

Base de données

Une **base de données** permet de stocker et de retrouver des données structurées, semi-structurées ou des données brutes ou de l'information, souvent en rapport avec un thème ou une activité ; celles-ci peuvent être de natures différentes et plus ou moins reliées entre elles.

Leurs données peuvent être stockées sous une forme très structurée (base de données relationnelles par exemple), ou bien sous la forme de données brutes peu structurées (avec les bases de données NoSQL par exemple). Une base de données peut être localisée dans un même lieu et sur un même support informatisé, ou répartie sur plusieurs machines à plusieurs endroits.

La base de données est au centre des dispositifs informatiques de collecte, mise en forme, stockage et utilisation d'informations. Le dispositif comporte un système de gestion de base de données (abréviation : SGBD) : un logiciel moteur qui manipule la base de données et dirige l'accès à son contenu. De tels dispositifs comportent également des logiciels applicatifs, et un ensemble de règles relatives à l'accès et l'utilisation des informations¹.

La manipulation de données est une des utilisations les plus courantes des ordinateurs. Les bases de données sont par exemple utilisées dans les secteurs de la finance, des assurances, des écoles, de l'épidémiologie, de l'administration publique (notamment les statistiques) et des médias.

Lorsque plusieurs objets nommés « bases de données » sont constitués sous forme de collection, on parle alors d'une **banque de données**.

Description

Une base de données est un « conteneur » stockant des *données*¹ telles que des chiffres, des dates ou des mots, pouvant être retraités par des moyens informatiques pour produire une information ; par exemple, des chiffres et des noms assemblés et triés pour former un annuaire. Les retraitements sont typiquement une combinaison d'opérations de recherches, de choix, de tri, de regroupement, et de concaténation².

C'est la pièce centrale d'un système d'information ou d'un *système de base de données* (ou *base de données* tout court), qui régit la collecte, le stockage, le retraitement et l'utilisation de données. Ce dispositif comporte souvent un logiciel moteur (cf. paragraphe suivant), des logiciels applicatifs, et un ensemble de règles relatives à l'accès et l'utilisation des informations¹.

Le système de gestion de base de données est une suite de programmes qui manipule la structure de la base de données et dirige l'accès aux données qui y sont stockées. Une base de données est composée d'une collection de fichiers ; on y accède par le SGBD qui reçoit des demandes de manipulation du contenu et effectue les opérations nécessaires sur les fichiers. Il cache la complexité des opérations et offre une vue synthétique sur le contenu. Le SGBD permet à plusieurs usagers de manipuler simultanément le contenu, et peut offrir différentes vues sur un même ensemble de données¹.

Le recours aux bases de données est une alternative au procédé classique de stockage de données, par lequel une application place des données dans des fichiers manipulés par l'application. Il facilite le partage des informations, permet le contrôle automatique de la cohérence et de la redondance des informations, la limitation de l'accès aux informations et la production plus aisée des informations synthétiques à partir des renseignements bruts. La base de données a de plus un effet fédérateur : dans une collectivité utilisant une base de données, une personne unique — l'administrateur de bases de données — organise le contenu de la base d'une manière bénéfique à l'ensemble de la collectivité, ce qui peut éviter des conflits dus à des intérêts divergents entre les membres de la collectivité³.

Une base de données nécessite généralement plus d'espace disque, le large éventail de fonctions offertes par les SGBD rend les manipulations plus complexes, et les pannes ont un impact plus large et sont plus difficiles à rattraper³.

Terminologie

modèle de données

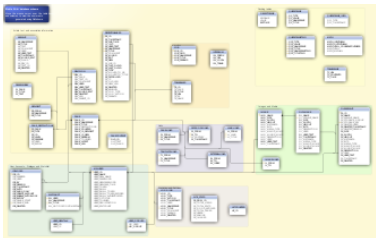
Le *schéma* ou *modèle de données*, est la description de l'organisation des données. Il se trouve à l'intérieur de la base de données, et renseigne sur les caractéristiques de chaque type de donnée et les relations entre les différentes données qui se trouvent dans la base de données. Il existe plusieurs types de modèles de données (relationnel, entité-association, objet, hiérarchique et réseau)⁴.

modèle de données *logique* et *physique*

Le modèle de données *logique* — ou *conceptuel* — est la description des données telles qu'elles sont dans la pratique, tandis que le modèle de données *physique* est un modèle dérivé du modèle logique qui décrit comment les données seront techniquement stockées dans la base de données⁴.

entité

Une entité est un sujet, une notion en rapport avec le domaine d'activité pour lequel la base de données est



Modèle de données relationnel.

utilisée, et concernant lequel des données sont enregistrées (exemple : des personnes, des produits, des commandes, des réservations...)^{3,5}.

attribut

Un attribut est une caractéristique d'une entité susceptible d'être enregistrée dans la base de données. Par exemple, une personne (entité), son nom et son adresse (des attributs). Les attributs sont également appelés des *champs* ou des *colonnes*³. Dans le schéma les entités sont décrites comme un lot d'attributs en rapport avec un sujet⁵.

enregistrement

Un enregistrement est une donnée composite qui comporte plusieurs champs dans chacun desquels est enregistrée une donnée. Cette notion a été introduite par le stockage dans des fichiers dans les années 1960⁵.

association

Les *associations* désignent les liens qui existent entre différentes entités, par exemple, entre un vendeur, un client et un magasin³.

cardinalité

La *cardinalité* d'une association — d'un lien entre deux entités A et B — est le nombre de A pour lesquelles il existe un B et inversement. Celle-ci peut être un-à-un, un-à-plusieurs ou plusieurs-à-plusieurs. Par exemple, un compte bancaire appartient à un seul client, et un client peut avoir plusieurs comptes bancaires (cardinalité un-à-plusieurs)⁴.

modèle de données relationnel

C'est le type de modèle de données le plus couramment utilisé pour la réalisation d'une base de données. Selon ce type de modèle, la base de données est composée d'un ensemble de tables (les *relations*) dans lesquelles sont placées les données ainsi que les liens. Chaque ligne d'une table est un *enregistrement*. Ces modèles sont simples à mettre en œuvre, fondés sur les mathématiques (la théorie des ensembles), ils sont très populaires et fortement normalisés⁴.

base de données relationnelle

Base de données organisée selon un modèle de données de type *relationnel*, à l'aide d'un SGBD permettant ce type de modèle.

modèle de données entité-association

Ce type de modèle est le plus couramment utilisé pour la conception de modèles de données *logiques*⁶. Selon ce type de modèle, une base de données est un lot d'entités et d'associations. Une *entité* est un sujet concret, un objet, une idée, pour laquelle il existe des informations. Un *attribut* est un renseignement concernant ce sujet — exemple le nom d'une personne. À chaque attribut correspond un *domaine* : un ensemble de valeurs possibles. Une *association* désigne un lien entre deux entités — par exemple, un élève et une école⁴.

modèle de données objet

Ce type de modèle est fondé sur la notion d'*objet* de la programmation orientée objet. Selon ce type de modèle, une base de données est un lot d'*objets* de différentes *classes*. Chaque objet possède des *propriétés* — des caractéristiques propres —, et des *méthodes* qui sont des opérations en rapport avec l'objet. Une *classe* est une catégorie d'objets et reflète typiquement un sujet concret⁴.

modèle de données hiérarchique

Ce type de modèle de données a été créé dans les années 1960 ; c'est le plus ancien modèle de données. Selon ce type de modèle, les informations sont groupées dans des *enregistrements*, chaque enregistrement comporte des *champs*. Les enregistrements sont reliés entre eux de manière hiérarchique : à chaque enregistrement correspond un enregistrement parent⁴.

modèle de données réseau

Ce type de modèle de données est semblable au modèle hiérarchique. Les informations sont groupées dans des *enregistrements*, chaque enregistrement possède des *champs*. Les enregistrements sont reliés entre eux par des pointeurs. Contrairement aux modèles hiérarchiques, l'organisation des liens n'est pas obligatoirement hiérarchique, ce qui rend ces modèles plus polyvalents⁴.

nul

Dans les modèles de données *relationnels*, un attribut peut avoir une valeur *nulle*, indiquant que la donnée est absente, non disponible ou inapplicable⁵.

clé primaire

Dans les modèles de données *relationnels*, la *clé primaire* est un attribut dont le contenu est différent pour chaque enregistrement de la table, ce qui permet de retrouver un et un seul enregistrement⁴.

clé étrangère

Dans les modèles de données *relationnels*, une *clé étrangère* est un attribut qui contient une référence à une donnée connexe — dans les faits la valeur de la clé primaire de la donnée connexe⁴.

intégrité référentielle

Dans les modèles de données *relationnels*, il y a situation d'*intégrité référentielle* lorsque toutes les données référencées par les *clés étrangères* sont présentes dans la base de données⁵.

bancarisation

La bancarisation de données est l'opération consistant à réunir des données dans une banque de donnée. Dans la plupart des pays, elle est encadrée par la loi, au moins dans le cas de données environnementales,

Activity Code	Activity Name
23	Patching
24	Overlay
25	Crack Sealing

Key = 24

Activity Code	Date	Route No.
24	01/12/01	I-95
24	02/08/01	I-66

Diagram illustrating a primary key (Activity Code) and a foreign key (Activity Code) relationship between two tables. The first table has columns Activity Code and Activity Name. The second table has columns Activity Code, Date, and Route No. An arrow points from the Activity Code column in the second table to the Activity Code column in the first table, labeled 'Key = 24'.

Date	Activity Code	Route No.
01/12/01	24	I-95
01/15/01	23	I-495
02/08/01	24	I-66

Clé primaire et clé étrangère.

publiques ou personnelles.

Typologie

L'usage qui est fait des données diffère d'une base de données à l'autre. Les bases de données peuvent être classifiées en fonction du nombre d'utilisateurs, du type de contenu, notamment s'il est faiblement ou fortement structuré, ainsi que selon l'usage qui est fait de la base de données, notamment l'utilisation *opérationnelle* ou à des fins d'*analyse* :

Les bases de données peuvent être classifiées en fonction du nombre d'utilisateurs — un seul, un petit groupe, voire une entreprise. Une base de données *de bureau* est installée sur un ordinateur personnel au service d'un seul utilisateur. Tandis qu'une base de données d'*entreprise* est installée sur un ordinateur puissant au service de centaines d'utilisateurs. Une base de données *centralisée* est installée dans un emplacement unique, tandis qu'une « base de données distribuée » est répartie entre plusieurs emplacements¹.

La manière la plus populaire de classer les bases de données est selon l'usage qui en est fait, et l'aspect temporel du contenu :

bases *opérationnelles* ou **OLTP** (de l'anglais *online transaction processing*)

sont destinées à assister des utilisateurs à tenir l'état d'activités quotidiennes¹. Elles permettent en particulier de stocker sur le champ les informations relatives à chaque opération effectuée dans le cadre de l'activité : achats, ventes, réservations, paiements. Dans de telles applications l'accent est mis sur la vitesse de réponse et la capacité de traiter plusieurs opérations simultanément².

bases d'*analyse* dites aussi **OLAP** (de l'anglais *online analytical processing*)

sont composées d'informations historiques telles que des mesures sur lesquelles sont effectuées des opérations massives en vue d'obtenir des statistiques et des prévisions. Les bases de données sont souvent des entrepôts de données¹ (anglais *datawarehouse*) : des bases de données utilisées pour collecter des énormes quantités de données historiques de manière quotidienne depuis une base de données opérationnelle. Le contenu de la base de données est utilisé pour effectuer des analyses d'évolution temporelle et des statistiques telles que celles utilisées en management. Dans de telles applications l'accent est mis sur la capacité d'effectuer des traitements très complexes et le logiciel moteur (le SGBD) est essentiellement un moteur d'*analyse*².

Les bases de données sont également parfois classifiées selon les caractéristiques du contenu :

- des données *non structurées* sont stockées à l'état brut, et nécessitent d'être retraitées en vue de produire de l'information/de la connaissance ;
- des données *structurées* sont formatées en fonction de l'usage qui va en être fait, en vue de faciliter le stockage, l'utilisation et la production d'information finie. Par exemple, un ensemble de factures peuvent être stockées brutes sous forme d'images bitmap, ce qui ne permettra pas de calculer des totaux et des moyennes, ou alors chaque facture peut être décomposée sous forme d'un tableau de chiffres sur lesquels il est alors possible d'effectuer des calculs¹.

Histoire

Les disques durs, mémoire de masse de grande capacité, ont été inventés en 1956. L'invention du disque dur a permis d'utiliser les ordinateurs pour collecter, classer et stocker de grandes quantités d'informations de façon plus souple et plus performante que le support antérieur : la bande magnétique.

Le terme *database* (base de données) est apparu en 1964 pour désigner une collection d'informations partagées par différents utilisateurs d'un système d'informations militaire⁷.

Les premières bases de données hiérarchiques sont apparues au début des années 1960. Les informations étaient découpées en deux niveaux de hiérarchie : un niveau contenait les informations qui sont identiques sur plusieurs enregistrements de la base de données. Le découpage a ensuite été étendu pour prendre la forme d'un diagramme en arbre⁷.

En 1965, Charles Bachman conçoit l'architecture Ansi/Sparc encore utilisée de nos jours. En 1969, il créa le modèle de données réseau au sein du consortium CODASYL pour des applications informatiques pour lesquelles le modèle hiérarchique ne convient pas⁷. Charles Bachman a reçu le prix Turing en 1973 pour ses « contributions exceptionnelles à la technologie des bases de données ».

En 1968, Dick Pick crée Pick, un système d'exploitation contenant un système de gestion de base de données « multivaluée » (SGBDR MV).

En 1970, Edgar F. Codd note dans sa thèse mathématiques sur l'algèbre relationnelle qu'un ensemble d'entités est comparable à une famille définissant une relation en mathématiques et que les jointures sont des produits cartésiens. Cette thèse est à l'origine des bases de données relationnelles⁷. Edgar F. Codd a reçu le prix Turing en 1981.

Le modèle entité-association a été inventé par Peter Chen en 1975⁸ ; il est destiné à clarifier l'organisation des données dans les bases de données relationnelles^{9,10}.

En 1990, la banque de données juridique LEADERS, avec mises à jour mensuelles sur ordinateur des clients, par disquettes dans un premier temps, CD-ROM par la suite (responsable René Janray), a été lancée en Belgique. Elle a été cédée en 2000 à la société Kluwer. Entretemps elle

avait rassemblé près de 1 000 clients dont les commerces et entreprises les plus importants. Un mémoire de fin d'étude à l'ISAT portait essentiellement sur ce produit et était intitulé *La banque de données Leaders : une percée dans le monde des éditions juridiques électroniques*. En 1998, dans un ouvrage consacré aux banques de données, l'éditeur MARABOUT a consacré tout un chapitre à la banque de données LEADERS.

Dans le modèle relationnel, la relation désigne l'ensemble des informations d'une table, tandis que l'association, du modèle entité-association, désigne le lien logique qui existe entre deux tables contenant des informations connexes.

Les premières bases de données étaient calquées sur la présentation des cartes perforées : réparties en lignes et colonnes de largeur fixe. Une telle répartition permet difficilement de stocker des objets de programmation ; en particulier, elles ne permettent pas l'héritage entre les entités, caractéristique de la programmation orientée objet.

Apparues dans les années 1990, les bases de données objet-relationnel utilisent un modèle de données relationnel tout en permettant le stockage des objets. Dans ces bases de données les associations d'héritage des objets s'ajoutent aux associations entre les entités du modèle relationnel¹¹.

Construction

Les étapes clefs du cycle de vie d'une base de données sont la conception et la mise en service¹².

Avant la conception, les utilisateurs et les producteurs des informations sont interviewés en vue de prendre connaissance des caractéristiques des informations, des relations entre les informations, ainsi que les caractéristiques du système informatique qui accueillera la base de données¹². L'objectif de cette étape est de recueillir les caractéristiques des informations dans la pratique, et les besoins des usagers, et de les formuler d'une manière simple, compréhensible autant par les usagers que les administrateurs de base de données¹².

Puis, sera créé un *schéma* d'ensemble du réseau d'informations et de relations, sous forme de diagramme comportant des entités, des attributs et des relations. Il existe différentes méthodes de modélisation des données, la plus connue étant le MCD (modèle conceptuel des données). Celui-ci permet notamment d'établir un *schéma* structuré de l'ensemble des données d'un système d'informations, compréhensible par tous ses acteurs (chef de produit, développeur, client s'il est averti...).

Ce plan est ensuite transformé en instructions formulées dans le langage de commande du SGBD et les instructions sont exécutées en vue de créer la structure de la base de données et la rendre opérationnelle¹².

La définition de l'organisation interne d'une base de données — son *modèle de données physique* — est l'étape finale de sa construction. Cette opération consiste tout d'abord à définir des *enregistrements* correspondant au modèle de données logique¹³. Les enregistrements sont stockés dans des fichiers, et chaque fichier contient typiquement un lot d'enregistrements similaires¹⁴. Lors de cette étape diverses techniques sont utilisées en vue d'obtenir un modèle qui aboutit à une vitesse adéquate de manipulation de données, tout en garantissant l'intégrité des données¹³.

La qualité du *modèle de données physique* a un impact majeur sur la vitesse des opérations sur la base de données¹³. Une simple amélioration peut rendre les opérations sur les données 50 fois plus rapides¹⁴, différence d'autant plus sensible qu'il y a une grande quantité de données¹⁴. Au début des années 2000, il existe des bases de données contenant plusieurs téraoctets (10^{12}) de données et des ingénieurs indépendants dont l'activité consiste uniquement à aider des clients à accélérer leurs bases de données¹⁴.

Une fois opérationnelle, des opérations de surveillance permettent de détecter des problèmes susceptibles de nécessiter des modifications du schéma. Des modifications peuvent également être apportées en cas de changement des besoins des utilisateurs¹².

Organisation interne

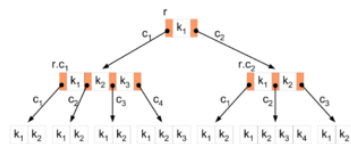
L'organisation interne d'une base de données — son *modèle de données physique* — comporte des *enregistrements* correspondant au modèle de données logique¹³, des pointeurs et des balises utilisées par le SGBD pour retrouver et manipuler les données¹⁴. Les enregistrements sont stockés dans des fichiers, et chaque fichier contient typiquement un lot d'enregistrements similaires¹⁴. L'organisation interne utilise diverses techniques visant à obtenir une vitesse adéquate de manipulation de données, tout en garantissant l'intégrité des données¹³.

L'organisation logique des données est indépendante de leur organisation physique. Ce qui signifie que la position des données dans les fichiers peut être entièrement modifiée sans que leur organisation sous forme d'enregistrements dans des tables ne soit touchée. Le SGBD organise les fichiers d'une manière qui accélère les opérations et qui diffère selon le matériel et le système d'exploitation pour lequel le système de gestion de base de données est conçu. Les enregistrements sont typiquement regroupés en grappes (anglais *cluster*), dont la taille est alignée sur une taille optimale pour le matériel (disques durs)¹³.

Les principales techniques utilisées dans le modèle de données physique sont les index, les vues matérialisées et le partitionnement. Le stockage des données se fait souvent par des dispositifs RAID¹⁴ et le SGBD utilise des techniques telles que les tables de hachage, les arbres B, les bitmaps ou les fichiers ISAM :

Index

Un index est un lot de données destiné à accélérer les opérations de recherche de données. La structure de l'index comporte des valeurs associées à des pointeurs où chaque pointeur permet de retrouver la donnée qui a cette



Arbre B, une structure souvent utilisée.

valeur¹⁴.

Vue matérialisée

Dans une *vue matérialisée*, le résultat d'une recherche / agrégation est enregistré dans la base de données, ce qui permet de l'utiliser plusieurs fois et d'économiser du temps. Cette technique est utilisée notamment dans les bases de données analytiques et les applications OLAP¹⁴.

Partitionnement

Dans la technique du partitionnement, le SGBD répartit les données entre plusieurs disques durs, ce qui accélère les opérations en diminuant la quantité de travail effectuée par chaque disque dur¹⁴.

RAID (de l'anglais *Redundant array of inexpensive disks*)

Dans la technique RAID, un dispositif matériel répartit les données entre plusieurs disques durs, ce qui accélère les opérations et évite la perte d'informations en cas de panne. Ce dispositif matériel est utilisé par le logiciel comme un disque dur ordinaire^{14, 15} ;

Table de hachage (en anglais *hashing*)

Dans cette organisation, une fonction de hachage est utilisée pour transformer, par calcul arithmétique, une valeur quelconque en un nombre entier. Le nombre obtenu est utilisé pour déterminer l'emplacement exact où sera enregistrée la donnée de cette valeur¹³.

Arbre B

L'arbre B est une structure en arbre où toutes les branches ont la même longueur, et chaque nœud comporte entre $N/2$ et N branches. C'est une structure très souvent utilisée pour les index¹³.

Bitmap

Les index en bitmap sont souvent utilisés pour des données où il n'existe que peu de valeurs possibles (exemple : genre M/F, jour de la semaine...). Pour chacune des valeurs possibles de la donnée l'index comporte un tableau de bits où le N ième bit est à « 1 » si la N ième donnée a la valeur en question¹³.

ISAM (de l'anglais *Indexed sequential access method*)

Dans cette organisation les enregistrements sont stockés triés selon la *clé primaire*, dans une structure découpée en cylindres et pistes de taille fixe. Chaque piste comporte un espace de réserve pour permettre l'insertion de nouveaux enregistrements. Une structure en arbre contient la valeur de la clé et un pointeur vers le premier enregistrement de chaque piste.

Journal

Un journal contient la liste des dernières opérations effectuées sur la base de données. Ce journal est utilisé par le moteur de base de données pour annuler les opérations, par exemple en cas de crash informatique, ou si les opérations comportent une erreur. Voir aussi Transaction informatique.

Système de gestion de bases de données

Un système de gestion de base de données est un ensemble de logiciels qui manipulent le contenu des bases de données. Il sert à effectuer les opérations ordinaires telles que rechercher, ajouter ou supprimer des enregistrements (*Create, Read, Update, Delete* abrégé *CRUD*), manipuler les index, créer ou copier des bases de données).

Les mécanismes du système de gestion de base de données visent à assurer la cohérence, la confidentialité et la pérennité du contenu des bases de données. Le logiciel refusera qu'un usager modifie ou supprime une information s'il n'y a pas été préalablement autorisé ; il refusera qu'un usager ajoute une information si celle-ci existe dans la base de données et fait l'objet d'une règle d'unicité ; il refusera également de stocker une information qui n'est pas conforme aux règles de cohérence telles que les règles d'intégrité référentielle dans les bases de données relationnelles.

Le système de gestion de base de données adapte automatiquement les index lors de chaque changement effectué sur une base de données et chaque opération est inscrite dans un journal contenu dans la base de données, ce qui permet d'annuler ou de terminer l'opération même en cas de crash informatique et ainsi garantir la cohérence du contenu de la base de données.

En 2009, IBM DB2, Oracle Database, MySQL, PostgreSQL et Microsoft SQL Server sont les principaux systèmes de gestion de base de données sur le marché.

Mise à disposition

Les bases de données sont de plus en plus souvent mises à disposition de leurs utilisateurs sur des serveurs, via l'internet ou sur des serveurs locaux pour une sécurité plus optimale. Les accès y sont plus ou moins sécurisés.

Certains producteurs de données (collectivités surtout) mettent certaines de leurs bases de données à disposition de tous et chacun, de manière libre. Les données dont ils sont légalement propriétaires ou dépositaires avec autorisation de diffusion (des photos par exemple) peuvent être dans ce cas accessibles en licence libre (ex CC-BY-SA), ou parfois l'ensemble de la base elle-même peut être accessible en licence libre (ODBL par exemple). Ces licences permettent de conserver la propriété intellectuelle sur les contenus, tout en autorisant le ré-

usage et la transformation, en citant la source. De nombreuses autres licences différentes, plus ou moins contraignantes existent.

Utilisations

La manipulation de données est une des utilisations les plus courantes des ordinateurs¹⁶. Les secteurs de la finance, des assurances, des écoles, de l'administration publique et les médias, secteurs majeurs de la société de l'information, qui offrent des services fondés sur des informations, utilisent des bases de données². Parmi les domaines d'utilisation pratiques, il y a les inventaires (stocks, magasins, bibliothèques), les outils de réservation (vols en avion, cinéma...), l'octroi de permis (de conduire, de chasse, de propriétaire de chien...), les ressources humaines, les salaires, la production industrielle (les machines de production sont souvent informatiques), ainsi que la comptabilité et la facturation².

Quelques exemples

Les applications informatiques de collecte de renseignements administratifs tels que dossiers médicaux, fiscaux, ou des permis de port d'arme font usage des bases de données.

Dans une application informatique de billetterie informatisée, les billets de voyage ou de concert sont enregistrés dans une base de données.

Dans un logiciel de comptabilité ou de trésorerie, les écritures ainsi que le plan comptable sont enregistrés dans une base de données ; le bilan est un rapport obtenu par synthèse automatique du contenu de la base de données. De même, dans une application informatique de gestion de la production assistée par ordinateur (abr. GPAO), l'état des stocks, les disponibilités du personnel et les délais sont enregistrés dans une base de données, et servent de base pour un rapport de planification.

Dans un logiciel d'aide au diagnostic médical, un ensemble de pathologies et de diagnostics est enregistré dans une base de données. Un moteur de recherche extrait les diagnostics qui correspondent le mieux aux pathologies choisies par l'utilisateur.

Le Système d'information Schengen est une application de base de données utilisée par les services de police et des douanes de l'Espace Schengen, en Europe pour collecter et s'échanger des renseignements judiciaires (mandat d'arrêts, empreintes digitales, interdictions de séjour...).

ITIS est une application informatique qui contient un catalogue de taxinomie des espèces vivantes (végétaux, animaux, champignons, microorganismes...).

La CIA, agence de renseignements aux États-Unis possède une des plus grandes banques de données au monde.

Le catalogue de la librairie en ligne amazon.com est une des plus grandes bases de données au monde avec plus de 250 millions d'ouvrages catalogués. La société Amazon.com est également propriétaire de la banque de données filmographique IMDb.

Les bibliothèques, notamment universitaires, mettent à disposition de leurs publics des accès à des bases de données¹⁷. Il s'agit en général de bases de données contenant des livres numériques et/ou des articles de la presse généraliste ou spécialisée, fournies par des prestataires extérieurs. Par exemple, la bibliothèque de l'École nationale supérieure des sciences de l'information et des bibliothèques (enssib¹⁸) propose un accès à 76 bases de données¹⁹ à ses usagers, dont A to Z (presse), Cairn (presse et e-books), Dawsonera (e-books), Électre (base de données bibliographique), etc. Certaines sont en accès libre et gratuit, d'autres en accès sur place uniquement, et d'autres en accès réservé sur place et à distance. Par ailleurs l'école (enssib) produit et met en ligne ses propres bases de données pour mieux desservir la communauté des professionnels des sciences de l'information et des bibliothèques : base de constructions de bibliothèques françaises depuis 1992²⁰, Service offres de stages et emplois de l'enssib²¹, Base Maguelone : base de données d'ornements typographiques²².

Dans les systèmes d'information géographique les informations de cartographie sont enregistrées dans une base de données. Des informations de cartographie du ciel sont disponibles dans les catalogues d'étoiles.

Dans les logiciels de forums ou de messagerie électronique, les messages sont souvent enregistrés dans une base de données.

Les logiciels antivirus utilisent souvent une base de données dans laquelle sont enregistrées les empreintes laissées par les virus informatiques.

La base de registre est une base de données qui contient les paramètres de configuration des systèmes d'exploitation Windows. Le Object Data Manager est une base de données similaire des systèmes d'exploitation AIX.

Dans un logiciel de gestion électronique de documents, des documents électroniques sont enregistrés dans une base de données. Dans un système de gestion de contenu ce sont des morceaux de page web qui sont enregistrés dans une base de données et dans un logiciel de commerce en ligne ce sont les annonces, ainsi que les annonceurs.

Dans une plate-forme d'apprentissage en ligne, les exercices, les examens, les cursus, ainsi que l'annuaire des enseignants et des apprenants sont enregistrés dans une base de données.

Un progiciel de gestion intégré tel que SAP ERP (de la société SAP AG) comporte un ensemble de logiciels qui utilisent tous la même base de données. Les différents logiciels de ce type de produit concernent des activités ordinaires des entreprises telles que la facturation, la comptabilité, les salaires, le suivi des commandes et des stocks.

Types d'utilisations

- Un moteur de recherche est un logiciel qui permet de retrouver des ressources (fichiers, documents, pages web) associées à des mots quelconques. La liste des mots et des ressources associées sont stockés dans une *base de données*.
- Une banque de données est une application informatique qui sert à collecter et permettre la publication d'un ensemble d'informations librement consultables et relatives à un domaine de connaissances. Par exemple des renseignements de *bibliographie*, de *linguistique* de justice, de *chimie*, d'*architecture*, de *cinématographie*, de *biologie*, d'*astronomie*, de *géographie*, de *médecine* ou de *jeu*. Les renseignements sont souvent stockés dans des *bases de données* et la banque de données équipée d'un moteur de recherche.
- Un système d'informations est une application informatique qui sert à collecter, classer, regrouper et modifier un ensemble d'informations relatives à une activité. Les systèmes d'informations sont d'usage courant dans les *entreprises* et les institutions telles que les *douanes*, les *hôpitaux*, la *marine*, les *transports* ou l'*armée*.
 - Un « *enterprise resource planning* » (ERP) est un système d'informations qui sert à la collecte et la maintenance des informations concernant l'activité d'une entreprise (ventes, achats, salaires...).
 - Un système d'information géographique est un système d'informations qui permet de collecter, manipuler des renseignements de cartographie (coordonnées de routes, de forêts, de plans d'eau). Ces informations sont utilisées par le système d'informations géographique pour créer des cartes géographiques.
- Un entrepôt de données est une base de données qui est utilisée pour collecter et stocker définitivement des informations historiques qui seront utilisées pour des statistiques et des analyses. Les informations sont souvent collectées à intervalle régulier depuis un système d'informations.
- L'informatique de gestion est un domaine d'activité et de connaissances orientées vers la manipulation de masse de grandes quantités d'informations (gestion de données), un secteur qui utilise des bases de données. Les L4G sont des environnements de développement des applications de base de données souvent utilisés en informatique de gestion. Ils sont composés d'un système de gestion de base de données et d'un langage de programmation.

Notes et références

- ^(en) Carlos Coronel, Steven Morris et Peter Rob, *Database Systems: Design, Implementation, and Management*, Cengage Learning - 2012, (ISBN 9781111969608)
- ^(en) Nick Dowling, *Database Design And Management Using Access*, Cengage Learning EMEA - 2000, (ISBN 9781844801091)
- ^(en) Philip J. Pratt et Joseph J. Adamski, *Concepts of Database Management*, Cengage Learning - 2011, (ISBN 9781111825911)
- ^(en) P.S. Gill, *Database Management Systems*, I. K. International Pvt Ltd - 2008, (ISBN 9788189866839)
- ^(en) Colin Ritchie, *Database Principles and Design*, Cengage Learning EMEA - 2008, (ISBN 9781844805402)
- ^(en) Sikha Bagui et Richard Earp, *Database Design Using Entity-Relationship Diagrams*, CRC Press - 2011, (ISBN 9781439861769)
- ^(en) « Data base technology, IBM » (<http://www.research.ibm.com/journal/rd/255/ibmrd2505O.pdf>) ^[PDF]
- Daniel Martin, *Bases de données : méthodes pratiques sur maxi et mini-ordinateurs*, Paris, Éditions Dunod, 1981, 253 p. (ISBN 2-04-011281-2, OCLC 024483887 (<https://worldcat.org/fr/title/024483887>), BNF 34658745 (<https://catalogue.bnf.fr/ark:/12148/cb34658745t.public>)).
- ^(en) « Entity-relationship logical design of database systems » (<http://ieeexplore.ieee.org/Xplore/login.jsp?url=http%3A%2F%2Fieeexplore.ieee.org%2Fiel2%2F197%2F538%2F00011854.pdf&authDecision=-203>)
- ^[PDF] ^(en) « Entity relationship modeling » (http://www.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf)
- ^(en) « Objet-relational database system » (<http://www.dcc.uchile.cl/~cgutierr/cursos/FDB/ordbms.html>)
- ^(en) Toby J. Teorey, Sam S. Lightstone, Tom Nadeau, H.V. Jagadish, *Database Modeling and Design: Logical Design, Fifth Edition*, Elsevier - 2011, (ISBN 9780123820204)
- ^(en) S. Sumathi et S. Esakkirajan, *Fundamentals of Relational Database Management Systems*, Springer - 2007, (ISBN 9783540483977)
- ^(en) Sam Lightstone, Toby J. Teorey et Tom Nadeau, *Physical Database Design*, Morgan Kaufmann - 2007, (ISBN 9780123693891)
- ^(en) John V. McLean, *Multi-platform Database Consulting: Database Useage On Mainframe, Mid-tier And Nt Servers*, AuthorHouse - 2004, (ISBN 9781418468798)
- ^(en) Allen G. Taylor, *Database Development For Dummies*, John Wiley & Sons - 2011, (ISBN 9781118085257)
- Bases de données (<http://www.enssib.fr/le-dictionnaire/bases-de-donnees>), in *Le Dictionnaire*, enssib, 2014.
- « Accueil » (<http://www.enssib.fr/>), sur *enssib*, 10 mars 2015 (consulté le 10 mars 2015)
- « Abonnements aux bases de données » (<http://www.enssib.fr/abonnements-aux-bases-de-donnees>), sur *enssib.fr*, 10 mars 2015 (consulté le 10 mars 2015)

20. « Constructions de bibliothèques françaises depuis 1992 » (<http://www.enssib.fr/constructions-de-bibliotheques/liste>), sur *enssib*, 10 mars 2015 (consulté le 10 mars 2015)
21. « Le service stages & emplois de l'enssib » (<http://www.enssib.fr/emploi/consulter-les-offres/emploi>), sur *enssib.fr*, 10 mars 2015 (consulté le 10 mars 2015)
22. « Accueil » (<http://maguelone.enssib.fr/>), sur *Maguelonne base d'ornements typographiques*, 10 mars 2015 (consulté le 10 mars 2015)

Annexes

Articles connexes

Notions techniques

- Base de données relationnelle
- Fauna Europaea
- Langage de requête
- NoSQL (« *Not Only SQL* »)
- Pick (système d'exploitation)
- Schéma conceptuel
- Structurae
- Structured Query Language (SQL)
- Système de gestion de base de données
- Table (base de données)
- Vue (base de données)

Notions juridiques

- Protection juridique des bases de données

Liens externes

- - Ressource relative à la santé :
 - (en) Medical Subject Headings (<https://meshb.nlm.nih.gov/record/ui?ui=D019991>)
 - Notices dans des dictionnaires ou encyclopédies généralistes :
 - Britannica* (<https://www.britannica.com/technology/database>)
 - Encyclopédie de l'Ukraine moderne* (http://esu.com.ua/search_articles.php?id=38829)
 - Gran Enciclopèdia Catalana* (<https://www.enciclopedia.cat/EC-GEC-0221295.xml>)
 - Store norske leksikon* (<https://snl.no/database>)
 - Notices d'autorité :
 - BnF (<http://catalogue.bnf.fr/ark:/12148/cb11931023c>)
 - (données) (<http://data.bnf.fr/ark:/12148/cb11931023c>)
 - LCCN (<http://id.loc.gov/authorities/sh86007767>)
 - GND (<http://d-nb.info/gnd/4113276-2>)
 - Japon (<http://id.ndl.go.jp/auth/ndlna/00865521>)
 - Israël (http://uli.nli.org.il/F/?func=find-b&local_base=NLX10&find_code=UID&request=987007532088905171)
 - Tchéquie (<http://aut.nkp.cz/ph114295>)
-

Ce document provient de « https://fr.wikipedia.org/w/index.php?title=Base_de_données&oldid=204090909 ».

-