

1- Étapes Modélisation et conception de bases de données

- *La modélisation des données* est la *première étape* du processus de conception de base de données. haut niveau (entité-association)
- Cette étape est parfois considérée comme une phase de *conception abstraite* de haut niveau, également appelée *conception conceptuelle*.
- L'objectif de cette phase est de décrire :
 - Les données contenues dans la base de données (par exemple, les entités : étudiants, professeurs, cours, matières)
 - Les relations entre les éléments de données (par exemple, les étudiants sont supervisés par des professeurs ; les professeurs enseignent des cours)
 - Les contraintes sur les données (par exemple, le numéro d'étudiant a exactement huit chiffres ; une matière a uniquement quatre ou six crédits)
- *La deuxième étape* consiste à exprimer les éléments de données, les relations et les contraintes en utilisant les concepts fournis par le modèle de données de haut niveau.
- Parce que ces concepts n'incluent pas les détails de mise en œuvre, le résultat du processus de modélisation des données est une représentation (semi-)formelle de la structure de la base de données.
- Ce résultat est assez facile à comprendre, il est donc utilisé comme référence pour s'assurer que toutes les exigences de l'utilisateur sont respectées.
- *La troisième étape* est la conception de la base de données.
- Au cours de cette étape, nous pouvons avoir deux sous-étapes :
 - Une appelée conception logique de la base de données, qui définit une base de données dans un modèle de données d'un SGBD spécifique, (interne, les 3 types, inclus relationnel)
 - et une autre appelée conception physique de la base de données, qui définit la structure de stockage interne de la base de données, l'organisation des fichiers ou les techniques d'indexation.
 - Ces deux sous-étapes sont les étapes de mise en œuvre de la base de données et de création des opérations/interfaces utilisateur.

2- Définition d'un Modèles de données

- Dans les phases de *conception de la base de données*, les données sont représentées à l'aide d'un certain modèle de données.
- Le modèle de données est une collection de concepts ou de notations pour décrire les données, les relations entre les données, la sémantique des données et les contraintes de données.

- La plupart des modèles de données incluent également un ensemble d'opérations de base pour manipuler les données dans la base de données.

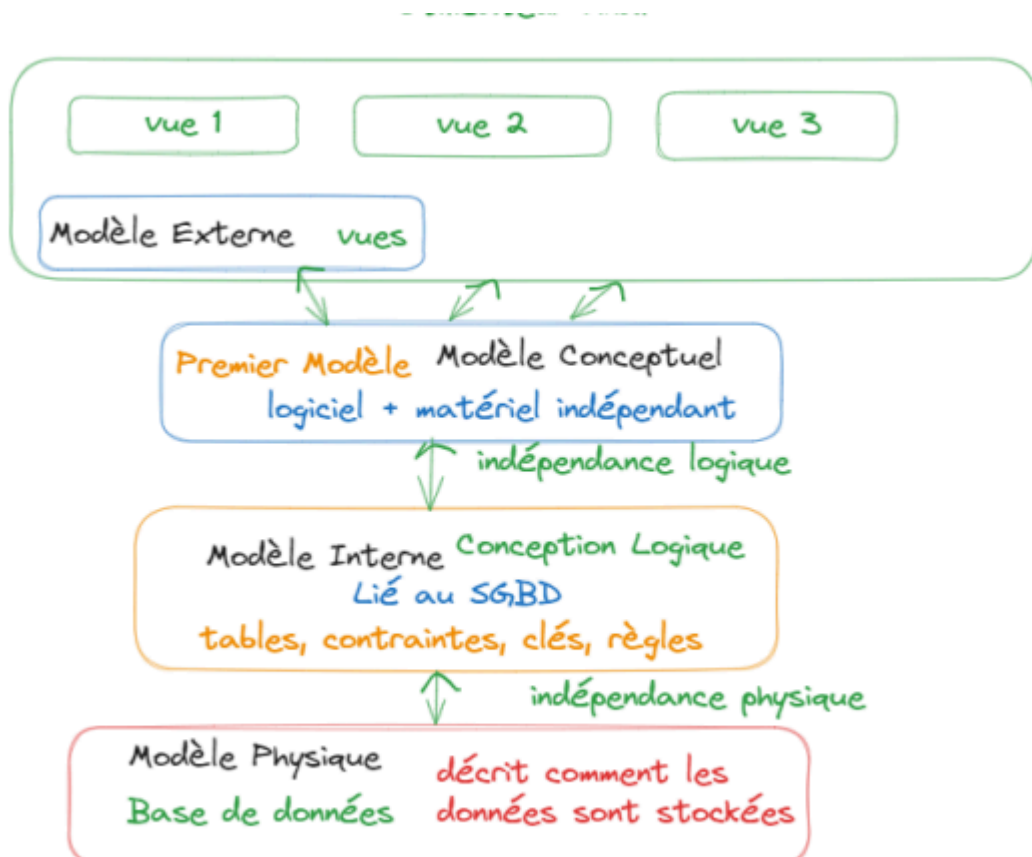
3- Degrés d'abstraction des données

Il existe 4 types de modèles de données : externes, conceptuels, internes, et physique

- Les modèles externes représentent la vue de l'utilisateur de la base de données.
- Les modèles conceptuels fournissent des capacités de structuration des données flexibles. (modèle de données relationnel, modèle de données réseau, modèle de données hiérarchique data model.)
- Les modèles internes considèrent une base de données comme une collection d'enregistrements de taille fixe.
- Modèles physiques: Représentent la représentation physique de la base de données. Ont le niveau d'abstraction le plus bas et décrivent comment les données sont stockée

4- Couche d'abstraction des données

- La couche d'abstraction des données permet de masquer la complexité des données aux utilisateurs de la base de données.
 - Elle fournit une interface cohérente aux données, indépendamment de la façon dont les données sont stockées et organisées physiquement.
 - La couche d'abstraction des données permet aux développeurs d'applications de se concentrer sur la logique métier de leurs applications, plutôt que sur les détails de stockage des données.



5- Types de modèles de données logiques basés sur des enregistrements

- Ces modèles fournissent des concepts que les utilisateurs peuvent comprendre, mais ils ne sont pas trop éloignés de la manière dont les données sont stockées dans l'ordinateur.
- Trois modèles de données bien connus de ce type sont les *modèles de données relationnels*, les *modèles de données réseau* et les *modèles de données hiérarchiques*.

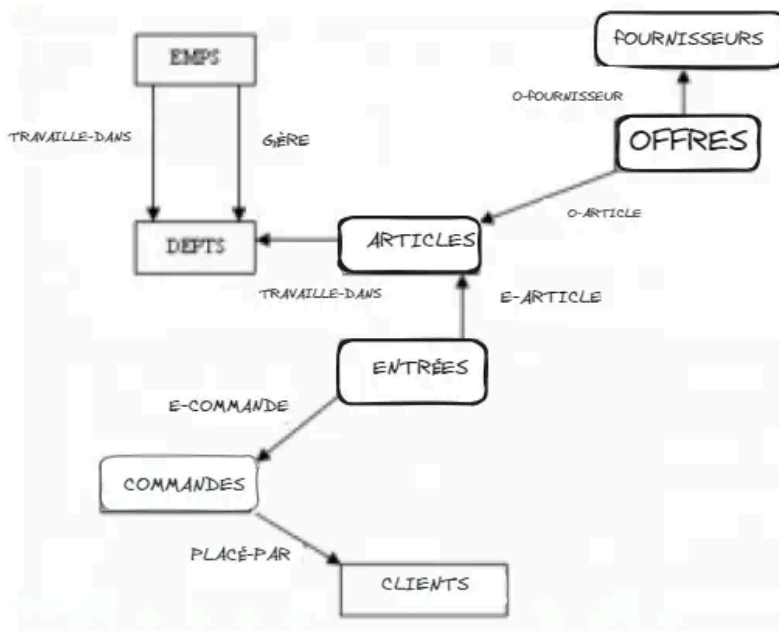
5.1 Modèle relationnel

- Ce modèle représente les données sous forme de relations ou de tables.
 - Par exemple, dans le système d'adhésion à Science World, chaque adhésion a de nombreux membres
 - L'identifiant d'adhésion, la date d'expiration et les informations d'adresse sont des champs dans l'adhésion.
 - Les membres sont des individus tels que Mickey, Minnie, Mighty, Door, Tom, King, Man et Moose.
 - Chaque enregistrement est dit être une instance de la table d'adhésion.
 - termes utilisés: relation ou table, champs-attributs , enregistrement: record (row, tuple) instance

customer_id	store_id	first_name	last_name	email
1	1	MARY	SMITH	MARY.SMITH@sakilacustomer.org
2	1	PATRICIA	JOHNSON	PATRICIA.JOHNSON@sakilacustomer.org
3	1	LINDA	WILLIAMS	LINDA.WILLIAMS@sakilacustomer.org
4	2	BARBARA	JONES	BARBARA.JONES@sakilacustomer.org
5	1	ELIZABETH	BROWN	ELIZABETH.BROWN@sakilacustomer.org
6	2	JENNIFER	DAVIS	JENNIFER.DAVIS@sakilacustomer.org
7	1	MARIA	MILLER	MARIA.MILLER@sakilacustomer.org

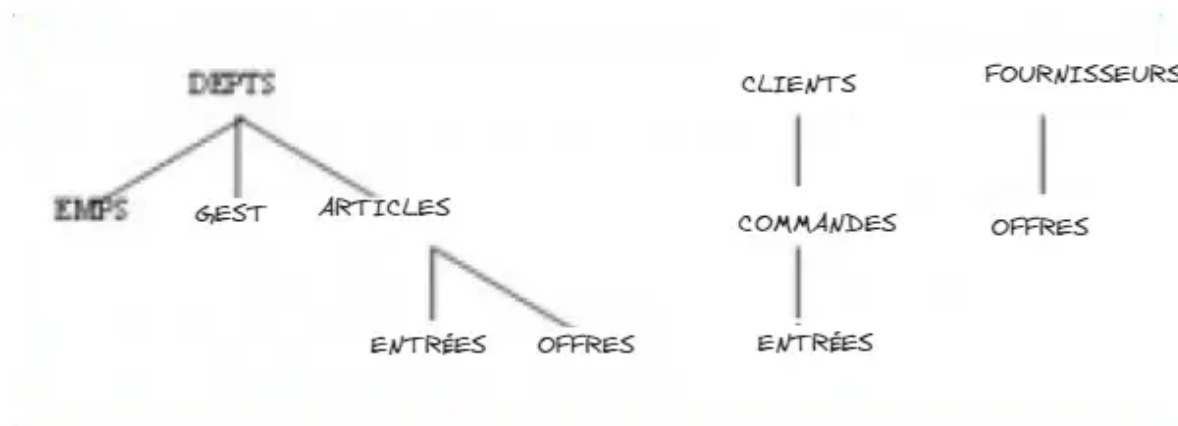
5.2 Modèle réseau

- Ce modèle représente les données sous forme de types d'enregistrement.
- Ce modèle représente également un type limité de relation un à plusieurs appelé type d'ensemble
- termes utilisés: enregistrement, ensemble



5.3 Modèle hiérarchique

- Ce modèle représente les données sous forme d'une structure arborescente hiérarchique.
- Chaque branche de la hiérarchie représente un certain nombre d'enregistrements liés.



6- Le modèle de données relationnel

Le modèle de données relationnel a été introduit par E. F. Codd en 1970. À l'heure actuelle, c'est le modèle de données le plus utilisé.

- Le modèle relationnel a fourni la base pour :

- La recherche sur la théorie des données/relations/contraintes
- De nombreuses méthodologies de conception de bases de données
- Le langage standard d'accès aux bases de données appelé langage de requête structuré (SQL)
- Presque tous les systèmes de gestion de bases de données commerciaux modernes
- Le modèle de données relationnel décrit le monde comme « un ensemble de relations (ou tables) interdépendantes ».

6-1 Concepts fondamentaux du modèle de données relationnel

- **Relation = Table**

- Une relation, également appelée table ou fichier, est un sous-ensemble du produit cartésien d'une liste de domaines caractérisée par un nom.
- Dans une table, chaque ligne représente un groupe de valeurs de données connexes.
- Une ligne, ou enregistrement, est également appelée un tuple.
- La colonne dans une table est un champ et est également appelés un attribut.
- *Définition formelle:*

- Étant donné n domaines
- Et r une relation définie sur ces domaines
- Alors: $r \subseteq D_1 \times D_2 \times \dots D_n$

- **Colonne:**

- La principale unité de stockage dans une base de données est appelée: colonne, champs, ou attribut.
- ils abritent les composants de base des données dans lesquels votre contenu peut être décomposé.
- pour décider quels champs créer, vous pouvez par exemple:
- sélectionner les composants communs des informations que vous stockerez dans la base de données
- éviter les détails qui distinguent un élément d'un autre.

- **Domaine**

- Un domaine est l'ensemble original de valeurs atomiques utilisé pour modéliser les données.
- valeur atomique signifie que chaque valeur dans le domaine est indivisible en ce qui concerne le modèle relationnel.
- exemples: Le domaine Prénom est l'ensemble des chaînes de caractères qui représentent les noms de personnes.
 - Le domaine État civil offre un ensemble de possibilités : Marié, Célibataire, Divorcé.

- **Enregistrements**

- Les enregistrements contiennent des champs qui sont liés, tels qu'un client ou un employé.
- Comme indiqué précédemment, un tuple est un autre terme utilisé pour l'enregistrement. (ou bien une ligne)

Nom d'attribut

customer_id	store_id	first_name	last_name	email
1	1	MARY	SMITH	MARY.SMITH@sakilacustomer.org
2	1	PATRICIA	JOHNSON	PATRICIA.JOHNSON@sakilacustomer.org
3	1	LINDA	WILLIAMS	LINDA.WILLIAMS@sakilacustomer.org
4	2	BARBARA	JONES	BARBARA.JONES@sakilacustomer.org
5	1	ELIZABETH	BROWN	ELIZABETH.BROWN@sakilacustomer.org
6	2	JENNIFER	DAVIS	JENNIFER.DAVIS@sakilacustomer.org
7	1	MARIA	MILLER	MARIA.MILLER@sakilacustomer.org

lignets

valeurs d'attributs

- Degré
 - Le degré est le nombre d'attributs dans une table. Dans l'exemple de la figure au dessus, le degré est 5

6-2 Propriétés d'une table

- Une table a un nom qui est distinct de toutes les autres tables de la base de données.
- Il n'y a pas de lignes en double ; chaque ligne est distincte.
- Les entrées dans les colonnes sont atomiques. La table ne contient pas de groupes répétés ou d'attributs à valeurs multiples.
- Les entrées des colonnes appartiennent au même domaine en fonction de leur type de données, notamment :
 - nombre (numérique, entier, float, smallint,...)
 - caractère (chaîne)
 - date
 - logique (vrai ou faux)
- Les opérations combinant différents types de données ne sont pas autorisées.
- Chaque attribut a un nom distinct.
- La séquence des colonnes n'est pas significative.
- La séquence des lignes n'est pas significative.

Références Bibliographiques

Watt, Adrienne, and Nelson Eng. *Database design*. 2nd Edition. BCcampus, 2014