

# MATT CRANE

---

mcrane@snapbug.geek.nz  
https://snapbug.geek.nz  
+1 707 764 6288

San Francisco  
CA 94107, USA

---

## WORK EXPERIENCE:

### Machine Learning Infrastructure Software Engineer, Instacart, CA, USA

November 2021–present

Building out the infrastructure to power the machine learning at Instacart. Instacart is built within the AWS ecosystem, so I gained a lot of experience with AWS tools and management/monitoring tools such as Terraform, Datadog etc. During my time at Instacart, I identified many areas for cost saving, including multi-million dollar savings by removing redundant logging.

**Vision**—I lead the development and writing of the teams founding vision document, which describes the steady-state multi-year final vision of what the machine learning infrastructure team would be building, how they would work together and support the needs of machine learning engineers in their current, immediate and future work. The team current consists of 15 engineers.

This vision document served as the basis for development of the team's planning and roadmap sessions, as well as core cross-company engineering initiatives.

**Feature Store**—Led the development of Instacart's feature store product. As part of this I lead a re-write initiative that improved p99 latency for requests by 10x, and ran on 10x fewer instances. Three engineers worked on the core feature store product, while dozens relied on it's functionality and robustness.

After the re-write the feature store was able to handle, with minimally deployed instances, traffic of two million keys per second at peak, while keeping an overall p99 latency of 12ms. Further, the service was engineered to allow infrastructure engineers to change underlying backends seamlessly.

### Research Scientist, Meta (*aka Facebook Inc.*), CA, USA

July 2018–November 2021

Focused on solving advertiser facing issues within the delivery system, including diagnosing systemic inefficiencies in ranking and delivery systems, large scale back-end migrations to unblock scaling of products, and development of new products. Served as an internal hiring point of contact for the team, as well as managing and mentoring interns and being a ramp-up buddy for new engineers, both senior and junior, to both the team and org.

**Ads Dynamism**—Led a team of seven engineers to migrate ads delivery to a new data model for a single ads product. This project spanned the whole ads delivery system and multiple XFN teams. Defined the project roadmap, timeline, and goals alongside XFN partner teams. Designed key infrastructure components, delegating effectively across this team.

This data model serves █% of Facebook ads revenue, and unblocked █% additional revenue through reducing serving infrastructure requirements by █%. The project serves as the basis for the entirety of ads delivery to migrate data models, unlocking a further █% additional revenue, and further reducing serving infrastructure requirements by █%.

**Knowledge Expression**—Developed a new rule-based delivery product that allows advertisers to express demographic constraints and hints for each asset within an ad. Led a small team of senior backend engineers, communicating engineering updates and issues encountered to both XFN partner teams, and area leadership. Alphas showed majority improved performance (█%), and overwhelming positive sentiment (█%).

Identified a key issue that affected █% of Facebook ads revenue, and was existential for multiple projects across multiple teams.

**Arbitrary Demand Experimentation Platform**—Designed and implemented an experimentation platform for arbitrary demand segmented experiments, that did not need budget controls. This platform supplemented, and bought advertiser experience metrics to, the standardised pre-existing supply segmented experimentation platform. It has since been used to support experiments that have shipped █% gains cumulatively.

**Continuity**—Improved the sustainability of an ad campaign performance by investigating all parts of delivery. Analyzed the effect of ranking model state/history, detected when campaigns had entered a sub-optimal state, and tested methods to break that state, and sustain performance.

**AA Variance**—Identified the source of large variance observed by advertisers on duplicated campaigns. After root-causing to ranking, and conducting a large-scale feature and model migration, an advertiser facing experiment showed revenue gains of █%, alongside other key business-sensitive advertiser experience metrics.

## EDUCATION:

---

**PhD — Computer Science**, University of Otago, New Zealand

**February 2012–March 2016**

Thesis: *Improved Indexing & Searching Throughput*

<http://ourarchive.otago.ac.nz/handle/10523/6223>

My doctoral research concentrated on investigating ways to make the processes of indexing and searching of web-scale collections more efficient without impacting on the effectiveness of the system.

During my candidature I was, and continue to be, an active member of the information retrieval community, and have presented at multiple SIGIR, ICTIR, and ADCS conferences and workshops, as well as CIKM.

**MSc (Thesis Only) with Distinction — Computer Science**, University of Otago

**October 2009–May 2011**

Thesis: *The New User Problem in Collaborative Filtering*

<http://otago.ourarchive.ac.nz/handle/10523/1938>

## SELECTED PUBLICATIONS:

---

A complete list is available on my website: <https://snapbug.geek.nz/#publications>.

Andrew Trotman and **Matt Crane**. “Micro-and Macro-optimizations of SaaS Search”. In: *Software: Practice and Experience* 49.5 (2019), pp. 942–950.

**Matt Crane**. “Questionable Answers in Question Answering Research: Reproducibility and Variability of Published Results”. In: *Transactions of the Association for Computational Linguistics* 6 (2018), pp. 241–252. URL: <https://transacl.org/ojs/index.php/tacl/article/view/1299>.

**Matt Crane**, J. Shane Culpepper, Jimmy Lin, Joel Mackenzie, and Andrew Trotman. “A Comparison of Document-at-a-Time and Score-at-a-Time Processing”. In: *Proceedings of the 10th ACM International Conference on Web Search and Data Mining. WSDM '17*. 2017, pp. 201–210.

**Matt Crane**, Andrew Trotman, and Richard O’Keefe. “Maintaining Discriminatory Power in Quantized Indexes”. In: *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management. CIKM '13*. 2013, pp. 1221–1224.

## ACTIVE INVOLVEMENT:

---

Journal Reviewing:

- Information Processing and Management (Elsevier)
- Information Processing Letters (Elsevier)
- Information Retrieval Journal (Springer)
- Transactions on Information Systems (ACM)
- Transactions on Knowledge and Data Engineering (IEEE)

Conference Reviewing:

- SIGIR 2021, 2020, 2019, 2018—International ACM SIGIR Conference on Research and Development in Information Retrieval
- CIKM 2019—ACM Conference on Information and Knowledge Management
- ICTIR 2017—ACM International Conference on the Theory of Information Retrieval
- ADCS 2019, 2018, 2017, 2016, 2015, 2014—Australasian Document Computing Symposium

Workshop Involvement:

- LIARR Workshop at SIGIR 2017—Lucene for Information Access and Retrieval Research
- RIGOR Workshop at SIGIR 2016—Reproducibility, Inexplicability, and Generalizability of Results

## SKILLS:

---

Programming languages; strong: Rust, Python, C++; moderate: many others, including SQL, Go, C, Typescript.

Systems: AWS, Terraform, Docker, Datadog, Snowflake

Troubleshooting and debugging of large, complex systems.