Lyndon White
Dept Electrical, Electronic
and Computer Engineering
The University of
Western Australia
35 Stirling Highway
Crawley, Perth

August 10, 2018

Alan Turing Instute
London NW1 2DB

Dear Dr Martin O'Reilly,

I am writing to you in support of my application for the position of research software engineer at
the Alan Turing Institute. I am soon to complete my PhD in natural language processing via
machine learning at the University of Western Australia. In 2014, I completed two concurrently-
studied Bachelor degrees: the first in Computer and Mathematical Science, majoring in Pure
Mathematics and Computation; and the second in Electrical and Electronic Engineering from
which I graduated with first class honours. The latter is a terminal degree in engineering. In 2011,
I briefly stepped away from academia to work as a software developer in the banking industry.
Here I worked as part of the agile transformation program, to help bring modern development
methodology to the bank's 800 programmers.

Reasons for my application
At a high level my primary interest is in *writing cool code, to solve cool problems*. This can
be broken down: *cool problems* are interesting, complex, and impactful; *cool code* is clean,
maintainable, understandable and yet powerful. At the Turing Institute, I would be able to do
this, while working with very clever people who are also driven to solve such *cool problems*.

I am excited about the work I see being done at the Turing Institute towards evolving our research
practices, and how society can benefit from data. Dr Tomas Petricek's work on the Gamma
project is an excellent example of this; as is the work being done by Prof Chris Williams' group
on Data Diff.

When I was a developer in industry, I was closely involved in the modernisation of software
practices used for the mainframe backend systems. This involved building tooling, infrastructure
for testing, version control, and continuous integration and educating specialists upon their use
and general best practices. During my PhD, I've put those skills to use and setup some similar
systems and tools for my own research software. As a long term goal, I would like to seem that
level of practical rigour in software brought widely to the academic world.

My experience and skills

My PhD research has been in natural language processing via machine learning. I also have
written a deep introductory book on this subject, which is due to be released in the next month.
During my studies, I have become familiar with a large variety of data science techniques. As
well as the supervised machine learning which has been crucial to my research; I also regularly
use a variety of analysis and problem solving techniques. From clustering and dimensionality
reduction, to constrained optimization. I also use a variety of methods for parallel programming;
with a preference for threads where possible to avoid communication overhead, but scaling out to
a cluster when required. Many examples of these techniques can be found in the tutorials on my
blog, and research papers. Both my blog and papers can be found at `http://white.ucc.asn.au/`.
I believe one of my most important skills is in problem and solution analysis. It is valuable to
consider this both on a theoretical and practical basis. For example, if a small problem could be
solved via mixed integer programming or via evolutionary algorithms, pragmatically it is likely
better to investigate the mixed integer programming solution first even though the theoretical

time complexity is worse.

I am an extremely strong Julia programmer. I am deeply familiar with the language and the ecosystem. It has been my language of choice during my PhD. I am also a strong programmer in Python, MatLab, C# and Java. I've worked on many projects in industry or in academia using these tool. I have some skill in large variety of other languages.

Outside of the directly technical pursuits I have some experience in producing a user front-end. In my work on developing testing tools in industry, we produced full desktop applications to ensure good and easy uptake by users. More recently at the Association of Computation Linguistics' conference, I presented a functional (if basic) web-application allowing anyone to make use of one of my research projects (`http://novelperspective.ucc.asn.au/`).

I am a passionate open source developer. I am involved in a great many projects, both projects under my own leadership, and collaborations with others. Most of my work has been in the Julia space, but I also contribute more broadly to python tools in my research area. A project I particularly excited about, and am in London to present on is DataDeps.jl. A preprint of my software paper for this is currently available (`https://arxiv.org/abs/1808.01091`). DataDeps.jl is a tool to improve replicability of research software that depends upon data, while also making it easier to write and run. DataDeps.jl also has a partner package DataDepsGenerators.jl, which I have suppervised a GSOC student in the development of, that is all about making reusing open data easy, while ensuring essential data provenance information is captured.

A good example of my expertise in collaborating to develop powerful software can been seen in my work as the co-maintainer of TensorFlow.jl (`https://github.com/malmaud/TensorFlow.jl`). TensorFlow.jl is used by many machine learning practitioners in the julia programming language. In some areas it is being superseded by new tools such as Flux.jl (which I also collaborate on), but TensorFlow.jl remains a connection to a cornerstone of the modern deep learning landscape. As well as my own direct contributes to the tool, largely in improving usability and making it easy to install, I have managed dozens of bug reports, and guided new contributors to our coding standards: in terms of code quality, documentation and testing. I'm also supervised two Google Summer of Code projects which require similar skills in teamwork and communication for code quality.

Overall, I believe I am an excellent candidate for a research engineering position at the Alan Turing Institute. I bring a deep understanding of academic research in the AI and data science space, together with strong and practical expertise in software development. I'd like to thank you for taking the time to read this application, and look forward to speaking to you more in the future.

Kind Regards
Lyndon White
BCM BEng(Hons) W.Aust.