

# Greg Bowyer

Apt 4, 1245 Washington Street  
San Francisco, 94108  
CA, United States

<https://github.com/GregBowyer>  
Email: [gbowyer@fastmail.co.uk](mailto:gbowyer@fastmail.co.uk)  
<http://gbowyer.freeshell.org/cv.pdf>

## Professional Experience

### Apache Lucene & Solr Committer

May 2012 - Present

I work in and on Lucene as a codebase, with a vested interest in index runtime performance.

### URX.

San Francisco

### Chief Scientist

Oct 2013 - Feb 2016

As the chief scientist my core responsibilities was to build and lead a data science and data engineering team to provide the core search technology powering a mobile centric contextual advert network.

- Managed a full data engineering & science team of 8 members
- Architected a complete terabyte scale data warehouse using PrestoDB & Spark
- Lead the implementation of a multivariate testing framework
- Developed a full search stack including query parsing & ontological inference
- Implemented a unbounded broad website crawler
- Managed and designed a semantic knowledge graph
- Developed a full elastic infrastructure on AWS, using Mesos & Ceph

### Shopzilla Inc.

Los Angeles

### Principal Specialist Engineer Search

May 2009 - Oct 2013

Working throughout the search eco-system of Shopzilla, migrating from a legacy in-house search engine towards a platform developed on Lucene.

- Developed color classification of images, with mean-shift clustering and visual bag of words
- Cleanroom implemented a full text classifier with 10ms inference performance
- Aggressively optimised in-house NLP tagger system, increasing performance tenfold
- Built a sponsored search auction, using game-theory and Latent Semantic Indexing
- Developed a JIT improving template performance, leading towards multi-million increases in profit
- Built performance testing tooling known as “dark-launch”, able to take production traffic and replay the same traffic live onto a target system
- Lead the architecture and development of the initial SOLR engine deployment, delivering the project ahead of time and on budget

### Miniweb Interactive Senior Java Engineer

Brentford - London  
December 2008 - May 2009

### eBay. Contract Software Engineer

Richmond - London  
July 2008 - December 2008

## OpenSource work

I work on and maintain various projects, all of which can be viewed on <https://github.com/GregBowyer>, some key examples include.

**Lucene-Solr** Various work on Lucene internals aiming for better performance.

**CLD2-CFFI** Bindings for the compact language detector to python.

**Planout Golang** Implementation of planout multi-variate testing in golang

**CMPH-CFFI** Bindings for using minimal perfect hashes in python

## Skills

**Languages:** Java 1.3+ Python 2.2+, C, C++, GoLang 1.4+, Javascript, Erlang, Fortran

**Information Retrieval:** Lucene Elastic Search, Solr, MG4J, Whoosh, IR-Modelling, Geo-Search, CBIR

**Machine Learning:** sklearn, spark-ml Semi Supervised and Unsupervised methods, Deep-Learning, Topic Models, Bag of Words models, Keras, Tensorflow, Theano

**Low level knowledge:** Bytecode manipulation, JNI, JPDA / JDWP, Hotspot internals, Java memory model, GC fundamentals, SIMD programming, Cache Aware & Oblivious programming, Linux kernel internals, x86 assembler, CUDA, OpenMP, Concurrency Kit, Blas & LAPack

**PyData stack:** Cython, PyCUDA, Numpy, IPython, Pandas, gensim, SciPy, Sympy, Matplotlib

**Data Engineering:** Hadoop, HBase, Spark, Mesos, MPI, Hive, PrestoDB, Apache Drill, PostgreSQL, Kafka

**Operating Systems:** Linux, Solaris, FreeBSD, Mesos, Ceph

**Misc:** Maven, Ant, CMake, AutoConf, L<sup>A</sup>T<sub>E</sub>X, Ansible

## Education

**Bsc. Hons Computer Science 1st class**

University of Greenwich  
Sep.2001 - June. 2004

*Undergraduate dissertation:* Peer to peer media streaming

**Pre graduate education**

MidKent College & Astor of Hever School  
1994 - 2001

Advanced GNVQ in IT & 11 GCSE's at grades A-C