

1 Présentation : bases de données

Le développement des traitements informatiques nécessite la manipulation de données de plus en plus nombreuses. Leur organisation et leur stockage constituent un enjeu essentiel de performance.

On a vu l'an dernier une manière de stocker et manipuler des données structurées à l'aide de différents formats (ex : CSV, JSON). Ces formats basés essentiellement sur du texte sont faciles à mettre en œuvre et à utiliser **mais ne sont pas adaptés au traitement d'un grand nombre d'informations**, en particulier lorsque celles-ci se trouvent réparties dans plusieurs tables ou fichiers.

Les premières bases de données sont apparues dans les années 1960 et se sont développées en même temps que l'informatique.

Le modèle relationnel a été introduit par Edgard Franck **Codd** en 1970 (prix Turing en 1981).

De nos jours les bases de données sont omniprésentes, en particulier sur le Web. La plupart des sites, en particulier dans le commerce en ligne, y font largement appel.

2 Mise en place des concepts avec un exemple

2.1 Exemple pratique

On considère la table suivante :

titre	nom auteur	prenom auteur	date nai auteur	langue	ann publi	thèmes
1984	Orwell	George	1903	anglais	1949	Totalitarisme, science-fiction, anticipation, Dystopie
Dune	Herbert	Frank	1920	anglais	1965	science-fiction, anticipation
Fondation	Asimov	Isaac	1920	anglais	1951	science-fiction, Economie
Le meilleur des mondes	Huxley	Aldous	1894	anglais	1931	Totalitarisme, science fiction, dystopie
Fahrenheit 451	Bradbury	Ray	1920	anglais	1953	science-fiction, Dystopie
Ubik	K. Dick	Philip	1928	anglais	1969	science-fiction, anticipation
Chroniques martiennes	Bradbury	Rey	1920	anglais	1950	science-fiction, anticipation
La nuit des temps	Barjavel	René	1911	français	1968	science-fiction, tragédie
Blade runner	K Dick	Philip	1928	anglais	1968	Intelligence artificielle, science fiction
Les robots	Asimov	Isaac	1920	anglais	1950	science fiction, Intelligence artificielle
La planète des singes	Boulle	Pierre	1912	français	1963	science fiction, Dystopie
Ravage	Barjavel	René	1911	français	1943	Science-Fiction, anticipation
Le maître du haut château	K.Dick	Philip	1928	anglais	1962	Dystopie, Uchronie
Le monde des A	Van Vogt	Alfred Elton	1912	anglais	1945	science fiction, IA
La fin de l'éternité	Asimov	Isaac	1920	anglais	1955	science-fiction, voyage dans le temps
De la Terre à la Lune	Verne	Jules	1828	français	1865	Science-Fiction, aventure



Imaginez quels problèmes peuvent se poser si on cherche à mettre en place un traitement automatisé pour les données de cette table.

Quelles solutions imaginer pour y remédier ?

2.2 Le modèle relationnel

Un des modèles de données le plus courant est le **modèle relationnel**.

Les principes de base de ce modèle sont les suivants :

- séparer les données dans plusieurs tables, ou **relations**.
- chaque relation contient des données relatives à un même sujet.
- éviter la redondance des données.
- ne pas stocker des données qui peuvent être calculées (exemple : jamais de ligne Total).
- chaque champ ne contient qu'une seule information.
- mettre les tables en relation par l'utilisation de clefs :
 - **clefs primaires** : leurs valeurs (souvent des entiers) permettent d'identifier un enregistrement de manière unique.
 - **clefs étrangères** : réfèrent une clef primaire d'une autre table.

2.3 Quelques éléments de vocabulaire

- Un **enregistrement** est composé de plusieurs informations distinctes appartenant à un même objet ou une même entité. *Pour faire simple, cela correspond à une ligne du tableau.*
- Un **attribut** est une information élémentaire appartenant à un enregistrement. *Les attributs se voient ici sur les intitulés de colonne.*
*Remarque : on parle aussi parfois de **champ**.*
- Un **attribut atomique** est un attribut ne contenant qu'une seule information.
- Chaque attribut possède un **type** : *le titre est du type chaîne de caractères et la note est du type entier.*
- Le **domaine** d'un attribut donné correspond à un ensemble (fini ou infini) de valeurs admissibles. *Le domaine pour note est l'ensemble des entiers de 0 à 10.*
- Le **schéma relationnel** d'une table décrit la structure de la relation : ensembles des attributs avec leur domaine en précisant la clef primaire et, le cas échéant, les clefs étrangères.

Exemples de présentation de **schémas relationnels** :

Classiquement, on souligne les attributs de la clef primaire, et note un # devant les attributs des clefs étrangères :

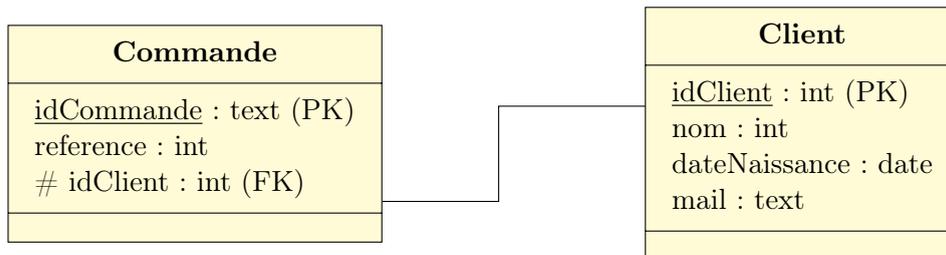
- relation Client (idClient : int, nom : text, prénom : text, mail : text)
- relation Commande (idCommande : text, référence : text, #idClient : int)

On peut aussi donner une représentation sous forme de table :

PK = Primary Key ; FK = Foreign Key.

Client	Commande
idClient (PK) : int	idCommande (PK) : text
nom : text	référence : text
prenom : text	idClient (FK) : int
dateNaissance : date	
mail : text	

Relation Client Relation Commande



Analyse de la table exemple :

1. Combien la table possède t-elle d'enregistrements ?
2. Quels sont tous les attributs de la table ?
3. Citer un attribut atomique et un attribut non atomique.
4. Pour chaque attribut, préciser le type et le domaine.



Construire une structure optimisée :

1. Identifier les éléments de la table qui sont en contradiction avec les principes du modèle relationnel énoncés ci-dessus.
2. Pour corriger ces problèmes qui transgressent les règles du modèle relationnel, redistribuer ces données dans 4 relations : Livre, Auteur, Theme, Langue.
3. Quels attributs proposer pour chacune de ces relations ?
4. Comment peut-on connaître l'auteur de 1984 et les thèmes abordés ?
5. Dans la relation Livre, quelle est la clef primaire ? quelles sont les clefs étrangères ?
6. Dans la relation Auteur, l'attribut date de naissance est-il une clef primaire ?
7. Quel peut être l'intérêt de mettre la langue dans une relation Langue ? Quel problème dans la table de départ cela résout-il ?
8. Écrire le schéma relationnel des 4 relations proposées.

2.4 En conclusion

On retiendra les principes généraux pour regrouper les données en plusieurs tables :

- Mettre dans une même table les données relatives à un même sujet.
- Chaque attribut ne contient qu'une seule valeur.
- Créer une clef primaire pour chaque relation.
- Créer de nouvelles relations pour :
 - éviter la redondance des données.
 - limiter les incohérences lors des mises à jour.
 - faciliter le traitement des données.

3 Extraire des informations dans un modèle relationnel

Extrait relation Clients :

IdClient	Nom	Adresse	Mail
1	Jean Bon	2 rue Jean Mermoz - Caen	jean.bon@free.fr
2	Alain Térieur	3 rue Paul Eluard - Hérouville	alain.terieur4@hotmail.com
3	Thérèse Etroit	16 rue de la porte - Mondeville	therese.etroit@orange.fr
4	Gilles Héjone	1 place de la bastille - Bénouville	gilledu14@sfr.fr
5	Hélène de Troie	2 rue Néper - Caen	ln23@laposte.net

Extrait relation Produits :

IdProduit	CodeProduit	NomProduit	Prix	Stock
1	12x24F	gel hydroalcoolique 100mL	3.21	2
2	21s53R	masque FFP2 x100	10,57	3
3	97D74S	visière de protection	0.50	10
4	10F36A	désinfectant industriel 10L	53,25	5

Extrait relation Commandes pour un jour J :

IdCmd	IdClient	IdProduit	Quantite	Expedie
1	4	1	1	1
2	2	2	2	0
3	1	1	1	1
4	3	3	4	0
5	1	4	3	0

1. Dans la table Commandes quelles est la clef primaire ? quelles sont les clefs étrangères ?
2. À combien s'élève le montant total des commandes de la journée J ?
3. Quels sont les noms des clients pour lesquels la commande a été expédiée ?
4. Quels sont les produits pour lesquels le stock ne sera pas suffisant ?