

Licence Topographie et géomatique
Cours n°1 : les bases de données et le SGBD
L'enseignant : Hadjira Lyes

I- L'origine des bases de données :

Le besoin de stocker et de manipuler de gros volumes de données existe depuis longtemps, et bien avant que l'informatique n'existe, on cherchait déjà à organiser de manière rationnelle le stockage et l'accès à ces données.

Avec les premiers ordinateurs, les informaticiens ont copié les organisations manuelles. A partir des fichiers manuels de l'SGF (système de gestion par fichier) on a créé des fichiers informatisés.

Définition d'un fichier : *un fichier est une collection structurée de fiches ou d'enregistrements.*

Exemple :

Code	Nom	Date d'inscription	Profession
001	Mohamed	1987	Agent
002	Sliman	1999	Surveillant
009	Souad	1982	Secrétaire
...

Ce fichier représente les employés d'une entreprise il contient une *fiche* ou un *enregistrement* par employé, chaque enregistrement est composée d'un ensemble de *champs*, correspondant aux caractéristiques que l'on désire mémoriser comme le nom de l'employé, son prénom, l'âge...

Remarque: Pour avoir une bonne idée d'un fichier, on peut l'imaginer à un tableau dans lequel les ligne correspondent aux enregistrement et les colonnes aux champs.

Le problème essentiel posé par le SGF est un problème d'organisation et d'accès aux enregistrements. Pour cette raison différentes méthodes de plus en plus sophistiquées ont été élaborées depuis le début de l'informatique. Nous allons en présenter les principes de base.

La situation avant les bases de données :

Dans l'organisation à base de fichiers les programmes qui ont besoins aux mêmes données utilisent des données physiquement distinctes. Cette organisation pré base de données pose de nombreux problèmes :

- *La redondance des données :* on a redondance si certaines données se trouvent stockées en plusieurs endroits. Ce qui entraînent un gaspillage d'espace disque et peut poser des problèmes d'incohérence lors de la mise à jour (il se peut qu'on modifie une données à un endroit et pas pour les autres).
- *La dépendance entre données et traitements :* si en modifie la structure d'une données il faut modifier tout les programmes qui l'utilisent. Et si on modifie une données il faut la modifie dans tout les endroits.
- *La complexité de la gestion des données :* en générale avec des dizaines ou des centaines d'applications il est difficile de gérer correctement les données d'une organisation.

La solution : pour éviter les problèmes évoqués ci-dessus, la première idée a été de *centraliser* les données pour mieux les partager en les rendant indépendantes des traitements. Ce travail a conduit depuis les années 70 jusqu'à nos jours, différent types d'organisation et de logiciels ont concrétisé peu a peu ces idées : il s'agit des bases de données et des systèmes de gestion de base de données SGBD.

II- Notion de base de données

Qu'est-ce qu'une base de données ?

Il est difficile de donner une définition exacte de la notion de base de données.

Définition 1 :

Une **Base de données est** *Un ensemble organisé d'informations avec un objectif commun.*

Peu importe le support utilisé pour rassembler et stocker les données (papier, fichiers, etc.), dès lors que des données sont rassemblées et stockées d'une manière organisée dans un but spécifique, on parle de base de données.

Définition 2

Base de données informatisée: *Une base de données informatisée est un ensemble structuré de données enregistrées sur des supports accessibles par l'ordinateur, représentant des informations du monde réel et pouvant être interrogées et mises à jour par une communauté d'utilisateurs.*

La gestion et l'accès à une base de données sont assurés par un ensemble de programmes qui constituent le *Système de gestion de base de données (SGBD)*.

III- Les avantages des bases de données et des SGBD

Les bases de données et les SGBD offrent les avantages suivants :

- *Une bonne représentation du monde réel* : la base doit fournir une image fidèle de la réalité avec des informations fiables et à jour.
- *La diminution de la redondance des données et des risques d'incohérence associés* : l'information contenue dans la base doit être unique d'un point de vue logique et physique
- *La sécurité et la confidentialité des données* : la sécurité des données est assurée par des procédures puissantes de sauvegarde physique, la confidentialité quant à elle est assurée par des procédures d'accès empêchant toute tentative non autorisée.
- *La rapidité des accès et des traitements* : l'organisation de la base offre amélioré les traitements et les temps d'accès
- *Une meilleure indépendance des données et des programmes d'application*. Il faut distinguer deux niveaux d'indépendance :
 - L'Indépendance physique : qui signifie qu'un changement concernant le stockage physique des données est transparent pour un utilisateur ou un programme d'application
 - L'Indépendance logique : signifiant qu'un changement de l'organisation logique globale est transparent pour les utilisateurs ou les programmes d'application existants.
- *Un accès aux données facilité* : grâce à la convivialité et à la puissance des outils d'interrogation intégrés dans les SGBD. Tout utilisateur non spécialiste peut facilement accéder à toutes sortes de données, même si ses besoins n'ont pas été prévus et intégrés dans un programme d'application.
- *Une meilleure gestion des données* : il est possible d'avoir une définition et une représentation fiable et claire des données pour toute l'organisation.

IV- Le cycle de vie d'une BD:

On appelle cycle de vie d'une BD la suite des phases suivantes

1. *La conception*: c'est la phase d'analyse qui aboutit à la détermination du contenu de la future BD c'est-à-dire son modèle de données
2. *L'implémentation* : consiste à décrire la base de données en langage de SGBD (LDD), et construire sa première version en utilisant le langage LMD.
3. *La phase de l'utilisation* : pendant cette phase la BD est prête à être utilisée par ces utilisateurs

V- les Modèles de base de données

Il existe quatre types de modèles des BD qui sont:

1. Modèle hiérarchique
2. Modèle réseau
3. Modèle relationnel
4. Modèle objet

VI- Le modèle relationnel

Dans ce modèle il n'existe qu'une seule structure qui est la **relation**. Une relation peut simplement être représentée sous forme d'une **table**.

Comme dans l'exemple suivant :

Code	Titre	Année	Type
001	Alien	1979	Science fiction
002	Vertigo	1958	Policier
003	Volte-face	1997	Action
...	...		

Un relation

La relation a un nom (**film**), et se compose de colonnes désignées par un nom d'attribut. Dans chaque colonne on trouve des valeurs d'un certain domaine. Dans ce modèle chaque ligne correspond à une entité.

Le schéma de la relation film est le suivant:

Film(code: nombre, titre: string, année: nombre, type: string).

6.1 Définitions

6.1.1 Domaine: un domaine de valeurs est un ensemble d'éléments d'un type élémentaire, par exemple : les entiers, les réels, les chaînes de caractères,...etc.

6.1.2 Attributs : les attributs nomment les colonnes d'une relation il servent à la fois à indiquer le contenu de cette colonne, et à la référencer quand on effectue des opérations. Un attribut est toujours associé à un domaine. Le nom d'un attribut peut apparaître dans plusieurs schémas de relations.

6.1.3 Le schéma d'une relation: c'est un nom d'une relation suivi de la liste des attributs où chaque attribut est associé à son domaine. La syntaxe est la suivante:

$R(A_1:\text{Domaine}_1, A_2:\text{Domaine}_2, \dots, A_n:\text{Domaine}_n)$

Où:

R: le nom de la relation

A_i : les noms des attributs

D_i : les domaines

6.1.4 le schéma relationnel d'une base de donnée : il est constitué d'un ensemble de schémas de relations qui décrivent, à l'aide des éléments présentés ci-dessus (domaine, attributs, nom de relation) le contenu d'une base de données.

Remarque: on peut trouver dans un schéma de relation plusieurs fois le même domaine, mais une seule fois le nom d'un attribut.

6.1.5 Clé de la relation: c'est un attribut spécifique de la relation, car il permet d'identifier chaque ligne de la relation d'une manière unique. La clé de la relation doit être signalé en gras ou en souligné

*Film(**code: nombre**, titre: string, année: nombre, type: string).*

Film(code: nombre, titre: string, année: nombre, type: string)

6.1.6 Le tuple: un tuple est une ligne dans la représentation d'une relation sous forme de table.

Exemple : (001, alian, 1979, science fiction) est un tuple de la relation film.

VII- Les opérations élémentaire sur une base de donnée

Il existe plusieurs opérations d'interrogations d'une BD parmi ces opérations on cite : la projection, la sélection, la jointure ainsi que d'autre opérations comme les opération de calcule.

La projection : consiste à choisir certains attributs d'une relation et d'éliminer les doublant (ligne ou tuples identique) dans le résultat. On a recours à une projection si on veut s'intéresser uniquement à certaines colonnes d'une table.

La sélection: comme son nom l'indique, est l'opération permettant de sélectionner les enregistrement qui répondent à certaines conditions.

La jointure: ou la composition, c'est une opération permettant de rapprocher deux relations (tables) à partir d'un attribut commun.

Ce type d'opération est nécessaire lorsque la réponse à une requête nécessite l'utilisation d'informations contenues dans plusieurs tables.

VIII- Les opération de mis-a-jour

Ces opération concernent :

- L'insertion d'une nouvelle information
- La modification d'un information
- La consultation d'une information
- La suppression d'une information

IX- Les utilisateurs des BD

Les domaines d'utilisation des BD sont nombreux, car tous les domaines ont besoins à organiser leurs données, on peut citer à titre d'exemple:

L'aménagement, la géographie, la gestion, l'environnement...etc.