Andrew Challis

Senior Data Scientist

☎ +61 482 043 277

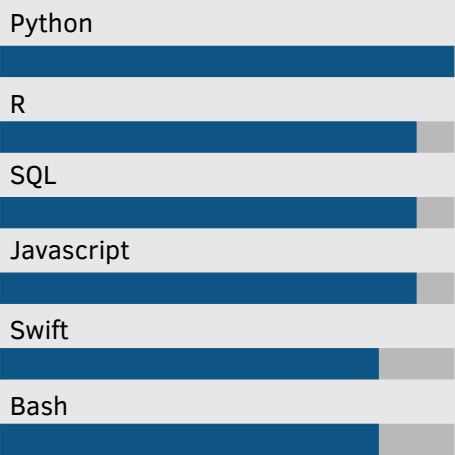🌐 andrewchallis.co.uk

@ andrewchallis@hotmail.co.uk

About me

Andy is extremely passionate about technology, particularly when it comes to investigating cutting-edge ideas and applications. Andy has a very strong analytical background which aids in his ability drive insightful change within a team. He also has strong leadership and teamwork skills as a result of playing lacrosse at a high level and captaining an undefeated team.

Interests

- Ⓦ Blogging on data science
- >_ Coding challanges
- 🔧 Playing with home automation
- ✈ Travelling the world
- 💚 Lacrosse
- 🍴 Cooking

Languages

Python

R

SQL

Javascript

Swift

Bash

Objective Statement

To design & deliver solutions using cutting-edge technologies that offer valuable insights to increase business value.

Education

2015-16	M.Sci. First class, magna cum laude	Lancaster University
	Majoring in Medical Statistics	
2012-15	B.Sc. First class, magna cum laude	Lancaster/Texas A&M
	Mathematics & Statistics	
2010-12	College A*AAB	Sale Grammar
	Specialising in Maths, Further Maths, Chemistry and Physics.	

Awards

2012	Lancaster University Academic Scholarship.
2014	Texas A&M Honours Student.

Experience

2019	Consultant Senior Data Scientist	Capgemini AUS

(Client: NAB)

The project was to built a staged deployment of an enterprise wide data lake stored in the cloud (AWS) that will manage the data and it's distribution across the bank. However, delivery speed was a key issue so I built an automation framework for ingesting over 250 data sources with meta data, using Python, AWS Lambda, Informatica BDM.

2018	Consultant Senior Data Scientist	Capgemini UK

(Client: HMRC)

Built and implemented a full scale solution for processing P11D (expense) forms using a custom tool leveraging Google Tesseract OCR and OpenCV.

Created a private light GUI client which would be used to verify and process the output of the models predictions by an agent.

At the end of the trial the result was over 25 times faster throughput of forms; which in turn saved HMRC staffing from peak numbers of 120 down to 10. The output of the validation also created training data to be used to create a more sophisticated model in the future.

(Client: Home Office)

Created a production ready containerised Facial Recognition and Analysis API using Python's OpenCV, DLib and Flask. This was used as part of an application process to access the quality of images. Specifically if they adhered to ICAO standards for machine readable images. During this project I also created unit tests and a full web front-end as well as a mobile application developed in Swift.

Developed a custom suite of Python libraries to speed up development, such as connecting to our Jira instance, logging/interacting with Rocketchat and connecting to our database & incorporating commonly used functions/data types.

Worked with the platform team to push Continuous Integration with GitLab using Jenkins and Docker containers.

Technical lead on a scenario to engineer full-scale automated pipelines from inception through to beta phase, presented to users in a front-end tool.

2017	Consultant Data Scientist	Capgemini UK

(Client: Thames Water)

Designed interactive mapping visualisations using both open source technologies for PoC and full-scale integration's with IBM's IOC for PoV which links the users' decisions to the operations.

Andrew Challis

Senior Data Scientist

☎ +61 482 043 277

🌐 andrewchallis.co.uk

@ andrewchallis@hotmail.co.uk

About me ——

Andy is extremely passionate about technology, particularly when it comes to investigating cutting-edge ideas and applications. Andy has a very strong analytical background which aids in his ability drive insightful change within a team. He also has strong leadership and teamwork skills as a result of playing lacrosse at a high level and captaining an undefeated team.

Interests ——

Ⓦ Blogging on data science

>_ Coding challanges

🔧 Playing with home automation

✈ Travelling the world

💔 Lacrosse

🍴 Cooking

Languages ——

Python

R

SQL

Javascript

Swift

Bash

Developed algorithms for detecting leaks in pipes using multiple data sources: pressure, flow, pipe attributes, environment and smart meters.

Mentoring junior staff and delivering 'lunch and learn' talks on hot topics.

(Client: DSTL)

Designed a data science competition (logo recognition in videos) for data-sciencechallenge.org which was sponsored by the client.

Created tutorials for ways in which to achieve an out-of-the-box baseline result using TensorFlow.

Curated images and videos for the competition from both paid-for and CC0 sources.

2016	Consultant Junior Data Scientist	Capgemini UK

(Client: Severn Trent)

Developed reservoir prediction algorithms for predicting how long water in underground reservoirs will last depending on demand. We overlaid a cost model for electricity and fines to find an optimal solution for the life cycle of reservoirs.

(Client: DWP)

Involved in architecting a data science platform that took advantage of Jupyter-Hub, Docker Swarm, Hadoop, AWS and multiple kernels (Python, R, Julia, Scala etc).