

SYSTEME D'INFORMATION ET MERISE

DU MODÈLE CONCEPTUEL DE DONNÉES A LA BASE DE DONNÉES

1. INTRODUCTION AUX SYSTEMES D'INFORMATION

1.1. La notion d'information

L'information constitue le support des connaissances et des communications humaines. On utilise parfois indifféremment les termes information et donnée. En fait, une donnée ne devient une information que lorsqu'elle est reçue par une personne qui en fait une interprétation. La signification d'une donnée brute n'est donc pas la même d'une personne à une autre.

Dans le contexte d'une entreprise, les données stockées dans le système informatique sont la traduction codée d'informations, ce qui explique que les deux termes, information et donnée, soient le plus souvent considérés comme synonymes.

1.2. Les caractéristiques de l'information

Une information possède un certain nombre de caractéristiques :

– **La forme** : l'information peut être écrite, orale, visuelle, olfactive, tactile ou encore gustative. Quelles que soient leurs formes, lorsque les informations sont le résultat d'opérations de traitement à partir d'informations brutes, on parle d'**informations structurées** (par exemple un bilan comptable) alors qu'on parle d'informations non structurées dès qu'elles sont obtenues seulement à partir des moyens d'expression naturels de l'être humain (par exemple, une simple note manuscrite portée sur un « Post-it »).

– **Le contenu sémantique** : une information peut être plus ou moins sélective (apporte-t-elle beaucoup de connaissances nouvelles ?), plus ou moins synthétique (est-elle très agrégée ou non ?) ou encore plus ou moins précise (est-elle totalement exhaustive ?).

Le coût et la valeur : on considère que le coût d'une information correspond à celui de sa manipulation, c'est-à-dire de sa recherche, de sa saisie, de son traitement, de son stockage voire de sa destruction. Il est donc difficile à déterminer précisément mais la valeur d'une information est encore beaucoup plus difficile à définir puisqu'elle peut être assimilée à sa capacité à améliorer la décision du dirigeant et donc à lui éviter de commettre des erreurs. En clair, c'est l'utilité de l'information pour celui qui la reçoit qui détermine sa valeur. En tout état de cause, le coût d'une information n'est justifié que s'il est inférieur à sa valeur.

La qualité d'une information

Toutes les informations ne sont pas forcément de bonne qualité. Il y a trois critères essentiels pour caractériser la qualité d'une information :

- sa fiabilité (est-elle exacte et à jour ?)
- sa disponibilité (parvient-elle au bon moment, aux bons destinataires et sous une forme directement et rapidement exploitable ?)
- sa pertinence (est-elle fidèle à la réalité, autorisée par la législation et ni redondante ni calculable à partir d'autres informations ?).

1.3. Le système d'information

Le système entreprise : Une entreprise est une organisation économique intégrée dans un environnement, dotée d'une autonomie juridique et combinant des facteurs de production pour produire des biens et services destinés à être vendus sur un marché.

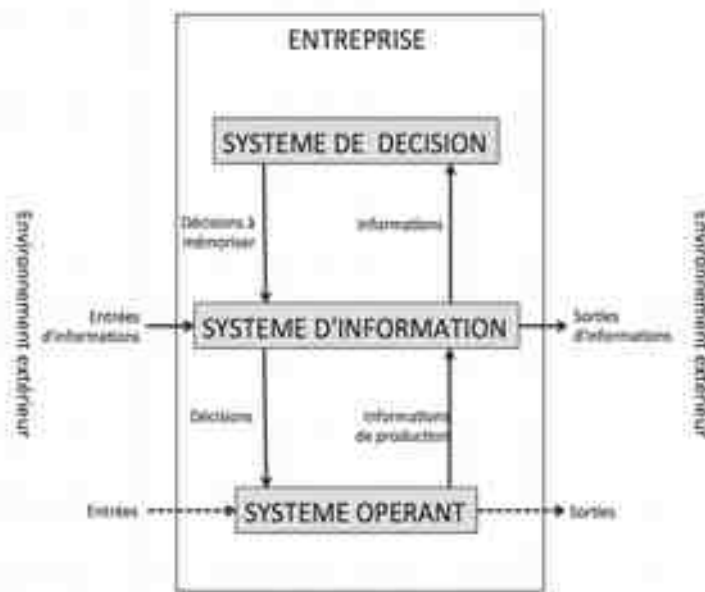
Comme tout système elle est composée de plusieurs sous-systèmes interagissant entre eux.

En l'occurrence, il y en a trois :

– **Le système opérant** : il est également nommé système opératoire ou système de production et réalise la production physique des biens et des services. Son activité est contrôlée par le système de décision. Il est relié à l'environnement par des flux physiques externes et aux autres sous-systèmes par des flux internes d'information.

– **Le système de décision** (également appelé système de pilotage ou système de management) finalise l'entreprise en lui assignant ses objectifs. Il analyse l'environnement et le fonctionnement interne de l'entreprise. Il assure le contrôle des tâches et assure la régulation du système. Il est relié aux autres sous-systèmes par des flux internes d'information.

– Le **système d'information** a un rôle central puisqu'il alimente l'entreprise en informations. Pour cela, il mémorise les informations, les traite et les communique aux deux autres sous-systèmes auxquels il est relié. Toutes les informations de l'entreprise, d'origine externe ou interne, passe donc par le Système d'Information.



1.4. Qu'est-ce qu'un système d'information ?

Définition 1 : « Ensemble des **informations formelles** circulant dans l'entreprise ainsi que les **procédures** et les **moyens** nécessaires pour les définir, rechercher, formaliser, conserver, distribuer. »

Christine Dumoulin (Management des systèmes d'information, EO)

Définition 2 : Le système d'information (SI) peut être défini comme un ensemble organisé de ressources (matériel, logiciel, personnel, données, procédures...) permettant d'acquérir, de stocker, de traiter, de communiquer des informations de toutes formes dans une organisation.

– Il y a donc tout d'abord **des individus** : ce sont toutes les personnes qui utilisent le système.

– Il y a également **des moyens matériels** : ce sont tous les dispositifs physiques permettant de recevoir, manipuler et émettre l'information ainsi que les supports de l'information, qu'ils soient papiers, magnétiques, optiques ou encore électroniques.

– Il y a ensuite **des logiciels et des procédures** : les logiciels correspondent à l'ensemble des programmes qui sont nécessaires au fonctionnement du Système d'Information (lorsqu'il est informatisé bien évidemment). Comme un système d'Information n'est que très rarement entièrement automatisé, les procédures décrivent comment sont articulés les traitements manuels et les traitements automatisés.

– Il y a enfin **les données** qui constituent la matière première des traitements. Elles sont soit saisies et dans cette hypothèse, correspondent à des événements nouveaux pour le Système d'Information, soit calculées et sont alors des résultats de traitement.

Le SI ne doit donc pas être assimilé au système informatique qui n'en est qu'un sous-ensemble.

Le système informatique constitue un support du SI qui prend en charge l'information numérisée et les traitements automatisés. D'une manière générale, ce sont la taille, le secteur d'activité, l'ancienneté de l'organisation mais aussi la stratégie des dirigeants qui déterminent le niveau d'automatisation d'un SI.

2. LA METHODE MERISE

Merise est un acronyme signifiant Méthode d'Étude et de Réalisation Informatique par les Sous-Ensembles ou pour les Systèmes d'Entreprise.

La méthode Merise a comme objectif d'aider et de guider les développeurs dans leurs phases d'analyse, de conception et le développement de l'applicatif.

La méthode Merise se caractérise par :

- une approche systémique en ayant une vue de l'entreprise en termes de systèmes,
- une séparation des données (le côté statique) et des traitements (le côté dynamique),
- une approche par niveaux.

2.1. Une approche par niveaux

Pour la conception d'un SI, il est nécessaire de considérer quatre niveaux d'étude :

- Le niveau conceptuel
- Le niveau organisationnel
- Le niveau logique
- Le niveau physique

2.2. Le niveau conceptuel

Le niveau conceptuel consiste à concevoir le SI en faisant abstraction de toutes les contraintes techniques ou organisationnelles et cela tant au niveau des données que des traitements.

Le niveau conceptuel répond à la question Quoi ? (le quoi faire, avec quelles données).

Le formalisme Merise employé sera :

Le **Modèle Conceptuel des Données** (MCD)

Le **Modèle Conceptuel des Traitements** (MCT)

2.3. Le niveau organisationnel

Le niveau organisationnel a comme mission d'intégrer dans l'analyse les critères liés à l'organisation étudiée. Le niveau organisationnel fera préciser les notions de temporalité, de chronologie des opérations, d'unité de lieu, définira les postes de travail, l'accès aux bases de données...

Les questions posées, au niveau des traitements, sont :

Qui ?

Où ?

Quand ?

Le formalisme Merise employé sera :

Le **Modèle Organisationnel des Données** (MOD)

Le **Modèle Organisationnel des Traitements** (MOT)

2.4. Le niveau logique

Le niveau logique est indépendant du matériel informatique, des langages de programmation ou de gestion des données. C'est la réponse à la question Avec quoi ?

Le formalisme sera :

Le **Modèle Logique des Données** (MLD)

Le **Modèle Logique des Traitements** (MLT)

2.5. Le niveau physique

Le niveau physique permet de définir l'organisation réelle (physique) des données. Il apporte les solutions techniques, par exemple sur les méthodes de stockage et d'accès à l'information. C'est la réponse au Comment ?

Le formalisme employé sera :

Le **Modèle Physique des Données** (MPD)

Le **Modèle Opérationnel et physique des Traitements** (MOpT)

2.6. Tableau récapitulatif

NIVEAU	DONNÉES	TRAITEMENTS	Objectifs	Préoccupation
CONCEPTUEL	MCD modèle conceptuel des données	MCT modèle conceptuel des traitements	QUOI ?	L'invariant du système d'information
ORGANISATIONNEL	MLD modèle logique des données	MOT modèle organisationnel des traitements	QUI ? OÙ ? QUAND ?	Les ressources dont il faut disposer
OPÉRATIONNEL	MPD modèle physique des données	MOpT modèle opérationnel des traitements	COMMENT ?	Les techniques à utiliser

2.7. Séparation des données et des traitements

Les traitements

Ils sont collectés via un processus d'interview et d'étude des documents.

Ils peuvent être de deux sortes : automatiques et manuels.

Ils sont déclenchés par l'arrivée d'évènements.

La gestion des traitements sert à identifier les fonctionnalités selon une approche qui va du général au particulier et qui définit leur découpage et leur enchaînement.

Les données (ou informations)

L'information est l'émission ou la réception de signaux oraux ou écrits, sonores, visuels ou multimédias dont le but est de déclencher les processus alimentant l'échange, base naturelle et indispensable de l'animation de l'organisation.

Les informations se recueillent à l'intérieur du domaine à étudier. La liste d'informations est constituée de plusieurs façons :

- l'interview
- l'étude des documents internes
- l'étude des documents externes
- cahier des charges

Pour le recueil d'informations, il est nécessaire de respecter certaines règles pour éviter des erreurs futures.

C'est le but de du modèle conceptuel de données qui a pour but de représenter les données qui seront utilisées par le système d'information.

A terme, il permettra à l'informaticien de concevoir et/ou modifier la base de données.

3. INTRODUCTION AUX BASES DE DONNEES

Aujourd'hui, la plupart des organisations gèrent et stockent leurs informations dans des bases de données.

Avant l'introduction massive des ordinateurs et des systèmes informatiques, les informations étaient organisées, triées sous forme de fiches, dossiers, classeurs...

Par exemple, un-e secrétaire tenait dans un classeur une fiche par client contenant tous les renseignements sur chaque client. Ce client était doté d'un numéro. Les clients étaient triés par ordre alphabétique (le plus souvent) ou par numéro.

Dans un autre classeur le-la secrétaire tenait un classeur de toutes les factures. Un classeur par année par exemple. Les factures étaient triées par date ou par numéro client....

Cette organisation des données est maintenant matérialisée dans une base de données.

Par exemple, lorsque vous consultez vos comptes sur Internet, il faut bien se rendre compte que les renseignements sur votre compte bancaire et votre identité sont stockés dans une base de données.

De même lorsque vous commandez un produit en ligne. Toutes les informations concernant le produit, son prix, ses références, sa quantité en stock... toutes ces informations sont stockées dans une base de données.

Afin de gérer au mieux toutes ses informations, les retrouver aisément et sans risque d'erreur, il est nécessaire de bien les organiser, de les STRUCTURER.

Une base de données est un ensemble de données organisé en vue de son utilisation par des programmes correspondant à des applications distinctes et de manière à faciliter l'évolution indépendante des données et des programmes.

3.1. SGBD

SGBD est l'acronyme de Système de Gestion de Base de Données. Il s'agit d'un ensemble de solutions logicielles permettant la manipulation des bases de données. Il sert notamment à effectuer des opérations ordinaires telles que consulter, modifier, construire, organiser, transformer, copier, sauvegarder ou restaurer des bases de données.

La plupart des SGBD permettant un accès simultané par plusieurs utilisateurs, les SGBD permettent également de sécuriser les données, de contrôler l'accès aux données, d'en assurer l'intégrité (la cohérence), etc...

3.2. Les différents modèles de bases de données

Les bases de données sont apparues à la fin des années 60, époque où la nécessité de gérer plus facilement et soupagement les données est apparu. Historiquement, un certain nombre de SGBD sont successivement apparus sur le marché. Ils se distinguent par la façon dont sont organisées les informations, ou, de manière plus précise, par le modèle de données sous-jacent utilisé.

On a vu apparaître (dans l'ordre) :

- **Les SGBD hiérarchiques** : les informations sont hiérarchisées de manière simple, entre elles : une information est reliée à une ou plusieurs sous-informations, et l'ensemble constitue un arbre.
- **Les SGBD réseau** : les informations sont reliées de manière simple ou multiple entre elles et l'ensemble constitue un graphe (ou réseau) plus ou moins maillé.
- **Les SGBD relationnels** : ce sont de loin les plus répandus sur le marché. Citons principalement MySQL, Oracle et SQL Server. Nous détaillerons le modèle et ses avantages dans la section suivante.
- **Les SGBD objet** : conceptuellement intéressants, puisque basés sur un modèle dans lequel données et traitement sont encapsulés, ils sont malheureusement assez peu sortis des laboratoires... Citons principalement le SGBD O2.
- **Les SGBD NoSQL** : ces types de SGBD sont de plus en plus utilisés ces dernières années, notamment pour effectuer des recherches sur Internet. Ils sont utilisés pour leurs performances par les sociétés leader du Web (Facebook, Google, Amazon...). Poussées par l'essor du Big Data, elles sont utiles pour les larges ensembles de données distribuées, et parfaites pour analyser des quantités importantes de données non structurées, ou stockées sur plusieurs serveurs cloud virtuels.

En janvier 2020, **les moteurs de bases de données les plus populaires** sont (*source <https://www.digora.com/fr/blog/top-10-des-bases-de-donnees-en-2020-par-popularite>*) :

1. Oracle (score = 1346)
2. MySQL (score = 1274)
3. Microsoft SQL Server (score = 1098)
4. PostgreSQL (score = 507)
5. MongoDB (score = 426)
6. IBM Db2 (score = 168)

7. Elasticsearch (score = 151)
8. Redis (score = 148)
9. Microsoft Access (score = 128)
10. SQLite (score = 122)

Les bases relationnelles sont de loin les bases de données les plus répandues.

3.3. Caractéristiques d'un SGBDR

Un SGBDR digne de ce nom doit disposer d'un certain nombre de caractéristiques minimales :

- **Gestion de gros volumes de données** (en consultation et en mise à jour).
- **Disponibilité** : pouvoir accéder aux données quasiment n'importe quand. Une disponibilité totale des données est possible (temps de reprise nul) il suffit de s'en donner les moyens logiciels et matériels...
- **Fiabilité** : des mécanismes de sauvegarde variés (physique, logique, off-line, on-line, totale, partielle, incrémentale), ainsi que des mécanismes de journalisation et de reprise permettent de restaurer une information sans pratiquement aucune perte, dans tous les cas de panne matérielle et/ou logicielle.
- **Confidentialité** : tout n'est pas accessible à tout le monde ! Se connecter à la base de données donne un certain nombre de droits, en fonction d'un profil défini et maintenu par un administrateur. La granularité d'accès peut aller jusqu'à la vision unique d'un champ d'un enregistrement d'une table particulière. Encore une fois, on ne manipule plus des fichiers...
- **Cohérence** : que les données soient réparties ou non - dans ce dernier cas, les mécanismes mis en jeu seront plus complexes - elles doivent être cohérentes. Cela sous-entend, d'une part que les accès concurrents d'utilisateurs (notamment lors des mises à jour) ne doivent pas compromettre l'intégrité des données, et d'autre part que ces données doivent satisfaire aux contraintes d'intégrité du modèle, mais aussi aux règles de gestion de l'entreprise.
- **Traçabilité** : on peut, notamment en cas de problème important ou en cas d'attaque du système, recourir à l'analyse de traces ou de logs du SGBD. Le niveau de détail de ces traces est paramétrable, et concerne soit les aspects système, soit le réseau, soit l'accès aux données élémentaires elles-mêmes.
- **Concurrence d'accès en lecture et en écriture** (avec une bonne granularité).
- **Gestion (efficace) des transactions.**
- **Portabilité, sur différents OS, des données et du code.**
- **Administrabilité** avec existence d'outils d'administration généraux.
- Existence **d'outils de surveillance en temps réel.**
- **Optimisation des performances.**

Il y a plusieurs approches pour définir ce qu'est une base de données relationnelle, selon le SGBD qui la gère :

pour un SGBD micro (ACCESS de MS, Approach, Open Office, Filemaker, 4D, etc) on peut voir la base comme l'ensemble des tables, requêtes, formulaires et états gérant un domaine d'activité de l'entreprise (gestion commerciale, gestion comptable, gestion de produits, gestion des ressources humaines, etc) ;

Ces SGBD sont plutôt utilisés pour gérer de petits volumes de données et/ou pour créer des applications rapides.

Pour un SGBD de grande ampleur (ORACLE, DB2 de IBM, SQL Server de MS, etc), la base de données est un ensemble plus global pouvant traiter de plusieurs domaines. Elle ne comporte que les données, les contraintes et les éléments de sécurité, mais est indépendante des éléments d'interface. Oracle occupe aujourd'hui une place prépondérante dans les entreprises (environ 40% de parts de marché).

4. TP MYSQL/MARIADB : Approche du modèle logique et conceptuel à travers la manipulation d'une base de données relationnelle

SOURCES

- Introduction aux systèmes d'information, Edition Ellipse
- documents de préparation au Capet, Michel Auguste
- Merise, Guide pratique de Jean-Luc Baptiste, Edition Eni