

OBSERVACIONES DE MEJORA DEL DOCUMENTO DE ARQUITECTURA BI

AREAS DE MEJORA:

1. Seguridad en el Acceso a Datos y APIs:

- 1.1 La documentación no detalla el uso de mecanismos como Managed Identities en Azure para asegurar las conexiones a servicios (OpenAI, SQL).

Sin Managed Identities, tendríamos que guardar claves o credenciales de Azure SQL o Azure OpenAI en configuraciones o variables de entorno, lo cual es un riesgo de seguridad.

Con Managed Identities, Azure Functions usaría su identidad asignada para conectarse a Azure SQL y OpenAI sin manejar credenciales. El servicio SQL o OpenAI verificaría su identidad mediante AAD y otorgaría acceso.

Esto elimina el riesgo asociado con la exposición de claves y asegura que las conexiones sean gestionadas de forma centralizada y segura.

Nota: Se debe aclarar este punto.

- 1.2 A la documentación le falta mencionar prácticas de encriptación de datos en tránsito y en reposo, como HTTPS y Transparent Data Encryption (TDE) en SQL.

2. Escalabilidad:

- 2.1 En la documentación no se describe cómo manejar grandes volúmenes de datos

3. Logging y Monitoreo:

- 3.1 En la documentación se menciona el manejo de errores, pero no se especifica si se integran herramientas como Azure Monitor o Application Insights para rastrear fallos o tiempos de respuesta en producción.

Es importante monitorear métricas del uso de tokens en Azure OpenAI, ya que influye directamente en costos.

4. Validación de Datos

- 4.1 Sería útil incluir y/o aclarar lo de la capa de validación previa en las transcripciones para garantizar que no contienen caracteres no válidos o datos incompletos antes de su procesamiento, ya que una validación previa reduce errores, mejora la eficiencia y asegura que el flujo de trabajo sea robusto frente a datos inconsistentes.

5. Resiliencia y Tolerancia a Fallos

- 5.1 No se menciona si las funciones de procesamiento tienen reintentos automáticos en caso de errores temporales.
- 5.2 No se menciona la Implementación de colas (por ejemplo, Azure Queue Storage) el cual podría mejorar la resiliencia, permitiendo un procesamiento asíncrono en lugar de depender únicamente de activaciones HTTP.

6. Detallado de Métricas de Eficiencia

- 6.1 En la documentación no hay indicación de métricas clave para medir el desempeño del sistema, como el tiempo promedio de procesamiento por fila o la efectividad de las extracciones

SUGERENCIAS Y CONSULTAS ESPECIFICAS:

Documentación Adicional: Por favor agregar diagramas de flujo o secuencia para detallar las interacciones entre componentes, con un énfasis en excepciones.

Integración de IA Responsable: Detallar cómo se manejan posibles sesgos en las respuestas de Azure OpenAI y si existen validaciones de calidad sobre los resúmenes generados.

Automatización de Implementación: Uso de plantillas de Azure Resource Manager (ARM) o Terraform para garantizar la consistencia en la infraestructura.

Es importante por parte de BIT-C suministrar las plantillas de implementación (ARM o Terraform) como parte de su entrega del proyecto, dado que son herramientas esenciales para garantizar que la infraestructura en Azure se configure de manera consistente, reproducible y acorde a las mejores prácticas.

El hipervínculo mencionado en el documento **Links a diagramas en MIRO**, no conduce a un sitio web, para revisarlos y/ o descargarlos.

Con estos ajustes y/o aclaraciones en la documentación, el sistema sería más robusto, escalable y estaría mejor preparado para evitar brechas en una plataforma de nube con Azure.