

Chapitre4 : Démarche de détermination d'une base de données

Introduction

La création d'une base de données doit être précédée par un travail de réflexion consistant à trouver la meilleure façon selon laquelle les différents sujets du monde réel doivent être traduits en structures de bases de données (tables, colonnes, clés primaires,...).

Ce travail de réflexion s'appelle la démarche pour "déterminer la structure" d'une base de données composée des étapes suivantes:

1) Délimiter le(s) domaine(s)

Cette première étape consiste à identifier le ou les domaines (à partir du monde réel) qui constituent l'objet de la base de données. Est-ce qu'il s'agit par exemple de gérer les données d'une bibliothèque publique, d'une agence de voyage, d'un lycée, etc.?

Une fois les domaines identifiés, on procède à une collecte d'informations qui feront l'objet de notre base de données à travers des entretiens avec les personnes concernées et/ou un recueil des documents décrivant l'existant (fiches, fichiers, feuilles de calcul, ...)

2) Déterminer les colonnes

Il s'agit de déterminer la liste des colonnes en respectant les règles suivantes :

- Une colonne doit représenter une information sous sa forme élémentaire, c'est-à-dire que pour la manipuler, on ne doit pas être obligé de la décomposer
 - Exemple : Le nom d'un élève doit être représenté par deux colonnes (nom et prénom).
- Une colonne ne doit pas correspondre à une information calculée ou dérivée à partir d'autres colonnes.
 - Exemple : On ne peut pas avoir une colonne « moyenne » qui est le résultat de deux colonnes « note » et « coefficient ».
- Des colonnes similaires ne doivent pas être présentes. Il faut garder une seule colonne.
 - Exemple : Si nous avons trois colonnes « note1 », « note2 » et « note3 », on doit les remplacer par une seule colonne « note ».
- Ne pas omettre (oublier) de colonnes. L'oubli d'une colonne entraîne un manque d'information dans la BD.
 - Exemple : Si nous omettons la colonne « coefficient » pour une matière, il sera impossible de calculer la note finale.

Pour élaborer la liste des colonnes, on peut utiliser un tableau ayant la structure suivante :

Liste des colonnes							
Nom	Description	Type de données	Taille	Obligation	Valeur par défaut	Valeurs autorisées	Sujet

3) Déterminer les tables

Il s'agit de déterminer la liste des tables qui vont constituer la base de données.

Cette étape est relativement facile car la liste des tables sera déduite à partir du tableau de colonnes : Chaque table correspond à un sujet de domaine étudié.

Cette liste peut être représentée sous forme du tableau suivant :

Liste des tables		
Nom table	Description	Sujet

4) Affecter les colonnes aux tables

Il s'agit d'une étape importante dans le processus de détermination de la structure de la base de données et qui décidera de la qualité de cette dernière.

Une colonne doit être affectée à une et une seule table sauf dans le cas d'une clé étrangère qui fait le lien entre table fille et table mère.

Une colonne qui n'est pas affectée à la bonne table peut entraîner les inconvénients suivants :

- Une perte d'espace dû à la redondance d'information.
- Un risque d'incohérence entre les données engendrées par la modification de cette information dans une ligne et pas dans les autres.
- Dégradation des temps de réponses dû à l'augmentation du volume de données.

• Exemple :

Placer la colonne **adresseClient** dans la table commande engendra la répétition de sa valeur 100 fois pour un client qui a passé 100 commandes.

5) Déterminer les clés primaires

- Chaque table doit avoir une seule clé primaire.
- La clé primaire permet d'identifier les autres colonnes de la table de façon unique.
- Une clé primaire peut être composée d'une seule colonne ou de plusieurs colonnes. On parle alors de clé composée.
- Dans le cas où aucune des colonnes de la table ne peut être utilisée comme une clé primaire, nous devons rajouter une nouvelle colonne et la prendre comme clé primaire.

6) Déterminer les liens entre les tables

Pour déterminer le lien entre deux tables A et B on doit poser deux questions :

- Pour une occurrence de A combien on a d'occurrence de B ?
- Pour une occurrence de B combien on a d'occurrence de A ?

Après avoir déterminé la table mère et la table fille on doit rajouter à la table fille une colonne qui correspond à la clé primaire de la table mère. Cette colonne rajoutée sera une clé étrangère.

- **Exemple :**

"Une commande est relative à un client" → cette phrase peut nous informer qu'il y a un lien entre la table commande et la table client. Un client peut avoir plusieurs commandes par contre une commande est relative à un et un seul client. D'où il faut ajouter la colonne "codeClient" qui est la clé primaire de la table Client comme clé étrangère dans la table Commande.

7) Analyser et affiner la structure de la base de données

Il s'agit de détecter les anomalies (oubli de colonnes, oubli de liens entre tables, mauvaise affectation de colonnes, redondance de colonnes, etc.) et les corriger (déplacer une colonne d'une table à une autre, créer une nouvelle table, éclater une colonne en plusieurs colonnes, etc.)

8) Applications

Voir pages : 72 → 77