

Juan Vásquez

(+52) 55-13-63-64-26 | juanmvs@proton.me | [Google Scholar](#) | [Github](#) | [Website](#)

EDUCATION

MS. in Computer Science with specialization in Artificial Intelligence August 2021 – Present
Applied Mathematics and Systems Research Institute (IIMAS) at UNAM Mexico City, Mexico

Thesis: *Feature engineering with λ -calculus for language models. (provisional)*

Adviser: Dr. Gemma Bel Enguix

BS. in Engineering Physics January 2015 – December 2019
Universidad Iberoamericana Mexico City, Mexico

Graduation project: *Data mining: Searching for Chaos in Polynomial Ordinary Differential Equations*

Adviser: Prof. Salvador Carrillo Moreno

NON-FORMAL EDUCATION

Summer school on Neurosymbolic Programming July 11 – 13, 2022
California Institute of Technology Pasadena California, USA

Data Science Workshop (six two-hour sessions) March – April, 2018
Universidad Iberoamericana Mexico City, Mexico

PEER-REVIEWED MANUSCRIPTS

Gemma Bel-Enguix¹, Gerardo Sierra¹, Helena Gómez-Adorno², Juan-Manuel Torres-Moreno³, Jesús-German Ortíz-Barajas⁴, and **Juan Vásquez**⁴. *Overview of PARMEX at Iberlef 2022: Paraphrase Detection in Spanish Shared Task, Procesamiento del Lenguaje Natural 69 (2022).*

Juan Vásquez, Gemma Bel-Enguix, Scott Thomas Andersen, Sergio-Luis Ojeda-Trueba. 2022. *HeteroCorpus: A Corpus for Heteronormative Language Detection*. In *Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP)*, pages 225–234, Seattle, Washington. Association for Computational Linguistics.

Gemma Bel-Enguix, Helena Gómez-Adorno, Karla Mendoza, Grigori Sidorov, and **Juan Vásquez**. *La #felicidad en Twitter: ¿qué representa realmente?*. *Linguamática* 14, no. 1 (2022): 3-15.

Juan Vásquez, Helena Gómez-Adorno, and Gemma Bel-Enguix. *Bert-based Approach for Sentiment Analysis of Spanish Reviews from TripAdvisor*. In *IberLEF at SEPLN*, pp. 165-170. 2021.

WORK IN PROGRESS

Gemma Bel-Enguix¹, Gerardo Sierra¹, Sergio-Luis Ojeda-Trueba¹, Helena Gómez-Adorno², **Juan Vásquez**³, and Scott Thomas Andersen³. *Overview of HOMO-MEX: at Iberlef 2023: Hate speech detection in Online Messages directed towards the MEXican spanish speaking LGBTQ+ population Shared Task*, *Procesamiento del Lenguaje Natural* (2023).

Juan Vásquez, Scott Thomas Andersen, Gemma Bel-Enguix, Braulio Santiago. *How Homo(phobic) is Mexico? Homo-MEX: A Novel Corpus for LGBT+Phobia Detection in Mexican Spanish*. 2023.

TEACHING EXPERIENCE

Introduction to NLP with Python (fifteen two-hour sessions) Mexico City, Mexico
Language Engineering Group (Grupo de Ingeniería Lingüística) at UNAM September 2021 – July 2022

- Created all the materials for the course.
- Presented each one of the fifteen sessions.
- Assisted the students through the course.

EMPLOYMENT

University of Arizona

External Consultant - NLP Engineer

March 2023 – currently

Mexico City, Mexico

- Performed automatic classification experiments on a corpus of crime-related news from Honduras, El Salvador, and Guatemala.
- Fine-tuned and evaluated different pre-trained multilingual language models with a manually annotated corpus of crime-related news from Mexico.
- Wrote various scripts for data cleaning and data pre-processing.

Beet Consulting

Data Consultant

June 2018 – January 2021

Mexico City, Mexico

- Wrote all the necessary code for the data pre-processing, data analysis, and data visualizations requested by the clients.
- Presented slides to various executives in order to help them have a better understanding of their HR departments.

GMEC Consulting

Business Analyst

October 2014 – July 2016

Mexico City, Mexico

- Contributed to the creation of various presentations for clients.
- Researched relevant data for various projects.

OTHER ACADEMIC ACTIVITIES

Reviewer

August 2022

- EasyChair paper subreview | Mexican International Conference on Artificial Intelligence 2022.

OTHER PROFESSIONAL ACTIVITIES

ABIOGENESIS Project | Fundación Cisneros

January 2021 – August 2021

Software Developer

- Planned, designed and worked on the computational aspects of the ABIOGENESIS project. The goal of this project was to digitize a custom painting, detect and modify its components using a machine learning algorithm, and create a new digital image that could be shown in an art installation that explored the intersection between art and artificial intelligence.
- Supervised the manual annotation of an image dataset, in order to train a custom object detection model.
- Wrote code that performed object detection on the digitized image created from the original painting. After the object detection, the objects changed their RGB values in order to simulate change in their physical properties.

VOLUNTEER EXPERIENCE

Plural Ibero | Universidad Iberoamericana

August 2017 – July 2018

- Helped organize, conduct, and moderate various talks for the *human diversity week* - five days in which the University allotted the Organization a budget that allowed us to create, promote, and execute various events that would create consciousness of the particular struggles faced by the LGBT+ population.
- Organized and conducted a weekly queer theory reading group with the members of the student association. Each session I would propose a reading, and I would encourage discussions around it.

SKILLS

Programming languages: Python (advanced), Java (coursework), C++ (coursework), R (coursework), Haskell (coursework), Coq (master's thesis work)

Libraries: pandas, NumPy, Matplotlib, spaCy, openCV, NLTK.

Frameworks: TensorFlow, PyTorch.

Spoken languages: Spanish (native), English (advanced), French (intermediate).