Talia Tseriotou

+44-7767650428 • t.tseriotou@qmul.ac.uk
https://www.linkedin.com/in/taliatseriotou • https://github.com/ttseriotou

Research Interests

Natural Language Processing, Temporal Representation Learning, LLMs, Memory-augmented Models, Personalisation, Transformer-based Models, Long-Context Modeling, Computational Social Media.

Education

Queen Mary, University of London **London, UK**
Ph.D. in Computer Science (Natural Language Processing) *Sep 2021 – Dec 2025 (expected)*

- Supervisors: Prof. Maria Liakata, Dr. Adam Tsakalidis
- Research Topic: Representation Learning for Temporal and Personalised User Modeling.
- Funding: DeepMind Ph.D. Studentship in AI (full scholarship)

Columbia University **New York, USA**
M.Sci. in Applied Mathematics *Sep 2015 – Dec 2016*

- Supervisor: Prof. Vineet Goyal
- Thesis: Optimization in Commercializing Seed Varieties
- Scholarship: Cyprus Children's Fund Scholarship

University of Cambridge **Cambridge, UK**
B.A./M.Eng. in Chemical Engineering *Oct 2011 – Jun 2015*

- Supervisors: Prof. Vassilios S. Vassiliadis, Dr. Antonio Del Rio Chanona
- Thesis: Lifting Effect on Global Nonlinear Optimization
- Scholarship: IKY Cyprus Scholarship

A-levels: Mathematics – A*, Further Mathematics – A*, Physics – A*, Chemistry – A* **Cyprus**

Selected Publications

4. **Tseriotou, T.**, Tsakalidis, A., and Liakata, M. (2024). Tempoformer: A transformer for temporally-aware representations in change detection. EMNLP Main Conference 2024.

3. Hills, A., **Tseriotou, T.**, Miscouridou, X., Tsakalidis, A., and Liakata, M. (2024). Exciting mood changes: A time-aware hierarchical transformer for change detection modelling. In Findings of ACL 2024.

2. **Talia Tseriotou**, Ryan Chan, Adam Tsakalidis, Iman Munire Bilal, Elena Kochkina, Terry Lyons, and Maria Liakata. 2024. Sig-networks toolkit: Signature networks for longitudinal language modelling. In Proceedings of EACL: System Demonstrations.

1. **Tseriotou, T.**, Tsakalidis, A., Foster, P., Lyons, T. and Liakata, M., 2023, July. Sequential Path Signature Networks for Personalised Longitudinal Language Modeling. In Findings of ACL 2023.

Professional Experience

Huawei Technologies R&D **London, UK**
Research Intern *Apr 2024 – Aug 2024*

- Supervisor: Dr. Fenia Christopoulou
- Conducted research on memory-augmented generative models for improved training.

Chubb Insurance **London, UK**
Senior AI Data Scientist *Oct 2020 – Aug 2021*

- Supervisor: Dr. Christopher M. Glaze
- Led advanced applied science production projects with visibility to the company's Chief Digital Officer.
- Implemented an NLP web-text 1100-class classification model in Australia that surpasses the agent's performance leveraging custom trained word2vec and name entity detection modules.
- Trained a BERT model with custom modules for advancing the performance of the classification model.
- Key contributions in reinforcing the Bing and Google business search for website matching and html parsing solutions achieving recall: 75% and precision: 95%.
- Implemented a POC on information extraction of year the business was established from its html using XGBoost.

AI Data Scientist *USA & Singapore, Mar 2018 – Oct 2020*

- Built an algorithm in Pyspark from published research which uses Expectation Maximization on Gaussian Mixture Models to impute the company's 135 million database.
- Advanced the algorithm by implementing in the solution a Dirichlet prior.
- Lifted the accuracy of a 1,500-class multi-label classification model by +10% through active learning.
- Derived and applied a novel regularization to penalize large coefficients.
- Advanced the same model to a hierarchical one that accounts for the class structure.
- Researched using simulations an algorithmic extension in the loss function to fix corrupted labels.
- Developed predictive models using random forests for the live pre-fill of 50% of Chubb's applications.
- Incorporated confidence scoring in the process by modeling the error using Gamma Regression.
- Performed a rigorous sensitivity analysis for stakeholders that boosted the live pre-fills by up to +54%.

Media Assembly **New York, USA**
Data Scientist *Apr 2017 – Mar 2018*

- Built the back-end non-linear constraint optimization algorithm of the company for budget allocation.
- Implemented a more computationally efficient algorithm by analytically deriving the gradients of the model and feeding them into the problem – Awarded Analytics Company Project for 2018.
- Advanced the Multi-Touch Attribution model by productionizing published research in bagging and tuning.
- Developed twelve models using multivariate linear regression with subset selection processes.

Medicover Genetics **Nicosia, Cyprus**
Bioinformatics Research Intern *Jun 2016 – Aug 2016*

- Constructed the sequencing pipeline in R and Python to detect abnormalities in pregnancies.
- Scripted a custom EM algorithm for fetal fraction estimation in maternal blood and compared result with MCMC runs.

Academic Experience

Media in the Digital Age - Special Interest Group **London, UK**
Co-organizer, The Alan Turing Institute *2023 – Onwards*

- Main organiser of the MEDIATE Workshop at ICWSM Conference (AAAI) - presenter, coordination of Program Committee, review process, speaker selection and communications.
- Co-organized the AI UK Fringe event - Computational Journalism in the GenAI era with 50 attendees.

Machine Learning Graduate Module **London, UK**
Teaching Assistant, Queen Mary University of London *Fall 2021*

- Assisted graduate students (class of 250) with technical questions and curriculum during the computing lab sessions on a weekly basis.
- Marked two cycles of assignments counting towards the final grade and facilitated solutions refinement.

Lifting Effect On Global Nonlinear Optimization **Cambridge, UK**
Graduate Researcher, University of Cambridge *2014 – 2015*

- Researched the effect of lifting using mathematical reformulations to reduce non-linearities on global

optimization of complex non-linear systems - Python and Pyomo .
- Formulated a total of 52 different problems using lifting and problem-specific transformations and achieved up to 87% increase in global optimum convergence in problems of 110 original variables.

Technical Skills

Programming Languages	Python (PyTorch, transformers, accelerate, beautifulsoup), R, PostgreSQL.
Frameworks	Spark, Databricks, AzureML
Other	Git, LaTeX, Bash, Unix, JIRA

Awards

DeepMind PhD Studentship **Sep 2021**

Media Assembly Awards Best Use of Analytics Award: Enhanced Attribution and Modeling **Jan 2018**

Dean's List Dedication to Student Success Award, Harlem Educational Activities Fund Inc. **Jun 2016**

Top 0.05% 5th Place among 9605 Candidates in Exams for University Entrance, Cyprus **Jun 2011**

Honorable Mention International Mathematical Olympiad, Kazakhstan **Jul 2010**

Bronze Medal European Science Olympiad, Sweden **Apr 2010**

National Representative in 4 different International Competitions (IMO, IOI, IChO) **2007 – 2011**

Leadership & Activities

Invited Speaker DeepMind x WiML Technical Onsite Event **2024**

Invited Speaker The Alan Turing Institute **2024**

Reviewer ARR (Aug '24, Feb '24), CLPsych '22, MEDIATE '23 **2022 – 2024**

Research Talk UKRI NLP Workshop, University of Edinburgh **2022**

Transformers Training For 40 data scientists on BERT implementation, Chubb **2021**

Event Organizer Women in Machine Learning and Data Science, NYC Chapter **2019**

Admissions & Community Intern Harlem Educational Activities Fund Inc, NY **2015 – 2017**

Department Representative Engineering Graduate Student Council, Columbia University **2015 – 2016**