

Fernando Alva Manchego

NATURAL LANGUAGE PROCESSING RESEARCHER

✉ f.alva@sheffield.ac.uk | 🏠 feralvam.github.io | 📧 feralvam | 🌐 feralvam

Education

University of Sheffield

PH.D. IN COMPUTER SCIENCE

Sheffield, UK

Sep. 2016 - Present

- Research Topic: Automatic Text Simplification
- Supervisor: Prof. Lucia Specia. Co-Supervisor: Dr. Carolina Scarton

University of São Paulo

M.Sc. IN COMPUTER SCIENCE

São Carlos, SP, Brazil

Jan. 2013

- Dissertation Title: Semi-Supervised Semantic Role Labeling for Brazilian Portuguese
- Supervisor: Prof. Dr. João Luís Garcia Rosa

Pontifical Catholic University of Peru

B.Sc. IN INFORMATICS ENGINEERING

Lima, Peru

Dec. 2008

- Dissertation Title: Information System for Plagiarism Detection in Digital Documents using Document Fingerprinting
- Graduated ranked 3 of 88

Experience

RESEARCH EXPERIENCE

University of Sheffield – Natural Language Processing Research Group

Sheffield, UK

POSTGRADUATE RESEARCH STUDENT

Sep. 2016 - Present

- Implemented novel algorithms to annotate word-level transformations performed in aligned original-simplified sentences.
- Collaborated in the implementation of a sentence simplification model based on sequence labelling.
- Implemented a software package to standardise automatic evaluation of sentence-level simplification output.
- Identified and quantified cross-sentence text transformations in professionally-simplified documents.
- Designed the data collection protocol to crowdsource manual simplifications involving multiple rewriting transformations.
- Collaborated in the creation, curation and analysis of a new dataset for assessing sentence simplification in English.
- Collected (via crowdsourcing) human judgements on the quality of automatic simplifications, using direct assessment.
- Implemented a multitask sentence simplification model that leverages data from related tasks to jointly perform multiple types of text transformations, namely: lexical paraphrasing, sentence splitting, and compression.
- Currently studying the capabilities of automatic metrics for assessing the quality of automatic sentence simplifications in scenarios with varying text rewriting transformations.

Amazon Research

Cambridge, UK

APPLIED SCIENTIST INTERN

Aug. 2018 - Nov. 2018

- Implemented a distant-supervision algorithm to annotate a large-scale resource with Entity labels.
- Implemented a multitask sequence-labelling model for Coreference Resolution and Named Entity Recognition.
- Explored multitask learning for Entity Linking, based on joint Named Entity Recognition, Coreference Resolution and Entity Disambiguation.

TEACHING EXPERIENCE

University of Sheffield – Department of Computer Science

Sheffield, UK

GRADUATE TEACHING ASSISTANT

Feb. 2017 - Jun. 2018

- Assisted students during the lab sessions of the Text Processing and Natural Language Processing modules.
- Marked lab reports following a pre-determined scheme set by the Course Leader.
- Designed and coordinated the course project for the Natural Language Processing module, which resulted in a publication in a top-tier conference.

- Course Leader for: Introduction to Computing, Algorithms, Programming Languages, and Text Mining.
- Planned lessons, developed teaching materials, designed assessment methodologies, and marked exams.
- Supervised 7 Informatics Engineering students on their final year dissertations. Of those, 2 projects resulted in publications in international conferences.

Publications

JOURNAL ARTICLES

- ALVA-MANCHEGO, FERNANDO**, C. SCARTON, L. SPECIA. 2020. *Data-Driven Sentence Simplification: Survey and Benchmark*. *Computational Linguistics* 46.1 (Mar. 2020), pp. 135–187
- I. VARGAS-CAMPOS, **ALVA-MANCHEGO, FERNANDO**. 2016. *SciEsp: Structural Analysis of Abstracts Written in Spanish*. *Computación y Sistemas* 20.3 (2016), pp. 551–558

CONFERENCE PAPERS

- ALVA-MANCHEGO, FERNANDO**, L. MARTIN, A. BORDES, C. SCARTON, B. SAGOT, L. SPECIA. 2020. *ASSET: A Dataset for Tuning and Evaluation of Sentence Simplification Models with Multiple Rewriting Transformations*. Accepted to ACL 2020
- ALVA-MANCHEGO, FERNANDO**, L. MARTIN, C. SCARTON, L. SPECIA. 2019. *EASSE: Easier Automatic Sentence Simplification Evaluation*. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP): System Demonstrations*, Hong Kong, China
- ALVA-MANCHEGO, FERNANDO**, C. SCARTON, L. SPECIA. 2019. *Cross-Sentence Transformations in Text Simplification*. *Proceedings of the 2019 Workshop on Widening NLP*, Florence, Italy
- P. FINNIMORE, E. FRITZSCH, D. KING, A. SNEYD, A. UR REHMAN, **ALVA-MANCHEGO, FERNANDO**, A. VLACHOS. 2019. *Strong Baselines for Complex Word Identification across Multiple Languages*. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota
- ALVA-MANCHEGO, FERNANDO**, J. BINGEL, G. PAETZOLD, C. SCARTON, L. SPECIA. 2017. *Learning How to Simplify From Explicit Labeling of Complex-Simplified Text Pairs*. *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Taipei, Taiwan
- G. PAETZOLD, **ALVA-MANCHEGO, FERNANDO**, L. SPECIA. 2017. *MASSAlign: Alignment and Annotation of Comparable Documents*. *Proceedings of the IJCNLP 2017, System Demonstrations*, Taipei, Taiwan
- A. QUISPESARAVIA, W. PEREZ, M. SOBREVILLA CABEZUDO, **ALVA-MANCHEGO, FERNANDO**. 2016. *Coh-Matrix-Esp: A Complexity Analysis Tool for Documents Written in Spanish*. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Portorož, Slovenia
- ALVA-MANCHEGO, FERNANDO**, J. L. G. ROSA. 2012. *Semantic Role Labeling for Brazilian Portuguese: A Benchmark*. *Proceedings of the 13th Ibero-American Conference on Artificial Intelligence (IBERAMIA 2012)*, Berlin, Heidelberg
- ALVA-MANCHEGO, FERNANDO**, J. L. G. ROSA. 2012. *Towards Semi-supervised Brazilian Portuguese Semantic Role Labeling: Building a Benchmark*. *Proceedings of the 10th International Conference on Computational Processing of Portuguese (PROPOR 2012)*, Berlin, Heidelberg

Scholarships

Faculty of Engineering Doctoral Academy Scholarship

Sheffield, UK

UNIVERSITY OF SHEFFIELD

Oct. 2016 – Sep. 2019

- Tuition fees and stipend for living expenses during Ph.D. studies at the University of Sheffield.

Research Internships Abroad Fellowship (BEPE) - Masters

SÃO PAULO RESEARCH FOUNDATION (FAPESP)

São Paulo, Brazil

Mar. 2012 – Jun. 2012

- Travel expenses, tuition fees and stipend for living expenses while a Visiting Postgraduate Research Student at The University of Edinburgh (supervised by Prof. Mirella Lapata).

Masters Fellowship in the State of São Paulo

SÃO PAULO RESEARCH FOUNDATION (FAPESP)

São Paulo, Brazil

Aug. 2011 – Dec. 2012

- Stipend for living expenses during my M.Sc. at the University of São Paulo.

Masters Fellowship

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES)

Brazil

Sep. 2010 – Jul. 2011

- Stipend for living expenses during my M.Sc. at the University of São Paulo.

Other Academic Activities

Workshop Co-Chair 2019 ACL Student Research Workshop (SRW)

Program Committee/Reviewer ACL, EMNLP, LREC, ACL-SRW, WinNLP, FEVER

Skills

Programming Fairseq, AllenNLP, spaCy, NLTK, PyTorch, Python

Languages Spanish (Native), English (C1), Brazilian Portuguese (Advanced)