

# Bases de données

## Cours 1 : Généralités sur les bases de données

**Odile PAPINI**

POLYTECH

Université d'Aix-Marseille

odile.papini@univ-amu.fr

<http://odile.papini.perso.luminy.univ-amu.fr/sources/BD.html>

# Plan du cours 1

- 1 Qu'est ce qu'une base de données
  - Exemples de bases de données
  - Différents types de bases de données
- 2 Qu'est ce qu'un système de base de données
  - Données
  - Matériel
  - Logiciel
  - Avantages de l'approche Base de Données
  - Exemples de SGBD
- 3 Quelques repères historiques
- 4 Place des BD dans l'informatique
- 5 BD et emploi

# Fonction d'un système de bases de données

## Assurer la conservation d'enregistrements informatiques

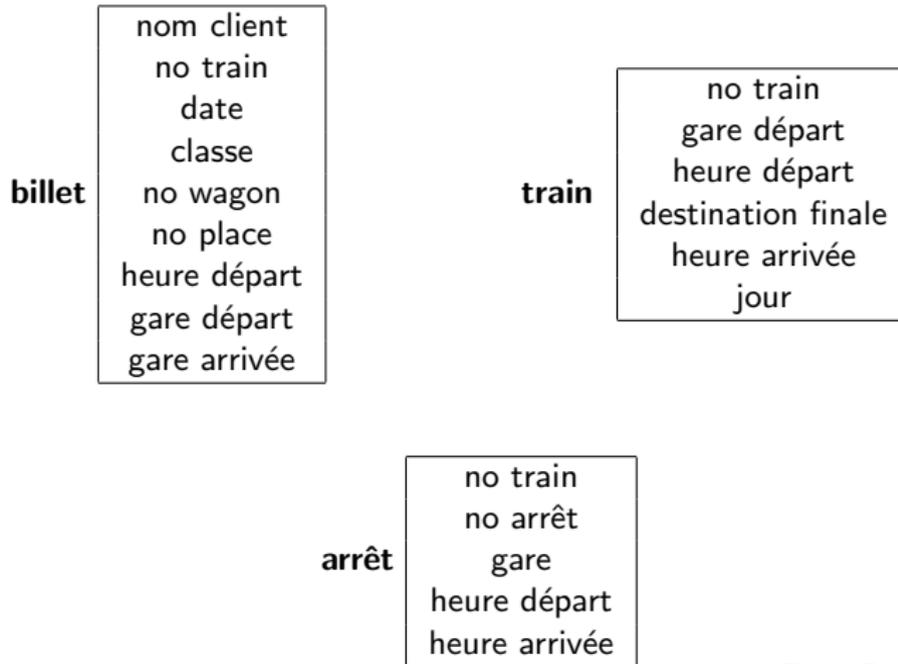
### Base de données (BD) :

- réservoir de fichiers de données informatiques
- opérations sur les fichiers

## Exemples de BD

- Gestion des personnels, étudiants, cours, inscriptions, ... de l'université
- Système de réservation de places d'avion chez Air France, de places de train à la SNCF
- Gestion des comptes clients de La Poste
- Gestion des commandes chez Amazon.com
- Gestion d'une bibliothèque
- Gestion des pages Web chez google.com
- ...

## Exemple de BD : réservation d'un billet de train



## Exemple de BD : cave à vin

### Base de données d'une cave

<i>NUM</i>	<i>VIN</i>	<i>PRODUCTEUR</i>	<i>ANNÉE</i>	<i>NB</i>
2	<i>CabernetSauvignon</i>	<i>ChateauduMayne</i>	1983	5
3	<i>Chardonnay</i>	<i>Jaboulet – Vercherre</i>	1984	7
6	<i>PinotNoir</i>	<i>Jaboulet – Vercherre</i>	1984	10
12	<i>Brunello</i>	<i>EmilioCostanti</i>	1977	6
15	<i>Silvaner</i>	<i>Guntrum</i>	1985	8
23	<i>Riesling</i>	<i>BalthasarRess</i>	1976	2

# Besoins pour la gestion d'une BD

- description
- manipulation
- interrogation
- cohérence
- garanties
- confidentialité
- efficacité

# Besoins pour la gestion d'une BD : description

- Description des données de l'application (trains, trajets, réservations) sans faire référence à une solution informatique particulière

## **modélisation conceptuelle**

- Elaboration d'une description équivalente pour le stockage des données dans le Système de Gestion de Bases de Données choisi

## **modélisation logique**

## **Langage de Description de Données (LDD)**

# Besoins pour la gestion d'une BD : manipulation

- Créer la base de données initiale avec les données représentant le réseau SNCF

## langage permettant l'insertion de données

- Créer au fur et à mesure les données sur les réservations.
- Modifier si besoin et éventuellement supprimer toute donnée déjà rentrée

## Langage de Manipulation de Données (LMD) (insertion, modification, suppression)

# Besoins pour la gestion d'une BD : interrogation

Répondre à toute demande d'information portant sur les données contenues dans la base.

- a) Le professeur Tournesol a-t-il une réservation pour aujourd'hui ? Si oui, donner les informations connues sur cette réservation.
- b) Quels sont les horaires des trains de Marseille à Cassis entre 9h et 10h le dimanche ?
- c) Donner les destinations au départ de Marseille sans arrêts intermédiaires.

**langage de requête (langage d'interrogation)**

## Besoins pour la gestion d'une BD : cohérence

Il faut pouvoir exprimer toutes les règles qui contraignent les valeurs pouvant être enregistrées de façon à éviter toute erreur qui peut être détectée.

- Il ne faut jamais donner la même place dans le même train à 2 clients
- Les arrêts d'un train sont numérotés de façon continue
- La date de réservation pour un train doit correspondre à un jour de circulation de ce train
- L'heure de départ d'une gare doit être antérieure à l'heure d'arrivée dans cette gare
- L'heure d'arrivée à un arrêt doit être postérieure à l'heure de départ de l'arrêt précédent

● ...

# Besoins pour la gestion d'une BD : garanties

- Il ne faut pas que les informations (par exemple, les réservations) soient perdues à cause d'un dysfonctionnement quelconque : erreur de programmation, panne système, panne de l'ordinateur, coupure de courant, etc

## garantie de fiabilité

- Il ne faut pas qu'une action faite pour un utilisateur (par exemple, l'enregistrement d'une réservation) soit perdue du fait d'une autre action faite simultanément pour un autre utilisateur (réservation de la même place).

## garantie de contrôle de concurrence

# Besoins pour la gestion d'une BD : confidentialité

- Toute information doit pouvoir être protégée contre l'accès par des utilisateurs non autorisés :

**en lecture**  
**en écriture**

- Interdire par exemple aux clients de modifier les numéros des trains ou les horaires ou leur réservation.

**garantie de confidentialité (privacy)**

## Besoins pour la gestion d'une BD : efficacité

Le temps de réponse du système doit être conforme aux besoins :

- en interactif : pas plus de 3 secondes
- en programmation : assez rapide pour assumer la charge de travail attendue (nombre de transactions par jour)

**mécanismes d'optimisation**

**éventuellement, répartition / duplication des données  
sur plusieurs sites**

## Différents types de BD

Il y en a pour tous les usages :

- Bases de données personnelles : MsAccess, ... : 10 Ko 100 Ko
- Bases de données professionnelles typiques : 100 Ko 100 Go
- Bases de données professionnelles très grandes : Very Large Databases (VLDB) : > 40 To

## Exemples de VLDB (taille) :

### Volume de quelques grandes bases de données

- ◆ Amazon (> 42 téraoctets)
- ◆ YouTube (> 45 téraoctets)
- ◆ AT&T (323 téraoctets)
- ◆ World Data Centre for Climate (220 téraoctets + 6 pétaoctets de données additionnelles)
- ◆ Google (... Po ? Eo ?)

1 gigaoctets (Go) = 1024 mégaoctets (Mo)

1 téraoctet (To) = 1024 gigaoctets (Go)

1 pétaoctet (Po) = 1024 téraoctets (To)

FIGURE: source : N. Durand

# Qu'est-ce qu'un système de base de données ?

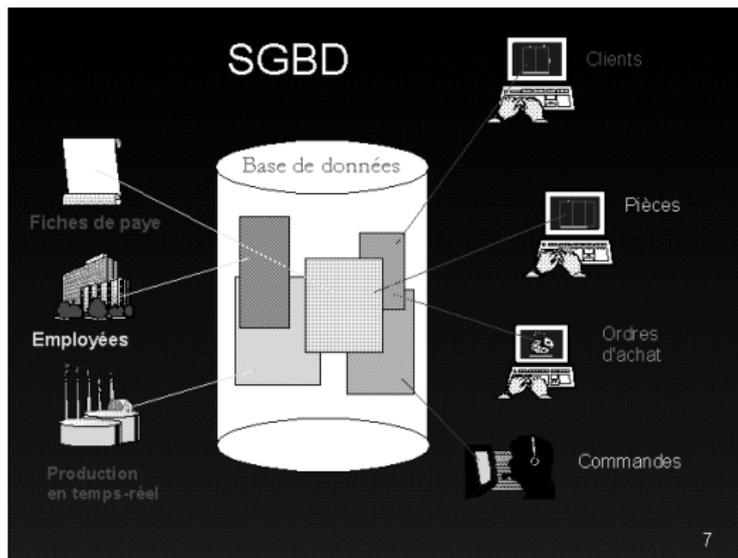
SGBD : système informatique dont le but est de maintenir des informations et les rendre disponibles à la demande

il est composé de :

- données
- matériel
- logiciel
- utilisateurs

# SGBD

## SGBD



# Les données

## Données intégrées

base de données vue comme l'unification de plusieurs fichiers

## Données partagées

base de données partagées entre plusieurs utilisateurs

## Le matériel

### Volumes de mémorisation secondaires

disques à tête amovible, supports d'enregistrements . . .

**Processeurs et mémoire principale** qui lui sont associés pour l'exécution du système de base de données

## Le logiciel

### Le système de bases de données (SGBD)

**Fonction** : protéger les utilisateurs de la BD des détails du niveau matériel

**Le SGBD** : le composant logiciel le plus important su système

# Les utilisateurs :

## 3 classes d'utilisateurs :

- les programmeurs d'applications
- les utilisateurs finaux
- l'administrateur de la base de données

# Qu'est-ce qu'une base de données ?

**une collection de données persistantes utilisées par les systèmes d'application d'une entreprise**

**données persistantes :**

données acceptées par le SGBD qui ne peuvent être enlevées que par une requête explicite envoyée au SGBD

**entité** : tout objet perceptible représenté dans la BD

**association** : relations entre les entités

les entités et les associations possèdent des **propriétés** simples ou complexes

# Qu'est-ce qu'une base de données ?

## Modèle de données :

définition logique indépendante et abstraite des **objets** et des **opérations** qui constitue une machine abstraite

**objets** : modélisent la structure de données

**opérations** : modélisent le comportement

## implémentation d'un modèle de données

réalisation physique sur une machine réelle des différents composants d'une machine abstraite

# Pourquoi une base de données ?

## Avantages :

- **compacité**
- **rapidité**
- **efficacité**
- **exactitude**
- **protection**
- **contrôle centralisé**

## Avantages de l'approche BD

- **données partagées**
- **redondance réduite**
- **incohérence évitée**
- **transactions gérées**
- **intégrité assurée**
- **sécurité appliquée**
- **conflits de besoins arbitrés**
- **normes appliquées**
- **indépendances des données obtenue**

## Exemple de Base de données

### Exemple : gestion des hôpitaux d'un département

"La direction départementale des Bouches du Rhône souhaite informatiser la gestion des hôpitaux du département. Chaque hôpital a des activités de soin dans les services médicaux et des activités de recherche dans les laboratoires. Les médecins employés sont obligatoirement rattachés à un seul hôpital. Ils ont le choix entre trois types de fonctions : consultant indépendant, médecin praticien, médecin chercheur. Les consultants indépendants ne sont rattachés à aucune structure interne de laboratoire ou de service. Les praticiens sont rattachés à un seul service. Les chercheurs sont rattachés à un laboratoire unique. Les fonctions de praticien et de chercheur peuvent être cumulées. Lorsqu'un médecin est consultant indépendant ou médecin praticien, il se voit confier la responsabilité d'un ou plusieurs patients. Dans certains cas nécessitant des traitements complexes, un même patient peut être suivi par plusieurs médecins.

Les hôpitaux, en plus de leurs nom et adresse, sont identifiés au moyen d'un code. De même, chaque service hospitalier et chaque laboratoire sont décrits par un code d'identification et un nom. On connaît, pour chaque médecin, son numéro matricule, ses coordonnées (nom, prénom, sexe, adresse, date de naissance) et sa spécialité (urologie, gynécologie, gériatrie, pédiatrie, etc.). Enfin, chaque patient possède un numéro de sécurité sociale, un nom, un prénom, une adresse, une date de naissance.

Les fonctions que le système devra effectuer sont : la gestion d'activité des médecins, la gestion du suivi des patients, les statistiques départementales (nombre de laboratoires par domaine, nombre de médecins, ...)."

24

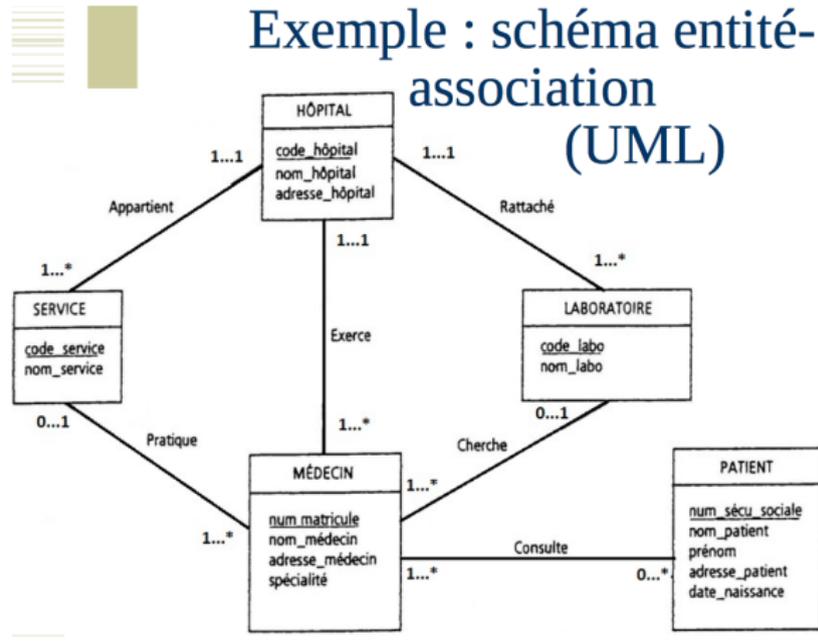
# Exemple de Base de données

## Schéma conceptuel

- ◆ Représentation du monde réel auquel se rapporte la BD
- ◆ Utilisation de *schémas aisément compréhensibles par tous*
- ◆ Modèle conceptuel : **entité-association**
- ◆ Défaut principal : devient rapidement illisible dans le cas de très importantes bases de données

FIGURE: source : N. Durand

## Exemple de Base de données



# Exemple de Base de données

## Schéma conceptuel

- ◆ Ensuite, modèle de représentation des données que le SGBD va comprendre et manipuler (modèle **logique**) :
  - **Modèle relationnel**
  - Modèle hiérarchique
  - Modèle réseau
  - Modèle objet
  - ...

FIGURE: source : N. Durand

# Exemple de Base de données

## Exemple Schéma relationnel

HOPITAL(code\_hopital, nom\_hopital, adresse\_hopital)  
LABORATOIRE(code\_labo, nom\_lab, #code\_hopital)  
MEDECIN(num\_matricule, nom\_medecin, adresse\_medecin, specialite)  
PATIENT(num\_secu\_sociale, nom\_patient, prenom, adresse\_patient,  
date\_naissance)  
CONSULTE(#num\_matricule, #num\_secu)  
...

FIGURE: source : N. Durand

# Exemple de Base de données

## Exemple Schéma relationnel

MEDECIN			
num_matricule	nom_medecin	adresse_medecin	specialite
12354	Dupont	Marseille	dermatologie
15789	Martin	Marseille	ophthalmologie
23574	Rousseau	Marseille	cardiologie
19355	Martin	Aix-en-Provence	dermatologie

PATIENT			
num_secu_social	nom	adresse	date_naissance
1860113156148	Bidule	Marseille	14/01/1986
1561183795267	Machin	Aubagne	26/11/1956

FIGURE: source : N. Durand

## Exemples de SGBD :

Nombreux SGBD sur le marché :

- ACCESS
- MySQL
- PostgreSQL
- ORACLE
- DB2

## SGBD capacité

Mainframe ou cluster de mainframes DB2, Oracle, Adabas, ...

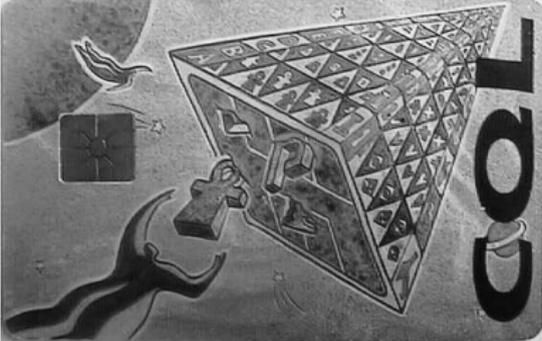
Machine base de données Teradata

Serveurs Unix & Windows 2000 DB2, Informix, Oracle, Sybase,  
Unify, SQL Server, Ingres,

Personnels MsAccess, Paradox

Carte de Crédit CQL

## BD et carte à puce



**LA CARTE A PUCE RELATIONNELLE**

*Avec CQL, GEMPLUS présente la première carte qui gère les données au format relationnel. Utilisant le langage CQL, elle communique avec les serveurs ORACLE de votre système d'information.*

ORACLE®

GEMPLUS

# Historique

## 1ère génération 1950 - 1965

SGF(Systèmes de Gestion de Fichiers)

Permettent de stocker et d'organiser les données sur mémoire secondaire

Différentes organisations :

- séquentielles
- séquentielles triées
- indexées
- séquentielles indexées

# Historique

## 2ème génération 1965 - 1970

SGBD navigationnel

Hierarchique (IMS)

Réseaux (Codasyl, IDS2)

Principe : relier les données provenant de différents fichiers

Caractéristiques : liens logiques entre les données

- hiérarchiques (modèle hiérarchique)
- hiérarchiques ++ (modèle réseau)

Liens physiques entre les enregistrements informatiques (pointeurs)

# Historique

## 3ème génération 1969 - ...

SGBD relationnel (DB2, Oracle, Informix, MsAccess)

Principe : langages assertionnels quoi au lieu de comment

SQL

- langage de définition de bases de données
- langage de manipulation de bases de données
- normes SQL1, SQL2, SQL3

# Historique

## **SGBD Orienté Objet (1990 - 1999)**

En pratique : une impasse (O2, Objectstore, Objectivity..)

SGBD relationnel objet (RO) 1993 - ...

évolution probable de tout SGBD relationnel

# Historique

## Base de données déductives 1980 - ...

- approche déclarative
- gestion de requêtes conditionnelles (si ... alors)
- gestion de requêtes récursives

extension de SQL99

DATALOG

LDL++

BD et Programmation logique : DES, DLV, XSB

## BD et internet

gestion de données :

- structurées
- semi-structurées
- hiérarchisées
- réparties

HTML → XML → RDF → OWL

## BD et Big data

BIG data = données massives provenant du Web

- gestion de données : SQL3, NOSQL, SPARQL , ...  
+
- analyse des données

But : Accéder en temps réel à des bases de données massives

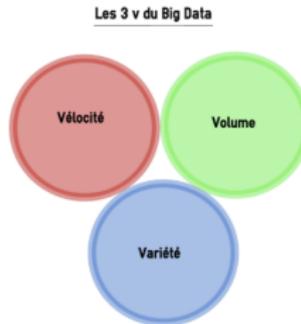


FIGURE: source : <http://www.lebigdata.fr>

# Place des BD dans l'informatique

domaines à enseigner dans un cursus d'informaticien (<http://www.ieee.org>) :

- Structures discrètes
- Fondements de la programmation
- Algorithmique et complexité
- Langages de programmation
- Architecture et organisation
- Systèmes d'exploitation
- Réseaux
- Interfaces Homme/Machine
- Informatique graphique
- **Systèmes intelligents**
- Management de l'information
- Génie Logiciel
- Problèmes socio-professionnels

# Le domaine des BD est vaste

- Modèles et systèmes d'information
- Outils de bases de données
- Modélisation des données
- Bases de données relationnelles
- Bases de données et langages de requêtes
- Conception de bases de données relationnelles
- Traitement transactionnel
- Bases de données distribuées
- Conception physique des bases de données
- Fouille de données
- Stockage et recherche d'information
- Hypertexte et hypermédia
- Information et systèmes multimédia
- Librairies digitales

# BD et emploi

Marché du travail : BD = gros pourcentage des offres d'emploi  
Exemples (extraits du site <https://www.indeed.fr/emplois>)

## **Data Manager**

Groupe La Poste - Paris (75)

Vous avez une appétence en marketing et avez une maîtrise des bases de données. Extraire les données et préparer les [fichiers...](#)

[Postuler directement](#)

[il y a 15 jours](#) - [sauvegarder](#) - [plus...](#)

## **Gestionnaire Base de Données**

[DOMINO](#) - [26 avis](#) - Paris 17e (75)

Les activités du gestionnaire de bases de données : Il aura pour objectif de coordonner entre les différents gestionnaires de bases de données des...

Officiel [interim](#) - [il y a 13 jours](#) - [sauvegarder](#) - [plus...](#)

## **Administrateur Base de Données / LILLE H/F**

[GFI](#) - [31 avis](#) - Lille (59)

Administrer les bases de données : Vous participez à la définition et à la mise en [oeuvre](#) des bases de données...

[il y a 1 jour](#) - [sauvegarder](#) - [plus...](#)

# BD et emploi

Marché du travail : BD = gros pourcentage des offres d'emploi  
Exemples (extraits du site <https://www.indeed.fr/emplois>)

Site <https://www.indeed.fr/emplois>

## **Gestionnaire base de données (H/F)**

[NutoXo](#) - 5 avis - Ivry-sur-Seine (94)

Gestionnaire Base de données (H/F). Poste basé à Ivry/Seine (94). Vous êtes garant de la qualité de la donnée de base, afin d'assurer le bon fonctionnement de...  
Grands Moulins de Paris - il y a 30+ jours - sauvegarder - plus...

## **Data Analyst H/F**

MANUTAN COLLECTIVITES - Niort (79)

Maîtrise de l'exploitation de bases de données de gros volumes. Rattaché(e) hiérarchiquement à la direction Marketing client, vous participez à la structuration...  
Apec.fr - il y a 2 jours - sauvegarder - plus...

## **DATA MANAGER H/F**

 ALTIZEM CRO - Lyon (69)

Base de données, scanning, archives), en particulier, \\*\*Gel de base de données (Database Lock). Coordination du développement et de l'implémentation de base...  
Postuler directement  
il y a 13 jours - sauvegarder - plus...

# BD et emploi

Tout informaticien peut être amené à

- Utiliser de grosses bases de données
- Concevoir de grosses bases de données
- Administrer de grosses bases de données