

Bases de données

Bases de données

Une base de données est un ensemble structuré d'informations.

- Dans le langage courant, elle peut désigner n'importe quelle source importante d'informations (dictionnaires, encyclopédies, etc.)
- En informatique, il s'agit d'informations stockées sous forme de fichiers et organisées de façon à être facilement manipulées.

Prenons par exemple ce tableau qui contient quelques informations sur des ouvrages...

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Azimov	1951	9
4	Le meilleur des mondes	Huxley	1931	9
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	6
7	Chroniques martiennes	Bradbury	1950	7
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	9
10	Les Robots	Azimov	1950	10
11	La Planète des singes	Boulle	1963	9
12	Ravage	Barjavel	1943	11
13	Le Maître du Haut Château	K.Dick	1962	7
14	La fin de l'éternité	Azimov	1955	9

Table : LIVRES

On peut l'imaginer comme un extrait d'une plus vaste collection d'informations...

Définition Vocabulaire

Cette **table** est une **relation** composée :

- d'une **en-tête** (libellés des attributs)
- d'un **corps** qui est composé :
 - de **t-uplets** (lignes)
 - d'**attributs** (colonnes)

id	titre	auteur	ann_publ	note	en-tête
1	1984	Orwell	1949	10	
2	Dune	Herbert	1965	8	
3	Fondation	Azimov	1951	9	
4	Le meilleur des mondes	Huxley	1931	9	
5	Fahrenheit 451	Bradbury	1953	7	
6	Ubik	K.Dick	1969	6	
7	Chroniques martiennes	Bradbury	1950	7	
8	La nuit des temps	Barjavel	1968	7	
9	Blade Runner	K.Dick	1968	9	
10	Les Robots	Azimov	1950	10	un t-uplet
11	La Planète des singes	Boule	1963	9	
12	Ravage	Barjavel	1943	11	
13	Le Maître du Haut Château	K.Dick	1962	7	
14	La fin de l'éternité	Azimov	1955	9	

attribut : titre corps

Remarque : Visuellement cette **table** est un tableau qui met en **relation** des données.

- Un t-uplet contient des éléments (exemple : (10,Les Robots, Azimov, 1950,10))
- Un attribut contient des éléments (exemple : l'attribut auteur contient (Orwell, Herbert,...))

Important : Pour chaque attribut d'une relation, il faut définir un **domaine**.

- Le domaine d'un attribut est l'ensemble (fini ou pas) des valeurs possibles (*entiers, flottants, chaînes de caractères, dates...*).

Exemple : le domaine de l'attribut note est : l'ensemble des entiers positifs.

celui de l'attribut auteur est : l'ensemble des chaînes de caractères.

(il faut rendre impossible l'entrée d'une note non entière lors de la création d'un t-uplet)

Fondamental

Unicité d'un t-uplet

Dans une relation, deux t-uplets identiques sont non autorisés.

Dans une relation, il faut pouvoir **identifier de façon unique un t-uplet**.

Il faut déterminer, parmi les attributs, lequel permet d'identifier de manière unique un t-uplet.

Cet attribut sera considéré comme la **clé primaire** de la relation.

Dans notre exemple :

- L'attribut note ne peut pas jouer le rôle de clé primaire (deux ouvrages pouvant avoir la même note)
- De même pour les attributs auteurs et ann_publ et titre.
- Il reste donc l'attribut "id" (pour identifiant), qui a été ajouté ici pour jouer le rôle de clé primaire.

Remarque : ici, nous avons créé artificiellement une clé primaire, car aucun des autres attributs ne pouvait convenir. (ce n'est pas toujours le cas)

id	titre	auteur	ann_publ	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Azimov	1951	9
4	Le meilleur des mondes	Huxley	1931	9
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	6
7	Chroniques martiennes	Bradbury	1950	7
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	9
10	Les Robots	Azimov	1950	10
11	La Planète des singes	Boule	1963	9
12	Ravage	Barjavel	1943	11
13	Le Maître du Haut Château	K.Dick	1962	7
14	La fin de l'éternité	Azimov	1955	9

Attention

Redondance

En y regardant de plus près, un certains nombre d'informations apparaissent plusieurs fois...

Par exemple : l'auteur "Azimov" apparaît 3 fois.

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Azimov	1951	9
4	Le meilleur des mondes	Huxley	1931	9
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	6
7	Chroniques martiennes	Bradbury	1950	7
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	9
10	Les Robots	Azimov	1950	10
11	La Planète des singes	Boulle	1963	9
12	Ravage	Barjavel	1943	11
13	Le Maître du Haut Château	K.Dick	1962	7
14	La fin de l'éternité	Azimov	1955	9

Table : LIVRES

Dans une base de donnée, on évite autant que possible **la redondance d'informations**.

Pour cela, on crée une autre relation AUTEURS que l'on lie avec la relation LIVRES.

id	titre	id_auteur	ann_publi	note
1	1984	1	1949	10
2	Dune	2	1965	8
3	Fondation	3	1951	9
4	Le meilleur des mondes	4	1931	9
5	Fahrenheit 451	5	1953	7
6	Ubik	6	1969	6
7	Chroniques martiennes	5	1950	7
8	La nuit des temps	7	1968	7
9	Blade Runner	6	1968	9
10	Les Robots	3	1950	10
11	La Planète des singes	8	1963	9
12	Ravage	7	1943	11
13	Le Maître du Haut Château	6	1962	7
14	La fin de l'éternité	3	1955	9

Relation : LIVRES

id	nom	premier	ann_naissance	langue_ecriture
1	Orwell	George	1903	anglais
2	Herbert	Frank	1920	anglais
3	Azimov	Isaac	1920	anglais
4	Huxley	Aldous	1894	anglais
5	Bradbury	Ray	1920	anglais
6	K.Dick	Philip	1928	anglais
7	Barjavel	René	1911	français
8	Boulle	Pierre	1912	français

Relation : AUTEURS

On a rajouté quelques informations supplémentaires sur les auteurs dans la relation AUTEURS.

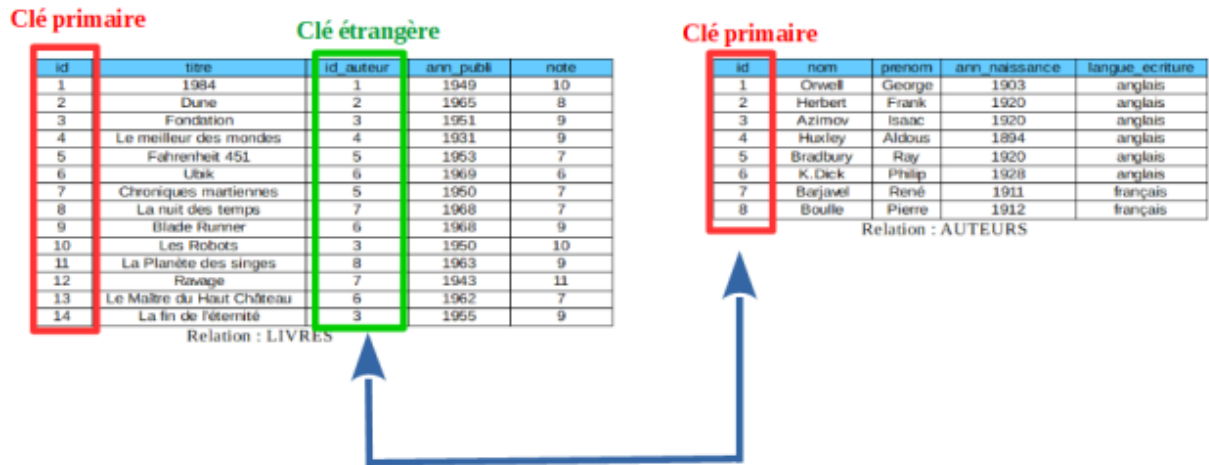
Fondamental

Clé primaire - Clé étrangère

Dans la relation AUTEURS, chaque auteur est identifié par l'attribut "id"(clé primaire de la relation)

- Dans la relation LIVRES, on a rajouté un attribut "id_auteur" qui est la **clé primaire** de la relation AUTEURS.
- L'attribut "id_auteur" est ce que l'on nomme une **clé étrangère** de la relation LIVRES, elle permet de faire le lien entre les deux relations.

Remarque : il peut y avoir plusieurs clés étrangères dans une relation.



Nous venons de définir les contours d'une base de données relationnelle.

Méthode Schéma relationnel

On appelle schéma relationnel l'ensemble des relations présentes dans une base de données.
c'est à dire qu'il faut fournir :

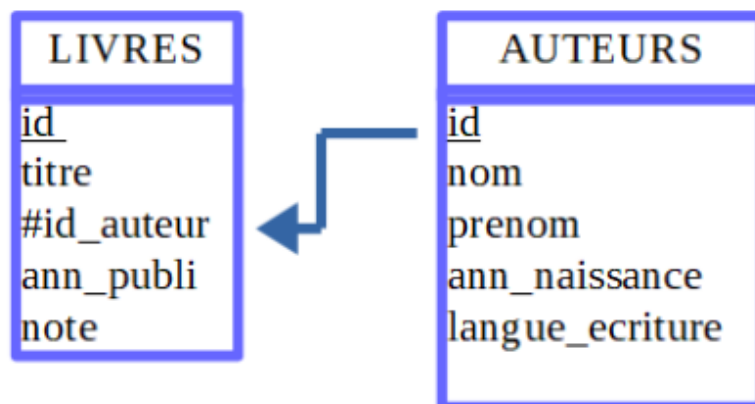
- les noms des relations
- la liste des attributs et de leurs domaines
- la clé primaire (on la souligne)
- les clés étrangères (précédées d'un #)

Pour notre exemple cela donne :

LIVRES (id :entiers positifs , titre : str , #id_auteur : entiers positifs, ann_publi : entiers positifs , note : entiers positifs)

AUTEURS (id :entiers positifs , nom : str, prenom : str , ann_naissance : entiers positifs , langue_ecriture : str)

Ce qui peut aussi se représenter avec ce schéma :



id, ann_publi, ann_naissance, note, id_auteur sont des entiers positifs
titre, nom, prenom, langue_ecriture sont des chaînes de caractères

