

I. Introduction

Le **data modeling** (modélisation des données) est une étape cruciale dans la conception de systèmes d'information. Il définit la structure, les relations et les contraintes des données pour répondre à des besoins métier spécifiques.

Objectif : Cette étude compare les principales techniques de modélisation, leurs avantages, inconvénients, et cas d'usage idéaux.

II. Techniques de Data Modeling

1) Modèle Relationnel

Ce modèle structure les données en tables liées par des clés primaires/étrangères.

- **Caractéristiques :**
 - Normalisation (1NF, 2NF, 3NF).
 - Langage SQL pour les requêtes.
 - Respect des contraintes ACID (*Atomicité, Cohérence, Isolation, durabilité*).
- **Avantages :**
 - Intégrité des données.
 - Flexibilité des requêtes complexes.
- **Inconvénients :**
 - Complexité à grande échelle.
 - Performances limitées sur les Big Data.
- **Cas d'usage :**
 - Systèmes transactionnels (ERP, CRM).
 - Applications bancaires.

2) Modèle Dimensionnel

Il est optimisé pour l'analyse via des schémas en étoile ou en flocon.

- **Caractéristiques :**
 - Tables de faits (métriques) et dimensions (contexte).
 - Dénormalisation partielle.
- **Avantages :**
 - Requêtes analytiques rapides.
 - Simplicité pour les utilisateurs métier.
- **Inconvénients :**
 - Peu adapté aux transactions OLTP.
 - Rigidité face aux changements de besoins.
- **Cas d'usage :**
 - Data Warehouses.
 - Tableaux de décision (BI).

3) Modèles NoSQL

C'est un modèles non relationnels adaptés aux données non structurées ou semi-structurées.

Type	Document (MongoDB)	Clé-Valeur (Redis)	Colonnes (Cassandra)	Graphe (Neo4j)
Structure	JSON/BSON	Paires clé-valeur	Colonnes dynamiques	Nœuds + relations
Avantages	Flexibilité schéma	Latence ultra-faible	Scalabilité horizontale	Requêtes relationnelles complexes
Inconvénients	Pas de jointures natives	Fonctionnalités limitées	Complexité de requêtes	Courbe d'apprentissage raide
Cas d'usage	Catalogues produits	Cache, Sessions utilisateur	IoT, Logs	Réseaux sociaux, Fraud detection

4) Data Vault

C'est un modèle hybride combinant le relationnel et l'historisation.

- **Caractéristiques :**
 - Hubs (entités), Links (relations), Satellites (historique).
 - Agilité face aux changements métier.
- **Avantages :**
 - Auditabilité complète.
 - Adapté aux entrepôts de données complexes.
- **Inconvénients :**
 - Complexité de mise en œuvre.
 - Requêtes lentes sans optimisation.
- **Cas d'usage :**
 - Entreprises régulées (santé, finance).
 - Data Lakes.

5) Modèle Entité-Attribut-Valeur (EAV)

Il Stocke les données en triplets (Entité, Attribut, Valeur).

- **Avantages :**
 - Schéma ultra-flexible.
 - Idéal pour les données hétérogènes.

- **Inconvénients :**
 - Requêtes complexes.
 - Pas de typage natif.
- **Cas d'usage :**
 - Systèmes de gestion de contenu (CMS).

III. Conclusion

Aucune technique n'est universelle :

- **Startups/Scale-ups :** Privilégiez le NoSQL pour sa flexibilité.
- **Grands groupes :** Optez pour un mélange Data Vault + Dimensionnel.
- **Secteur régulé :** Combinez Relationnel (OLTP) et Data Vault (audit).