ADAM NEW
Big Data Engineer

ABOUT ME

I am hard-working and take a professional approach towards work, performing the roles I am given with confidence, maturity and reliability. My time spent working as a team member in an intense lab environment has taught me the importance of team coordination whilst maintaining efficiency and precision, knowing when to lead and when to follow.

I was first inspired to focus on data analytics in my future research placements and my career when I observed the folding@home project at university. It uses large scale simulation and analytics to predict previously unknown protein structures. These ranged from, misfolded proteins implicated in Alzheimer's disease to mutated proteins which may increase cancer risk, and are now being targeted in new cancer therapies. This scientific progress really opened my eyes to the value that can be extracted from large data sources.

Outside of work, I have a deep passion for music.

CONTACT US

Kubrick Group
T: 020 3866 4620
E: consultants@kubrickgroup.com
W: kubrickgroup.com

OVERVIEW

My primary experiences working with Big Data occurred in my research projects during university. My third-year research project revolved around large scale data evaluation and stratification work. I worked with patient serum to build a representative sample of antibody loop lengths relative to patient age. A suite of Python based tools were used to create molecular models of these structures, which underwent simulated stress tests.

During my masters year I decided to take a more data driven approach to research. I worked at the Randall Division of Cell and Molecular Biophysics, with one of my responsibilities being the collection and analysis of data derived from exposure of crystals to synchrotron radiation*.

I was delighted to be able to join Kubrick to now solely focus on data and data analytics in my career.

WORK

BIG DATA ENGINEER – KUBRICK GROUP, LONDON
At Kubrick Group I gained experience and qualified in: Professional and Communication Skills, IOC (Investment Operations Certificate) Level 1, Advanced Excel, Web Application development (UI/UX, JavaScript, CSS and HTML), Agile and SQL.
I then specialised in the Big Data Engineering course leading to accreditation as a Cloudera CCA Developer.

October 2016 – present
I worked on Big Data projects which exercised my skills in data ingestion and processing using Hadoop EcoSystem tools such as Sqoop, Flume and OpenSource Python Libraries including Pandas and Twython. We cleansed and munged data after exploratory data analysis through Python libraries including Plotly and Pandas then moving onto Spark to process, transform, analyse and stage the data into HDFS.

Projects included:
MetOffice Weather Data: involving Python API and Twitter processing into HDFS followed by further processing in the Hadoop Ecosystem.
Police Crime Data Analysis involving SSIS processing, SQL Analysis, and presentation through custom web front end.

STUDENT RESEARCHER, THE STEINER GROUP, KINGS COLLEGE LONDON
2016

STUDENT RESEARCHER, THE FRATERNALI GROUP, KINGS COLLEGE LONDON
2015

EDUCATION

MSCI BIOCHEMISTRY 2:1 2012 – 2016
King's College London

Haberdashers Monmouth Boys School 2008 – 2012

A LEVELS:
MATHEMATICS (A), FURTHER MATHEMATICS (A), BIOLOGY (A),CHEMISTRY (A)

ENGLISH LITERATURE AS LEVEL (A)

KEY SKILLS

Python Libraries (Anaconda):
PyMongo, Pandas, Numpy, scikit learn, Plotly and Twython
Cloudera Hadoop version 5.4:
HDFS, Yarn, Sqoop, Flume, Avro, Parquet, Impala, Hive and Hue
NoSQL: **SQL:**
MongoDB. API with PyMongo   SSIS, Advanced T-SQL and Stored Procedures
Apache Spark:
RDD, DataFrames, SparkSQL, PySpark with some exposure to and understanding of Scala
Source Control and IDE:
Git, SourceTree and Git Extensions, PyCharm
Other Skills:
Linux CentOS 6, SSH, Putty, Agile, JSON and BSON, Python Pickle and Shelve

*performed on the I-24 high brilliant beamline at the diamond synchrotron in Oxfordshire.