

# Structurer l'information

Rappel : dans un système d'information de type base de données, l'information réside dans :

- des "objets" ou "entités" qui jouent un rôle distinct au sein du système d'information (*enregistrements* regroupés dans des *tables*), et les propriétés qui les caractérisent (*rubriques*).
- les relations qui s'établissent ou non entre ces objets (*liens*).

## Règles d'or : une table = un type d'objet, une rubrique = une propriété de l'objet

- Chaque table doit héberger les informations correspondant à un seul type d'"objet" ou "entités" ;
- Chaque rubrique d'une table doit correspondre à *une propriété unique de cet objet* (l'information correspondante n'est disponible que si l'objet existe + on ne peut saisir qu'une seule donnée dans la rubrique).

Le non-respect de ce principe entraîne des effets pervers : redondance, risque de perte d'informations, risque d'incohérence des données, etc.

Il faut donc arriver, lorsque l'on établit le *dictionnaire des données*, à cerner les contours des objets qui entrent en jeu dans le système d'information : en résumé, cela consiste à regrouper les informations qui caractérisent un même objet, en se demandant *de quoi dépend chaque donnée prise en compte*.

La difficulté : il arrive souvent que ce que l'on perçoit spontanément comme "une propriété d'un objet" recouvre en réalité deux objets mis en relation. Par exemple :

### Exemple 1 : dictionnaire du cinéma

La notion de générique recouvre en fait des films et des individus mis en relation (La personne A intervient dans le film B en tant que C...)

### Exemple 2 : scolarité à l'Université

La notion d'inscription pédagogique recouvre en fait des étudiants (Objet A) et des EP (Objet B) qui sont mis en relation (A est inscrit à B).

Autrement dit : même des informations que l'on considérerait dans la vie de tous les jours comme faisant partie des "propriétés" naturelles d'un objet, devront, dans une base de données, être considérées comme des objets distincts et être placées dans une table à part.

## Comment voir s'il faut une ou plusieurs tables ?

Considérer que l'on a deux objets autonomes, et qu'il faut donc les mettre dans des tables distinctes, dès que :

- *Le même objet A peut être en relation avec plusieurs objets B différents ou vice versa* : car si l'on réunissait A et B en une seule table, il y aurait redondance des informations concernant l'objet A dans toutes les fiches B qui le concernent, avec à la clé d'inévitables erreurs de saisie et un encombrement inutile.

La question à se poser :

**L'objet A peut-il correspondre à plusieurs objets B ou vice versa ?**

Un acteur joue dans plusieurs films / un film à plusieurs acteurs à son générique.

Un étudiant est inscrit dans plusieurs EP / un EP accueille plusieurs étudiants

- *On veut pouvoir avoir des infos sur A, même s'il ne correspond à aucun B* : dans ce cas A doit être représenté dans une table à part, sinon il y a risque de perte d'info en cas de suppression (ou de non saisie) de B.

La question à se poser :

**Est-ce que A doit pouvoir exister indépendamment de B (A est un objet autonome) ou bien A n'est-il pertinent que si B existe (A n'est qu'une propriété de B) ?**

Un film peut ne pas avoir d'acteurs (dessins animés, documentaires...), il faut quand même pouvoir mémoriser et retrouver les infos qui le concernent. Toutes les infos concernant le film lui-même (titre, genre, etc.) constituent donc un objet indépendant.

(On a besoin de 2 tables : films / acteurs)

En revanche, il ne peut y avoir d'information "Date de sortie du film" sans film : "Date de sortie" est donc une pure propriété du film.

Un EP peut ne pas avoir d'inscrits à un moment de l'année (avant la rentrée par exemple), il faut pourtant pouvoir accéder aux infos qui le concernent (code, libellé, règlement de scolarité...). Ces informations constituent donc un objet indépendant (EP).

(On a besoin de 2 tables : EP / étudiants)

En revanche, il ne peut y avoir d'information "libellé EP" sans EP et cette information n'existe que si l'EP existe : "libellé EP" est donc une simple propriété de l'EP.

## ■ Comment voir à quoi rattacher une propriété ?

Il n'est pas toujours évident de voir de quel objet dépend une propriété. Si cela ne pose pas de problème pour des informations du genre *Prénom*, *Âge* etc., certaines données sont plus délicates à cerner parce qu'elles ne sont pas propres à un objet mais bien à l'association même de deux objets distincts :

Cas spécifique : la donnée considérée est une propriété non pérenne, c'est-à-dire une information qui caractérise non pas un objet mais un lien entre deux objets.

La question à se poser :

**Est-ce qu'il suffit que l'objet A existe pour que la propriété P existe ? Ou bien P n'a-t-elle de pertinence que si l'objets A est en relation avec B ?**

L'information "rôle interprété" (P) n'est propre ni à un film (le même rôle se retrouve dans plusieurs films en cas de "remake"), ni à un acteur (un acteur interprète plusieurs rôles au cours de sa carrière) :

Cette information n'est pertinente que dans le cadre de l'association d'un acteur et d'un film (L'acteur A joue le rôle P dans le film B).

Dans ce cas, si les deux objets sont en relation 1-n (ou n-n) la rubrique devra être dans la table intermédiaire (celle qui enregistre les associations A-B

L'information "Validé/Échec" (P) n'est propre ni à l'EP (valeur différente selon les étudiants), ni à l'étudiant (valeur différente selon les EP qu'il présente).

Cette information n'a de sens que par l'association d'un étudiant et d'un EP. (L'étudiant A obtient le résultat P à l'EP C).