

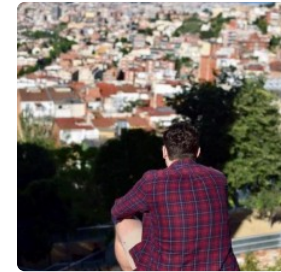
# Andy Challis

Data Scientist

Melbourne, AU

[andrewchallis.co.uk](http://andrewchallis.co.uk) [andrewchallis@hotmail.co.uk](mailto:andrewchallis@hotmail.co.uk) [+61482043277](tel:+61482043277)

[achallis](#) [ghandic](#)



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.

## SKILLS

### Python

Master

PyTorch TensorFlow  
FastAPI OpenCV  
scikit-learn Pandas  
Flask Dash mkdocs

### Full Stack

Master

NodeJS React  
KonvaJS CSS SASS  
HTML5 Hugo Traefik

### DevOps

Master

AWS GCP Docker  
Kubernetes Git Github  
GitLab Bitbucket

### Databases

Master

PostgreSQL MySQL  
Elastic Hive sqlite

## WORK EXPERIENCE (3)

**Manager** at Capgemini Invent

**Jun 2019 - Current**

**Melbourne, AU**

Capgemini Invent combines strategy, technology, data science and creative design to solve the most complex business and technology challenges.

- Designed, built and deployed multiple open source deep learning systems into production with 99.9% availability
- Solution Lead for a \$4.5 million-dollar account, where we scaled up the project team from 7 to 32 members within 12 months
- Spearheaded the [CapgeminiInventIDE](#) open source project on GitHub, which has boosted our external visibility in the space massively, allowing applicants to see the level we are performing at and the ability to contribute alongside us
- Designed and ran an 8-week beginner/intermediate Python course for the whole of Capgemini Australia. The course curriculum was built and designed by myself and incorporated any additional topics that participants wanted to cover.
- Lead the Python community of practice inside of Capgemini where I actively mentored individuals
- Primary interviewer for building high preferment teams, graded over 40 take home assessments, and both face-to-face and remote technical interviews
- Produced and maintained 5 IP rich deep learning microservices with a demo front end written in React to showcase to potential clients, leading to ~ \$1 million dollars in project leads

**Senior Consultant** at Capgemini - Insights & Data

**Jan 2018 - Jun 2019**

📍 **Manchester, UK**

Capgemini I&D aims to Drive digital transformation through AI, Analytics & Platform.

- Involved with every facet of Data Science. Scoping, architecting, and leading data science projects & implementations
- Technical lead for product development teams of around 8 people
- Responsible for the designing and development of advanced models, as well as high-value use cases
- Worked with vague (and complicated) project requirements from clients, while understanding how to evaluate and discover the best approaches to modelling

**Consultant** at Capgemini - Insights & Data

**Oct 2016 - Jan 2018**

📍 **Manchester, UK**

Capgemini I&D aims to Drive digital transformation through AI, Analytics & Platform.

- Developed a number of fast-moving PoCs where I had to consolidate my skills in a business context
- Put into practice the theoretical knowledge of statistics and data science I had gained from university.
- Practiced and proven in delivering no-compromise work under extreme pressure.
- Used a range of techniques including predictive analytics, machine learning (ML), traditional statistics and geo spatial analysis through Agile methodologies.
- Developed algorithms for leak detection for leading UK water utility companies.
- Collaboratively worked with internal and external consultants to deliver impactful insights to clients.

## PROJECTS (8)

---

**Cognitive Document Handling** at Services Australia

**Jul 2019 - Current**

📍 **Melbourne, AU**

AWS Python Docker Kubernetes TDD OpenCV PyTorch ONNX FastAPI TensorFlow Dash React BitBucket JIRA

In this project I performed a highly technical role with an emphasis on leadership, which has been paramount to its technical excellence and rigor, I pushed new boundaries within Services Australia and set the bar high for quality output. We were challenged to meet strict SLA's for which I architected the end to end solution with sign off from the client.

- Solution Lead for a \$4.5 million-dollar account
- Provided technical direction for a fully remote team of 35 people, plus client stakeholders
- Deployed a production deep learning system in centerlink with an error rate of < 0.1% serving hundreds of thousands of Australians per day
- Managed the delivery of 6 products, 100% on time and beyond the client's expectations, adapted to the client's requests and changing stake holder landscape by rapidly iterating on feedback and improvements showing the flexibility and adaptability of our team
- Organized weekly lunch and learns, pair programming sessions, and out of hours mentorship
- Considered a role model project, both internally and on the client side during corona virus pandemic, demonstrating how the remote working model can scale

**Data Hub** at NAB

**Jan 2019 - Jul 2019**

📍 **Melbourne, AU**

AWS Python Kafka Hive Jenkins Terraform Docker TDD WatchDog Flask GitHub JIRA

The project was to build a enterprise wide data platform in the cloud (AWS) that would manage the data and it's distribution across the bank. Delivery speed was a key issue so Andy assembled a team to build an automation framework for ingesting over 250 data sources along with the representative meta data.

- Architected and built a self service micro-service for registering data sources to a data platform specifically for files being sent via SFTP.
- Built a horizontally scalable (EC2/Docker) Python daemon that monitored file-system events (EFS) and triggered custom data flows for each object, such as; moving to specified dynamic S3 paths (S3), triggering a custom transformation job (Jenkins), automated un-compressing, clearing from file system (EFS)
- Advised team members on deployment strategies (Terraform), unit testing (PyTest), alerting/logging (SNS + Splunk), documentation standards and containerisation
- Lead an automation squad consisting of 8 people within NAB to help deliver a centralised self service data platform.
- Had a major impact on software practices within squads by; creating reusable distributed packages for other teams to take advantage of, building tactical tools to increase the work-flow and automation of manual data ingestion – effectively cutting ingestion time from 6 weeks down to 2 (allowing for business approvals and procedures).
- Built user interfaces (Flask, Javascript, HTML, CSS) for other squads to use that speed up their validation efforts for data ingestion.

**Automated Form Processing** at HMRC

**Jun 2018 - Aug 2018**

📍 **Telford, UK**

AWS Python Docker Javascript Frontend TDD TensorFlow OpenCV GitLab JIRA

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.

**Facial Recognition & Analysis API** at Home Office

**May 2017 - Feb 2018**

📍 **Croydon, UK**

Python Docker Swift Javascript Frontend TDD Jenkins PostgreSQL TensorFlow OpenCV BitBucket JIRA

Created a production ready containerized 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.

- 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.

### Automated Leakage Detection at Thames Water

Feb 2017 - Apr 2017

📍 London, UK

IBM Bluemix Python R Time Series Geo Spatial Docker Folium BitBucket JIRA

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.

- 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.

### Design of Data Science Competition at DSTL

Dec 2016 - Jan 2017

📍 Manchester, UK

Python TensorFlow OpenCV Docker GitHub JIRA

Designed a data science competition (logo recognition in videos) for [datasciencechallenge.org](http://datasciencechallenge.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.

### Predicting Reservoir Demand at Thames Water

Nov 2016 - Nov 2016

📍 London, UK

R BitBucket JIRA

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.

### Data Science Platform at Department of Work and Pensions

Oct 2016 - Nov 2016

📍 London, UK

Docker GitHub Kubernetes JIRA

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

## EDUCATION (3)

---

### Master of Science - Medical Statistics at Lancaster University

2015 - 2016

Grade: First class, magna cum laude

### Bachelors of Science - Mathematics & Statistics at Lancaster University + Texas A&M

2012 - 2016

Grade: First class, magna cum laude

### A Levels at Sale Grammar School

2010 - 2012

Grade: A\*AAB

## AWARDS

---

**Imagination Paragon Award** at ISG

2020

Solution lead for the project where we were recognized our innovative and impactful collaboration for 'Document Management Modernization through Intelligent Automation'

**Inventor of the Quarter** at Capgemini Invent

2020

Gained recognition from senior leadership within Capgemini Invent for being 'a role model with impact'

**Academic Scholarship** at Lancaster University

2012

**Honors Student** at Texas A&M University

2014

## LANGUAGES

---

**English**

*Native Speaker*



**German**

*Intermediate*



## INTERESTS

---

**Tech**

LinkedIn Posts

Website Blogs

Lacrosse

Rugby

Opensource

**Sports**

## REFERENCES

---

“ Andy's leadership on the Services Australia project has been incredible, continually going above and beyond for the work we are delivering for the client, resulting in client feedback such as 'we never have a doubt that you guys can deliver' due to Andy's tech leadership. Andy also has shown incredible leadership focusing on building people and the team to make sure everyone succeeds, by building skillsets and fostering a culture & environment of innovation. With the inclusion of client teams and constant client involvement in technical aspects of the project, Andy has excelled on inclusively working with the client to upskill and bring them along the journey of this project. ”

**Jeremiah Mannings | Associate Director**

“ Throughout my 25+ year career, I can categorically state that he is one of the most impressive professionals I have had the pleasure to work with. Any project Andy is involved in will be a success and I look forward to working with him again ”

**Benjamin Moretti | Senior Manager**

“ Andy is a very high calibre data scientist. Andy is hard working and committed on projects. He is quick to learn and is able to apply his learnings and technical skills into real life scenarios. He has a strong technical skillset both in terms of data science but also front end development. He is an effective team player and good communicator. ”

**Oliver Cochrane | Director**

“ Andy is a very enthusiastic, committed and talented Data Scientist, working with Andy on technically challenging OCR project revealed all of those facets in great depth. Andy was able to join an established team, quickly get up to speed and not only contribute to the development, he improved the accuracy and also the quality of the work being undertaken. I was very impressed with his skill set and dedication and would hope to work together with Andy again in the future, a genuine asset to any employer. ”

**Paul Sylvester | Senior Engineer**

“ Andy joined my team as a critical point in delivery of a high-profile project. He was quickly able to get up to speed with the approach we'd taken to date and improve the solution we were developing. He brought in ideas and worked hard to prove them. He has excellent technical knowledge, is able to adapt what he knows to customer projects quickly, can identify key areas for focus, and works hard to deliver them. I would fully recommend Andy to any data science / machine learning project, and would actively look to have him on my team again ”

**Mark Strefford | Engagement Manager**