

Cours de bases de données,  
<http://sql.bdpedia.fr>

Qualité d'un schéma relationnel

# Qualité d'un schéma

Cette session discute des bons et mauvais schémas relationnels.

Les notions introduites sont :

- Les anomalies et incohérences dues à un schéma défectueux
- Les dépendances fonctionnelles
- Les clés primaires et étrangères
- La **normalisation**

Ces diapositives correspondent au support en ligne disponible sur le site <http://sql.bdpedia.fr/>

# Pourquoi pas une unique table ?

titre	année	nomMES	prénomMES	annéeNaiss
Alien	1979	Scott	Ridley	1943
Vertigo	1958	Hitchcock	Alfred	1899
Psychose	1960	Hitchcock	Alfred	1899
Kagemusha	1980	Kurosawa	Akira	1910
Volte-face	1997	Woo	John	1946
Pulp Fiction	1995	Tarantino	Quentin	1963
Titanic	1997	Cameron	James	1954
Sacrifice	1986	Tarkovski	Andrei	1932

Risques d'anomalies et d'incohérences en **insertion**, **modification** et **suppression**.

# Notion de dépendance fonctionnelle

Il y a **dépendance fonctionnelle**  $A \rightarrow B$  entre deux attributs  $A$  et  $B$  d'une relation  $R$  quand la connaissance de la valeur de  $A$  implique la connaissance de la valeur de  $B$ .

Exemple : le schéma (nom, prénom, noSS, dateNaissance, adresse, email)

Les dépendances :

- $email \rightarrow nom, prénom, noSS, dateNaissance, adresse$
- $noSS \rightarrow email, nom, prénom, dateNaissance, adresse$

Les DF servent à analyser la qualité d'un schéma.

# Les manuscrits et les experts

Un éditeur fait évaluer tous les manuscrits qu'on lui soumet par un expert.

Voici le schéma contenant tous les attributs.

Manuscrit (*id\_manuscrit*, *auteur*, *titre*, *id\_expert*, *nom*, *commentaire*)

Et les dépendances **minimales** et **directes** :

- *id\_manuscrit* → *auteur*, *titre*, *id\_expert*, *commentaire*
- *id\_expert* → *nom*

Ce schéma a le même type de problème que celui des films.  
Nous allons l'analyser avec les DF.

# Un exemple de la relation Manuscrit

<b>id_m</b>	<b>auteur</b>	<b>titre</b>	<b>id_exp</b>	<b>nom</b>	<b>commentaire</b>
10	Serge	L'arpète	2	Philippe	Une réussite
20	Cécile	L'art du chant grégorien	2	Sophie	Bravo
10	Serge	L'arpète	2	Philippe	Une réussite
10	Philippe	SQL	1	Sophie	La référence

- $id\_expert \rightarrow nom$  n'est pas respectée par le premier et deuxième nuplet.
- $id\_manuscrit \rightarrow auteur, titre, id\_expert, commentaire$  n'est pas respectée par le premier et quatrième nuplet.

Les DF sont respectées par le premier et le troisième, mais du coup ce sont des **doublons**. Il faut **éviter** ces anomalies.

# Notion de clé

## Clé

Une clé d'une relation  $R$  est un sous-ensemble **minimal**  $C$  des attributs tel que tout attribut de  $R$  dépend fonctionnellement de  $C$ .

- $id\_manuscrit$  est une clé pour la relation *Manuscrit*
- $(id\_manuscrit, auteur)$  n'est pas une clé car non minimal

On peut trouver plusieurs clés. On en choisit une comme **clé primaire**. Les autres sont clés secondaires.

# Normalisation

## Forme normale (troisième)

Un schéma de relation  $R$  est **normalisé** quand, dans toute dépendance fonctionnelle  $S \rightarrow A$  sur les attributs de  $R$ ,  $S$  est une clé.

Le schéma précédent n'est pas normalisé car il existe une dépendance fonctionnelle  $id\_expert \rightarrow nom$ , alors que l'attribut  $id\_expert$  n'est pas une clé.

On peut **toujours** se ramener à des relations normalisées.

# La bonne solution

Voici le schéma de nos manuscrits, **normalisé**

- Manuscrit (id\_manuscrit, auteur, titre, id\_expert, commentaire)
- Expert (id\_expert, nom)

Elles sont toutes deux en troisième forme normale (3FN) **et** on n'a pas perdu d'information

**Important** : l'attribut id\_expert dans Manuscrit est une **clé étrangère**.

# Pour n'a-t-on pas perdu d'information ?

Parce qu'on peut reconstituer **exactement** la relation complète

- Connaissant `id_expert`, je connais toutes les autres informations sur l'expert.
- Connaissant `id_manuscrit`, je connais `id_expert`
- Par **transitivité**, connaissant `id_manuscrit`, je connais `id_expert`, et connaissant `id_expert`, je connais toutes les autres informations sur l'expert

L'association de la clé primaire et de la clé étrangère est le mécanisme de liaison des nuplets en relationnel.

## Avec un exemple

id_expert	nom	adresse
1	Sophie	rue Montorgueil, Paris
2	Philippe	rue des Martyrs, Paris

id_manuscrit	nom	titre	id_expert	Commentaire
10	Serge	L'arpète	2	Une réussite
20	Cécile	Un art du chant grégorien sous Louis XIV	1	Bravo

La valeur d'une clé étrangère est **toujours** celle d'une clé de la table référencée.

# Schéma normalisé et contraintes

On normalise les schémas, et on demande au système de garantir deux contraintes.

## Contrainte d'unicité

Une valeur de clé ne peut apparaître **qu'une fois** dans une relation.

## Contrainte d'intégrité référentielle

La valeur d'une clé étrangère doit **toujours** être également une des valeurs de la clé référencée.

On obtient des schémas sans anomalie.

# À retenir

Un schéma de BD correct :

- Toute relation a une clé
- Tous les attributs dépendent **directement** de la clé (forme normale)
- On peut lier les nuplets les uns aux autres avec des clés étrangères.
- Le système évite les anomalies en garantissant la **contrainte d'unicité** (clé) et la **contrainte d'intégrité référentielle** (clé étrangère)

Comment trouve-t-on le bon schéma ? Nous y reviendrons.