

LA MINUTE DU PSIR

Préparation des données structurées

La Minute du PSIR de la semaine passée, avait présenté les différents types de base de données (BdD) et les services spécialisés offerts par le PSIR en la matière. Cette semaine la Minute va aborder le sujet de la conception d'une base de données relationnelle comme MySQL ou PostgreSQL.

Une base de données relationnelle est une collection de données classées dans des tables structurées sous la forme de colonnes de variable/valeur. Cette organisation permet de produire facilement des extractions, des vues et d'ajouter des index pour améliorer les temps de réponse des requêtes d'interrogation. Les logiciels qui permettent de créer, utiliser et maintenir des bases de données relationnelles sont des systèmes de gestion de base de données relationnelle (SGBDR).

Pour la plupart, ils utilisent le langage SQL. **L'intérêt étant les possibilités d'interrogations croisées plus ou moins complexes des données.**

La conception d'une base de données relationnelle se déroule en 3 étapes : **la modélisation conceptuelle, la modélisation logique et l'implémentation physique.**

La modélisation conceptuelle permet de décrire le contenu de la base de données en termes conceptuels, indépendamment de toute considération informatique, par des représentations graphiques. Dans ce modèle, chaque type d'objet (les articles d'une revue par exemple) est représenté par un dessin graphique (appelé « Entité »), souvent un rectangle, avec son nom (« Article » pour notre exemple) et une liste d'informations que l'on veut stocker avec cet objet (par exemple : Titre, Résumé, Nom auteur, Prénom auteur, Date de publication, Nom éditeur, adresse éditeur etc.). A cette étape, le modèle n'est pas encore formel, donc certaines représentations peuvent être équivoques. **Pour que notre base soit plus efficace en termes de stockage, il faut éviter les données redondantes.** Par exemple, pour éviter de répéter les informations d'auteurs et d'éditeurs à chaque article, nous pouvons les stocker dans deux tables « Auteur » et « Éditeur » distinctes. Puis on crée des références dans la table « Articles » pour associer un article avec ses auteurs et son éditeur contenus dans les autres tables. **Le terme relationnel prend tout son sens ici.**

La modélisation logique est le résultat de la traduction du modèle conceptuel en un schéma propre et formel à un type de base de données, mais indépendamment des choix d'implémentation spécifiques. Pour la base de données relationnelle comme MySQL, les entités du modèle conceptuel sont transformées en tableaux à deux dimensions : les informations de chaque entité deviennent les colonnes de la table et chaque ligne de la table représente les données complètes d'un objet (article, auteur, éditeur, date, etc.). **L'ensemble des tables et de leurs colonnes ainsi obtenues devient la structure de la base de données.** Les liaisons entre les objets associés sont appelées « Relation » dans le modèle logique.

L'implémentation physique correspond aux choix techniques, en termes de SGBDR, et à leur mise en œuvre (programmation, optimisation...). Cette étape est réalisée généralement par la traduction directe du modèle logique à l'aide du logiciel spécifique au SGBDR choisi.

Vous pouvez utiliser le logiciel gratuit « MySQL Workbench » pour réaliser le modèle conceptuel et le transformer jusqu'à obtenir une vraie base de données relationnelle.

<https://dev.mysql.com/downloads/workbench/>

Pour aller plus loin :

<https://cyril-gruau.developpez.com/merise/>

NIVEAU DIFFICULTÉ



THEMATIQUE

Données

PERSONNES CONCERNÉES

