

# Análisis de Caso: Arquitectura de Datos –InfoHealth

---

## Diagnóstico de la situación actual

La empresa InfoHealth enfrenta múltiples desafíos en la gestión de datos:

- Fuentes de datos heterogéneas: Historias clínicas, sensores IoT, formularios, emails y hojas de cálculo, sin integración centralizada.
- Duplicidad y falta de trazabilidad: Información repetida, sin control de versiones ni rastro de origen.
- Riesgos de seguridad: Posible exposición de datos sensibles por falta de políticas robustas.
- Desconfianza en los reportes: Directivos y médicos cuestionan la calidad y puntualidad de los datos.
- Procesos manuales: Analistas dedican tiempo excesivo a depurar información.

## Propuesta de arquitectura base por capas

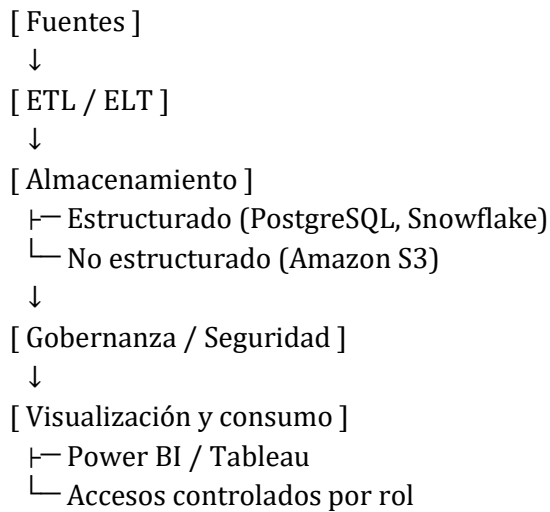
### Principios de arquitectura aplicados (según DAMA-DMBOK)

1. Calidad de datos: Toda arquitectura debe garantizar precisión, completitud consistencia.
2. Gobierno de datos: Se requiere control, trazabilidad y responsabilidad sobre los datos.
3. Seguridad de la información: La protección de los datos personales de pacientes es crítica.

### Justificación del diseño

- Es escalable, ya que utiliza servicios cloud y desacopla fuentes, almacenamiento y visualización.
- Mejora la eficiencia en el trabajo de analistas mediante automatización de ETL.
- Facilita la interoperabilidad clínica a través de estándares como HL7/FHIR.
- Responde a exigencias del sector salud con herramientas que soportan trazabilidad, seguridad, y cumplimiento normativo (HIPAA, GDPR).

## Esquema visual de la solución



## Conclusión

InfoHealth necesita una arquitectura moderna, modular y gobernada. Esta propuesta le permitirá integrar sus múltiples fuentes, mejorar la confianza en los datos, y ofrecer acceso seguro y eficiente a todos sus usuarios, desde médicos hasta gerencia.

## Resumen de capas y herramientas sugeridas

Capa	Descripción propuesta	Herramientas sugeridas
Fuentes de datos	Integración de EHR, sensores IoT, formularios digitales, correo y archivos compartidos	HL7/FHIR, MQTT, Formularios web, APIs REST
Almacenamiento	Separación entre datos estructurados y no estructurados; almacenamiento escalable en la nube	PostgreSQL, Snowflake, Amazon S3, Azure Blob
Procesamiento y limpieza	ETL/ELT para consolidar y transformar datos, eliminando duplicados	Apache NiFi, Talend, dbt, Python
Acceso y visualización	Dashboards con acceso controlado por perfil	Power BI, Tableau, Metabase
Seguridad y gobernanza	Políticas de acceso, catálogo de datos, encriptación y auditoría	Apache Atlas, AWS Lake Formation, IAM, DLP