

# Andy Twigg

[andy@atwigg.com](mailto:andy@atwigg.com)  
6502454007

CA 94002

I'm a computer scientist with expertise in large-scale data processing and ML.

I currently lead ML efforts at Lacework, a cloud security company treating security as a big data problem. I have been an entrepreneur (CTO of 2 acquired startups Acunu, C9) and an academic (PhD Cambridge, and I was a Fellow in Computer Science at Oxford University 2008-13).

## Experience

*2020-present: Distinguished engineer, Lacework*

Lacework is a cloud security company treating security detection as a big data problem. I lead the core "[polygraph](#)" technology – large-scale data processing and ML pipelines. We process several billion msgs/hr from agent and cloud logs. Team of ~20 engineers, 3 TL and 2 EM. Lacework raised ~\$1.8B at a ~\$8.3B valuation, and was acquired by Fortinet in 2024.

- rewrote implementation from golang to pyspark, EC2 savings of ~\$2m/yr, ~10x reduction in SEVs
- led new models using inductive GNNs with pytorch-geometric, streaming density estimators

*2018-2020: EIR (entrepreneur in residence), Milliways Ventures*

Exploring ideas around deep RL. [One project](#) was to try to train a model to learn how to trade using RL. I collected several TBs of L3 data from a large crypto exchange and built an event-driven simulator that allows better queue length estimation (eg due to cancellations). I supervised Stanford CS246 students who worked on it as coursework. code: <https://github.com/andytwigg/deeptrade>

*2014-2017: CTO, C9 and Chief Scientist, Insidesales.com*

C9 built one of the first systems to [apply ML to improve sales efficiency](#) via bottom-up forecasting, etc. C9 was [acquired by insidesales.com](#) in 2015. I led the data science and machine learning teams.

*2013: Founder, Featurestream.io*

I built a [streaming random forest](#) on spark streaming and experimented with offering it via an API. Code: <https://github.com/featurestream/>

*2009-13: Cofounder, CTO, Acunu (acquired)*

We built a streaming analytics system based on Cassandra, sketching algorithms, and [Stratified B-trees](#), which are a fully-versioned LSM/fractal tree (see [here](#), [here](#)). We started the London Big Data meetup. Acquired by a big tech company.

*2008-13: Fellow in Computer Science, St Johns College, University of Oxford*

Academic post (elected by open competition); took a sabattical to found Acunu.

*2006-7: Microsoft Research (Cambridge) and Technicolor Research (Paris)*

Developed [algorithms for P2P streaming](#) problems with optimal throughput/latency tradeoffs.

## Education

*2006: PhD Computer Science, Cambridge University (King's College)*

*Thesis:* [Approximate graph routing with failures](#). Nominated for BCS Best Dissertation Award.

*1999-2002: BSc Computer Science, Warwick University (top 1st)*

## Teaching

I have taught various courses at Oxford & Cambridge including Randomized Algorithms, Data Structures and Algorithms, Probability, Complexity Theory.

## *Interests*

I enjoy DIY, drumming, golf. I rowed for Cambridge Lightweights and King's College men's 1st VIII

## **Selected Publications**

[\*Persistent Cache-oblivious Streaming Indexes\*](#), arxiv, abs/1707.08186, 2017

[\*Locality-preserving allocations problems and coloured bin packing\*](#) with E Xavier., J. Theoretical CS, 2015

[\*Stratified B-trees and versioned dictionaries\*](#). Twigg et al, HotStorage 2011

[\*Constrained-path labellings on graphs of bounded clique-width\*](#), with B Courcelle, *Theory Comput. Syst.*, 2010

[\*Epidemic live streaming: optimal performance trade-offs\*](#), Bonald et al, *SIGMETRICS*, 2008.

[\*Worst-case time decremental connectivity and k-edge witness problems\*](#). ArXiv,abs/0810.5477, 2008

[\*Connectivity checking in 3-connected planar graphs with obstacles\*](#). Courcelle et al., Notes in Disc Math, 2008

[\*Rate-optimal schemes for peer-to-peer live streaming\*](#) Massoulie, Twigg., J. Perf Eval , 65(11-12):804–822, 2008

[\*Randomized decentralized broadcasting algorithms\*](#) with Massoulie et al, INFOCOM, pages 1073–1081, 2007

[\*Forbidden-set labelling on graphs\*](#). With Courcelle et al. PODC (LOCALITY), 2007

[\*Compact forbidden-set routing\*](#). Bruno Courcelle and Andrew Twigg. STACS 2007.

[\*The complexity of fixed point models of trust in distributed networks\*](#). with K Krukow, J Theoretical CS, 2007

[\*Compact forbidden-set routing \(PhD Thesis\)\*](#). Technical report UCAM-CL-TR-678, 2006

[\*Provably optimal decentralized broadcasting algorithms\*](#). With Massoulie et al, MSR-TR- 2006-105