

BDD tuto 1: Modélisation

Diapo1

Dans la vidéo précédente, j'ai montré l'intérêt d'utiliser des systèmes sur plusieurs tables. Pour que ce type de système fonctionne, il faut analyser la situation et aboutir à un schéma de la modélisation. C'est cette étape essentielle que je vais tenter de décrire ici.

Diapo2

Pour que l'analyse soit complètes, il faut passer quatre étapes.

**** Combien de tables sont nécessaires ?**

Il faut recueillir un maximum d'information sur la situation : quelle est la fonction des personnes qui vont accéder aux données ? De quelles informations auront-ils besoin, et sous quelles formes ? Le mieux est de tenter d'obtenir un échantillon des données qui existent peut-être déjà. Si ce n'est pas possible il faut faire une mise en situation en inventant des données qui correspondent à la situation.

Après un premier classement, on détermine les tables, et les attributs de chaque objet que ces données doivent décrire. Il faudra créer au moins une table par type d'objet

Diapo3

Dans notre exemple du centre aéré, ça va aller vite, on a déjà fait l'inventaire dans les vidéos précédentes des personnes qui allaient utiliser la base, et des tableaux dont ils auraient besoin. Pour les objets, il y a des parents qui ont comme attributs nom, prénom, adresse ...etc, tandis que les enfants ont nom, prénom et date de naissance.

Comme les parents et les enfants n'ont pas les mêmes attributs, il faut au moins deux tables: Parent et Enfant.

Diapo4

**** retrouver les relations un à plusieurs**

Précisons avant ça ce qu'est une relation.

Les relations associent les tables deux à deux. On peut trouver des relations de trois types:

La relation un à un: à un élément de la table A, il ne peut y avoir qu'un élément de la table B qui correspond, et réciproquement

La relation un à plusieurs: à un élément de la table A, il peut y avoir une ou plusieurs éléments de la table B

La relation plusieurs à plusieurs: à plusieurs éléments de table A, il peut y avoir plusieurs éléments de la table B

Diapo5

Lorsqu'on soupçonne l'existence d'une relation un à un, cela indique dans la plupart des cas qu'une des tables est inutile. En effet, comme c'est une relation un à un, cela signifie qu'il y a autant d'élément dans les deux tables. On peut très bien enregistrer nos informations dans la première table

et se passer de la deuxième. Si attribue des casiers à chaque enfant pour qu'ils rangent leurs affaires. On pourrait enregistrer cette information comme une caractéristique de l'enfant. On a donc pas besoin de table avec les casiers.

Diapo6

Au même titre, quand on soupçonne l'existence d'une relation de plusieurs à plusieurs entre deux tables, c'est qu'il y a forcément une table entre les deux, avec deux relations un à plusieurs. En fait, techniquement, la relation plusieurs à plusieurs n'est pas réalisable.

On est contraint de trouver cette table intermédiaire parmi les tables qu'on a inventoriées, ou en ajouter une.

Diapo7

Dans notre exemple, on associe un enfant systématiquement à un seul parent, ce sont les données de départ. On peut déjà écarter la relation plusieurs à plusieurs.

Diapo8

Ça c'était dans le sens enfant-parent, mais dans l'autre sens: est-ce qu'on associe un parent systématiquement à un seul enfant? La réponse est non, donc ce n'est pas une relation un à un.

On a bien une relation de un à plusieurs.

Diapo9

**Ajout des colonnes clés

Concrètement, la relation entre deux tables se concrétise par un lien d'égalité entre deux colonnes de ces tables, qu'on appelle des colonnes "clé". Une des colonnes clé correspond forcément à la clé primaire d'une des tables.

L'autre colonne, qu'on appelle clé "étrangère", va contenir des numéros qu'on pourra trouver aussi dans la clé primaire, (c'est des matricules.). Contrairement à la clé primaire, la clé étrangère pourra contenir des doublons

Dans notre exemple, à un parent, on associe un ou plusieurs enfants(s), le lien est donc entre la clé primaire de la table parent, et une autre colonne qu'il faut ajouter à la table enfant, qui contient les matricule des parents.

Diapo10

Pour bien comprendre, suivons le cheminement que devra prendre le logiciel. Imaginons une requête très simple comme "quel est le nom du parent de kévín", le logiciel devra d'abord trouver la ligne concernant kévín dans la table enfant, puis aller dans la colonne du matricule, lire la valeur et rechercher cette valeur dans les matricules des parents.

Une fois trouvée, il lui faudra aller dans la colonne du nom pour lire la valeur et envoyer le résultat. On va bien avec cette animation que le logiciel à besoin de savoir quelles sont les colonnes clés de la relation pour pouvoir appliquer le principe d'égalité, qu'on appelle aussi l'intégrité référentielle.

Diapo11

Pour aider à l'analyse, et également pour implanter notre base de donnée dans le logiciel, il faut

faire un schéma.

Pour cela, il faut tout d'abord faire une ligne par colonne de la table, comme ceci.

Par la suite, il faut placer les types de données qui sont attendus dans chaque colonne dans une autre colonne du schéma.

Le matricule, c'est un nombre écrit automatiquement par le logiciel, les noms et prénom sont du texte. On peut préciser que le texte ne dépassera pas une certaine taille. L'adresse c'est du texte un peu plus long, le code postal est un nombre entier, les montants sont des nombres décimaux...

Pour la table des enfants: matricule de l'enfant: nombre auto, les nom prénom sont du texte avec la même longueur maxi que pour les parents (c'est logique), la date de naissance, date, le matricule du parent: nombre entier. Et oui: le contenu de la colonne matricule du parent sera forcément rempli de numéro de matricule de parents (qui ne seront pas automatique).

Ensuite, il faut placer un symbole de clé en face des clés primaires, donc ici c'est la colonne matricule.

Pour matérialiser la relation qu'il y a entre ces tables, je trace un trait entre les deux colonnes clé et je précise le sens de la relation un à plusieurs: 1 matricule de parent apparaît une fois ici, tandis qu'il peut apparaître plusieurs fois ici. L'infini veut dire plusieurs (parfois certain écrit un n)

Voilà le schéma est conçu. Cet exemple était volontairement très simplifié, malheureusement la réalité est souvent beaucoup plus complexe.

Diapo12

Dans les vidéos suivantes on verra comment ou pourrait compléter notre base de données de centre aéré, puis on abordera d'autres situations qui nécessitent plus de tables. À bientôt