

Guía de estudio 2

Modelos de datos y lenguajes de consulta

1. Propósito del estudio

Analizar cómo los **modelos de datos** y los **lenguajes de consulta** influyen de manera decisiva en la arquitectura de los sistemas, en los patrones de acceso a la información, en la complejidad del software y en la capacidad de evolución a largo plazo.

2. Modelo de datos como abstracción fundamental

Contenidos clave

- El modelo de datos define cómo se estructura la información.
- Determina qué operaciones son naturales, costosas o inviables.
- Condiciona el diseño del software y las interfaces de consulta.
- No existe neutralidad: toda elección implica trade-offs.

Preguntas de evaluación

1. ¿Por qué el modelo de datos se considera una decisión arquitectónica y no solo técnica?
2. ¿Cómo influye el modelo de datos en la expresividad de las consultas?
3. ¿Qué riesgos implica elegir un modelo de datos inadecuado para el dominio?

3. Modelo relacional

Contenidos clave

- Organización en tablas con filas y columnas.
- Uso de claves primarias y foráneas.
- Normalización para reducir redundancia.
- Lenguajes declarativos para consultas complejas.
- Separación entre modelo lógico y físico.

Fortalezas

- Manejo natural de relaciones complejas.
- Consultas ad hoc potentes.
- Madurez conceptual y operativa.

Limitaciones

- Fricción con estructuras jerárquicas.
- Desajuste con modelos orientados a objetos.
- Coste cognitivo de joins extensivos.

Preguntas de evaluación

4. ¿Qué ventajas ofrece la normalización en el modelo relacional?
5. ¿Por qué los joins son una fortaleza y al mismo tiempo una fuente de complejidad?
6. ¿En qué tipos de dominios el modelo relacional resulta especialmente adecuado?

4. Modelo documental

Contenidos clave

- Datos representados como documentos autocontenidos.
- Estructuras jerárquicas y anidadas.
- Acceso por agregados completos.
- Menor necesidad de joins.

Fortalezas

- Alta localidad de datos.
- Reducción del desajuste objeto–persistencia.
- Simplicidad para ciertos patrones de acceso.

Limitaciones

- Dificultad para modelar relaciones complejas.
- Riesgo de duplicación de datos.
- Complejidad creciente cuando aumentan las referencias cruzadas.

Preguntas de evaluación

7. ¿Qué se entiende por agregado y por qué es central en el modelo documental?
8. ¿En qué escenarios el anidamiento es preferible a la normalización?
9. ¿Qué problemas aparecen cuando un dominio altamente relacional se modela como documentos?

5. NoSQL y persistencia políglota

Contenidos clave

- NoSQL como familia heterogénea de sistemas.
- Respuesta a necesidades de escalabilidad y flexibilidad.
- Convivencia de múltiples modelos en un mismo sistema.
- Uso de cada tecnología según su patrón de acceso.

Preguntas de evaluación

10. ¿Por qué NoSQL no debe entenderse como una alternativa única al modelo relacional?
11. ¿Qué implica adoptar una estrategia de persistencia políglota?
12. ¿Qué riesgos arquitectónicos surgen al usar múltiples modelos de datos?

6. Desajuste entre modelos y aplicaciones

Contenidos clave

- Diferencia entre estructuras en memoria y persistencia.
- Necesidad de capas de traducción (ORM, repositorios).
- Coste cognitivo y técnico del mapeo.

Preguntas de evaluación

13. ¿Qué es el desajuste objeto–modelo de datos?
14. ¿Por qué los ORMs no eliminan completamente este problema?
15. ¿Cómo influye este desajuste en el diseño de la aplicación?

7. Relaciones y complejidad estructural

Contenidos clave

- Relaciones simples favorecen modelos documentales.
- Relaciones many-to-many favorecen modelos relacionales o de grafos.
- La densidad relacional condiciona la elección del modelo.

Preguntas de evaluación

16. ¿Cómo afecta la densidad de relaciones a la elección del modelo de datos?
17. ¿Por qué algunos modelos se degradan al aumentar la complejidad relacional?

18.Dé un ejemplo de dominio donde las relaciones sean el factor dominante.

8. Convergencia de modelos

Contenidos clave

- Relacionales con soporte para datos semiestructurados.
- Documentales con agregaciones y referencias.
- Aparición de modelos híbridos.

Preguntas de evaluación

19.¿Qué significa la convergencia entre modelos relationales y documentales?

20.¿Qué ventajas ofrecen los enfoques híbridos?

21.¿Qué nuevas complejidades introducen estos enfoques?

9. Lenguajes de consulta

9.1 Lenguajes imperativos

Contenidos clave

- Describen cómo obtener el resultado paso a paso.
- Control total desde la aplicación.
- Menor capacidad de optimización automática.

Preguntas de evaluación

22.¿Cuáles son las ventajas y desventajas de los lenguajes imperativos?

23.¿Por qué estos lenguajes trasladan la responsabilidad de optimización a la aplicación?

9.2 Lenguajes declarativos

Contenidos clave

- Describen qué resultado se desea.
- Separación entre intención y ejecución.
- Mayor capacidad de optimización y paralelización.

Preguntas de evaluación

24.¿Por qué los lenguajes declarativos favorecen la escalabilidad?

25.¿Qué ventajas ofrecen frente a cambios en volumen o distribución de datos?

10. Procesamiento distribuido

Contenidos clave

- Procesamiento por etapas (map/reduce).
- Adecuado para grandes volúmenes de datos.
- Evolución hacia lenguajes y DSLs de mayor nivel.

Preguntas de evaluación

26.¿En qué escenarios es adecuado el procesamiento distribuido tipo MapReduce?

27.¿Por qué surgen lenguajes declarativos sobre estos modelos?

11. Modelos de grafos

Contenidos clave

- Representación basada en nodos y aristas.
- Consultas por patrones de relación.
- Adecuado para dominios altamente conectados.

Preguntas de evaluación

28.¿Cuándo un modelo de grafos es preferible a uno relacional?

29.¿Qué ventajas ofrece la consulta por patrones?

30.Compare grafos con propiedades y almacenes de triples.

12. Lenguajes lógicos y reglas

Contenidos clave

- Uso de reglas declarativas.
- Consultas composable.
- Potente para inferencia y dominios complejos.

Preguntas de evaluación

31.¿Qué tipo de problemas se benefician del uso de lenguajes basados en reglas?

32.¿Por qué estos lenguajes no suelen usarse para consultas simples?

13. Esquema y validación

Contenidos clave

- El esquema siempre existe.
- Decisión clave: validación en escritura vs lectura.
- Impacto en consistencia y evolución.

Preguntas de evaluación

33.¿Qué diferencias existen entre schema-on-write y schema-on-read?

34.¿Cómo afecta esta decisión a la calidad de los datos?

14. Síntesis integradora

Contenidos clave

- No existe un modelo de datos universalmente óptimo.
- La elección depende del dominio, los patrones de acceso y la evolución esperada.
- El diseño consciente evita complejidad innecesaria.

Preguntas de evaluación

35.¿Qué criterios deberían guiar la selección de un modelo de datos?

36.Analice un sistema real e identifique por qué se eligieron sus modelos de datos.