

# Danae Sanchez Villegas, Ph.D.

✉ danaesavi@gmail.com  
in danae-sanchez-villegas  
🌐 <https://danaesavi.github.io/>




## Employment History

- Feb 2024 – Present     **Postdoctoral Researcher**, NLP Section, Copenhagen University (Dr. Desmond Elliott). Copenhagen, DK.  
Researching multimodal models for video understanding and NLP.  
Teaching Natural Language Processing, covering both foundational concepts and state-of-the-art advancements.  
Supervising MSc projects on medical applications of multimodal models.
- Aug. 2023 – Jan 2024     **Research Associate**. University of Sheffield, NLP Group, (Prof. Nikolaos Aletras). Sheffield, UK.  
Researched multimodal models for NLP in political science applications.  
Focused on social media political ads collection, campaign strategy analysis, and large-scale multimodal ad analysis.
- Jun 2022 – Dec 2022     **Applied Scientist Intern**. Amazon, Alexa Shopping. London, UK.  
Developed generative models to improve customer satisfaction prediction.
- Mar 2017 – Sep 2018     **Data Scientist**. Deep dive. Mexico City, MX.  
Developed data science solutions for financial, insurance and real estate companies using NLP, social media analysis and geospatial analysis.

## Education

- 2019 – 2023     **Ph.D. Computer Science, University of Sheffield** Sheffield NLP, UK.  
Thesis: *Beyond Words, Analyzing Social Media with Text and Images*.  
Supervisor: Prof. Nikolaos Aletras.
- 2018 – 2019     **M.Sc. Computer Science (Distinction), University of Sheffield**. Sheffield, UK.  
Thesis: *Machine Comprehension Using Commonsense Knowledge*.  
Supervisor: Prof. Rob Gaizauskas.
- 2012 – 2016     **B.Sc. Computer Engineering, Instituto Tecnológico Autónomo de México**. Mexico City, MX.  
Thesis: *Time Series Modeling for Stock Market Prediction Using Machine Learning Methods*.  
Supervisor: Prof. Fernando Esponda.

## Skills

- |                 |   |
|-----------------|---|
| Coding          |  Python, Java, Bash, SQL.                                      |
| NLP/ML Software |  PyTorch, Tensorflow, HuggingFace Transformers, NLTK, Keras.   |
| Languages       |  English (Fluent), Spanish (Native), Danish (A2), German (A1). |

## Skills (continued)

### Research



**Research & Publishing:** Expertise in NLP, multimodal models, and computational social science.

**Applied Research:** Skilled in translating academic research into practical applications, including multimodal models for political science, medical diagnostics, and customer-focused products.

**Teaching & Mentorship:** Experienced in teaching NLP and supervising MSc research projects.






**Communication:** Delivered talks and workshops at conferences and events.

## Research Publications





1. **Sanchez Villegas D.**, Preotiu-Pietro D, Aletras N.  
*Improving Multimodal Classification of Social Media Posts by Leveraging Image-Text Auxiliary tasks.*  
*Findings of the Association for Computational Linguistics: EACL 2024*, pages 1126–1137, 2024.
2. **Sanchez Villegas D.**, Goanta C., Aletras N.  
*A Multimodal Analysis of Influencer Content on Twitter.* **Area Chair Award – Society & NLP**  
*Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AACL)*, pages 225–240, 2023.
3. Gow-Smith, **Sanchez Villegas D.**  
*Sheffield’s Submission to the AmericasNLP Shared Task on Machine Translation into Indigenous Languages.*  
**First Place Submission**  
*Proceedings of the Workshop on Natural Language Processing for Indigenous Languages of the Americas (AmericasNLP)*, pages 192–199, 2023.
4. Ao X., **Sanchez Villegas D.**, Preotiu-Pietro D, Aletras N.  
*Combining Humor and Sarcasm for Improving Political Parody Detection.*  
*Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 1800–1807, 2022.
5. **Sanchez Villegas D.**, Mokaram S., Aletras N.  
*Analyzing Online Political Advertisements*  
*Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3669–3680, 2021.
6. **Sanchez Villegas D.**, Aletras N.  
*Point-of-Interest Type Prediction using Text and Images.*  
*Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7785–7797, 2021.
7. Maronikolakis A., **Sanchez Villegas D.**, Preotiu-Pietro D, Aletras N.  
*Analyzing Political Parody in Social Media*  
*Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 4373–4384, 2020.
8. **Sanchez Villegas D.**, Preotiu-Pietro D, Aletras N.  
*Point-of-Interest Type Inference from Social Media Text.*  
*Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (AACL-IJCNLP)*, pages 804–810, 2020.

## Professional and Academic Activities




### Awards and Achievements

- 2023      **Area Chair Award Society and NLP**, A multimodal analysis of influencer content on twitter, AACL.
- 2023      **First Place in the AmericasNLP Shared Task**, Sheffield's submission to the AmericasNLP shared task on machine translation into indigenous languages, AmericasNLP.
- 2019-2023      **Scholarship** UKRI Centre for Doctoral Training in Speech and Language Technologies. Tuition fees and stipend for living expenses during Ph.D. studies.
- 2018-2019      **Scholarship** Consejo Nacional de Ciencia y Tecnología. Stipend for living expenses during my M.Sc. at the University of Sheffield
- 2018      **Honorable Mention** Ex ITAM Research Award. Research award for my BSc. dissertation. Stipend for living expenses during my M.Sc. at the University of Sheffield

### International Exchanges and Summer Schools

- 2024      **DisAI Summer School on trustworthy, multilingual and multimodal AI**. Bratislava, SK.
- 2023      **Lisbon Machine Learning Summer School**. Instituto Superior Técnico, Lisbon, PT.
- 2021      **Oxford Machine Learning School** Oxford, UK.
- 2016      **Research Internship at Nanyang Technological University** Singapore, SG.



### Reviewing and Committee duties

- 2024      **LatinX in AI Workshop Co-organizer** Workshop on diverse NLP topics including multilingual and low-resource language models.
- 2024      **Area Chair**: LREC-COLING Social Media Track, 2024
- 2020-Present      **Reviewer**: NAACL, EMNLP, EACL, ACL, AACL.

### Invited Talks

- 2024      **CLASP Seminar, Gothenburg University**. *Beyond Words: Analyzing Social Media with Text and Images*. Gothenburg, SE.
- 2024      **Delve Deep Learning School**. *Multimodal Models for Natural Language Processing*.
- 2024      **Mexican NLP Summer School**. *Exploring Transformers and Limitations in Language Modeling*. Mexico City, MX.
- 2023      **CoAStAL NLP, Copenhagen University**. *Analyzing Social Media with Text and Images*. Copenhagen, DK.
- 2022      **Maastricht University**. *NLP for Computational Social Science*. Maastricht, NL.
- 2022      **Utrecht University**. *Introduction to machine learning and annotations*. Utrecht, NL.
- 2021      **ITAM Research Seminar**. *Machine Learning Models for Analyzing Social Media Text*. Mexico City, MX.
- 2020      **Sheffield NLP Seminar**. *NLP for Computational Social Science*. Sheffield, UK.

### Supervisory Roles

- 2024-Present      **Co-supervision of MSc. thesis** (with Dr. Desmond Elliott): *Challenges in Chest X-Ray Image Classification*. Copenhagen University, DK.
- 2022      **Co-supervision of MSc. thesis** (with Prof. Nikolaos Aletras): *Combining Humor and Sarcasm for Improving Political Parody Detection*. University of Sheffield, UK.