Stephen Merity

Harvard University (2013 – 2014)

Master of Science in Computational Science and Engineering

University of Sydney (2008 – 2011)

Bachelor of Information Technology (First Class Honours + University Medal)

smerity@smerity.com / (857) 600 0763

EXPERIENCE

Downloads the Internet — July 2014 – Present

Common Crawl, San Francisco

The Common Crawl Foundation produces snapshots of the Web on a regular basis, hundreds of terabytes in size, and works to encourage start-ups, companies, and academics to use this massive dataset. As part of my work, I run regular full web crawls using a modified Apache Nutch platform, speak at conferences, evangelize open data and perform large scale experiments with Hadoop and AWS.

First Employee — Jan 2013 – Aug 2013

Grok Learning, Sydney

First employee of Grok Learning, a startup aimed at teaching high school students about computer science and programming. Helped the co-founders with startup strategy and implementation of the online learning interface. Launched in July, sustainable and profitable with thousands of happy customers.

Consultant (Data Analytics & Machine Learning) — Apr 2012 – Dec 2013

Implemented and provided guidance on large-scale data analytics & machine learning projects for start-ups and businesses. Projects ranged from writing real-time analytics dashboards to providing machine learning implementation advice on feature hashing.

Data Analytics and Interaction Mechanics — Dec 2010 – Apr 2012

Freelancer.com, Sydney

TEAM LEAD

Application of advanced data mining methods, statistical testing and machine learning techniques to increase the overall customer satisfaction and revenue on *Freelancer.com*.

Google App Engine — Dec 2009 – Mar 2010

Google Sydney

INTERNSHIP

Helped triage and fix multiple bugs in live production code and implemented a feature to enable rapid querying of existing user data. Work was performed in three languages (Java, C++ and Python) with code reviews and unit tests required for each commit.

PUBLICATIONS

Measuring the impact of Google Analytics: http://smerity.com/cs205_ga/ — 2013

[Paper] *Frontier Pruning for Shift-Reduce* CCG *Parsing* — 2011

[Thesis] *Integrated Tagging and Pruning via Shift-Reduce* CCG *Parsing* — 2011

TECHNICAL SKILLS

- Programming: Proficient in Python and C, familiar with using C++, Java, Go, HTML, CSS and shell scripting

- Particular expertise in the field of Natural Language Processing including tagging, parsing and the underlying machine learning classifiers and algorithms used to carry out these tasks