

Juan Vásquez

+1 (720) 668-7342 | [juan\[at\]vasqueza.com](mailto:juan[at]vasqueza.com) | [Google Scholar](#) | [Github](#) | [Website](#)

EDUCATION

Ph.D. in Computer Science

University of Colorado Boulder

Aug 2023 – Present

Boulder, USA

Adviser: Dr. Maria Leonor Pacheco

Co-adviser: Dr. Alexis Palmer

MS. in Computer Science with specialization in AI

Applied Mathematics and Systems Research Institute (IIMAS) at UNAM

Aug 2021 – June 2023

Mexico City, Mexico

Thesis: *Detection of digital LGBT+phobia in Mexican Spanish with neural learning algorithms and a formal-semantics-based approach using lambda calculus*

Adviser: Dr. Gemma Bel-Enguix

Co-adviser: Dr. Karla Ramírez-Pulido

BS. in Engineering Physics

Universidad Iberoamericana

Jan 2015 – Dec 2019

Mexico City, Mexico

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

Adviser: Dr. Salvador Carrillo-Moreno

NON-FORMAL EDUCATION

The 34th European Summer School on Language, Logic and Information

University of Ljubljana, Faculty of Computer and Information Science

July 31st – Aug 11st 2023

Ljubljana, Slovenia

The 13th Lisbon Machine Learning School

Instituto Superior Técnico

July 14th – 20th, 2023

Lisbon, Portugal

Summer school on Neurosymbolic Programming

California Institute of Technology

July 11st – 13rd, 2022

Pasadena, CA, USA

Data Science Workshop (six two-hour sessions)

Universidad Iberoamericana

March – April, 2018

Mexico City, Mexico

PEER-REVIEWED MANUSCRIPTS

Gómez-Adorno, Helena, Gemma Bel-Enguix, Hiram Calvo, Sergio Ojeda-Trueba, Scott Thomas Andersen, **Juan Vásquez**, Tania Alcántara, Miguel Soto, and Cesar Macias. “[Overview of HOMO-MEX at IberLEF 2024: Hate Speech Detection Towards the Mexican Spanish speaking LGBT+ Population](#)”. *Procesamiento del Lenguaje Natural* 73 (2024): 393-405.

Ginn, Michael, Ali Marashian, Bhargav Shandilya, Claire Post, Enora Rice, **Juan Vásquez**, Marie Mcgregor, Matthew Buchholz, Mans Hulden, and Alexis Palmer. “[On the Robustness of Neural Models for Full Sentence Transformation](#)” *Proceedings of the 4th Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP 2024)*, pp. 159-173. 2024.

Andersen, Scott, Segio-Luis Ojeda-Trueba, **Juan Vásquez**, and Gemma Bel-Enguix. “[The Mexican Gayze: A Computational Analysis of the Attitudes towards the LGBT+ Population in Mexico on Social Media Across a Decade](#)”. *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*, pp. 178-200. 2024.

Bel-Enguix, Gemma, Helena Gómez-Adorno, Gerardo Sierra, **Juan Vásquez**, Scott Thomas Andersen, and Sergio Ojeda-Trueba. “[Overview of HOMO-MEX at Iberlef 2023: Hate speech detection in Online Messages directed Towards the MEXican Spanish speaking LGBTQ+ population.](#)” *Procesamiento del Lenguaje Natural* 71 (2023): 361-370.

Osorio, Javier, and **Juan Vásquez**. “[Classifying Organized Criminal Violence in Mexico using ML and LLMs.](#)” *Proceedings of the 6th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text*, pp. 1-10. 2023.

Vásquez, Juan, Scott Andersen, Gemma Bel-Enguix, Helena Gómez-Adorno, and Sergio-Luis Ojeda-Trueba. “[HOMO-MEX: A Mexican Spanish Annotated Corpus for LGBT+phobia Detection on Twitter](#)”. *The 7th Workshop on Online Abuse and Harms (WOAH)*, pp. 202-214. 2023.

Bel-Enguix, Gemma, Gerardo Sierra, Helena Gómez-Adorno, Juan-Manuel Torres-Moreno, Jesus-German Ortiz-Barajas, and **Juan Vázquez**. “Overview of PAR-MEX at Iberlef 2022: Paraphrase Detection in Spanish Shared Task.” *Procesamiento del Lenguaje Natural* 69 (2022): 255-263.

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

Vázquez, Juan, Gemma Bel-Enguix, Scott Thomas Andersen, and Sergio-Luis Ojeda-Trueba. “HeteroCorpus: A Corpus for Heteronormative Language Detection.” *Proceedings of the 4th Workshop on Gender Bias in Natural Language Processing (GeBNLP)*, pp. 225-234. 2022.

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

TEACHING EXPERIENCE

CSCI 4622: Machine Learning - Teaching Assistant ♡†‡ Spring 2025
University of Colorado Boulder - Department of Computer Science Boulder, CO, USA

CSCI 3155: Principles of Programming Languages - Teaching Assistant ♡†‡ Fall 2024
University of Colorado Boulder - Department of Computer Science Boulder, CO, USA

CSCI 5832: Natural Language Processing - Teaching Assistant †‡ Spring 2024
University of Colorado Boulder - Department of Computer Science Boulder, CO, USA

CSCI 1300: Starting Computing - Teaching Assistant ♡†‡ Fall 2023
University of Colorado Boulder - Department of Computer Science Boulder, CO, USA

Introduction to Programming with Python - Main Lecturer ♣ Spring 2023
Prepa Ibero Mexico City, Mexico

- Taught the materials through the semester to two groups of around 25 students each.
- Graded the students.

Introduction to NLP with Python - Main Lecturer ♣ Sept 2021 – July 2022
Language Engineering Group (Grupo de Ingeniería Lingüística) at UNAM Mexico City, Mexico

- Presented each one of the fifteen sessions.
- Assisted the students through the course.

♡ Assisted in teaching the course materials in weekly recitations

† Assisted in creating the course materials

‡ Hosted weekly office hours

♣ Created all the course materials

EMPLOYMENT

University of Arizona March 2023 – Aug 2023
Consultant - NLP Engineer Mexico City, Mexico

- Performed and evaluated various automatic text classification experiments on a *corpus* of crime-related news from Honduras, El Salvador, and Guatemala.
- Wrote various scripts for data cleaning and data pre-processing.
- Published a paper with the findings from this project.

Beet Consulting June 2018 – Jan 2021
Data Consultant 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 Oct 2014 – July 2016
Business Analyst Mexico City, Mexico

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

OTHER ACADEMIC ACTIVITIES

Conference organizer 14th out in STEM (oSTEM) conference Portland, OR, USA	Oct 17th-28th, 2024
Workshop organizer 2024 Queer in AI Workshop at NAACL ^o Mexico City, Mexico	June 2024
Reviewer 2024 MICA ⁱ * Puebla, Mexico	May 2024
Reviewer 2024 Annual Conference of the NAACL ^o Mexico City, Mexico	May 2024
Reviewer 2024 LREC-COLING ^o Torino, Italy	Jan 2024
Keynote Talk 2023 Queer in AI Workshop at ACL ⁺ Toronto, Canada	July 2023
• Title: <i>The State of NLP in Latin America</i>	
Reviewer 2022 MICA ⁱ * Monterrey, Mexico	Oct 2022
First place 2021 Rest-Mex co-located at Iberian Languages Evaluation Forum (IberLef) Málaga, Spain	July 2021
• Shared task - <i>Sentiment Analysis for Mexican Touristic Places</i>	
Talk The Language Engineering Group (GIL), UNAM Mexico City, Mexico	May 2021
• Title: <i>Applied Artificial Intelligence for Sentiment Analysis of Online Reviews in Spanish</i>	

^o The International Conference on Computational Linguistics, Language Resources and Evaluation

^o Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics

⁺ The Annual Meeting of the Association for Computational Linguistics

^{*} Mexican International Conference on Artificial Intelligence

OTHER PROFESSIONAL ACTIVITIES

Research Assistant	May 2024 – Aug 2024
<i>University of Colorado Boulder - Department of Computer Science</i>	<i>Boulder, CO, USA</i>

- Parsed a *corpus* with *tweets* using an Abstract Meaning Representation (AMR) tool, to generate semantic representations of the *tweets*.
- Trained various Graph Neural Networks (GNNs) with the generated AMRs.
- Implemented a novel neurosymbolic architecture combining AMR features with GNNs.

ABIOGENESIS Project Fundación Cisneros	Jan 2021 – Aug 2021
<i>Software Developer</i>	

- 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 changes in their physical properties.

VOLUNTEER EXPERIENCE

Plural Ibero Universidad Iberoamericana	Aug 2017 – July 2018
<ul style="list-style-type: none">• Helped organize, conduct, and moderate various talks for the <i>human diversity week</i> - 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 LGBTQ+ 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), Scala (coursework), Coq (master's thesis work).

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

Frameworks: TensorFlow, PyTorch.

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