

DI

DEVELOPPEUR INFORMATIQUE

Modèle Conceptuel des données

Démarche de construction

Dossier d'informations

<u>Module</u>	1
<u>Séquence</u>	4
<u>Capacité</u>	3

L'objectif de ce dossier est d'illustrer, par l'exemple, la démarche de construction d'un schéma conceptuel de données au travers de 3 étapes :

1. L'établissement du dictionnaire des données à partir de l'analyse du système d'informations
2. La structuration des données par l'étude des dépendances fonctionnelles
3. La construction du schéma conceptuel et sa validation.

1-L'établissement du dictionnaire des données à partir de l'analyse du système d'informations

La première étape, dans la construction d'une base de données, consiste à réunir toutes les données que l'on souhaite modéliser en particulier à partir des documents représentatifs de ces données.

Le terme "documents" doit être ici pris dans un sens large. Il peut s'agir en effet soit de documents papier, soit de documents magnétiques (enregistrement d'entretiens, disque ou disquette informatique)... soit de tout autre support utilisable pour conserver des informations.

Pour illustrer les différentes étapes de conception d'une base de données nous utiliserons, comme exemple type, la gestion d'étudiants qui suivent les différents enseignements d'un diplôme.

Les documents de travail utilisés sont les suivants :

Liste des étudiants			
N° étudiant	Nom de l'étudiant	Date de naissance	Sexe
1	Dupont, Charles	18-03-1981	H (Homme)
2	Dubois, Jules	02-11-1982	H (Homme)
3	Favier, Isabelle	02-02-1979	F (Femme)
...	...	...	...

Relevé de notes			
N° étudiant : 1 Dupont, Charles			
N° matière	Nom	Coefficient	Note / 20
1	Mathématiques	3	10
2	Informatique	2	9
3	Sociologie	2	12,5
4	Histoire	1	13
5	Géographie	1	7
Moyenne Générale			10,3

Liste des enseignants				
N° enseignant	Nom	Grade	Ancienneté	Matière enseignée
10	Bertrand, Pierre	ASS	2	Sociologie
11	Dupont, Auguste	MCF	3	Mathématiques
...	...	...	...	...
15	Simon, Etienne	ASS	5	Histoire Géographie

Règles de gestion :

- Une matière est enseignée que par plusieurs enseignants.
- Un enseignant peut enseigner plusieurs matières
- Un étudiant n'a qu'une note par matière.

Dictionnaire des données

Partant des documents, la phase d'analyse consiste à en extraire les informations élémentaires (non décomposables) qui vont constituer la future base de données.

La réunion de l'ensemble des données élémentaires, que l'on appelle des propriétés ou des attributs, constitue le dictionnaire des données. Chaque attribut du dictionnaire de données peut être caractérisé par les propriétés suivantes :

Propriété	Signification
Mnémonique	Abréviation du nom de l'attribut.
Libellé	Libellé contenant la signification précise et le rôle de l'attribut.
Type de donnée	Type de l'attribut : entier, réel, chaîne de caractères, date...
Contraintes d'intégrité	Liste des contraintes sur les valeurs possibles de l'attribut
Règle de calcul	Règle de calcul (d'obtention) de l'attribut correspondant.

Le dictionnaire des données relatif aux documents précédents est le suivant :

Mnémonique	Libellé	Type	Contraintes
Ancien	Ancienneté dans le grade	Entier	≥ 0
Cdsexe	Code sexe	Chaîne(1)	H ou F
Coeff	Coefficient de la matière	Entier	≥ 0 et < 6
Dtnaiss	Date de naissance	Date	
Grade	Grade de l'enseignant	Chaîne(3)	ASS ou MCF ou PR
Lbsexe	Libellé du sexe	Chaîne(7)	Homme ou Femme
Moyenne	Moyenne au diplôme	Réel	≥ 0 et ≤ 20
Nomat	Nom de la matière	Chaîne(15)	
Nomens	Nom de l'enseignant	Chaîne(20)	
Nometu	Nom de l'étudiant	Chaîne(20)	
Note	Note obtenue à la matière	Réel	≥ 0 et ≤ 20
Numat	N° de la matière	Entier	≥ 0
Numens	N° de l'enseignant	Entier	≥ 0
Numetu	N° de l'étudiant	Entier	≥ 0

2-La structuration des données par l'étude des dépendances fonctionnelles

Les informations contenues dans le dictionnaire de données doivent être structurées en "paquets" homogènes (entités, tables) dans lesquels la répétition (redondance) d'informations doit être minimale.

Les attributs qui sont calculés (qui ont une règle de calcul) *ne doivent pas* être pris en compte dans la phase de structuration des données. Ils sont en effet obtenus à partir d'autres attributs, et les inclure dans la phase de structuration introduirait une forme de redondance et donc des possibilités d'incohérence. Dans notre exemple, l'attribut "Moyenne" appartient à cette catégorie.

On dira qu'un ensemble d'attributs est homogène si tous les attributs qui le composent ont un lien "direct". La redondance à l'intérieur d'un ensemble doit être minimale car la redondance est source d'incohérence. En effet, si la même information est répétée plusieurs fois, cela signifie :

- Qu'elle devra être saisie plusieurs fois, ce qui multiplie les risques d'erreurs de frappe.
- Que lorsqu'elle devra être mise à jour, les modifications devront intervenir à plusieurs endroits d'où des risques de mise à jour partielle.

Pour regrouper les attributs du dictionnaire de données on utilise un élément structurant qui s'appelle la dépendance fonctionnelle (DF).

Notion de dépendance fonctionnelle (DF)

On dit qu'il existe une dépendance fonctionnelle entre un attribut A1 et un attribut A2, on note $A1 \rightarrow A2$, si connaissant une valeur de A1 on ne peut lui associer qu'une seule valeur de A2. On dit aussi que A1 détermine A2. A1 est la source de la dépendance fonctionnelle et A2 le but.

Exemple de DF

Dans notre exemple, connaissant un numéro d'étudiant, on connaît de manière unique le nom de cet étudiant et, entre autres, sa date de naissance. On a donc :

Numetu $\rightarrow$ Nometu, Dtnaiss.

En revanche, l'expression : Numetu $\rightarrow$ Note est fausse.

En effet, connaissant un numéro d'étudiant on peut connaître les notes qu'il a obtenues dans chaque matière. La connaissance du numéro d'étudiant ne permet donc pas de connaître une note particulière et il n'y a donc pas de DF.

DF simple :

On dira qu'une DF est simple si sa source n'est composée que d'un seul attribut. Une DF simple caractérise une *entité (table)* dont la source est la clé et dont les propriétés sont constituées par le but de la DF.

DF composée :

On dira qu'une DF est composée si sa source est composée par la réunion de plusieurs attributs.

Une DF composée caractérise une *association entre entités (tables)* dont la source est la clé et dont les propriétés sont constituées par le but de la DF. Il ne doit pas y avoir d'attributs superflus dans la source d'une DF composée. Par exemple si nous avons les 2 DF composées :

1- A1, A2, A3 → A4

2- A1, A2 → A4

La DF numéro 1 est incorrecte puisque A3 est superflu. La DF composée "correcte" est :
A1, A2 → A4

Pour représenter les entités et associations nous utiliserons la notation suivante :

NOM_DE_L'ENTITE (Source, But ...). La clé est soulignée.

Soit par exemple :

DF	Entités
A1 → A2, A3, A4	ENTITE1 (<u>A1</u> , A2, A3, A4)
A3 → A4	ENTITE2 (<u>A3</u> , A4)
A1, A5 → A6	ASSOC1 (<u>A1</u> , <u>A5</u> , A6)

Démarche de structuration des données

La structuration des données du dictionnaire des données s'effectue en 5 étapes :

1. Détermination de la liste des DF simples
2. Prise en compte des attributs non classés dans l'étape 1 et détermination des DF composées
3. Construction des entités à partir des DF simples
4. Construction des associations porteuses de données à partir des DF composées
5. Construction des associations non porteuses de données qui ne sont pas déduites directement du dictionnaire des données mais de l'observation du système d'informations

L'application de cette démarche à notre exemple nous donne les résultats suivants :

1-DF simples de la base de données ETUDIANTS

Numetu → Nometu, Dtnaiss,

Cdsexe → Lbsexe

Numens → Nomens, Grade, Ancien

Numat → Nomat, Coeff

2-Attributs non classés et DF composées de la base de données ETUDIANTS

L'attribut "note" n'a pas été classé dans l'étape précédente, il intervient donc dans une DF composée. Cette DF est la suivante :

Numetu, Numat -> Note

3-Construction des entités

ETUDIANT (Numetu, Nometu, Dtnaiss)

SEXE (Cdsexe, Lbsexe)

ENSEIGNANT (Numens, Nomens, Grade, Ancien)

MATIERE (Numat, Nomat, Coeff)

4-Construction des associations

NOTES (Numetu, Numat, Note)

5-Construction des associations non porteuses de données

On a vu précédemment que les dépendances fonctionnelles composées sont en fait des associations entre entités. La structuration des données inclut également des associations sans attributs rattachés.

Dans notre exemple, il est nécessaire de connaître :

- Le sexe de chaque étudiant
- Les matières enseignées par un enseignant

Nous rajoutons donc 2 associations :

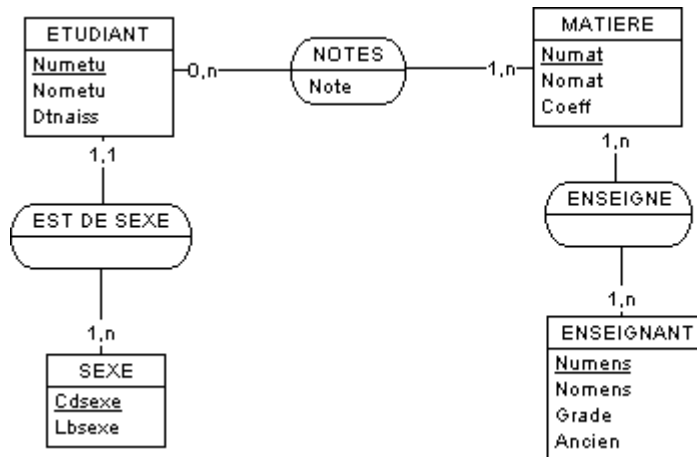
EST DE SEXE(Numetu, Cdsexe)

ENSEIGNE(Numens, Numat)

3-La construction du schéma conceptuel et sa validation.

La dernière étape consiste à schématiser, par le modèle entité-association, les éléments recensés dans l'étape précédente et à affecter les cardinalités pour chaque entité, en tenant compte des règles de gestion listées tout au long de l'étude.

Représentation graphique du schéma conceptuel



Cette représentation est le schéma conceptuel des données du domaine "ETUDIANTS".
C'est la structure de représentation optimale des données de notre problème : Cohérence maximale et redondance minimale.