

Terms of References

Position Title: Data Informatics Manager

The Data Informatics Manager (DIM) will lead the design, implementation, and management of data centre of Khyber Medical University. The DIM will provide a role in architecting efficient data systems, ensure data availability and integrity, provide technical leadership in data-related initiatives and leverage data-driven insights.

Responsibilities

Develop data infrastructure strategy

Formulate a strategy for building, managing and maintaining data infrastructure with particular focus on:

- Data governance and policies: Define roles, responsibilities, and data ownership. Develop policies and procedures for data collection, storage, and sharing.
- Data access and security: Outline and establish data security measures, including encryption, hashing, access controls, sharing protocols, and authentication.
- Data storage and management: Define procedures and tools for managing the data. Determine the mode and nature of data storage. Develop data quality standards.
- Advanced data analytics: Explore the adoption of tools for data analytics and procedures for sharing the insights.

Design and implement data architecture

- Design and implement data architectures, databases, and data pipelines for efficient data collection, storage, and retrieval.
- Assess the hardware and software requirements and define data requirements and data flow processes.
- Install appropriate servers and software dependencies/packages.
- Create user roles, access controls and permissions on the server(s).
- Install and configure the appropriate database systems (MySQL, PostgreSQL, MongoDB, MariaDB) on the server(s).
- Create database schemas, tables, relationships, and constraints.
- Define access controls and permissions at the table and column level to restrict data access based on user roles.
- Write documentation for the database schema, including table structures, data types, relationships, primary keys, foreign keys, and constraints.
- Create data pipelines by extracting data, performing transformations (such as cleaning, normalization and filtering), and loading the processed data into locally hosted databases.

Data and databases management

- Cleanse and validate datasets to remove inconsistencies, errors, duplicate records.

- Normalize data by standardizing formats (e.g., dates, missing fields, long/wide format).
- Implement access controls to restrict unauthorized access to sensitive data on system(s) and databases and create context specific queries for limited access to databases.
- Develop and implement strategies for data privacy and security.
- Implement regular data backups of the locally hosted databases. and ensure data integrity during recovery processes.
- Prepare datasets for surveys
- Design databases to facilitate data analysis and reporting.

Servers and network services setup and management

- Migrate infrastructure from existing servers.
- Create and configure hostname, IP addresses, domains/subdomains.
- Configure DNS for domains/subdomains and point them to apps, including ports mapping with docker-proxy and SSL proxy where needed.
- Install SSL certificates to domains for secure connections
- Set up monitoring systems to track the health, performance, and availability of the server components, storage, and memory.
- Take security measures to prevent intrusions.
- Perform routine maintenance of the server.

Develop web portals and user management systems.

- Build user-friendly graphical user interfaces (GUIs) for data access, download, upload and interaction adhering to the best practices in user experience design.
- Create multiple roles for the user of the web portals such as admin role, staff role, researcher role, student role, enumerator role.
- Integrate in real-time the data from the data collection tools into the web data portals.
- Anonymise the data where required.

Build and deploy applications

- Deploy data collection tools.
- Build apps for automated update of datasets for data collection tools.
- Deploy, configure and manage secured data-sharing clouds.
- Build dashboards for monitoring of surveys and visualization of survey and other datasets.
- Build apps for generation of automated reports of survey data and other datasets.
- Build apps for transforming raw data into structured formats suitable for analysis and reporting.

Data Analytics

- Utilize statistical techniques to derive insights from data, contributing to data-driven decision-making and enhancing the organization's knowledge repository.

- Collaborate with data professionals to explore methods and support innovation.
- Build data visualization dashboards and automate reports creation.

Provide technical leadership

- Lead a team of professionals engaged in data management, analysis, and technology-related tasks.
- Work with the team to formulate the security framework for internal and external data sharing.
- Provide technical solutions to staff members to address gaps and inconsistencies in the infrastructure.
- Provide technical support to the organization staff involved in data management and IT operations.

Data centre operations and management

- Oversee the development, operation, management of the data centre infrastructure.
- Provide strategic guidance and leadership to the teams engaged in data infrastructure.
- Manage day-to-day operations of the centre.
- Monitor data centre performance, uptime, and availability to ensure seamless operations.
- Forecast data centre capacity needs and plan for scalability to accommodate future growth.
- Optimize resource utilization and allocation to ensure cost-effective operations.

Qualifications

- Master's degree or Bachelor's degree in Economics, Data Science, Computer Science, Information Systems or a related technical field with a 5 years relevant experience in data management, data analysis, data infrastructure and report writing.
- Experience with ETL (extract, transform and load) design, implementation and maintenance.
- Experience working with major database platforms such as MySQL, Postgres SQL, Oracle server.
- Experience with Linux command line, writing bash scripts, cronjob or task scheduler, and virtualization technologies such as Docker.
- Proficiency in data analysis tools and programming languages (e.g., R, Python, SQL).
- Proficiency in data visualization tools (e.g., R Shiny, Python Dash, Tableau, Power BI).
- Strong analytical skills with the ability to interpret complex data sets and generate actionable insights.
- Proficiency in report writing tools (R Markdown, Jupyter Notebook).
- Excellent teamwork and interpersonal skills.