WILLIAM GORDON
Big Data Engineer

ABOUT ME

I have always enjoyed problem solving, logic puzzles and I've played chess from a young age. This passion for puzzles led me to choose a programming project during my degree.

My time spent at a theoretical physics group working on my own project confirmed how much I enjoy programming and working on larger projects.

I have now come to Kubrick to develop my IT skills and am looking forward to applying the skillset that Kubrick have provided in a professional environment.

CONTACT US

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

OVERVIEW

I recently graduated from Bath University in physics. My masters project was based on writing simulations of materials and I thoroughly enjoyed writing my own program and solving the problems in my code which is why I decided to pursue a career in IT.

I find the field of Big Data particularly interesting as it's real power seems relatively untapped. The opportunity to be a part of the next generation of data engineers and utilise the power of Big Data is one I am embracing and I look forward to realising the potential of Big Data as I continue in my IT 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.

August 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.

EDUCATION

UNIVERSITY OF BATH, MASTERS DEGREE IN PHYSICS (MPHYS) 2:1 CLASSIFICATION (2012-2016)

My masters research project was a 6 month theoretical project undertaken at the ESRF in Grenoble. It involved writing my own simulation of magnetic materials from scratch using Matlab. My research was in the relatively recent field of ultrafast magnetization dynamics where I was capable of catching up with the latest level of simulations and suggest future research.

ADDITIONAL TECHNOLOGY SKILLS

MatLab
Mathematica and Maple, a basic understanding from use in physics coursework

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

kubrick
A NEW FORCE IN TECHNOLOGY CONSULTING