt at Unreliable News Detection in Swedish*”. Rest-UP 2 Workshop, co-located with LREC 2022. ([link](#))
- Seraphina Goldfarb-Tarrant*, Rebecca Marchant*, **Ricardo Muñoz Sánchez***, Mugdha Pandya*, Adam Lopez. “*Intrinsic Bias Metrics Do Not Correlate with Application Bias*”. ACL-IJCNLP 2021. ([link](#))

* equal contributions

OTHER PAPERS

- Maria Irena Szawerna, Simon Dobnik, **Ricardo Muñoz Sánchez**, Therese Lindström Tiedemann, Elena Volodina. “*Detecting Personal Identifiable Information in Swedish Learner Essays*”. CALD-Pseudo Workshop, co-located with EACL 2024. ([link](#))
- Dimitrios Kokkinakis, **Ricardo Muñoz Sánchez**, Mia-Marie Hammarlin. “*Scaling-up the Resources for a Freely Available Swedish VADER (svVADER)*”. NoDaLiDa 2023. ([link](#))
- Dimitrios Kokkinakis, **Ricardo Muñoz Sánchez**, Sebastianus Bruinsma, Mia-Marie Hammarlin. “*Investigating the Effects of MWE Identification in Structural Topic Modelling*”. 19th Workshop on Multiword Expressions, co-located with EACL 2023. ([link](#))

OTHER PUBLICATIONS

- Elena Volodina, David Alfter, Simon Dobnik, Therese Lindström Tiedemann, **Ricardo Muñoz Sánchez**, Maria Irena Szawerna, Xuan-Son Vu (Eds.). “*Proceedings of the Workshop on Computational Approaches to Language Data Pseudonymization (CALD-pseudo 2024)*”. Association for Computational Linguistics, St. Julian's, Malta. ([link](#))
- Stian Rødven-Eide, **Ricardo Muñoz Sánchez**. “*Detecting fake papers with the latent algorithm for recursive search*” in Elena Volodina, Dana Dannélls, Aleksandrs Berdicevskis, Markus Forsberg, Shafqat Virk (Eds.) *LIVE and LEARN - Festschrift in honor of Lars Borin* (2022). ([link](#))
- Tanja Seppälä, **Ricardo Muñoz Sánchez**, Márton András Tóth. “*Neighbours with different practices - Doctoral education in Finland and in Sweden in the linguistics*” in *Kielingua* (December 10th, 2021). ([link](#))

CERTIFICATIONS AND TRAINING

DATA CAMP

- Data Analysis with Python (2018), Python Programming (2018)

COURSERA

- Miracles of Human Language: An Introduction to Linguistics (2021)

ACTIVITIES

CODE AGAINST HATE

Winning Team ([presentation](#))

Online
Sep 2020

- Challenge: propose a way to reduce online hate speech without censoring freedom of speech in 48 hours.
- Proposed Hate.no, a plugin that would alert a user if a message could be hate speech, providing an opportunity to change said message before posting.
- Assessed and determined the backend architecture of the app through discussions with the rest of the technical team and facilitated communication between the technical and the conceptual teams.
- Reviewed the cost of moderating comment sections and determined that this solution would slash moderation costs by at least 20%.

HELLO WORLD HACKATHON 2020

Runner-up of the Quantum Black challenge

Edinburgh, United Kingdom
Feb 2020

- Challenge: set up a pipeline for a ML project with Quantum Black's kedro package in less than 24 hours.
- Proposed an AI system that predicts whether an epidemic of an airborne disease will turn into a pandemic.
- Examined possible concepts and recommended creating a new dataset (with less than 30 datapoints) by brainstorming with the rest of the team.
- Implemented the pipeline and the neural network through pair programming, achieving 60% accuracy.

TECHNICAL SKILLS

- Python (Expert). Packages mastered include: pandas, Numpy, scikit-learn, xgboost, PyTorch, NLTK, Gensim, spaCy, matplotlib, seaborn, jupyter.
- Git / Github (Expert)
- SQL (Advanced)
- C++ (Advanced)

LANGUAGES

- Spanish (Native Speaker)
- English (IELTS 7.5; C1 equivalent)
- French (DELF B1)