

# Joel Mackenzie

---

📍 Melbourne, Victoria, Australia  
🌐 <https://jmmackenzie.io>  
✉ [joel.mackenzie@unimelb.edu.au](mailto:joel.mackenzie@unimelb.edu.au)

## Education

**Doctorate of Philosophy (Computer Science).** February 2016 – October 2019  
PhD Thesis: *Managing Tail Latency in Large Scale Information Retrieval Systems*.  
RMIT University, Melbourne, Australia.

**Bachelor of Computer Science, Hons.** February 2012 – November 2015  
Honours Thesis: *Efficiency-Effectiveness Trade-offs for Location Aware Search*.  
First class honours.  
RMIT University, Melbourne, Australia.

## Work Experience

**University of Melbourne** – Postdoctoral Research Fellow February 2020 – Present  
Working with Prof. Alistair Moffat on the ARC Discovery Project *Efficient Compression and Querying Techniques for Massive Text Collections*.

**RMIT University** – Research Assistant October 2019 – February 2020  
Worked with Prof. Shane Culpepper on a number of Information Retrieval research projects.

**Microsoft Research** – Research Intern June 2018 – September 2018  
Worked with the SQL and the Language and Information Technologies teams to understand user behavior in the email search domain. Involved analysis of user behavior logs from a set of human intelligence tasks related to email search.

**RMIT University** – Programmer May 2017 – September 2017  
Data preparation and scripting for the RMIT Computing department for the Australian Research Council's "Excellence in Research for Australia" rankings. Involved writing tools to automatically extract citation information from both Google Scholar and Microsoft Academic, as well as further data preparation and analysis.

**Envato** – Assistant Search Engine Analyst (Intern) January 2016 – April 2016  
Worked with the Discovery team to continue improving search effectiveness across some of Envato's most popular marketplaces. Involved creating tooling for analysis of query logs to gain insights into the search behaviour of users on Envato market.

## Teaching Experience

**RMIT University** July 2014 – July 2019

- ♦ Database Systems – Tutor and lab assistant.
  - ♦ Semester 1, 2015.
  - ♦ Semester 1, 2019.
- ♦ Information Retrieval – Tutor.
  - ♦ Semester 1, 2017.
  - ♦ Semester 1, 2018.

- ♦ Programming Fundamentals – Head tutor.
  - ♦ Semester 1, 2018.
- ♦ Computing Theory – Tutor.
  - ♦ Semester 1, 2015.
- ♦ Operating Systems Principles – Tutor and lab assistant.
  - ♦ Semester 2, 2014.

## Research Interests

I am primarily interested in the discipline of Information Retrieval, including areas such as efficient indexing and retrieval, ranking, evaluation, and user behavior. I also have a broad interest in data structures and algorithms, and how we can continue to efficiently handle extremely large data.

## Publications

J. Mackenzie, M. Petri, and A. Moffat: **Faster Index Reordering with Bipartite Graph Partitioning**. In *Proceedings of the 44th International ACM Conference on Research and Development in Information Retrieval (SIGIR 2021)*, July, 2021. Short paper.

R. Benham, J. Mackenzie, J. S. Culpepper, and A. Moffat: **Different Keystrokes for Different Folks: Visualizing Crowdsourcing Querying Behavior**. In *Proceedings of the 6th ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR 2021)*, March, 2021. Demo paper.

J. Mackenzie, and A. Moffat: **Examining the Additivity of Top- $k$  Query Processing Innovations**. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM 2020)*, October, 2020. Long paper.

J. Mackenzie, R. Benham, M. Petri, J. R. Trippas, J. S. Culpepper, and A. Moffat: **CC-News-En: A Large English News Corpus**. In *Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM 2020)*, October, 2020. Resource paper.

J. Mackenzie, Z. Dai, L. Gallagher, and J. Callan: **Efficiency Implications of Term Weighting for Passage Retrieval**. In *Proceedings of the 43rd International ACM Conference on Research and Development in Information Retrieval (SIGIR 2020)*, July, 2020. Short paper.

J. Lin, J. Mackenzie, C. Kamphuis, C. Macdonald, A. Mallia, M. Siedlaczek, A. Trotman, and A. de Vries: **Supporting Interoperability Between Open-Source Search Engines with the Common Index File Format**. In *Proceedings of the 43rd International ACM Conference on Research and Development in Information Retrieval (SIGIR 2020)*, July, 2020. Demo paper.

R. Benham, J. Mackenzie, A. Moffat, and J. S. Culpepper: **Boosting Search Performance Using Query Variations**. In *ACM Transactions on Information Systems (TOIS)*, October, 2019. Journal paper.

A. Mallia, M. Siedlaczek, J. Mackenzie, and T. Suel: **PISA: Performant Indexes and Search for Academia**. In *Proceedings of the Open-Source IR Replicability Challenge, co-located with SIGIR (OSIRRC 2019)*, July, 2019. Workshop paper.

M. Petri, A. Moffat, J. Mackenzie, J. S. Culpepper, and D. Beck: **Accelerated Query Processing Via Similarity Score Prediction**. In *Proceedings of the 42nd International ACM Conference on Research and Development in Information Retrieval (SIGIR 2019)*, July, 2019. Long paper.

J. Mackenzie, K. Gupta, F. Qiao, A. H. Awadallah, and M. Shokouhi: **Exploring User Behavior in Email Re-Finding Tasks**. In *Proceedings of the 2019 World Wide Web Conference (WWW 2019)*, May, 2019. Long paper.

J. Mackenzie, A. Mallia, M. Petri, J. S. Culpepper, and T. Suel: **Compressing Inverted Indexes with Recursive Graph Bisection: A Reproducibility Study**. In *Proceedings of the 41st European Conference on Information Retrieval (ECIR 2019)*, April, 2019. Long paper (reproducibility track).

L. Gallagher, J. Mackenzie, and J. S. Culpepper: **Revisiting Spam Filtering in Web Search**. In *Proceedings of the 23rd Annual Australasian Document Computing Symposium (ADCS 2018)*, December, 2018. Short paper.

R. Benham, L. Gallagher, J. Mackenzie, B. Liu, X. Lu, F. Scholer, A. Moffat, and J. S. Culpepper: **RMIT at the 2018 TREC CORE Track**. In *Proceedings of the 27th Text REtrieval Conference (TREC 2018)*, November, 2018. Notebook paper.

R. Benham, J. S. Culpepper, L. Gallagher, X. Lu, and J. Mackenzie : **Towards Efficient and Effective Query Variant Generation**. In *Proceedings of the First Biennial Conference on Design of Experimental Search and Information Retrieval Systems (DESIRES 2018)*, August, 2018. Long paper.

J. Mackenzie, C. Macdonald, F. Scholer, and J. S. Culpepper: **On the Cost of Negation for Dynamic Pruning**. In *Proceedings of the 40th European Conference on Information Retrieval (ECIR 2018)*, March, 2018. Short paper.

J. Mackenzie, J. S. Culpepper, R. Blanco, M. Crane, C. L. A. Clarke, and J. Lin: **Query Driven Algorithm Selection in Early Stage Retrieval**. In *Proceedings of the 11th International ACM Conference on Web Search and Data Mining (WSDM 2018)*, February, 2018. Long paper.

J. Mackenzie, F. Scholer, and J. S. Culpepper: **Early Termination Heuristics for Score-at-a-Time Index Traversal**. In *Proceedings of the 22nd Annual Australasian Document Computing Symposium (ADCS 2017)*, December, 2017. Long paper. Runner up best paper award.

L. Gallagher, J. Mackenzie, R. Benham, R-C. Chen, F. Scholer, and J. S. Culpepper: **RMIT at the NTCIR-13 We Want Web Task**. In *Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR 2017)*, December, 2017. Notebook paper.

R. Benham, L. Gallagher, J. Mackenzie, T. T. Damessie, R-C. Chen, F. Scholer, A. Moffat, and J. S. Culpepper: **RMIT at the 2017 TREC CORE Track**. In *Proceedings of the 26th Text REtrieval Conference (TREC 2017)*, November, 2017. Notebook paper.

J. Mackenzie: **Managing Tail Latencies in Large Scale IR Systems**. In *Proceedings of the 40th International ACM Conference on Research and Development in Information Retrieval (SIGIR 2017)*, August, 2017. Extended abstract.

M. Crane, J. S. Culpepper, J. Lin, J. Mackenzie, and A. Trotman: **A Comparison of Document-at-a-Time and Score-at-a-Time Query Evaluation**. In *Proceedings of the 10th International ACM Conference on Web Search and Data Mining (WSDM 2017)*, February, 2017. Long paper.

J. Mackenzie, R-C. Chen, and J. S. Culpepper: **RMIT at TREC 2016 LiveQA Track**. In *Proceedings of the 25th Text REtrieval Conference (TREC 2016)*, November, 2016. Notebook paper.

J. Mackenzie, F. M. Choudhury, and J. S. Culpepper: **Efficient Location-Aware Web Search**. In *Proceedings of the 20th Annual Australasian Document Computing Symposium (ADCS 2015)*, December, 2015. Long paper. Best paper award.

## Awards and Funding

**RMIT Vice Chancellors List Award:** Awarded for ranking in the top 2% of higher education university-wide, 2014 (BSc) and 2015 (BSc, Hons).

**Australian Postgraduate Award:** PhD scholarship from 2016 to 2019.

**SIGIR Travel Award:** WSDM 2017, SIGIR 2017.

**RMIT TIGER/BDDA Group Student Paper Award:** For “*Query Driven Algorithm Selection in Early Stage Retrieval*” (WSDM 2018).

**RMIT CIDDA CReD Research Excellence Award:** For “*Exploring User Behavior in Email Re-Finding Tasks*” (WWW 2019).

**Best Paper Award:** ADCS 2015, ADCS 2017 (Runner up).

**University of Melbourne Early Career Seed Funding (2021):** \$20,000.

## Service

### Peer Reviewing

Australasian Document Computing Symposium (ADCS): 2016–2017.

International ACM Conference on Research and Development in Information Retrieval (CIKM): 2016–2017.

The International Journal on Very Large Databases (VLDB): 2017.

The Journal of Information Science: 2019.

International Symposium on String Processing and Information Retrieval (SPIRE): 2019.

Empirical Methods in Natural Language Processing (EMNLP): 2020.

ACM Transaction on Information Systems (TOIS): 2020

### Program Committee Member

Australasian Document Computing Symposium (ADCS): 2018–2019.

International ACM Conference on Information and Knowledge Management (CIKM): 2017–2021.

International ACM Conference on Research and Development in Information Retrieval (SIGIR): 2019–2021.

International ACM Conference on Web Search and Data Mining (WSDM): 2021.

European Conference on Information Retrieval (ECIR): 2021.

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD): 2021.

Empirical Methods in Natural Language Processing (EMNLP): 2021.

### Proceedings Chair

International ACM Conference on Web Search and Data Mining (WSDM): 2019.

ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR): 2021.

### SIGIR Artifact Evaluation Committee (AEC)

Member: 2021–Present.