

## Bases de données

(généralités)

### 1. Introduction

Une base de données permet d'organiser, de stocker, de mettre à jour et d'interroger des données structurées volumineuses utilisées simultanément par différents programmes ou différents utilisateurs. Cela est impossible avec les représentations tabulaires étudiées en classe de première.



Exemples et tailles de bases de données (données 2025, fiabilité ?)

Google : 100Po

Amazon : Plus de 175 Zo de données (produits et clients) ;

600 Po pour le CERN uniquement pour le LHC (accélérateur de particules) et 10Po chaque heure pour certains télescopes

1 kilo octet (Ko) = 1024 octets  $\rightarrow 2^{10}$  bits

1 tera octet (To) = 1099511627776 octets  $\rightarrow 2^{40}$  bits

1 petaoctet (Po) = 1.1258999068426e15 octets  $\rightarrow 2^{50}$  bits

1 zetaoctet (Zo) = 1e21 octets

### 2. Système de gestion de base de données

#### A. Définitions

Un Système de Gestion de Base de Données (SGBD) est un logiciel (ou un ensemble de logiciels) permettant de manipuler les données d'une base de données. Manipuler, c'est-à-dire *sélectionner* des informations tirées de cette base, *modifier* des données, en *ajouter* ou en *supprimer* (ce groupe de quatre opérations étant souvent appelé "CRUD", pour *Create, Read, Update, Delete*).

Le SGBD veille à :

- La gestion des accès concurrents : que faire si deux personnes accèdent et modifient en même temps la base de données?
- L'efficacité du traitement des requêtes.
- La sécurisation des accès : qui a le droit de lire, d'écrire dans la base de données.

Il existe plusieurs types de modèles de base de données (relationnel, entité-association, objet...).

Un Système de Gestion de Base de Données Relationnel (SGBDR) est un SGBD qui implémente la théorie *relationnelle*.



mongoDB

ORACLE



PostgreSQL



SQL Server



cassandra

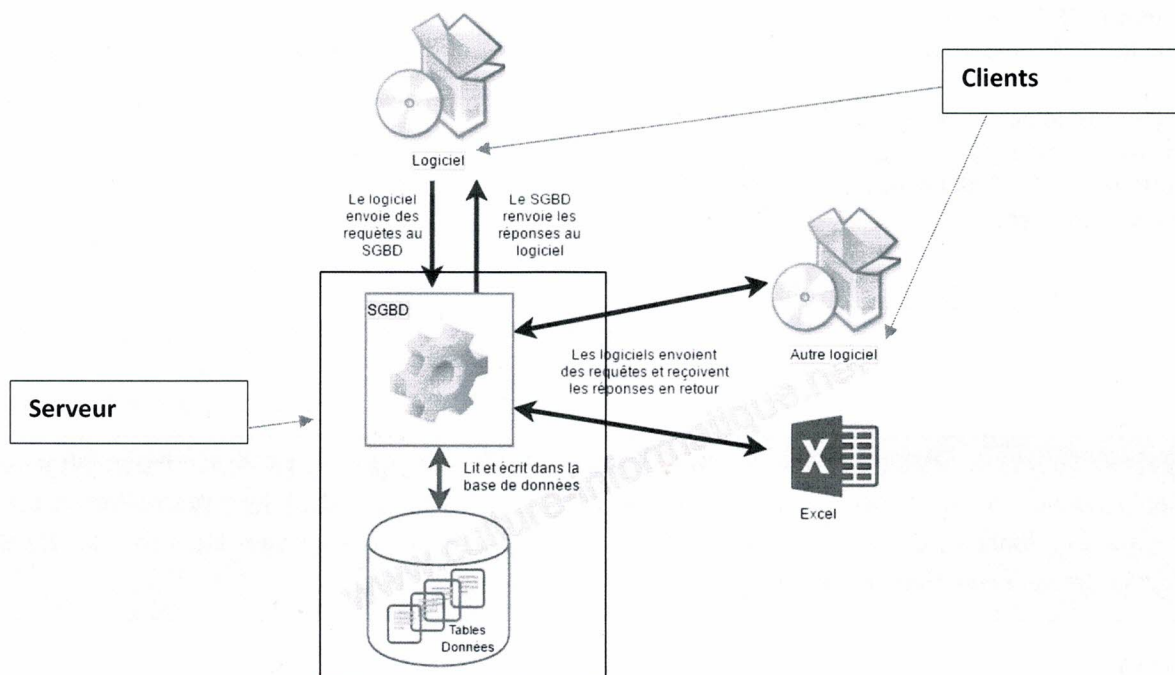


GraphQL

CouchDB  
relax

Un paradigme est synonyme d'une représentation, d'une vision, d'un modèle.

La plupart des SGBD sont basés sur un modèle **Client - Serveur**. C'est-à-dire que la base de données se trouve sur un **serveur** qui ne sert qu'à ça, et pour interagir avec cette base de données, il faut utiliser un logiciel "**client**" qui va interroger le **serveur** et transmettre la réponse que le **serveur** lui aura donnée. Le **serveur** peut être installé sur une machine différente du **client** ; c'est souvent le cas lorsque les bases de données sont importantes. Ce n'est cependant pas obligatoire.



#### Exemple de requêtes :

- Demande d'ajout d'un élément dans une base de données
- Demande de suppression d'un élément
- Recherche de données suivant plusieurs critères
- Ajouter un utilisateur avec mot de passe et privilèges associées

Pour envoyer ces **requêtes** et obtenir des réponses, on a besoin d'un langage pour discuter entre le **client** et le **serveur**. Le langage utilisé est le **SQL** (Structured Query Language), soit le langage de **requêtes** structurées.

Sur ce schéma, aucun **client** n'accède directement à la base de données. Tous les accès se font via le SGBD, en utilisant les requêtes **SQL** (tous les langages de programmation (PHP, python, etc.) sont à même d'envoyer des requêtes **SQL**).

### 3. Base de données relationnelle

#### A Introduction

La base de données est composée d'un ensemble de tables (les relations) dans lesquelles sont placées des données. Ces modèles sont simples à mettre en œuvre et fondés sur les mathématiques (la théorie des ensembles), ils sont très populaires et fortement normalisés.

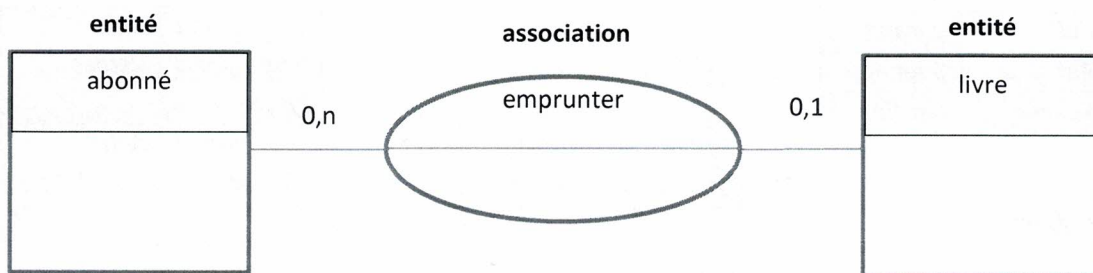
#### B Modèle entité association

Le modèle **entité-association** est souvent utilisé en premier pour concevoir la base de données. Ensuite, on l'utilise pour en faire un **modèle relationnel**.

Une **entité** est un objet, un sujet, une notion en rapport avec le domaine d'activité pour lequel la base de données est utilisée, et concernant celui pour lequel des données sont enregistrées (exemple: des personnes, des produits, des commandes, des réservations, ...).

Une **association** est un lien entre deux ou plusieurs **entités**.

Exemple : base de données d'une librairie, un abonné (**entité**) emprunte (**association**) un livre (**entité**).



Des couples de chiffres apparaissent alors entre chaque **entité** et l'**association**. Ces couples sont des **cardinalités**.

Il faut lire la cardinalité comme suit :

Un abonné peut emprunter de 0 à n (plusieurs livre) → }  
Un livre peut être emprunté par 0 ou 1 abonné ← } rem de lecture

Remarque : ce modèle n'est exploitable que si l'on autorise qu'un seul emprunt d'un livre possible par abonné (Pourquoi ? aurez-vous la réponse à la fin de ce cours ?)

Une **entité** et une **association** sont composées de deux parties, l'**en-tête** et le **corps**.

**a) L'en-tête** : Il est composé de plusieurs **attributs**. Les **attributs** sont également appelés des champs ou des colonnes.

Un **attribut** est une caractéristique d'une **entité** ou d'une **association** susceptible d'être enregistrée dans la base de données. Une **relation** est l'ensemble des **attributs** nécessaires pour décrire une **entité** ou une **association**.

Exemple : le nom et le prénom d'un abonné sont des **attributs**. Les dates d'emprunt et de retour sont des **attributs** d'emprunter. Le titre du livre, le format (livre de poche, livre grand format, édition numérotée), le prix du livre sont des **attributs** du livre.

| abonné               |
|----------------------|
| nom<br>prénom<br>... |

relation

Un **attribut** a un **type**.

Un **attribut** peut avoir un **domaine de valeurs**. Le **domaine de valeur** est l'ensemble des valeurs que peut prendre un **attribut**.

Exemples : Le nom d'un abonné est une chaîne de caractères et il n'a pas de **domaine de valeurs**. Le format d'un livre a pour **domaine de valeurs** « livre de poche, livre grand format, édition numérotée ».

**b) Le corps** : Il contient les **données** (occurrences, enregistrements) proprement dites.

| abonné   |        |
|----------|--------|
| prénom   | nom    |
| Paul     | Durand |
| Béatrice | Dupont |
| Michelle | Martin |

| emprunter      |                |
|----------------|----------------|
| Date d'emprunt | Date de retour |
| 1/2/2019       | 1/3/2019       |
| 25/3/2020      | 26/3/2020      |
| 9/9/2020       | 30/10/2020     |
| 10/10/2020     | 12/10/2020     |

| livre                         |              |
|-------------------------------|--------------|
| titre                         | auteur       |
| 20 milles lieux sous les mers | Jules Vernes |
| le journal d'Anne Frank       | Anne Franck  |

**c) L'identifiant**

Chaque occurrence (une donnée, un enregistrement d'un tableau) doit être identifiée d'une manière unique.

L'identifiant d'une entité ou d'une association est constitué par un ou plusieurs de ses attributs.

Pour une entité, quand cela n'est pas possible, on attribue un numéro unique à chaque occurrence (tuple, donnée, enregistrement).

→ cle primaire (PK)

L'identifiant d'une association peut être obtenu en juxtaposant les identifiants des entités qui participent à la liaison.

| abonné    |          |        |
|-----------|----------|--------|
| id_abonné | prénom   | nom    |
| 1         | Paul     | Durand |
| 2         | Béatrice | Dupont |
| 3         | Michelle | Martin |

| livre     |                               |                 |
|-----------|-------------------------------|-----------------|
| id_auteur | titre                         | auteur          |
| 1         | 20 milles lieux sous les mers | Jules Vernes    |
| 2         | le journal d'Anne Frank       | Anne Franck     |
| 3         | L'École de la Dualité         | Julia Narantian |

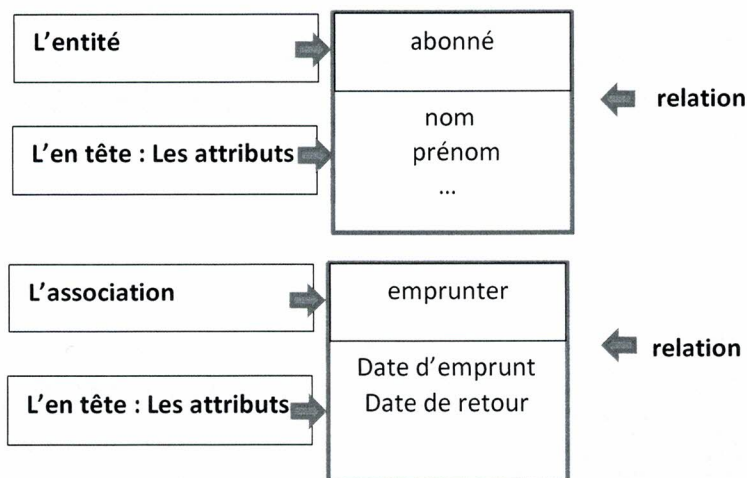
identifiant de l'association

| emprunter |           |                |                |
|-----------|-----------|----------------|----------------|
| id_abonné | id_auteur | Date d'emprunt | Date de retour |
| 1         | 1         | 1/2/2019       | 1/3/2019       |
| 2         | 1         | 25/3/2020      | 26/3/2020      |
| 1         | 3         | 9/9/2020       | 30/10/2020     |
| 2         | 2         | 10/10/2020     | 12/10/2020     |

FK (cle secondaire)

La base de données est composée d'un ensemble de tables (les relations) dans lesquelles sont placées des données. Les **entités** et les **associations** sont transformées en tables.

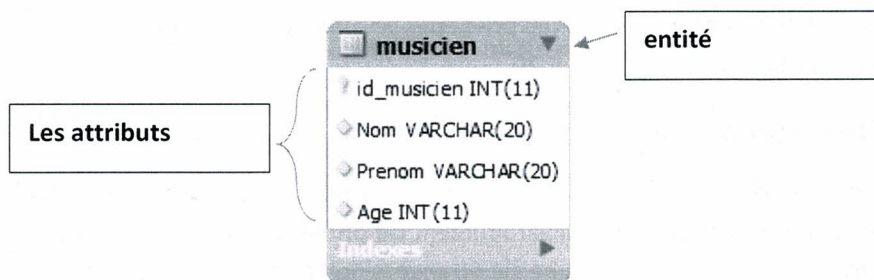
Exemple de la partie B :



Chaque enregistrement est un **p\_uplets** (tuple).

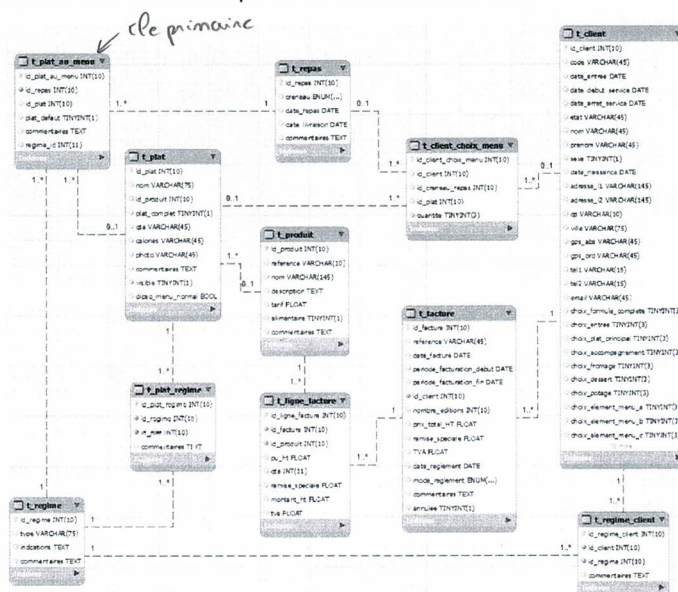
Avec un outil de gestion de base de données on peut choisir de ne représenter que **l'en tête** de la relation. Dans ce schéma **l'entité** ou **l'association** sont décrites comme un lot **d'attributs**.

Example:



**les relations** (tables) d'une base de données peuvent être nombreuses.

Example:



Les contraintes d'intégrité sont des règles assurant la cohérence des données dans la base de données.

#### - La contrainte de relation

Chaque relation liée à une **entité** possède un **attribut** particulier permettant d'identifier sans ambiguïté un **tuple** (enregistrement, donnée). Cet attribut particulier se nomme **clé primaire**. Sa valeur est unique et non nulle. C'est l'identifiant du modèle **entité association** d'une **entité**.

Ce numéro est un entier unique attribué automatiquement à chaque nouvel enregistrement. il est incrémenté automatiquement (à préciser lors de la création d'une relation). Il permet de distinguer sans ambiguïté chaque **p-uplet (tuple)** (enregistrement, donnée) de chaque relation.

Exemple : base de données d'une librairie (voir partie B)

| abonné       |          |        |
|--------------|----------|--------|
| clé primaire | prénom   | nom    |
| 1            | Paul     | Durand |
| 2            | Béatrice | Dupont |
| 3            | Michelle | Martin |

| livre        |                               |              |
|--------------|-------------------------------|--------------|
| clé primaire | titre                         | auteur       |
| 1            | 20 milles lieux sous les mers | Jules Vernes |
| 2            | le journal d'Anne Frank       | Anne Franck  |

#### - La contrainte de domaine (déjà vu c'est le domaine de valeur d'un attribut)

Une contrainte de domaine définit la plage des valeurs prises par un **attribut**. L'**attribut** note d'un élève doit, par exemple, être un nombre décimal compris entre 0 et 20.

#### - La contrainte de référence (ou contrainte d'intégrité référentielle)

Pour relier deux **entités** entre elles via une **association**, il faut utiliser pour la **relation** liée à l'**association** une ou plusieurs **clés étrangères**. Une **clé étrangère** est obligatoirement **clé primaire** d'une **relation** liée à une **entité**. La contrainte de référence est que la valeur prise par une **clé étrangère** doit correspondre à la valeur prise par la **clé primaire** dans l'autre **relation**.

Pour que les données restent utilisables et **cohérentes**, il ne faut pas que l'on puisse détruire des données qui dépendent les unes des autres. C'est le rôle de l'**intégrité référentielle** de protéger ces **relations**.

Exemple : base de données d'une librairie (voir partie B)

| abonné       |          |        |
|--------------|----------|--------|
| clé primaire | prénom   | nom    |
| 1            | Paul     | Durand |
| 2            | Béatrice | Dupont |
| 3            | Michelle | Martin |

| livre        |                               |              |
|--------------|-------------------------------|--------------|
| clé primaire | titre                         | auteur       |
| 1            | 20 milles lieux sous les mers | Jules Vernes |
| 2            | le journal d'Anne Frank       | Anne Franck  |

| emprunter            |                      |                |                |
|----------------------|----------------------|----------------|----------------|
| clé étrangère abonné | clé étrangère auteur | date d'emprunt | date de retour |
| 1                    | 1                    | 1/2/2019       | 1/3/2019       |
| 1                    | 2                    | 25/3/2020      | 26/3/2020      |
| 3                    | 1                    | 9/9/2020       | 30/10/2020     |
| 2                    | 2                    | 10/10/2020     | 12/10/2020     |

La juxtaposition des deux clés étrangères de la table emprunter identifie d'une façon unique une donnée (occurrence, enregistrement, tuple).

## E La normalisation relationnelle

La normalisation relationnelle représente des règles à respecter pour construire correctement le modèle relationnel. Il existe plusieurs formes normales. Cette partie débordant du programme, nous ne parlerons ici que de la première forme normale:

une relation suit la première forme normale si :

- Elle a au moins une **clé primaire**.
- Chacun de ses attributs a une valeur **atomique**. C'est-à-dire que son attribut n'est ni une liste de valeurs ni composé.

Exemple 1: Ici la **relation** livre ne respecte pas la première forme normale car on peut décomposer l'**attribut** auteur en deux **attributs** nom et prénom.

Exemple 2: Une adresse qui aurait une valeur de ce type 27 rue du Ballet - BP 60105 44001 NANTES Cedex 1 n'est pas un **attribut** qui respecte la première forme normale. Il faudrait décomposer l'adresse en plusieurs **attributs** (numéro 27, nom de la voie rue du Ballet, boîte postale BP 60105, code postal 44001 et ville NANTES Cedex 1).

## F Plusieurs relations, pourquoi ?

| Films               | annee | realisateurs | roles    | acteurs            |         |          |                   |               |
|---------------------|-------|--------------|----------|--------------------|---------|----------|-------------------|---------------|
|                     |       | nom          | prénom   |                    | nom     | prénom   | date de naissance | date de décès |
| Les scilliens       | 1969  | Henri        | verneuil | Roger Sartet :     | Delon   | Alain    | 08/11/1935        |               |
| Les scilliens       | 1969  | Henri        | verneuil | Vittorio Manalese: | Gabin   | Jean     | 17/05/1904        | 15/11/1976    |
| Intouchables        | 2011  | Eric         | toledano | philippe:          | Cluzet  | François | 21/09/1955        |               |
| Intouchables        | 2011  | Eric         | toledano | driss:             | Sy      | Omar     | 20/01/1978        |               |
| Intouchables        | 2011  | Olivier      | Nakache  | philippe:          | Cluzet  | François | 21/09/1955        |               |
| Intouchables        | 2011  | Olivier      | Nakache  | driss:             | Sy      | Omar     | 20/01/1978        |               |
| Nos jours heureux   | 2006  | Eric         | toledano | joseph:            | Sy      | Omar     | 20/01/1978        |               |
| Nos jours heureux   | 2006  | Olivier      | Nakache  | joseph:            | Sy      | Omar     | 20/01/1978        |               |
| Parents d'élèves    | 2020  | Noémie       | Saglio   | nora portel:       | Jordana | Camilla  | 15/09/1992        |               |
| La famille bélier   | 2014  | Eric         | Lartigo  | paula bélier:      | Louane  | Emera    | 26/11/1996        |               |
| La famille bélier   | 2014  | Eric         | Lartigo  | gigi bélier:       | Viard   | Karin    | 24/01/1966        |               |
| Je suis un assassin | 2004  | Thomas       | vincent  | ben castelano:     | Cluzet  | François | 21/09/1955        |               |
| Je suis un assassin | 2004  | Thomas       | vincent  | suzy Castelano:    | Viard   | Karin    | 24/01/1966        |               |

Ici il y a beaucoup de duplications d'informations. A partir d'un modèle entité-associations on obtiendra un modèle relationnel qui va les éviter. Ceci fera l'objet d'un travail dirigé.

## 4. Ressources

Principale : merci à M Hourdin Thomas professeur en BTS SN au campus Saint-Felix Lasalle qui a mis ces supports de cours et TP à disposition et dont une partie des connaissances et définitions sont soit extraites, soit reformulées ou ont aidées à la structuration de ce cours d'introduction aux bases de données.

#### Autres ressources:

[https://www.culture-informatique.net/comment-marche-base-de-donnees/#:~:text=Une%20base%20de%20donn%C3%A9es%20est,en%20reparle%20juste%20en%20dessous\).](https://www.culture-informatique.net/comment-marche-base-de-donnees/#:~:text=Une%20base%20de%20donn%C3%A9es%20est,en%20reparle%20juste%20en%20dessous).)

[https://fr.wikipedia.org/wiki/Syst%C3%A8me\\_de\\_gestion\\_de\\_base\\_de\\_donn%C3%A9es#/media/Fichier:Table\\_relationnel.png](https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_gestion_de_base_de_donn%C3%A9es#/media/Fichier:Table_relationnel.png)

<https://openclassrooms.com/fr/courses/918836-concevez-votre-site-web-avec-php-et-mysql/913655-quest-ce-quune-base-de-donnees>

<https://openclassrooms.com/fr/courses/1959476-administrez-vos-bases-de-donnees-avec-mysql/1970063-transactions>

<https://laurent-audibert.developpez.com/Cours-BD/?page=langage-sql>

[https://fr.wikipedia.org/wiki/Mod%C3%A8le\\_relationnel](https://fr.wikipedia.org/wiki/Mod%C3%A8le_relationnel)

<http://www.eaton.fr/FR/EatonFR/marches/DataCenter/Solutions/Typedecentrededonnees/index.htm>

<https://www.base-de-donnees.com/plus-grosse-base-de-donnees/>

<https://www.ionos.fr/digitalguide/hebergement/aspects-techniques/systeme-de-gestion-de-base-de-donnees-sgbd/>

[https://fr.wikipedia.org/wiki/Forme\\_normale\\_\(bases\\_de\\_donn%C3%A9es\\_relationnelles\)#1FN\\_%E2%80%93\\_Premi%C3%A8re\\_forme\\_normale](https://fr.wikipedia.org/wiki/Forme_normale_(bases_de_donn%C3%A9es_relationnelles)#1FN_%E2%80%93_Premi%C3%A8re_forme_normale)

<https://www.maxicours.com/se/cours/le-modele-relationnel/>

<https://stph.scenari-community.org/bdd/nor1-lin/co/norUC023.html>

<https://openclassrooms.com/fr/courses/4055451-modelisez-et-implementez-une-base-de-donnees-relationnelle-avec-uml/4458402-optimisez-votre-modele-relationnel-avec-les-formes-normales>