

JOB DESCRIPTION

APPENDIX C

Before completing this form, please read the BTP 'Guide to writing job descriptions for Police Staff roles' Appendix B to the SOP.

A. POST DETAILS:

Job Title:	Data Developer	Current Grade:	B004
Department:	Analytics & Insight	Area:	FHQ
Reports To:	Data Science Delivery Manager	No of Posts:	1
Level of vetting:	MV	Post Number:	

PURPOSE OF THE POST: *Why the post exists and what it has to achieve*

Responsible for the technical data processing that will enable the data process from collection, cleansing, integration and modelling data to aid insights and predictive modelling within the data science team.

Working with a large, rich datasets that have a huge potential in providing untapped insights. This large data needs to be cleaned and maintained to a demandingly high standard, which will allow the data science team to enact their functions effectively.

Using programming and ETL processes to transform an extensive variety of datasets (ranging from crime-related, weather data to corporate organisational data) to aid the internal functions and decisions being made and assisting the BTP to develop its 'Big Data' vision (including processing and transformation of audio and visual data to enable predictive analytics).

The postholder will manage data workflows, build analytics tools that utilise the data pipelines, and ETL processes to provide actionable insights into key areas of operational and strategic policing decision.

C. DIMENSIONS OF THE POST *The key statistics associated with the post*

Financial – Direct or Non-Direct

Non Direct

Staff Responsibilities – Direct or Non-Direct

Direct

Any Other Statistical Data – Direct Responsibility

Manage the technical data platform for the Data science team that will enable core reporting for over 8 core system and lead to the provision of approximately 150 management reports delivering accurate and timely data from the frontline to the Chief Officer Group. Enabling the vision for self service and active use of Power BI within the BTP.

D. PRINCIPAL ACCOUNTABILITIES: *What the job is accountable for and required to deliver*

- Transform data using ETL (Extract, Transform and Load) processes to ensure the data is easily accessible, efficiently built to optimize the Force's data aspiration and optimize data retrieval and manage end to end data processing for the data science team from source, accessibility, security and storage
- Applying ETL tools and/or associated programming tools including scripts, XML, TSQL/PLSQL, Power BI, R and Python to transform data in the data science platform to enable integration, linkage and promote data quality. For example, where a variety of datasets exist, the Data Developer must identify the key data points that link all disparate data streams and utilise the appropriate programming to or script to link the data in the database. This will present itself seamlessly in the Cloud environment ready for analysis to be undertaken
- Execute ETL processing in the technical environment, over a variety of datasets so they can be used efficiently for reporting or analysis by data scientists.
- To actively seek out and collect a variety of data inputs from internal and external sources to inform predictive policing and enhance analytical insight. Inspect and migrate those data sources and link the various data outputs to enable advanced insights to be drawn from the data. For example, harvest open source data published via TfL re busiest periods and overlay against record crime, incidents, texts etc to establish patterns and trends associated with footfall
- Responsible for data quality within the data and enabling automated processes to ensure the data is usable and accurate for use by Data Scientists and Analysts ensuring that assurance, validation and linkage is maintained. For example, ensuring locations are mapped consistently using same naming conventions to increase data quality which will assist in accurate analysis.
- To build relationships and influence external data teams within other Forces and organizations to enable enhanced sharing of data in a secure environment. For example, visiting other forces and industry data teams to identify new data points and work with them to integrate and broaden BTP data repository.
- Work closely with Business Intelligence functions in Technology ensure the data structure in Microsoft Azure is efficient and accurate to enable a useful data environment for Data Scientists and Analysts to work within (Microsoft Azure is a cloud service for building, deploying and managing intelligent applications through a global network of data centres)
- To ensure that the data pipeline – the acquisition and processing of data – is working efficiently for the Data scientists and insight teams. For example, creating tests to capture any ETL failures or inefficient processing with a view to improve.
- Develop algorithms / pipelines to help make raw data more useful to the Data Science and Analytical teams
- Working closely with Technology, Information Governance and other cross functional teams to align architecture with business requirements and create end-to-end data solutions that are secure and support our analysts and business teams.
- Identify ways to improve data reliability, efficiency and quality within the data area used by Data Scientists
- Work with cross-functional teams to identify, design, and implement internal process improvements:

automating manual processes, optimizing data delivery, re-designing infrastructure for greater scalability, etc.

- Excellent communications skills are required including verbal and written. For example, the postholder may be required to compile recommendation papers for Chief Officer Group decision making
- Excellent interpersonal and organisational skills with a proven ability and confidence to network, negotiate and persuade stakeholders in a professional manner.
- Working with stakeholders and customers to ensure the successful delivery of change, implementation of standards and to influence decision-making.

E. DECISION MAKING:

Make decisions

-

Significant say in decisions

- Significant say in decisions of the design, management and organization of data within the data science platform. This will include designing the data science data platform and providing recommendations to Technology to improve data design
- Outputs of the products created by the Data Developer will be used to make strategic decisions at Chief Officer, BTPA and DfT level. For example, Data scientists regularly provide modelling scenarios to the Chief Officer Group, the data used for this will be the responsibility of the Data developer to maintain and ensure accuracy
- Outputs will significantly influence financial decision making and help to aid against reputational risk. For example the integration of data set by the Data developer will be used to identify efficient resourcing profiles

F. CONTACT WITH OTHERS: *The frequent contacts the post holder has with others and for what purpose*

Internal

Extensive liaison with Data scientists, Performance and Insight analysts, Information Governance, Business Intelligence Developers, Cyber Security Team (Tech), Police officers and staff

External

British Transport Police Authority (BTPA) – identifying ways to maintain data quality and integrity

Train Operating Companies and rail industry staff – building relationships to identify common data sources that be integrated into the Data Science data platform

Rail Delivery Group (RDG) – building relationships to identify common data sources that be integrated into the Data Science data platform

RSSB – building relationships to identify common data sources that be integrated into the Data Science data platform

Home Office – building relationships to identify common data sources that be integrated into the Data Science data platform and identifying best practise in data development for data science

Department for Transport (DfT)

Her Majesty's Inspectorate of Constabulary (HMIC) – identifying best practise in terms of data processing and data quality with a view to embed new and improved practice in the BTP

G. REQUIREMENTS: *The skills, knowledge, experience, qualifications and training required to perform the job.*

Essential Criteria:

Qualifications and Training:

- Minimum of Degree required in a Computer Science or related quantitative field.

Experience:

- Strong experience in the application of ETL tools and/or associated programming tools including scripts, XML, TSQL/PLSQL, Power BI, R and Python.
- Applied knowledge and experience of effective data management; sources, accessibility, security and storage approaches
- Ability to conduct data quality assurance, validation and linkage
- Proven experience of using technical solutions to solve complex data problems
- Strong experience in database interrogation and analysis tools, such as SQL and SAS
- Experience using MySql, Oracle and Mongo Db
- Strong experience of supporting and manipulating complex data sets within an Azure environment.
- Experience in using data to discover tasks that can be automated
- Ability to interpret user requirements and present data in a clear and compelling format using data visualisation tools such as Power BI
- Strong experience of working with business users and having the confidence to contribute positively within a team environment.
- Strong experience of working with cloud-based solutions and the Microsoft product suite

Skills:

- Strong skills in applying Hadoop, MapReduce, Hive, Pig, Data streaming, NoSQL, SQL, programming to new and big data
- Excellent communication skills
- Enhanced ability to build relationships and create networks across multi-disciplinary agencies to expand data repositories
- Ability to analyse and design data and data processes
- Ability to interpret and complete business requirements
- Attention to detail

Knowledge:

- Knowledge of how to develop dashboards, reports and other visualizations for stakeholders
- Robust understanding of data security and how to enforce this in a cloud environment
- Statistical methodologies and data analysis techniques

Desired Criteria:

Qualifications and Training:

- A qualification in a relevant discipline, preferably with a statistical, mathematical, computer technology or analytical focus

Experience:

- Experience of using Power Apps, BOXI & Power BI
- Deploy sophisticated analytics programs, machine learning and statistical methods
- Prepare data for predictive and prescriptive modelling

Knowledge:

- Knowledge of the wider technology and data developments
-

H. ANY ADDITIONAL INFORMATION: *Information relevant to the role, including any particularly challenging/ difficult aspects of the job. If competencies have been developed for this post, these can be listed here.*

I. AUTHORISATION DETAILS

Prepared By:	Vanita Patel	Date:	17/09//2020
Area Commander /FHQ	Vanita Patel	Date:	17/09/2020
HoD:		Date:	17/09/2020
Evaluation Panel:	Rewards Team	Date:	25/09/2020

REWARD