

Ressource R1.05 : Introduction aux bases de données et SQL

Modélisation de Bases de Données

Mathieu PALOSSE

Professeur de Numérique Sciences Informatiques

IUT d'informatique de Rodez

mathieu.palosse@iut-rodez.fr

BUT 2022/2023

C4

R1.05

1/36

Organisation

Savoirs étudiés :

- Approche de la conception des bases de données : modèle conceptuel de données et traduction vers le modèle relationnel + contrainte simple + redondance
- Algèbre relationnelle

Volume horaire :

- 4 CM de 1h
- 10 TP de 1h15
- 2 évaluations sur table de 1h : semaine 42 et 47
- Tests et QCM ponctuels

S.A.E. S1.04 : Création d'une base de données

- Semaines 39 à 48
- Soutenances semaine 48

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

2/36

Références

C. CHRISMENT, K. PINEL-SAUVAGNAT, O. TESTE, M. TUFFERY, *Bases de données relationnelles : concepts, mise en oeuvre & exercices*, Hermes-Lavoisier. **2008**, ISBN 978-2-7462-2086-7

G. GARDARIN, *Bases de données*, Eyrolles, **2003**, ISBN 2-212-11281-5

M. CLOUSE, *Algèbre relationnelle : Guide pratique de conception d'une base de données relationnelle normalisée*, Editions ENI, **2008**, ISBN 978-2-7460-4154-7

Cours de l'équipe Bases de Données de l'Institut de Recherche en Informatique de Toulouse (IRIT – CNRS – INP – UT3 – UT1 – UT2J)

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

3/36

Plan du cours

I. Introduction aux Bases de Données

II. Modélisation Conceptuelle

III. Le Modèle Relationnel

IV. Algèbre relationnelle

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

4/36

I. Introduction aux Bases de Données

- 1) Pourquoi les bases de données (BD) ?
- 2) BD et SGBD
- 3) Repères historiques
- 4) Conception d'une BD

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

5/36

I. 1) Pourquoi les BD ?

Dans une structure (entreprise, université, etc), le système d'information (SI) est possède :

Des informations qui apparaissent :

- À des instants différents
- À des endroits différents
- Sous des formes différentes

Une gestion de l'informations automatisée, avec une multiplicité de :

- Fichiers
- Applications
- Langages
- Supports physiques de stockages
- Normes
- Equipes de développement
- ...

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

6/36

I. 1) Pourquoi les BD ?

Exemple : le SI de l'Université, vu par...



Service comptabilité



Enseignant



Secrétariat

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

7/36

I. 1) Pourquoi les BD ?

Exemple (suite) : Approche par fichiers (1)

Nom: Jean
Duroux
Année: L2
Chimie
Frais: 500
euros



Service comptabilité

Nom: Jean
Duroux
Année: L2
Chimie
Informatique
: 15/20
Physique:
16/20



Enseignant

Nom: Jean
Duroux
Année: L2
Chimie
Modules
choisis:
informatique,...



Secrétariat

Problème :

Plusieurs formats de fichiers



difficulté de gestion

BUT 2022/2023

C4




R1.05

Modélisation de bases de données

8/36

I. 1) Pourquoi les BD ?

Exemple (suite) : Approche par fichiers (2)

<p>Nom: Jean Duroux Année: L2 Chimie Frais: 500 euros</p> 	<p>Nom: Jean Philippe Duroux Année: L2 Chimie Informatique: 15/20 Physique: 16/20</p> 	<p>Nom: Jean Duroux Année: L2 Chimie Modules choisis: informatique,....</p> 
Service comptabilité	Enseignant	Secrétariat

Problème :

Redondance des données → Incohérences possibles

BUT 2022/2023

C4




R1.05

Modélisation de bases de données

9/36

I. 1) Pourquoi les BD ?

Exemple (suite) : Approche par fichiers (3)

<p>Nom: Jean Duroux Année: L2 Chimie Frais: 500 euros</p> 	<p>Nom: Jean Philippe Duroux Année: L2 Chimie Informatique: 15/20 Physique: 16/20</p> 	<p>Nom: Jean Duroux Année: L2 Chimie Modules choisis: informatique,....</p> 
Service comptabilité	Enseignant	Secrétariat

Problème :

Interrogation des données → Redondance de code
Coût élevés
Maintenance difficile

BUT 2022/2023

C4

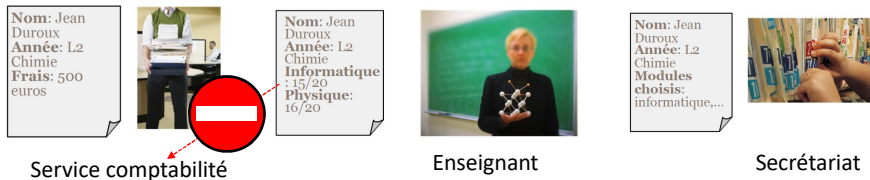
R1.05

Modélisation de bases de données

10/36

I. 1) Pourquoi les BD ?

Exemple (suite) : Approche par fichiers (4)



Problème :

Partage des données → Confidentialité ?

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

11/36

I. 2) BD et SGBD

L'approche Base de Données :

Toute organisation repose sur un SI dont une des fonctionnalités majeures concerne la gestion des données. Il lui faut pouvoir facilement *archiver*, *retrouver*, *traiter* et *mettre à jour* les données, qui sont mémorisées au sein d'une Base de Données.

Définition :

Une *Base de Données (BD)* est un ensemble structuré d'informations agrégées ou élémentaires accessible par une communauté d'utilisateurs.

Définition :

Un *Système de Gestion de Base de Données (SGBD)* est un logiciel qui permet d'interagir avec une base de données et implante les fonctions citées précédemment.

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

12/36

I.2) BD et SGBD

Objectifs fondamentaux d'un SGBD :

- Centraliser l'information : supprimer la redondance, centraliser les contrôles
- Assurer l'indépendance données-traitement
- Permettre des liaisons entre ensembles de données
- Intégrité : fiabilité et cohérence (contraintes, format, etc)
- Sécurité : reprise après panne
- Confidentialité : identification, authentification et autorisations (utilisateurs)
- Partage de données : mise à jour sous forme de *transactions*

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

13/36

I.2) BD et SGBD

Critères de classification des SGBD :

- Le modèle logique des données : Relationnel, objet-relationnel, graphe,...
- Le nombre d'utilisateurs : Un seul (BD de bureau) / des centaines (serveur de BD d'une entreprise)
- L'usage qui en est fait :
 - **OLTP (On Line Transaction Processing)**
 - Cible des SGBD depuis leur existence
 - Banque, réservation en ligne
 - Très grand nombre de transactions en parallèle
 - Transactions simples
 - **OLAP (On Line Analytical Processing)**
 - Entrepôts de données, Data cube, Data mining
 - Faible nombre de transactions
 - Transactions très complexes

BUT 2022/2023

C4

R1.05

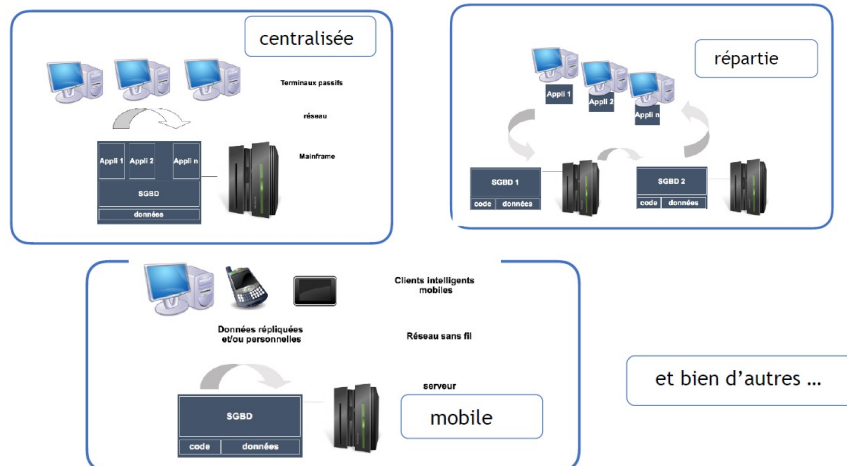
Modélisation de bases de données

14/36

I.2) BD et SGBD

Critères de classification des SGBD (suite):

- L'architecture retenue



BUT 2022/2023

C4

R1.05

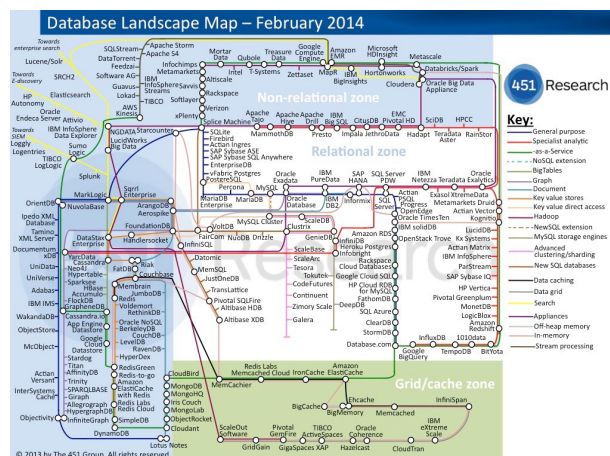
Modélisation de bases de données

15/36

I.2) BD et SGBD

Critères de classification des SGBD (suite):

- Une caractéristique technique précise :
 - In-memory database
 - Cloud database
 - Embedded database
 - ...
- Ou encore... →



BUT 2022/2023

C4

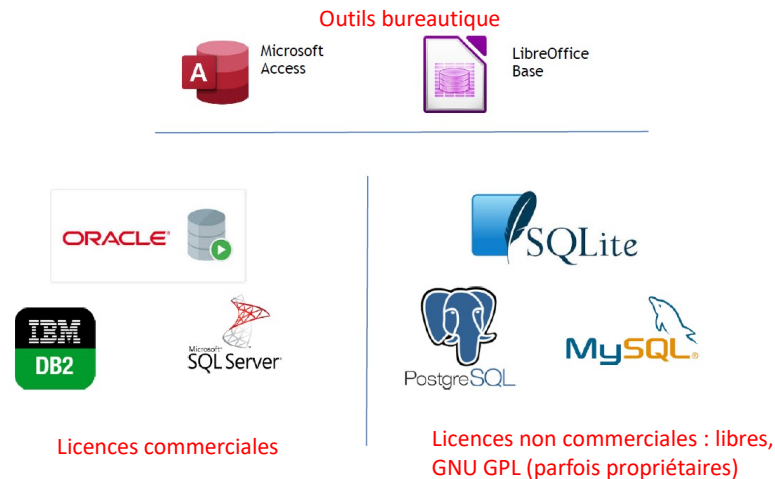
R1.05

Modélisation de bases de données

16/36

I.2) BD et SGBD

SGBD relationnels sur le marché :



BUT 2022/2023

C4

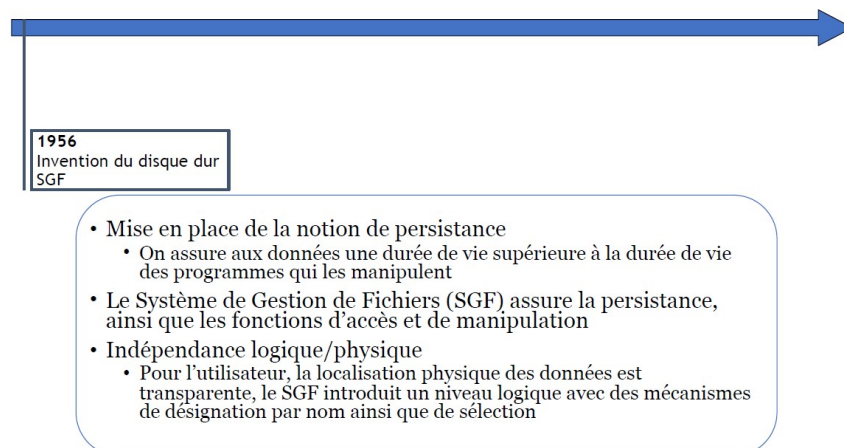
R1.05

Modélisation de bases de données

17/36

I.3) Repères historiques

Systèmes de Gestion de Fichiers



BUT 2022/2023

C4

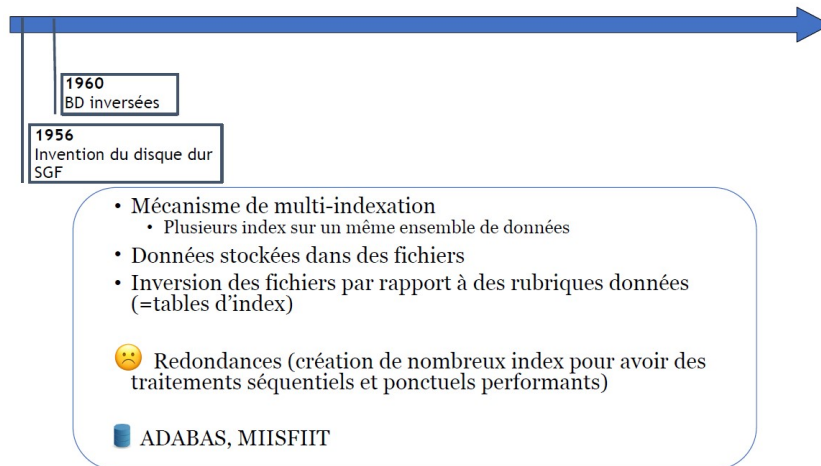
R1.05

Modélisation de bases de données

18/36

I.3) Repères historiques

Bases de données inversées



BUT 2022/2023

C4

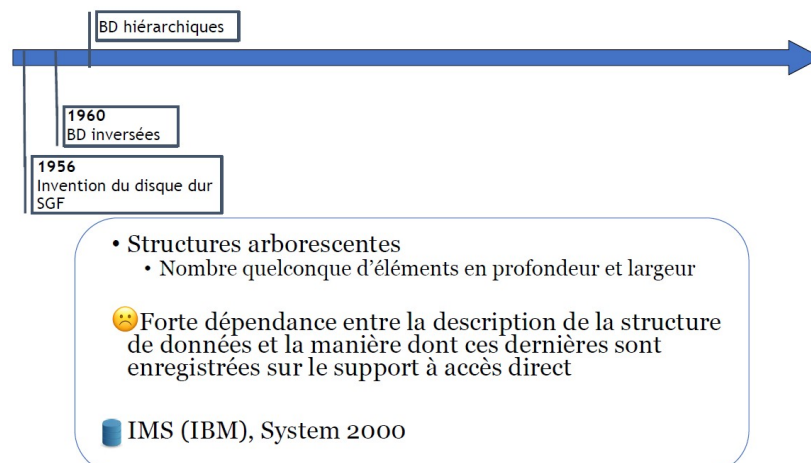
R1.05

Modélisation de bases de données

19/36

I.3) Repères historiques

Bases de données hiérarchiques



BUT 2022/2023

C4

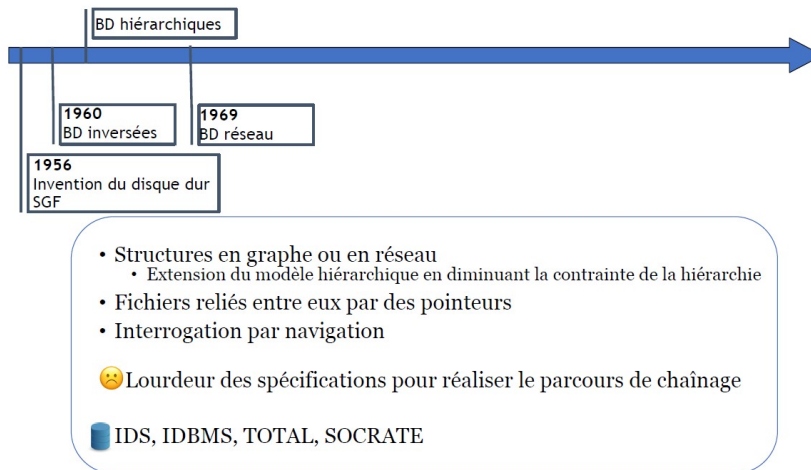
R1.05

Modélisation de bases de données

20/36

I.3) Repères historiques

Bases de données réseau



BUT 2022/2023

C4

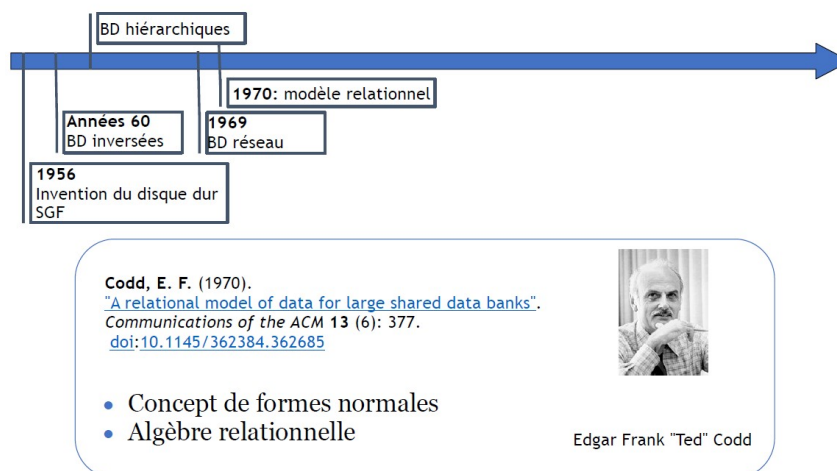
R1.05

Modélisation de bases de données

21/36

I.3) Repères historiques

Bases de données relationnelles (1)



BUT 2022/2023

C4

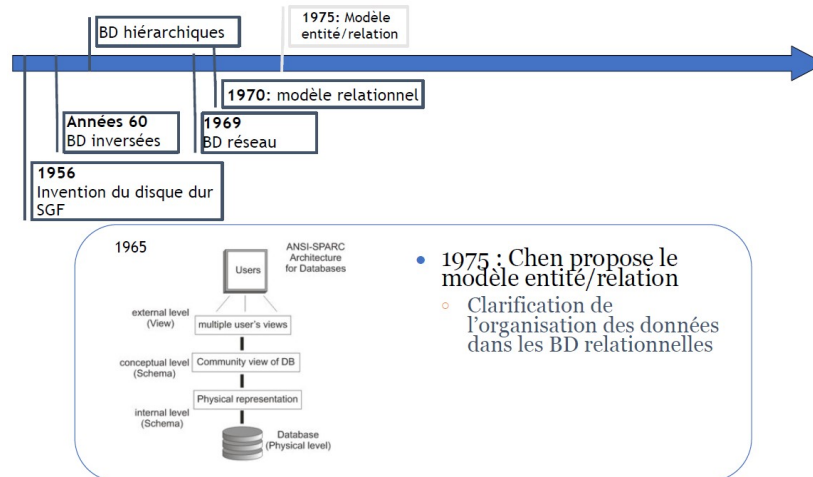
R1.05

Modélisation de bases de données

22/36

I.3) Repères historiques

Bases de données relationnelles (2)



BUT 2022/2023

C4

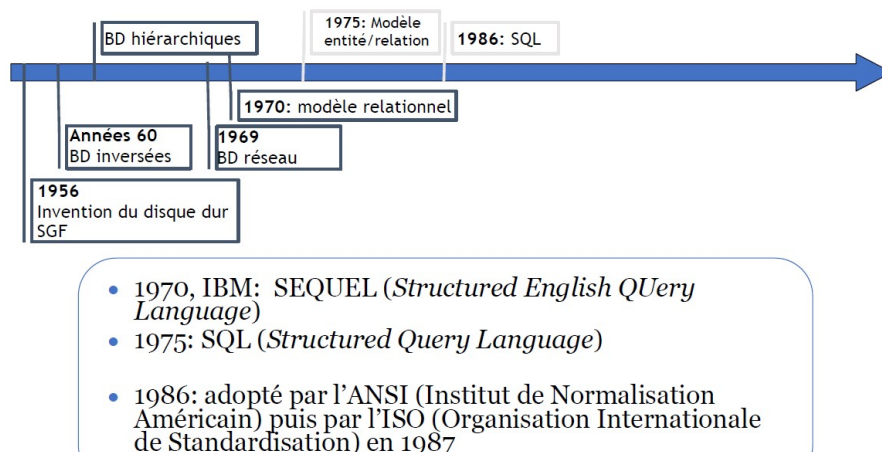
R1.05

Modélisation de bases de données

23/36

I.3) Repères historiques

Bases de données relationnelles (3)



BUT 2022/2023

C4

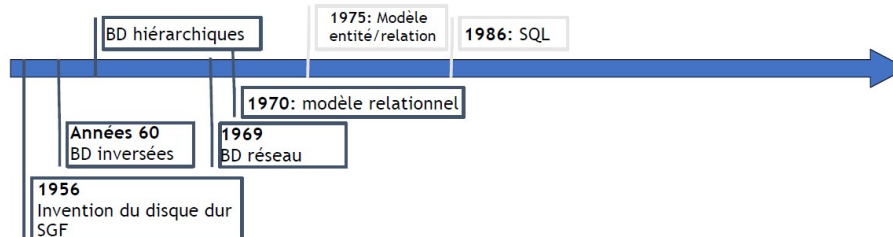
R1.05

Modélisation de bases de données

24/36

I.3) Repères historiques

Bases de données relationnelles (4)



- Fin des années 1980: démocratisation des SGBD relationnels



Oracle, Ingres, Sybase, INFORMIX, MS SQL Server, ...

BUT 2022/2023

C4

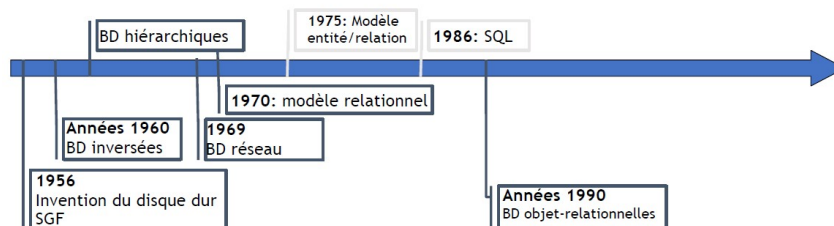
R1.05

Modélisation de bases de données

25/36

I.3) Repères historiques

Bases de données objet-relationnelles



- Approche NF2 (Non First Normal Form)
- Fonctionnalités orientées-objet: encapsulation, classes, héritage, surcharge



O2, Gesmstone, Ontos, Object Store

BUT 2022/2023

C4

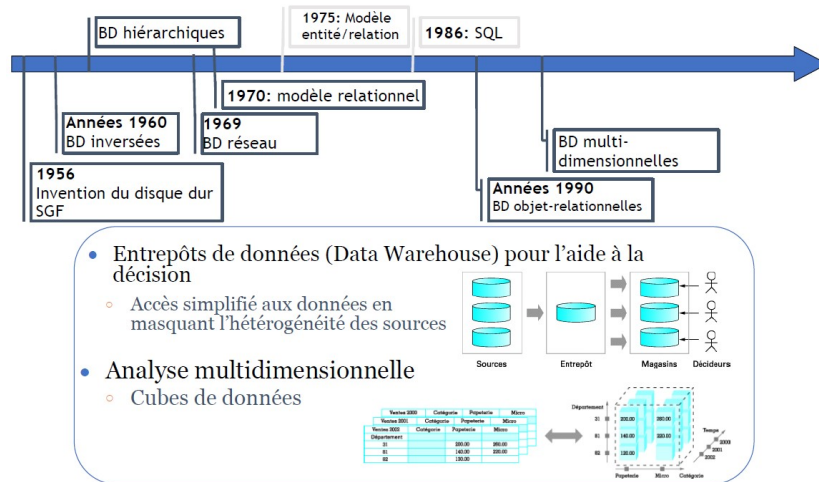
R1.05

Modélisation de bases de données

26/36

I.3) Repères historiques

Bases de données multidimensionnelles



BUT 2022/2023

C4

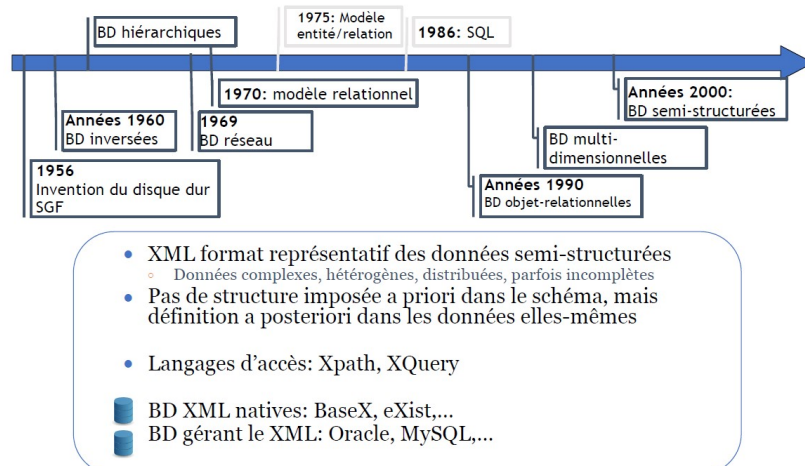
R1.05

Modélisation de bases de données

27/36

I.3) Repères historiques

Bases de données semi-structurées



BUT 2022/2023

C4

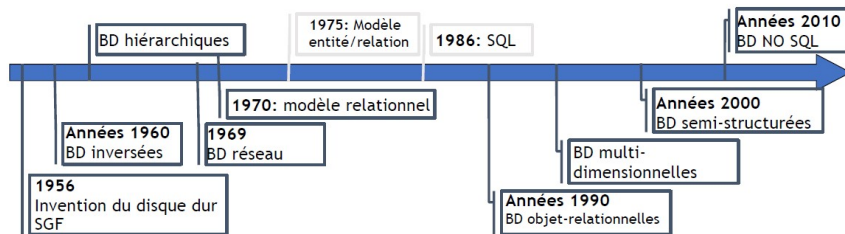
R1.05

Modélisation de bases de données

28/36

I.3) Repères historiques

Bases de données No SQL



- No SQL = Not Only SQL
- Moins de formalisme, plus de pragmatisme
 - Grands volumes de données (Big Data)
- Bases NO-SQL:
 - Orientées clé/valeur
 - Bases de document
 - BD orientées colonne
 - BD graphes

Amazon Dynamo
 CouchDB, MongoDB
 BigTable (Google)
 Neo4J

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

29/36

I.4) Conception d'une BD

Modélisation du réel

Réel					
Modèle conceptuel	<ul style="list-style-type: none"> • Indépendant du SGBD • Diagramme de classes • Schéma E/A 				
Modèle logique	<ul style="list-style-type: none"> • Dépendant du modèle de données • Indépendant du SGBD 	Réseau	Relationnel	Objet	XML
Modèle physique	<ul style="list-style-type: none"> • Dépendant du modèle de données • Dépendant du SGBD 	<ul style="list-style-type: none"> • Organisation physique des données • Structures de stockage des données • Structures accélératrices (index) 			

BUT 2022/2023

C4

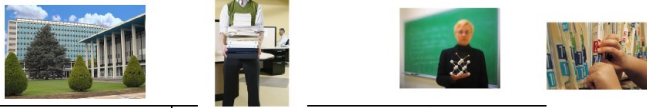

R1.05

Modélisation de bases de données

30/36

I.4) Conception d'une BD

Modélisation du réel

Réel					
Modèle conceptuel	<ul style="list-style-type: none"> •Indépendant du SGBD •Diagramme de classes •Schéma E/A 				
Modèle logique	<ul style="list-style-type: none"> •Dépendant du modèle de données •Indépendant du SGBD 	Réseau	Relationnel	Objet	XML
Modèle physique	<ul style="list-style-type: none"> •Dépendant du modèle de données •Dépendant du SGBD 	<ul style="list-style-type: none"> •Organisation physique des données •Structures de stockage des données •Structures accélératrices (index) 			

BUT 2022/2023

C4

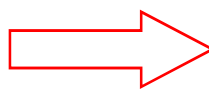
R1.05

Modélisation de bases de données

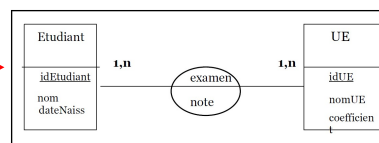
31/36

I.4) Conception d'une BD

Etapes de conception d'une BD relationnelle :

**Réel** (monde extérieur)

Modélisation


Modèle Conceptuel :
Schéma Entité/Association
Modification du
modèle conceptuel

Normalisation ?

Règles de
passageMise en œuvre
de la BD**Modèle physique**

Langage SQL

 Etudiant (idEtudiant, nom, dateNaiss)
 Examen (#idEtudiant, #idUE, note)
 UE (idUE, nomUE, coefficient)
Modèle Logique : schéma relationnel

BUT 2022/2023

C4

R1.05

Modélisation de bases de données

36/36