

miermontoto /
tfg

<> Code

Issues 3

Pull requests

Actions

Projects 1

Security

In

tfg / doc / requisitos.md



miermontoto listado inicial de requisitos

yesterday



20 lines (19 loc) · 3.27 KB

Preview

Code

Blame

Raw



Listado inicial de requisitos

Requisitos funcionales

- **Integración de Datos:** El datalake ha de ser capaz de integrar datos provenientes de las **bases de datos internas** de la plataforma base, así como de fuentes externas, incluyendo aplicaciones de terceros y otras bases de datos.
- **Transformación y Limpieza de Datos:** Debe ofrecer funcionalidades para transformar y **limpiar los** datos, incluyendo la eliminación de duplicados, la corrección de formatos y la estandarización de datos para asegurar su integridad y calidad.
- **Gestión de Metadatos:** El sistema debe gestionar metadatos de manera eficaz, permitiendo a los usuarios entender el origen, el contenido y el contexto de los datos almacenados en el mismo.
- **Automatización del Flujo de Datos:** Automatización en la ingestión, transformación, y carga de datos (ETL) para asegurar la actualización y disponibilidad constante de los mismos.
- **Soporte para Análisis de Datos:** Debe permitir la realización de análisis de datos **complejos, incluyendo el procesamiento** de grandes volúmenes de datos.
- **Interfaz de Consulta Flexible:** El datalake debe ofrecer **una interfaz de** consulta flexible y potente, permitiendo a los usuarios realizar búsquedas complejas y obtener insights de los datos integrados. El sistema ha de estar pensado para poder ofrecer diferentes salidas de datos como pueden ser una intranet y consultas para clientes de la plataforma base.
- **Integración:** Debe poder ofrecer un sistema que permita la integración con otras aplicaciones, facilitando la extracción, carga y manipulación de datos desde

sistemas externos.

- **Gestión de la Calidad de Datos:** Herramientas y procesos para monitorear y mejorar continuamente la calidad de los datos almacenados en el datalake.

Requisitos no funcionales

- **Escalabilidad:** Capacidad de escalar horizontal y verticalmente para manejar incrementos en el volumen de datos sin degradar el rendimiento o la disponibilidad.
- **Disponibilidad:** Alta disponibilidad del sistema, incluyendo redundancia y mecanismos de recuperación ante desastres para asegurar el acceso continuo a los datos.
- **Seguridad y Privacidad de los Datos:** Implementación de medidas de seguridad avanzadas, incluyendo control de acceso basado en roles, encriptación de datos en reposo y en tránsito, auditorías de seguridad y protección contra amenazas internas y externas.
- **Rendimiento:** Capacidad de minimizar los tiempos de respuesta atendiendo a grandes volúmenes de información y/o consultas complejas.
- **Interoperabilidad:** Integración con otros sistemas y tecnologías, incluyendo plataformas de análisis de datos, herramientas de visualización y aplicaciones empresariales.
- **Eficiencia de Costos:** Optimización del uso de recursos para minimizar los costos operativos y de almacenamiento, incluyendo el uso eficiente del almacenamiento y los costes de despliegue asociados a los servicios cloud.
- **Mantenibilidad:** Facilidad para actualizar, configurar y mantener el sistema, incluyendo la documentación adecuada y el soporte para la detección y solución de problemas.
- **Cumplimiento Normativo:** Asegurar que el sistema cumple con todas las leyes y regulaciones relevantes con la GDPR, así como las exigencias marcadas por la plataforma sobre la que se origina.