Solution Architect (Data), Job Profile

|  |
| --- |
| Purpose of role |
| A **Software Architect - Data** is responsible for design, implementation, and management of data solutions within organization. This role involves architecting scalable and cost-effective data solutions that harness cloud-native services for storage, processing, and analytics. Responsibilities include designing efficient data pipelines and storage solutions, ensuring adherence to data security and compliance standards, and optimizing performance for environments. Additionally, the job role demands close collaboration with cross-functional teams to understand business requirements and deliver robust data solutions that align with organizational objectives. |
| Key responsibilities |
| **Main:**   * Responsible for designing the overall data architecture that meets the organization's needs. This involves understanding business requirements, defining data models, selecting appropriate technologies, and designing data flows. * Create and maintain conceptual, logical, and physical data models that represent the organization's data assets. This includes defining data entities, relationships, attributes, and ensuring data integrity and consistency. * Design solutions for integrating data from various sources, such as databases, applications, and external systems. They may use ETL (Extract, Transform, Load) processes, data pipelines, or real-time integration techniques to ensure data flows smoothly across different systems. * Ensure data governance policies and security measures are in place to protect sensitive data and comply with regulations. * Define data access controls, encryption methods, and audit mechanisms to safeguard data assets. * Optimize data solutions for performance and scalability by tuning database configurations, optimizing queries, and implementing caching mechanisms. They also monitor system performance and adjust the configurations as needed. * Understand cloud-native data services and architect solutions that leverage cloud security, scalability, and flexibility, leveraging services like AWS, Azure, Google Cloud Platform, snowflake. * Collaborate with cross-functional teams including data engineers, software developers, business analysts, and stakeholders to understand requirements and deliver effective solutions. They also communicate technical concepts to non-technical stakeholders in a clear and understandable manner. * Integrate continuous integration and continuous deployment (CI/CD) practices into data solution development processes. This involves automating testing, deployment, and monitoring of data pipelines and solutions, ensuring faster delivery cycles and higher quality deployments. * Involve in project management activities such as project planning, resource allocation, and tracking project progress to ensure timely delivery of data solutions. * Documenting data architecture designs, standards, and guidelines to communicate architectural decisions and promote consistency across projects. This involves creating data dictionaries, data flow diagrams, and architecture diagrams to facilitate understanding and collaboration among stakeholders. * Providing technical leadership and guidance to team, including data engineers, developers, and business analysts. Staying up to date about emerging trends, technologies, and best practices in data management and architecture. Evaluating new tools and techniques, experimenting with innovative solutions, and continuously improving data architecture practices to drive business value and competitive advantage. |
| Specialist skills and experience |
| **Essential:**   * Deep understanding of major cloud platforms such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), and their respective services for compute, storage, networking, databases, and analytics. * Expertise with modern data warehousing solutions like Snowflake (preferred), Amazon Redshift, Google Big Query, including their features, capabilities, and best practices for data security, storage, querying, and analytics. * Strong hands-on expertise in big data technologies - Hadoop, Apache Spark, Apache Kafka for handling large volumes of structured and unstructured data, as well as real-time data processing and analytics. * Proficiency in data integration tools and platforms such as Apache NiFi, Apache Airflow, Informatica, AWS Glue for building and managing data pipelines to ingest, transform, and load data from various sources into data warehouses and lakes. * Understanding of data modelling techniques and tools such as ER diagrams, dimensional modelling, and data modelling tools like ERwin, ER/Studio, for designing and documenting data structures and relationships. * Knowledge of relational databases (Snowflake, Teradata) as well as NoSQL databases (MongoDB, Cassandra) for different data storage and processing requirements. * Expertise in programming languages commonly used in data engineering and analytics such as Python, SQL, shell scripting. * Experienced in implementation of machine learning and AI technologies and frameworks for building predictive analytics, recommendation systems, and other advanced analytics solutions on cloud platforms. * Hands-on with DevOps principles and practices for automating deployment, monitoring, and management of data solutions using tools like Git, Jenkins, Terraform. * Proficiency in containerization technologies like Docker and container orchestration platforms like Kubernetes for deploying and managing containerized applications and microservices in cloud environments.   **Good to have:**   * Experience in Data Cataloging and Metadata management. * Understanding of graph database concepts and technologies like Neo4j * Knowledge of emerging data architecture practices like Data Mesh for decentralized data ownership, domain-oriented data architecture. * Serverless data processing pipelines. * Knowledge of Data virtualization.   **Desired:**   * Certifications – TOGAF/DAMA/IBM   Relevant Data engineering/Analytics certifications from GCP/AWS/Microsoft |
| Competencies / Behaviors |
| * Team engagement: Builds effective working relationships. * Collaboration: Communicates effectively with peers, supervisors, and cross functional teams. * Learn and Grow:  Continuously invest in learning and upskilling technology and business acumen. * Analytical thinking: Thinks critically, providing well-reasoned insights through a data driven approach. * Agility: Quickly adapts and remains flexible while managing risks * Drives performance: Continually raises the bar for oneself (and others) * Acts with integrity:  Embrace Organizations values as an integral part of daily interactions and does the right thing. * Delivers results: Exhibits strong time management skills, and commitment to delivering tasks on time to the highest quality standards. |