**RACHEL PEAKE, MBCS**

mob. 07423 540068

email: rachelpeakedba@outlook.com

**Summary**

I am a PRINCE2 qualified Data Engineer and member of the British Computer Society with 10 years + experience in front and back-end development using a windows stack. I have worked with companies such as the NHS, the University of Oxford and the Myers Briggs Company and have skills in SQL Server 2008R2, 2012 and 2016, Hadoop, HDFS, Hive, Spark, AWS, GCP and Python. I have supported the Leeds Teaching Hospital Trust Real-World Data team during the pandemic assisting with data modelling, governance and ETL processes to support COVID-19 research. I delivered several high-quality applications for the University of Oxford and the NHS including student feedback dashboards, Pharmacy Services dashboards, clinical research data collection tools and medical image sorting/anonymisation apps. I am skilled in requirements gathering and systems analysis, development, testing and deployment, data extraction and visualisation.

**Continuing Professional Development/Technical Skills**

|  |  |  |  |
| --- | --- | --- | --- |
| **Certifications/**  **Courses attended** | **Server/Cloud experience** | **Programming/**  **Mark up languages** | **Data tools** |
| PRINCE2 Foundation and Practitioner  Azure Fundamentals AZ-900  Azure Data Fundamentals DP-900  COMPTIA Server+  ITIL Foundations  SSIS  SQL for Data Science  Python for Data Science  R Intermediate | Windows Server  Apache  SQL Server 2012/2016  Google Cloud Platform  Azure  PowerShell | C  Java  PHP  XML  HTML  CSS  Python with Pandas, NumPy, PyMongo, PyDicom, tkinter | T-SQL  MySQL  PostgreSQL  Hive  Hadoop/HDFS  Spark/PySpark  Databricks  R  Google BigQuery  SSRS/SSIS |

**Experience**

**10/2021 – current Data Engineer**

ASDA

**Highlights**

* Migrating data from Walmart systems to new Azure platform and establishing new ETL pipelines to support BAU and new reporting requirements using SQL and PySpark.
* Establishing new ETL and ELT pipelines from on-prem data sources to Azure Synapse.
* Optimisation of PySpark jobs to ensure best performance when reporting to 3rd party customers.
* Ensuring technical design documents are up to date when a pipeline is deployed.
* Analysis of current data sets to identify gaps at source and rectify before the pipeline is scheduled for deployment.
* Supporting new starters to get development and production environments set up.
* Completed Azure training (AZ-900/DP-900), gaining more solid experience in Azure Storage, Data Factory and Synapse.

**10/2020 – 10/2021 Research Informatics Manager**

Leeds Teaching Hospitals Trust – Leeds

**Highlights**

* Rapid development of data pipelines to feed COVID-19 Research projects being delivered during the pandemic, meeting SLAs for 90-day turnaround of projects from application to data release.
* Rapid development of ETL scripts to make the right data accessible for the team pulling together datasets from structured and unstructured sources.
* Managing a small team of Data Managers working on collaborative data projects including building metadata for [DATA-CAN](https://www.data-can.org.uk/).
* Delivery of public training course, “How to use healthcare data for research” hosted by LTHT, covering regulatory compliance, extraction, cleaning, de-identification techniques, storage and publication of research involving routinely collected healthcare data.
* Working with colleagues in Information Governance, Research and Innovation and Informatics to ensure the correct approvals are in place for the collection and delivery of datasets for research projects.
* Working alongside researchers from Leeds Teaching Hospitals Trust, the University of Leeds, University of Oxford, and other partner organisations, providing a high quality of customer service whilst working to deliver complex datasets in agreed timeframes.

**04/2017 – 10/2020 Database Manager**

Leeds Biomedical Research Centre - Leeds

**Highlights**

* Rapidly fixed security issues, hard drive failures and created documentation for an inherited server farm which hosted a bespoke data collection system for clinical research data. This was achieved with minimal disruption to service users.
* Successfully delivered over 30 clinical research databases using SQL and MYSQL, pulling data into Power Bi and SSRS for reporting and visualisation.
* Managed the migration from physical to hybrid cloud infrastructure (Azure) to improve efficiency and reduce costs.
* Built and managed key data pipelines to source data from the NHS warehouse, developing ETL processes to pull data into our own bespoke research systems.
* Automated key tasks using shell, python, and SQL scripts.
* Implemented a suite of data management policies and procedures (based on DAMA-DMBOK framework) including audit documentation, data catalogues and metadata repository to document data assets.
* Successfully underwent a Medicines and Health Regulatory Agency (MHRA) inspection where the systems were audited against the regulatory standards for medicines trials.
* Worked with other key staff in the department to deliver novel trial designs and open them to recruitment using industry standard tools such as MACRO and REDCap.
* Designed and delivered training to researchers, nurses, and support staff to ensure data collection and management were performed in line with department procedures.

**08/2015 – 04/2017 Online Support Assistant / Course Design and Delivery Assistant**

University of Oxford - Oxford

**Highlights**

* Successfully delivered a dashboard to track helpdesk issues by creating ETL processes from ticketing system (trac) using PostgreSQL and Power Query. This data was then used to re-design the [online help website](https://onlinesupport.conted.ox.ac.uk/) and reduced the number of helpdesk issues per term by an average of 34% across the first year.
* Created data pipelines from student feedback data and Google Analytics to a data warehouse using PostgreSQL and Power Query. I designed a series of automated data visualisations and presented the results to the Head of Continuing Education who requested the results be shared each term with the course authors and design team to feed into course design decisions and marketing strategy.

**07/2010 – 07/2015 Operations Coordinator**

The Myers Briggs Company - Oxford

**Highlights**

* Design and implementation of data pipelines to feed sales and marketing initiatives.
* Automation of collation and analysis of customer feedback data using Excel VBA tools, reporting to Head of UK Sales.

**05/2008 – 07/2010 Personal Assistant to Lead IT Pharmacist**

Oxford Radcliffe Hospitals NHS Trust - Oxford

**Highlights**

* Developed SharePoint 2007/2010 intranet site for HR data reporting & performance dashboards.
* Developed data visualisations for Pharmacy Services re-design project.
* Design and build of data collection tools to assess the current state of pharmacy services across the South-Central region, and analysis of data sets with results reported to the South-Central Regional Chief Pharmacists group.

**Education**

2020-2021 Postgraduate Module: Data Management Open University

2020-2021 Postgraduate Module: Software Development Open University

2015-2016 Advanced Diploma in Data and Systems Analysis University of Oxford Distinction

2002 - 2006 BA (Hons) Modern History (with Computing) Oxford Brookes University

2:1

2000 - 2002 BTEC National Diploma in Computer Studies Tile Hill College, Coventry

Distinction