Sian Gooding

www.linkedin.com/in/sian-gooding

✉ siangooding@gmail.com

☎ 07510997252

Education

- **PhD in Computer Science** — *Current position*
  University of Cambridge — *2019 – 2022*
  My PhD focuses on building machine learning models for automatic text simplification, with a particular emphasis on the integration of eye tracking data to create humanistic models of text comprehension.

- **MPhil Advanced Computer Science** — *Distinction (86%)*
  University of Cambridge — *2017 – 2018*
  Ranked in the top 5% of the year and won the Google research thesis award. I specialised in Natural Language Processing, Deep Learning and Machine Learning.

- **BSc in Computer Science** — *First Class Honours (80%)*
  University of Birmingham — *2013 – 2016*
  Elective focus: Machine Learning, Natural Language Processing, Nature Inspired Optimisation, Compilers.

Scholarships & Awards

- **Harding Distinguished Postgraduate Scholarship** — *2019*
  Fully-funded Cambridge Scholarship for the most talented PhD students from the UK and worldwide

- **Google Best Thesis Award** — *2018*
  Award for highest ranked Cambridge Advanced Computer Science MPhil thesis

- **Cambridge Graduate Tutors' Award** — *2018*
  Graduate award for distinguished performance at MPhil

- **ACL Shared Task Winner** — *2018*
  Developed winning complex word identification system in an international shared task

- **Cambridge Computer Laboratory MPhil Scholarship** — *2017*
  One of eight students to receive a 2017 scholarship for the MPhil course

Publications

- Gooding, S., and Kochmar, E. Recursive context-aware lexical simplification. 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)* (2019) *Link*

- Gooding, S., and Kochmar, E. Complex word identification as a sequence labelling task. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (2019) *Link*

- Gooding, S., Kochmar, E., Sarkar, A., and Blackwell, A. Comparative judgments are more consistent than binary classification for labelling word complexity. In *Proceedings of the 13th Linguistic Annotation Workshop* (2019) *Link*

- Gooding, S., and Briscoe, T. Active learning for financial investment reports. In *Proceedings of the Second Financial Narrative Processing Workshop (FNP 2019)* (2019) *Link*

- Gooding, S., and Kochmar, E. CAMB at CWI shared task 2018: Complex word identification with ensemble-based voting. In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications* (2018) *Link*

Professional Experience

- **Data Science Teacher**
  Cambridge Spark *Present*
  Cambridge Spark provides data science training for professionals, I co-teach the Natural Language Processing with Python on the Advanced Data Science course.

- **Research Assistant**
  University of Cambridge *2018*
  My research was focused on active learning techniques for machine learning algorithms, applied to the financial domain. I additionally published papers at multiple top tier venues on text simplification.

- **Technology Graduate Scheme**
  Tesco Plc *2017*
  During the one year graduate scheme I specialised in data science leadership. I worked on a wide variety of machine learning projects, including a placement in Bangalore as a data science engineer.

- **Machine Learning Engineer**
  Lloyd's Bank *2015*
  I spent four months as a machine learning engineer at Lloyd's Bank. During this time, I designed and implemented a system to predict customer energy usage, using semi-supervised learning, which could generated personalised saving strategies for customers.

Selected Talks

Neural Text Simplification - *Cambridge Computer Laboratory TechTalk 2019*

Contextual Text Simplification - *Cambridge Assessment 2019*

Active Learning: Data Quality over Quantity - *Cambridge Data Science Symposium 2018*

Complex Word Identification - *Oxbridge Women in Computer Science 2018*

Academic Roles

- **Conference Reviewer**
  Reviewed for top-ranked international conferences and workshops, including ACL, NAACL and BEA

- **Practical Demonstrator**
  Data Science: principles and practice, Part II CST course, Computer Laboratory, University of Cambridge, 2019
  Overview of Natural Language Processing, MPhil course, Computer Laboratory, University of Cambridge, 2018

- **Course Supervisor**
  Formal Models of Language, Part IB CST course, Computer Laboratory, University of Cambridge 2019

- **Member of Cambridge Computer Lab Faculty board**
  The faculty board oversees the work of the Computer Laboratory, I acted as MPhil representative in 2018

- **Fitzwilliam Entrepreneur Society (FES) Committee**
  FES organises entrepreneurial events and talks for five of the Cambridge Hill Colleges

Programming Skills

- **Languages**: Python, Java, SQL, MATLAB **Technologies**: Hadoop, Android, Tensorflow, PyTorch

References

Professor Ted Briscoe
✉ ejb@cl.cam.ac.uk
Computer Laboratory
University of Cambridge

Dr. Ekaterina Kochmar
✉ ek323@cam.ac.uk
Computer Laboratory
University of Cambridge