Danae Sánchez

Email: danaesavi@gmail.com
Website: danaesavi.github.io
LinkedIn: danae-sanchez-villegas
GitHub: github.com/danaesavi
Phone: (+45) 50281180

EDUCATION

The University of Sheffield — Sheffield, UK
Ph.D. in Computer Science. Supervisor: Professor Nikolaos Aletras — 2019–2023
- Natural Language Processing Group (SheffieldNLP)
- Focus: multimodal social media analysis, data collection, language modelling

The University of Sheffield — Sheffield, UK
M.Sc. in Computer Science, (Distinction). Supervisor: Professor Rob Gaizauskas — 2018–2019
- Dissertation: Machine Comprehension Using Commonsense Knowledge

Instituto Tecnológico Autónomo de México — Mexico City, MEX
B.Sc. in Computer Engineering, GPA. 8.91/10. Supervisor: Dr. Fernando Esponda — 2012–2016
- Dissertation: Time Series Modeling for Stock Market Prediction Using Machine Learning

PROFESSIONAL EXPERIENCE

University of Copenhagen — Copenhagen, DK
Postdoctoral Researcher at CoASTaL Group, NLP Section — Feb 2024 – Now

University of Sheffield – Department of Computer Science — Sheffield, UK
Research Associate at the Natural Language Processing Group — Aug 2023 – Jan 2024

Amazon — London, UK
Applied Scientist Intern at Alexa Shopping Science UK — Jun 2022 – Dec 2022

Deep Dive Data Science — Mexico City, MEX
Data Scientist — 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.

Nanyang Technological University — Singapore, SGP
Research Intern with the SenticNet Team — 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
- **German** Basic (A1)

Teaching & Academic Activities

- **LatinX in AI Workshop Organizer** Volunteer Chair of LXAI Workshop at NAACL 2024   2024–Now
- **Reviewer** ACL, NACL, EACL, ARR, NLLP Workshop, W-NUT Workshop   2020–Now
- **Teaching Assistant** Assist students during practical sessions and grade assessments following a marking criteria set by the Course Leader.   2020–2023
- **LxMLS Monitor** Tutor students during the lab sessions of the Lisbon Machine Learning School.   2023
- **MSc. Dissertation Co-Supervisor**. Co-supervised Xiao Ao (with Daniel Preoţiuc-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 2021**. Attended the Oxford Machine Learning Summer School   2021
- **LxMLS 2020** Attended the Lisbon Machine Learning Summer School   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

- **CoAStAL NLP, University of Copenhagen**. Analyzing Social Media with Text and Images   Oct 2023
- **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

Publications

[1] **Sánchez Villegas, Danae**, D. Preoţiuc-Pietro, and N. Aletras, "A multimodal analysis of influencer content on twitter", in *EACL (accepted)*, Association for Computational Linguistics, 2024.

[2] 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 (AmericasNLP). **First Place***, Association for Computational Linguistics, 2023.

[3] **Sánchez Villegas, Danae**, C. Goanta, and N. Aletras, "A multimodal analysis of influencer content on twitter", in *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. **Area Chair Award Society and NLP***, Association for Computational Linguistics, 2023.

[4] X. Ao, **Sánchez 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*, Association for Computational Linguistics, 2022.

[5] **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*, Association for Computational Linguistics, 2021.

[6] **Sánchez Villegas, Danae**, S. Mokaram, and N. Aletras, "Analyzing Online Political Advertisements", in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, Association for Computational Linguistics, 2021.

[7] 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*, Association for Computational Linguistics, 2020.

[8] **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*, Association for Computational Linguistics, 2020.

Scholarships and Awards

- **Area Chair Award Society and NLP** for the paper: *A multimodal analysis of influencer content on twitter* in AACL   2023
- **First Place** in the AmericasNLP Shared Task described in the paper: *Sheffield's submission to the AmericasNLP shared task on machine translation into indigenous languages*   2023
- **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