

# Fernando Emilio Alva Manchego

NATURAL LANGUAGE PROCESSING RESEARCHER

✉ f.alva@sheffield.ac.uk | 🏠 feralvam.github.io | 📄 feralvam | 🎓 Fernando Alva-Manchego

## Education

---

### University of Sheffield

Sheffield, UK

PH.D. IN COMPUTER SCIENCE

Apr. 2021

- Thesis: *Automatic Sentence Simplification with Multiple Rewriting Transformations*
- Supervisors: Prof. Lucia Specia and Dr. Carolina Scarton

### University of São Paulo

São Carlos, SP, Brazil

M.Sc. IN COMPUTER SCIENCE

Jan. 2013

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

### Pontifical Catholic University of Peru

Lima, Peru

B.Sc. IN INFORMATICS ENGINEERING

Mar. 2009

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

## Experience

---

### RESEARCH EXPERIENCE

#### University of Sheffield – Bergamot

Sheffield, UK

RESEARCH ASSOCIATE IN MACHINE TRANSLATION

Oct. 2020 - Present

- Researching approaches for adequacy-oriented word-level Quality Estimation (QE) of Machine Translation.
- Implementing `deepQuest-py`, a library for deep learning based QE that includes state-of-the-art models for sentence-level and word-level QE based on fine-tuning pre-trained Transformer-based language models.

#### University of Sheffield – APE-QUEST

Sheffield, UK

RESEARCH ASSOCIATE (PART-TIME)

Feb 2020 - Sep. 2020

- Designed and conducted a study on the trade-off between speed, cost and quality of incorporating Quality Estimation models in a computer-aided translation and post-editing workflow in a translation company.

#### University of Sheffield – Natural Language Processing Research Group

Sheffield, UK

POSTGRADUATE RESEARCH STUDENT

Sep. 2016 - Sep. 2020

- Implemented novel algorithms to annotate word-level transformations performed in aligned original-simplified sentences.
- Implemented EASSE, a software package to standardise automatic evaluation of sentence-level simplification output.
- Designed the data collection protocol to crowdsource manual simplifications involving multiple rewriting transformations.
- Collaborated in the creation, curation and analysis of ASSET, a new dataset for assessing Sentence Simplification in English.
- Crowdsourced Simplicity-DA, a dataset with direct assessments on the quality of automatic sentence simplifications.
- Conducted the first meta-evaluation of automatic evaluation metrics for Sentence Simplification.
- Implemented a multitask Sentence Simplification model to jointly perform multiple types of text transformations.

#### Amazon Research

Cambridge, UK

APPLIED SCIENTIST INTERN

Aug. 2018 - Nov. 2018

- 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

MSc. DISSERTATION CO-SUPERVISOR

Feb. 2021 - Present

- Co-supervising (with Dr. Carolina Scarton) a MSc student from the Advanced Computer Science program with the project *Match the Word Count: An Essay Editing Assistant*.

## Research Experience for Peruvian Undergraduates (REPU)

Online

MENTOR

Jan. 2021 - Mar. 2021

- Supervised a visiting student on implementing pipeline models for cross-lingual Text Simplification of COVID-19 information, and evaluating their capabilities using both readability metrics and human judgements.

## 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.

## Pontifical Catholic University of Peru – Department of Engineering

Lima, Peru

LECTURER (TEACHING)

Mar. 2013 - Jul. 2016

- Course Leader for: Introduction to Computing, Algorithms, Programming Languages, and Applications of Computer Science.
- 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

---

Natural Language Processing is a conference-driven field. The main venues with peer-reviewed proceedings are ACL (CORE A\*), EMNLP (CORE A) and NAACL (CORE A). Relevant journals include Computational Linguistics (CORE A\*).

### JOURNAL ARTICLES

**FERNANDO ALVA-MANCHEGO**, CAROLINA SCARTON, LUCIA SPECIA. 2021. *The (Un)Suitability of Automatic Evaluation Metrics for Text Simplification*. *Computational Linguistics* (Accepted 2021)

**FERNANDO ALVA-MANCHEGO**, CAROLINA SCARTON, LUCIA SPECIA. 2020. *Data-Driven Sentence Simplification: Survey and Benchmark*. *Computational Linguistics* 46.1 (Mar. 2020), pp. 135–187

IRVIN VARGAS-CAMPOS, **FERNANDO ALVA-MANCHEGO**. 2016. *SciEsp: Structural Analysis of Abstracts Written in Spanish*. *Computación y Sistemas* 20.3 (2016), pp. 551–558

### CONFERENCE PAPERS

**FERNANDO ALVA-MANCHEGO**, LUCIA SPECIA, SARA SZOC, TOM VANALLEMEERSCH, HEIDI DEPRAETERE. 2021. *Validating Quality Estimation in a Computer-Aided Translation Workflow: Speed, Cost and Quality Trade-off*. In *Proceedings of Machine Translation Summit 2021 Volume 2: MT Users and Providers Track*, (Accepted)

AMIT GAJBHIYE, MARINA FOMICHEVA, **FERNANDO ALVA-MANCHEGO**, FRÉDÉRIC BLAIN, ABIOLA OBAMUYIDE, NIKOLAOS ALETRAS, LUCIA SPECIA. 2021. *Knowledge Distillation for Quality Estimation*. In *Findings of the Association for Computational Linguistics: ACL 2021*, (Accepted)

MOUNICA MADDELA, **FERNANDO ALVA-MANCHEGO**, WEI XU. 2021. *Controllable Text Simplification with Explicit Paraphrasing*. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 3536–3553, Online

KERVY RIVAS ROJAS, **FERNANDO ALVA-MANCHEGO**. 2021. *IAPUCP at SemEval-2021 Task 1: Stacking Fine-Tuned Transformers is Almost All You Need for Lexical Complexity Prediction*. In *Proceedings of SemEval-2021*, (Accepted)

**FERNANDO ALVA-MANCHEGO**, LOUIS MARTIN, ANTOINE BORDES, CAROLINA SCARTON, BENOÎT SAGOT, LUCIA SPECIA. 2020. *ASSET: A Dataset for Tuning and Evaluation of Sentence Simplification Models with Multiple Rewriting Transformations*. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 4668–4679, Online

**FERNANDO ALVA-MANCHEGO**, LOUIS MARTIN, CAROLINA SCARTON, LUCIA SPECIA. 2019. *EASSE: Easier Automatic Sentence Simplification Evaluation*. In *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*, pp. 49–54, Hong Kong, China

**FERNANDO ALVA-MANCHEGO**, CAROLINA SCARTON, LUCIA SPECIA. 2019. *Cross-Sentence Transformations in Text Simplification*. In *Proceedings of the 2019 Workshop on Widening NLP*, pp. 181–184, Florence, Italy

- PIERRE FINNIMORE, ELISABETH FRITZSCH, DANIEL KING, ALISON SNEYD, ANEEQ UR REHMAN, **FERNANDO ALVA-MANCHEGO**, ANDREAS VLACHOS. 2019. *Strong Baselines for Complex Word Identification across Multiple Languages*. In *Proceedings of the 2019 Conference of the North American Chapter of the ACL (NAACL)*, pp. 970–977, Minneapolis, Minnesota
- FERNANDO ALVA-MANCHEGO**, JOACHIM BINGEL, GUSTAVO PAETZOLD, CAROLINA SCARTON, LUCIA SPECIA. 2017. *Learning How to Simplify From Explicit Labeling of Complex-Simplified Text Pairs*. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (IJCNLP)*, pp. 295–305, Taipei, Taiwan
- GUSTAVO PAETZOLD, **FERNANDO ALVA-MANCHEGO**, LUCIA SPECIA. 2017. *MASSAlign: Alignment and Annotation of Comparable Documents*. In *Proceedings of the IJCNLP 2017, System Demonstrations*, pp. 1–4, Taipei, Taiwan
- ANDRE QUISEPESARAVIA, WALTER PEREZ, MARCO SOBREVILLA CABEZUDO, **FERNANDO ALVA-MANCHEGO**. 2016. *Coh-Matrix-Esp: A Complexity Analysis Tool for Documents Written in Spanish*. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pp. 4694–4698, Portorož, Slovenia
- FERNANDO ALVA-MANCHEGO**, JOÃO LUÍS G ROSA. 2012. *Semantic Role Labeling for Brazilian Portuguese: A Benchmark*. In *Proceedings of the 13th Ibero-American Conference on Artificial Intelligence (IBERAMIA 2012)*, pp. 481–490, Berlin, Heidelberg
- FERNANDO ALVA-MANCHEGO**, JOÃO LUÍS G ROSA. 2012. *Towards Semi-supervised Brazilian Portuguese Semantic Role Labeling: Building a Benchmark*. In *Proceedings of the 10th International Conference on Computational Processing of Portuguese (PROPOR 2012)*, pp. 210–217, Berlin, Heidelberg

## Grants

---

### EAMT Sponsorship of Activities

EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION

Jan. 2021 – Dec. 2021

- Project: *Multilingual Translation with Readability-Controlled Output Generation* (Total: €2,678)
- Co-lead with Dr. Matthew Shardlow from Manchester Metropolitan University.

## Scholarships

---

### Faculty of Engineering Doctoral Academy Scholarship

UNIVERSITY OF SHEFFIELD

Sheffield, UK

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 while a Visiting Postgraduate Research Student at The University of Edinburgh.

### 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.

## Invited Talks

---

Jul. 2021 **BeyondMinds**, Multi-Operation Automatic Text Simplification

Online

Mar. 2021 **SheffieldNLP Seminar**, The (Un)Suitability of Automatic Metrics for Sentence Simplification

Online

Nov. 2020 **NLP with Friends**, Challenges in Automatic Evaluation of Text Simplification

Online

## Other Academic Activities

---

<b>Program Committee/Reviewer</b>	ACL, NAACL, EACL, EMNLP, LREC, ACL-SRW, EACL-SRW, AmericasNLP, WinNLP, FEVER
<b>Workshop Co-Chair</b>	2019 ACL Student Research Workshop (SRW)
<b>Student Representative</b>	EACL Student Board (2018-2019)

## Skills

---

<b>NLP/ML Software</b>	Fairseq, HuggingFace Transformers, AllenNLP, spaCy, NLTK, PyTorch
<b>Programming</b>	Python, Shell scripting
<b>Languages</b>	Spanish (Native), English (C1), Brazilian Portuguese (Advanced)