

# Danae Sánchez Villegas

Email: [danaesavi@gmail.com](mailto:danaesavi@gmail.com)  
Website: [danaesavi.github.io](https://danaesavi.github.io)  
LinkedIn: [danae-sanchez-villegas](https://www.linkedin.com/in/danae-sanchez-villegas)  
GitHub: [github.com/danaesavi](https://github.com/danaesavi)  
Phone: (+44) 7517605321

## EDUCATION

---

### The University of Sheffield

Ph.D. in Computer Science. Supervisor: Dr. Nikolaos Aletras

Sheffield, UK

2019–Present

- Natural Language Processing Group (SheffieldNLP)
- Focus: multimodal social media analysis, data collection, language modelling
- Expected Graduation: September 2023

### The University of Sheffield

M.Sc. in Computer Science, (Distinction). Supervisor: Professor Rob Gaizauskas

Sheffield, UK

2018–2019

- Dissertation: Machine Comprehension Using Commonsense Knowledge

### Instituto Tecnológico Autónomo de México

B.Sc. in Computer Engineering, GPA. 8.91/10. Supervisor: Dr. Fernando Esponda

Mexico City, MEX

2012–2016

- Dissertation: Time Series Modeling for Stock Market Prediction Using Machine Learning

## PROFESSIONAL EXPERIENCE

---

### Amazon

Applied Scientist Intern at Alexa Shopping Science UK

London, UK

Jun 2022 – Dec 2022

### University of Sheffield – Department of Computer Science

Teaching Assistant

Sheffield, UK

Sep. 2020 – Now

- Text Processing COM 3110/4115 (Autumn 2020, 2021).
- Natural Language Processing COM6513 (Spring 2020, 2021, 2022, 2023).

### Deep Dive Data Science

Data Scientist

Mexico City, MEX

Mar 2017 – Sep 2018

- Developed data science solutions for financial, insurance and real estate companies using natural language processing, social media analysis and geospatial analysis.

### Resuelve

Digital Marketing Trainee

Mexico City, MEX

Nov 2016 – Feb 2017

- Developed a system to help analysts create reports that support their decision-making.

### Nanyang Technological University

Research Intern with the SenticNet Team

Singapore, SGP

Summer 2016

- Explored sentiment analysis models using data in English and Spanish.

## SKILLS

---

- **NLP/ML Software:** PyTorch, Tensorflow, HuggingFace Transformers, NLTK, Keras
- **Programming:** Python, Bash, Java, SQL

## LANGUAGES

---

- **English** Fluent
- **Spanish** Native

## TEACHING & ACADEMIC ACTIVITIES

---

- **Teaching Assistant** Assist students during practical sessions and grade assessments following a marking criteria set by the Course Leader. 2020–Now
- **Reviewer** ACL, NACL, EACL, ARR, NLLP Workshop, W-NUT Workshop 2020–Now
- **MSc. Dissertation Co-Supervisor.** Co-supervised Xiao Ao (with Daniel Preotiuc-Pietro and Nikolaos Aletras), a MSc. student from Sheffield University with the project: *Combining Humor and Sarcasm for Improving Political Parody Detection* 2022
- **Equality, Diversity, and Inclusion Analysis** Data analysis on the gender pay gap and awarding gap in the Computer Science Department of Sheffield University. 2021
- **OxML.** Attended the Oxford Machine Learning Summer School (remote) 2021
- **LxMLS** Attended the Lisbon Machine Learning Summer School (remote) 2020
- **Code First Girls Instructor** Delivered an introduction to Python programming course (remote) for students who identify as a woman or non-binary and are complete beginners in programming. 2020

## INVITED TALKS

---

- **IDS Research Talk, Maastricht University.** NLP for Computational Social Science Apr 2022
- **Annotation hackathon, Utrecht University** Introduction to machine learning and annotations Apr 2022
- **VVA** Multidisciplinary discussion on the impact of influencers on advertising and consumer protection Nov 2021
- **AI4Dignity Counterathon 2021, LMU** Multidisciplinary discussion on datasets and NLP models for extreme speech detection Jul 2021
- **ITAM Research Seminar** Machine Learning Models for Analyzing Social Media Text Mar 2021
- **Sheffield NLP Seminar** NLP for Computational Social Science Dec 2020

## SCHOLARSHIPS AND AWARDS

---

- **Scholarship,** UKRI Centre for Doctoral Training in Speech and Language Technologies 2019–2023  
*Tuition fees and stipend for living expenses during Ph.D. studies at the University of Sheffield.*
- **Scholarship,** Consejo Nacional de Ciencia y Tecnología (Conacyt) 2018–2019  
*Stipend for living expenses during my M.Sc. at the University of Sheffield.*
- **Honorable Mention,** Ex ITAM Research Award 2018  
*Research award for my BSc. dissertation.*

## PUBLICATIONS

---

- [1] E. Gow-Smith and **Sánchez Villegas, Danae**, “Sheffield’s Submission to the AmericasNLP Shared Task on Machine Translation into Indigenous Languages”, in *Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023.
- [2] X. Ao, **Sanchez Villegas, Danae**, D. Preotiuc-Pietro, and N. Aletras, “Combining humor and sarcasm for improving political parody detection”, in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Seattle, United States: Association for Computational Linguistics, Jul. 2022, pp. 1800–1807.
- [3] **Sánchez Villegas, Danae** and N. Aletras, “Point-of-Interest Type Prediction using Text and Images”, in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 7785–7797.

- [4] **Sánchez Villegas, Danae**, S. Mokaram, and N. Aletras, “Analyzing Online Political Advertisements”, in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Online: Association for Computational Linguistics, Aug. 2021, pp. 3669–3680.
- [5] A. Maronikolakis, **Sánchez Villegas, Danae**, D. Preotiuc-Pietro, and N. Aletras, “Analyzing Political Parody in Social Media”, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online: Association for Computational Linguistics, Jul. 2020, pp. 4373–4384.
- [6] **Sánchez Villegas, Danae**, D. Preotiuc-Pietro, and N. Aletras, “Point-of-Interest Type Inference from Social Media Text”, in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*, Online: Association for Computational Linguistics, Dec. 2020, pp. 804–810.