Systèmes d’information

Programme

Chap 0 : Introduction

Chap 1 : L’architecture et les processus metiers en entreprise : des données en production

Chap 2 : L’information décisionnelle

Chap 3 : Conception d’une BD décisionnelle.

Chap 4 : e-Reporting : Analyse des données et transmission.

Chap 0 : Introduction

1. Definition

C’est un système qui permet :

* Collecter
* Stocker
* Traiter
* Transmettre

En vue de prise de décision.

Remarque : Les systèmes d’information existaient avant l’informatique ; l’informatique n’étant finalement qu’un catalyseur.

Ex :

* + La vendeuse de divers qui note ses achats et ventes(collecte) dans un cahier(stockage) ; et qui à la fin de la semaine peut faire un point(traitement) et décider quoi acheter la prochaine fois(transmission).

1. Des applications et formes de systèmes d’information :

Il existe plusieurs formes de S.I :

* Les S.I géographiques : ce sont des SI qui utilisent des informations de localisation
* Les SI territoriales : sont beaucoup utilisés pour permettre une gestion efficace du territoire.

Les SI sont au cœur du fonctionnement des processus métiers de l’entreprise. A travers les échanges, le système s’enrichit, et on peut toutefois faire des recherches, en vue de proposer des alternatives de solution.

Système opérant

Système exécutant

SI

1 2

3 4

Exemple :

1. Comment je fais pour augmenter mon chiffre d’affaires de 30%
   1. Niveau SI :
      * Analyse : vente augmente de 45% en Decembre
2. Transmission de l’information au système opérant
3. Collecte des infos nécessaires
4. Retour des infos au SI
5. Les systèmes d’information et l’informatique

L’essor de l’informatique et par la suite la digitalisation, ont permis de démocratiser et d’opérationnaliser les systèmes d’informations dans tous ses composants.

3.1. La collecte

La collecte de données aujourd’hui se fait à partir de 3 sources :

* Les enquêtes :
  + Les enquêtes traditionnelles sur fiche
  + Les enquêtes numériques (Google forms, Kobocollect)
* Internet et internet des objets :
  + Les réseaux sociaux collectent chaque seconde, des informations de tous genres sur les utilisateurs. La puissance de cette collecte se mesure à la puissance des GAFAM aujourd’hui.
  + Les Iot qui sont ces petits systèmes embarqués souvent reliés à un cloud, permettent grâce aux nombreux capteurs qu’ils portent, de collecter des informations de tout genre sur l’environnement, mais aussi sur l’être humain (wearables).
* Les banques de données : Ex : INSTAD (ex INSAE)

3.2. Le stockage

Le stockage se faisait traditionnellement sur des registres, fiches ou cahiers ; mais évidemment, cela à très vite montré ses limites (difficulté de stockage, recherche, agrégat). De nos jours, grâce à la numérisation, les informations sont stockées sur des ordinateurs, serveurs, datacenters. Ces données sont organisées en bases de données notamment SQL(base de donnée de production), NOSQL(base de donnée massive), base de donnée multidimensionnelle (les entrepôts de données)

3.3 L’analyse :

Traditionnellement, l’analyse se limitait à faire quelques statistiques, jusqu’à la limite du faisable sur des fiches de cahier. De nos jours, l’analyse se fait sur des ordinateurs à travers des méthodes et algorithmes d’analyse partant de simples statistiques ( min, max, avg, count, sum), jusqu’à des méthodes plus élaborées de Clustering ( clustering, data mining)

3.4 La transmission

Elle consiste à présenter les résultats d’analyse (généralement sous 3 formes) :

* Des tableaux de données ( forme tabulaire)
* Des graphes (sankey)
* Les tableaux de bord (powerBi, tableau)

1. Les types de systèmesd’information :

Il existe 3 types de systèmes d’info :

* SI de traitement de transaction : ce sont les systèmes utilisés en production ou dans la gestion du quotidien de l’entreprise. Au cœur de ce genre de système se trouvent les bases de données transactionnelles orientées CRUD (Ex : traitement des données de vente, d’achat, de stockage, production des factures, bons de livraison)

SI d’infos de gestion : il apporte un support numérique aux gestionnaires. Ce système est souvent sur les données générées par les SI (Ex : il permet de produire des rapports qui résument la performance de l’entreprise). Ce genre des systèmes est souvent supporté par un tableau de gestion qui regroupe au sein de la même page ou de plusieurs composants les indicateurs de performance clés (KPI). Le but de ce genre de tableau est de mettre en évidence très rapidement, des problèmes, des exceptions, des écarts de tendance

* SI d’aide à la décision : Ce sont des systèmes qui permettent à l’être humain de prendre des décisions ou simplement qui aident l’être humain là où il n’est pas suffisamment bon. Ils sont parfois assimilés à des systèmes d’experts. Le nom est prêt a à l’intelligence artificielle, et représente un système de base de connaissance, du genre un oracle.

Remarque : Quand on parle des rôles du système d’information, on fait souvent référence aux 4 composants du système d’information.

1. Conception d’un système d’information

Un système d’information se concoit presque de la même manière que les systèmes en informatique de facon générale. Cette conception passe éssentiellement par deux étapes :

* La modélisation : au cours de cette phase qui se fait avec UML ou Merise, on passe le problème de sa forme imprécise et compréhensible par l’être humain, à une forme plus technique, qui permet de passer à une implémentation précise et compréhensible par l’ordinateur. Ex : Besoins du client – specifications – conceptions(generale et detaillée)
* L’implémentation est constituée necessairement deux parties :
  + Phase de developpement
  + Phase de test : Il existe 3 types de tests :
    - Les tests unitaires : tester chacunes des fonctions
    - Les tests d’intégration : souvent, le dvp se fait module par module. Il faut ensuite connecter les modules et s’assurer qu’ils marchent.
    - Les tests globaux : on fait les tests de fonctionnalités pour s’assurer que le logiciel marche dans son ensemble. A ce niveau, on peut faire appel à des beta testeurs. Ces testeurs sont soumis à plusieurs scénarios.

La démarche pour la conception d’un système d’information se résume souvent en 3 cycles :

* Le cycle de vie (conception, implémentation, maintenance)
* Le cycle de décision : c’est tout ce qui touche à l’analyse des données et à l’interprétation des résultats obtenus.
* Le cycle d’abstraction : C’est renvoyer le système d’information en niveau conceptuel, logique et technique.

Chap 1 :

1. Architecture des entreprises

Au cœur des entreprises, se trouvent les systèmes d’information, que ce soit (STT, SIGEST ou SIAD).

Les entreprises fonctionnent à travers leur processus métier, du système opérant au système exécutant, en passant par les systèmes d’information.

Au cœur du fonctionnement des systèmes d’informations se trouvent les données.

1. Les bases de données en production

Encore appelées bases de données transactionnelles, sont utilisées pour gérer le quotidien des entreprises. Pour ce faire, elles sont organisées et respectent des règles pour faciliter leur traitement.

Definitions

Base de données : c’est une collection de données organisées et structurées en vue de faciliter les traitements (CRUD efficace). Organisé suppose un certain rangement dans une certaine logique. Structuré signifie : respecte un ensemble de règles

SGBD : c’est un logiciel qui permet d’effectuer des traitements sur les bases de donnés. Ex : MySQL, PostGres, SQL server, ORACLE.

Pour qu’une base de données soit correcte et bien formée, il faut qu’elle soit au moins en troisième forme normale :

1fn--2fn--3fn--4fn--5fn--6fn (inclusion)

Dans les SGBD, le traitement se fait grâce à un langage de manipulation, souvent le langage sql

Data warehouse : Définition

Le Data Warehouse est une collection de données orientées sujet, intégrées, non volatiles, historisées, organisées pour le support d’un processus d’aide à la décision.

Orienté sujet : parcequ’on a des objectifs, on a construit la base de données. On ne construit pas pour remplir

Intégré :

Non volatile : il n’y a pas de delete ou update dans un entrepot de données

Historisé : l’aspect temps est toujours utilisé, et l’onformation s’enregistre selon cet aspect

Modélisation des ED

Ter

…

Cette structuration des tables respecte strictement les règles de modélisation définies à travers les formes normales : 1FN—2FN—3FN—4FN—5FN----

Une base de donnée est bien formée si elle est au moins en 3eme forme normale.

Toute tables est constituée d’attributs. Dans une représentation graphique qu’on appelle souvent le schéma de relation, on représente les tables comme un rectangle portant le nom de la table et l’ensemble des attributs. Chaque attribut à un type. La plupart des types qui existent en programmation existent aussi pour les bases de données. Mais on retient essentiellement les types numériques, Les types string, les types DATE.

Il existe des attributs particuliers

* clé primaire : c’est un attribut (ou ensemble) qui permet d’identifier de façon unique chaque ligne d’une table. Il est identifiable par un soulignement.
* Clé étrangère ou secondaire : C’est un clé primaire d’une table qui est présente dans une autre table. Elle est identifiable par la présence d’un # devant.

Exemple : Soit la base de données représentée par le schéma de relations suivant :

* Produit, client et vente sont des tables.
* Produit et Client sont des entités
* Vente est une association
* IDprod, Nomprod, NomCli, TelCli… sont des attributs
* IdProd, IdCli, (IdProd, IdCli, Date) sont des clés primaires
* IdProd et IdCli dans la table ‘ventes’ sont des clés étrangères.

La représentation tabulaire des données consiste à représenter une table comme une matrice ou les colonnes sont de attributs et les lignes représentent les données.

Ex :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| idProd | nomProd | Prix | DatePer | catégorie |
| 1 | Banane | 200 |  | dessert |
| 2 | Cafe | 800 |  | PD |
| 3 | Sucre | 300 |  | PD |
|  |  |  |  | PD |
|  |  |  |  |  |
|  |  |  |  |  |

On appelle cardinalité, le nombre d’attributs d’une table. On le note souvent Card(nom de la table).

Le domaine d’un attribut représente le domaine de définition de cet attribut. Si c’est un type numérique, cela se présente sous la forme d’un intervalle, et de modalité dans le cas contraire.

EX :

Card(produit)=5

Dom(Produit.Prix)= [0-100.000]

Dom(Produit.cat) = [Dessert, PD, Divers,..]

1FN == impose généralement que tout attribut soit atomique, c.a.d non composite( ne peut contenir plus d’une information)

Ex : adresse qui contient : le n de rue, d’appartement, le nom du quartier, de la ville, du pays, ne peut etre un attribut valide. Nous sommes obligés de le décomposer en 5 différents attributs.

* : Il ne peut pas y avoir d’attribut « prixTotal » (attribut calculé) puisqu’on peut l’obtenir à partir des attributs qte et P.U.

2FN== Impose que toute table doit contenir une clé primaire, et que tous les autres attributs présents dans cette table doivent dépendre complètement et fonctionnellement de la clé primaire c.a.d qu’il ne peut pas y avoir d’autres attributs non clé primaire qui identifient d’autres attributs de la table. Ex : si nous partons de l’idée qu’un étudiant participe aux activités d’un et un seul club, et que les activités d’un club se déroulent dans une salle donnée, la relation suivante n’est pas valide en 2FN : etudiant(mat,nom,age,tel,club,salle) parceque matricule ne détermine pas fonctionnellement la salle qui ne dépend que du club. La modélisation valide en 2FN sera : etudiant(mat,nom,age, tel,#idclub,#idsalle) ; club(idclub,nomclub,salle), salle(idsalle,nomsalle,nbplace)

3FN== impose qu’il n’y ait pas de dépendance fonctionnelle transitive

Ex : si A—B et B—C , on ne doit plus avoir A—C

Dimension : représente un angle d’analyse du pb.(Table)

Fait : représente l’objet de l’analyse.

Mesures : Ce sont les attributs de la table de fait. Représente les qtés/éléments à mesurer

Représentation : Les dimensions et les faits se représentent souvent graphiquement exactement comme des tables.

Exemple : Analyser les ventes suivant la région et le produit :

1. Les topologies (logique)

Il y a 3 topologies :

* Topologie en étoile : Une seule table de fait au centre, entourée de plusieurs dimensions
* Topologie en Flocon de Neige : Topologie en étoile ou les dimensions sont éclatées s’il faut en sous dimensions
* Topologie en constellation : C’est une topologie qui regroupe plusieurs autres topologies reliées par leurs dimensions

Exercice

1. Faire un tableau qui présente 5 différences entre les bases de données traditionnelles et les entrepôts de données
2. Quelles sont les différences entre les 3 formes de topologies et études de cas ( dans quel cas utilise on chaque type de topologie
3. Expliquez la diff entre OLTP ET OLAP
4. Difference entre les opérateurs Cube et Rollup
5. Soit la table suivante :

(idproduit, description, marque, sous catégorie, catégorie, département, poids, taille, couleur)

1. Identifier tous les groupes d’attributs formant une hiérarchie dimensionnelle (granularité) dans cette table
2. Comment serait modelisé cette table dans une base de données transactionnelle classique. ?

Exercice 2

1. En supposant que vous travaillez dans une entreprise de production cinématographique, il vous est demandé d’analyser les productions en vue de rentabiliser la boîte
2. Il vous est demandé d’analyser et d’identifier les facteurs qui contribuent à l’échec ou à la réussite des étudiants en milieu universitaire ( Supposez que vous pouvez avoir accès à toutes les données)
3. En supposant que vous avez une entreprise de vente de cours en ligne, Et en étant spécialiste des entrepôts de données, proposez une modélisation qui permettra de décider des promotions à faire en fonction de la performance des cours, et d’évaluer ensuite l’impact de ces promotions.
4. Pour chacune des questions, donner des exemples, ajouter quelques opérations pour extraire les données, et ce qu’on veut en faire

Exercice 3