Sebastian Strug

sebstrug.com   strugsebastian@gmail.com   07415 061 884   github.com/sebstrug

EXPERIENCE

Data Engineer, Confidential   **January 2020 – Present, London**
- Managing over 50 web scraping & parsing pipelines in Python including continuously running jobs and batch jobs.
  - Interacting with Postgres databases, AWS S3, Docker, Kubernetes.
- Aiding in the monitoring and maintenance of a pipeline scraping over 100,000 pages daily.

Data Analyst, PwC   **September 2018 – January 2020, Edinburgh**
- Data lineage and data cleansing project for large public sector client.
  - Built a change data capture process in SSIS to migrate and transform data daily from over 20 different historic Oracle databases into a centralised SQL server.
  - Communicated with and managed third-party consultants on behalf of the client, building a strong relationship with the client to ensure that the delivery provided value.
- Analysis and automation work for various financial services clients.
  - Performed analysis of financial services data in SQL and C#, with a dashboard in QlikView.
  - Spearheaded the automation of 430 FTE hours of routine work performed each year by creating Python scripts to query a third-party API.
  - Managed three FTE staff, providing guidance, and allocating and reviewing work.

Researcher, Nanoelectronics and Materials, McGill University   **January 2017 – July 2017, Montreal**
- Automated a set of microscope imaging tests used to determine the thickness of quasi-2D materials with Python.
- Independently researched a new material to be used for next-generation semiconductors using a variety of experimental techniques.

PROJECTS & HOBBIES

HowFarCanIGo   **July 2019**
- Built a tool in Python querying the Google Maps API to show how far you can travel for set times using various methods from any point.
- Motivated by a desire to understand how accessible London is from my home, I was able to generalise this to any location.
- Published my methodology and results online: 'Creating a Modern Day Travel Map of London' (Towards Data Science).

Open Source Software   **Ongoing**
- Contributed to the documentation for the Pandas library in Python, a library for data analysis used by over 150,000 developers.
- Understood the Git pipeline and workflow for a project with over 200,000 lines of code.

MeetMeHalfway   **April 2018**
- Created a website to find the cheapest flights for you and an international friend to meet 'halfway'.
- Built as part of a hackathon with a team of four in Python with Flask using the SkyScanner API.

EDUCATION

Physics MPhys (Hons), 1st Class, University of Manchester   **2014 — 2018, Manchester**
- Masters project on using convolutional neural networks to identify Chinese calligraphy, an as yet unsolved problem due to the variety of calligraphy. Made use of TensorFlow to produce neural networks, and generative adversarial networks, using TensorBoard to visualise results.
- Commended by the Chief Examiner for achieving an overall average above 85%.
- Study abroad year at McGill University, Montreal, achieving a GPA of 3.8/4.

St. Benedict's School, Ealing   **2000 — 2014, London**
- 5 A*s at A-Level in Mathematics, Further Mathematics, Physics, Economics, EPQ.
- 13 A*s at GCSE.

SKILLS, ACTIVITES & INTERESTS

Programming languages: Python (*experienced*), SQL (*experienced* – Postgres and T-SQL), Scala (*basic*), C++ (*basic*).

Other technical skills: AWS, Git, Docker, Azure, Agile workflows, CI/CD, Excel.

Languages: English (*native*), Polish (*mother tongue*), Spanish (*basic*).
- Varsity athlete with McGill rowing team, winning bronze at the Ontario University championships.
- Keen public speaker, having won Best Speaker at London Athenians Toastmasters and Didsbury Toastmasters.