Christopher Davis

chris.davis90@gmail.com

PROFESSIONAL EXPERIENCE

July 22 - Present	**Postdoctoral Researcher, Computer Laboratory, University of Cambridge**
Sep. 20 - Dec. 20	**Applied Scientist Intern, Amazon Alexa AI**
Apr. 20 - Aug. 20	**Machine Learning Consultant**

Apr. 20 - Aug. 20 — **Machine Learning Consultant**
- Worked with researchers from Imperial College London on a project to detect sustainability initiatives in company documents.

July 17 - Sept. 17 — **Machine Learning Research Intern, Computer Laboratory, University of Cambridge**
- Investigated events around the Arab Spring using the GDELT dataset of global news events.
- Collaborated with researchers from CRASSH to identify problems from using an automated event extraction dataset.

Oct. 14 - Aug. 16 — **Software Engineer, AVEVA Solutions Ltd., Cambridge, UK**
AVEVA Solutions Ltd, a FTSE 250 company, is the world's leading engineering software provider to the plant, power and marine industries.
- Working across the full stack in a team of 6 to develop an enterprise integration platform.
- Provided support and training to incoming graduates.
- Led scrum meetings and agile sprint retrospectives.

Oct. 12 - Oct. 14 — **Technical Graduate, AVEVA Solutions Ltd., Cambridge, UK**

Jul. 10 - Jul. 11 — **Professional Training Year, Mondrian Investment Partners Ltd., London**

EDUCATION

2017 - 2023 — **University of Cambridge, PhD in Computer Science (NLP)**
Supervised by Prof. Paula Buttery; funded by an EPSRC Doctoral Training Studentship.
Focused on developing and analysing language models for grammatical error detection, with applications to automated assessment and feedback for second language learners.

2016 - 2017 — **University of Cambridge, MPhil in Advanced Computer Science**
Achieved 1st Class with Distinction (equivalent to 4.0 GPA)
Supervised by Dr. Ekaterina Shutova
Thesis: Modelling Visual Context in Multimodal Semantics
Modules Include: Introduction to NLP, Syntax and Parsing, Machine Learning for NLP, Machine Learning and Algorithms for Data Mining, Biomedical Information Processing.

2008 - 2012 — **University of Surrey, BSc Computer Science w/ Industrial Experience**
Achieved 1st Class Honours (equivalent to 3.8 GPA).

PUBLICATIONS

2024 — **Prompting open-source and commercial language models for grammatical error correction of learner text**
Christopher Davis, Andrew Caines, Øistein Andersen, Shiva Taslimipoor, Helen Yannakoudakis, Zheng Yuan, Christopher Bryant, Marek Rei, Paula Buttery.
arXiv preprint

2023 — **CLIMB – Curriculum Learning for Infant-inspired Model Building**
Richard Diehl Martinez, Zébulon Goriely, Hope McGovern, Christopher Davis, Andrew Caines, Paula Buttery, Lisa Beinborn
*BabyLM Challenge @ CoNLL 2023 -- **"Most Interesting Paper Award"***

2023 — **On the application of Large Language Models for language teaching and assessment technology**
Andrew Caines, Luca Benedetto, Shiva Taslimipoor, Christopher Davis, Yuan Gao, Øistein E. Andersen, Zheng Yuan, Mark Elliott, Russell Moore, Christopher Bryant, Marek Rei, Helen Yannakoudakis, Andrew Mullooly, Diane Nicholls and Paula Buttery.
In Proceedings of AIED 2023 Empowering Education with LLMs: the Next-Gen Interface and Content Generation, to appear

2022	**Probing for targeted syntactic knowledge through grammatical error detection**
	Christopher Davis, Christopher Bryant, Andrew Caines, Marek Rei, Paula Buttery.
	Proceedings of the 26th Conference on Computational Natural Language Learning (CoNLL-2022)
	~20% acceptance rate. Offered oral presentation.

2021 **Multi-class grammatical error detection for correction: A tale of two systems**
Zheng Yuan, Shiva Taslimpoor, Christopher Davis, and Christopher Bryant.
In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP-2021)

2019 **Deconstructing multimodality: visual properties and visual context in human semantic processing**
Christopher Davis, Luana Bulat, Anita Vero, Ekaterina Shutova.
Proceedings of the Eighth Joint Conference on Lexical and Computational Semantics (*SEM 2019)

Accurate modelling of language learning tasks and students using representations of grammatical proficiency.
Ahmed Zaidi, Andrew Caines, Christopher Davis, Russell Moore, Paula Buttery and Andrew Rice.
Educational Data Mining (EDM) 2019

CAMsterdam at SemEval-2019 Task 6: Neural and graph-based feature extraction for the identification of offensive tweets.
Guy Aglionby, Christopher Davis, Pushkar Mishra, Andrew Caines, Helen Yannakoudakis, Marek Rei, Ekaterina Shutova and Paula Buttery.
In Proceedings of the International Workshop on Semantic Evaluation 2019 (SemEval 2019)

2018 **Modelling visual properties and visual context in multimodal semantics.**
Christopher Davis, Luana Bulat, Anita Vero and Ekaterina Shutova
Workshop on Visually Grounded Interaction and Language, NeurIPS 2018, Montreal, Canada

TALKS

2021 Probing grammatical error detection models.
Institute for Automated Language Teaching and Assessment. Cambridge, UK.

2020 Investigating grammatical error detection in context.
Institute for Automated Language Teaching and Assessment. Cambridge, UK.

2019 Improving grammatical error detection for learner English text using first-language specific priors.
Language Sciences Research Symposium for Early-career Researchers. Cambridge, UK.

REVIEWING

AI4Ed @ AAAI 2024; EMNLP 2020, 2022; LREC 2023, 2024; Rep4NLP 2020, 2023; ViGIL 2021

TEACHING

Easter 2020 Supervisor for Formal Models of Language *(Part IB/II)*

Lent 2020 Supervisor, demonstrator, ticker for Machine Learning & Real-World Data (MLRD) *(Part IA/IB)*

Lent 2019 Supervisor, lead demonstrator, ticker for MLRD *(Part IA/IB)*

Jan. 2019 *(External teaching)* Tutor at Cambridge Spark, for the Machine Learning and NLP Modules

Michaelmas 2018 Demonstrating for Natural Language Processing (*Part II*)

Lent 2018 Co-lead demonstrator & ticker for MLRD (*Part IA/IB*)

SKILLS

- Algorithms and data structures
- Deep Learning, Natural Language Processing
- Python, PyTorch, HuggingFace, LLM
- Scikit-learn, C#
- Software Engineering: OOP, Agile, Unit Testing
- Stanford Machine Learning course (Coursera)

LANGUAGES

- English (native, professional).
- Spanish (lower intermediate).
- Polish (learning).