

DOMINIC SMITH

Data Engineer

+32 471 76 24 02

<https://www.linkedin.com/in/dominic-smith-373470112/>

@ domlucasmith@gmail.com

Brussels, Belgium



SUMMARY

As an accomplished Lead Data Engineer with a product-oriented mindset, I possess a proven track record of driving innovation and optimising data infrastructure to deliver impactful solutions.

With a strong sense of ownership over my work and a commitment to excellence and accountability, I have lead successful projects while ensuring seamless collaboration across teams.

My communication and interpersonal skills facilitate productive relationships with colleagues and stakeholders at all levels, fostering an environment of shared vision and mutual understanding.

EXPERIENCE

Lead Data Engineer

Jetpack.AI

01/2019 - Present Brussels, BE

Delivered software projects individually and collaboratively, covering a wide spectrum of technologies and data

- Lead data engineering, science and SRE solutions across different Cloud Providers
- Mentored & coached colleagues to deliver consistent results & aligned business objectives with technological solutions
- Project management: lead team of developers, project management, adhering to safety standards (GDPR, PII)
- Designed and developed numerous ETL pipelines in Python or Scala to suit client needs
- Developed specific tools, to assure data quality, validity & governance

Data Engineer

Sentiance

01/2018 - 01/2019 Antwerp, BE

Integrated and scaled complex machine learning models, in Python, across AWS resources

- Developed and deployed systems to improve CI/CD, technical debt and testing and deployment of microservices
- Developed additional services, both as a microservice and a batch job, to implement in an existing Lambda architecture
- SRE experience in deploying load balancers & network configurations
- Redeveloped the data streaming platform, in Python, by introducing asynchronous calls to reduce latency and guarantee real-time processing

Software Engineer

CERN

09/2016 - 12/2017 Geneva, CH

Lead the upgrade, commissioning and maintenance of the software with Python & C++ and hardware of one of the key sub-systems of the CMS detector at CERN

- Assigned contact person for the sub-system, requiring daily monitoring and regular alignment with fellow software engineers
- Regular detector on-call shifts: a high-level duty, requiring 24h on-call availability for the complete CMS detector

SKILLS

Python	Spark	Cloud Infra	Docker
Java	C++	Scala	SQL/ NoSQL
K8s	Terraform	Airflow	APIs

KNOWLEDGE

Data Architecture

Experience in building, maintaining and evolving, scalable, resilient applications. Distributed systems and event driven architectures. Data Warehouse methodologies.

Cloud Based Infrastructure

Extensive experience with AWS and Azure services, and Infrastructure as Code tools.

Data Management & Reporting

Experience implementing dashboards, management & reporting technologies. Knowledge of analysing metrics about system health and performance.

SRE

Experience & knowledge of dev-ops, distributed systems, load balancing & networking.

Proficiency with languages

Proven record of delivering projects written in many languages using design patterns, testing & CI/ CD

ACHIEVEMENTS

Coached and mentored junior colleagues

Colleagues became confident & autonomous to handle tasks by themselves.

Developed and implemented real time data processing platform

By implementing a micro batch processing architecture, data was analysed in real time and results presented to the client.

Implemented new real time data streaming architecture to improve speed

By switching to an alternative storage architecture, the processing times were greatly reduced and the SLAs met their criteria.

Designed a data quality governance tool

Data quality issues were blocking many pipelines. By developing a data governance tool, issues were caught and repaired before affecting other systems.

EXPERIENCE

Data Scientist

CERN

09/2014 - 12/2017 Geneva, CH

Extensive use of Python and software libraries to extract, clean and analyse raw data from the CMS detector

- Utilised modern, analytical and batch processing tools (w. Python & C++) on data to statistically optimise results
- Deployed custom-made algorithms across distributed systems to process large datasets
- Designed, developed and deployed cutting edge tools (w. Python & R) to perform advanced analytics on raw data. Projects -Research

Software Engineer

CERN

01/2016 - 01/2017 Geneva, CH

Side Project during PhD CERN

- Developed alternative and novel algorithms in Python to reduce statistical noise in data obtained at CERN
- Main author of Python software that generated simulated data used to develop and test models

EDUCATION

PhD Experimental Particle Physics

Vrije Universiteit Brussel & University of Bristol

2013 - 2017 Brussels, BE & Bristol, UK

MSci Physics

University of Glasgow

01/2008 - 2013 Glasgow, UK

CERTIFICATION

Amazon Web Services Solutions Architect Associate

Amazon Web Services

LANGUAGES

English

Native



French

Native



Swedish

Intermediate



PUBLICATIONS

Alternative angular variables for suppression of QCD multi jet events in new physics searches with missing transverse momentum at the LHC

Arxiv

Dominic Smith, Tai Sakuma, Henning Flaecher,

03/2018 <https://arxiv.org/abs/1803.07942>

Publication Description