



## *Département Informatique*

### *BUT 1*

#### **Ressource R1.05 : Introduction aux bases de données et SQL**

26 octobre 2021

### **Cours et exercices**

**Objectif :** *Le cours et les TD / TP dans ce cahier sont choisis pour vous permettre d'acquérir les deux compétences critiques : AC1 et AC3 dans le cadre de votre formation : AC1 concerne la création, la consultation et la mise à jour d'une base de données; AC3 concerne la conception d'une base de données relationnelle à partir d'un cahier des charges. Tous les exercices présents dans ce cahier d'exercices doivent être traités (chez soi ou en classe) avec le soutien de l'enseignant chargé de votre groupe de TD / TP.*





# Table des matières

<b>1</b>	<b>Système d'information et bases de données</b>	<b>3</b>
1	Approche systémique . . . . .	3
2	Limites des fichiers . . . . .	6
3	Bases de données . . . . .	7
4	Architecture ANSI / SPARC . . . . .	8
5	Conception de bases données . . . . .	9
6	Modèles et systèmes . . . . .	10

# 1 Système d'information et bases de données

## 1 Approche systémique

La systémique a pour objectif l'étude des systèmes dans un sens large. Elle a été largement étudiée et formalisée par *J.L. Le Moigne* dans les années 70 avant d'être reprise par d'autres auteurs en particulier dans la méthode **Merise**. Cette discipline permet la définition globale des informations du système et sa confrontation avec des acteurs du système. Dans cette approche l'entreprise est modélisée comme une boîte noire qui assure une fonction principale de transformation des ressources (des flux d'entrée en flux de sorties).

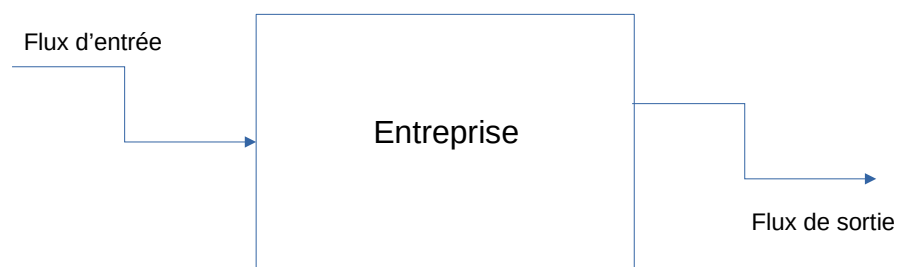


FIGURE 1.1: Système

Pour la représentation systémique, en 1977 Le Moigne, propose principalement la décomposition suivante :

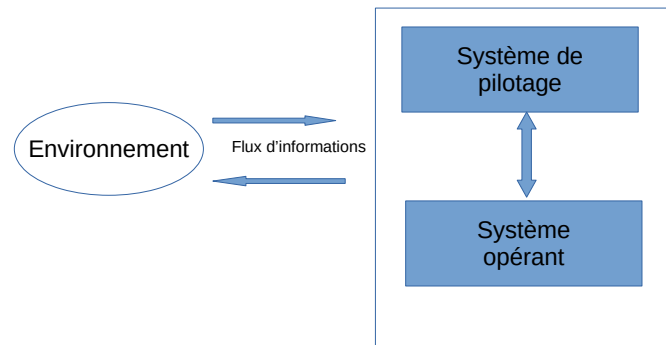


FIGURE 1.2: Système

- Système de pilotage : A ce niveau se déroulent les fonctions de contrôle et la prise des décisions
- Système opérant : A ce niveau se trouvent les fonctions de transformations des flux d'entrée en flux de sortie
- Le système et son environnement interagissent à travers des flux d'informations qui circulent à l'intérieur du système afin de permettre d'analyser son propre fonctionnement.

Pour fonctionner, un système a besoin aussi de mémoriser, de diffuser des informations et surtout d'effectuer tous les traitements nécessaires au bon fonctionnement du système.

L'approche systémique ajoute entre les deux sous\_systèmes un système d'information qui regroupe des informations utilisées tant par les fonctions de transformation que par les fonctions de pilotage

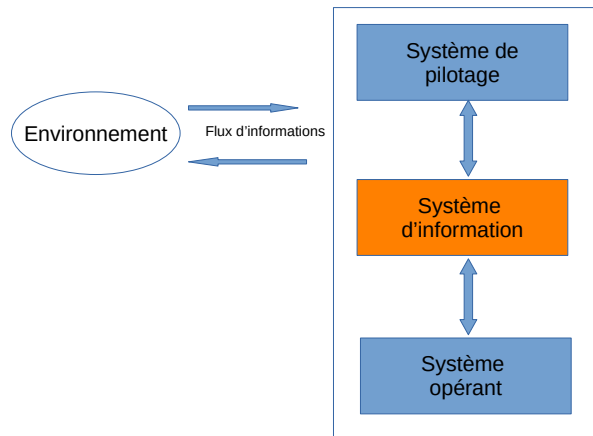


FIGURE 1.3: Approche systémique

Le rôle du système d'information est de collecter, traiter, mémoriser et de diffuser l'information.

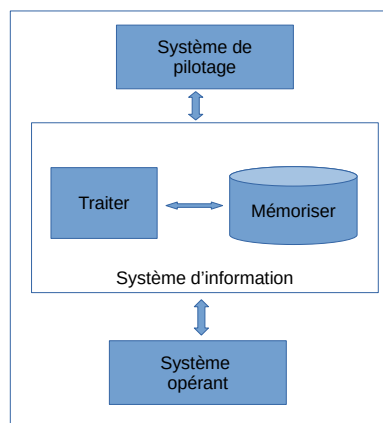


FIGURE 1.4: Système d'information

La figure suivante illustre un exemple d'un système d'information concernant des clients, des fournisseurs et des flux d'informations d'entrées/sorties.

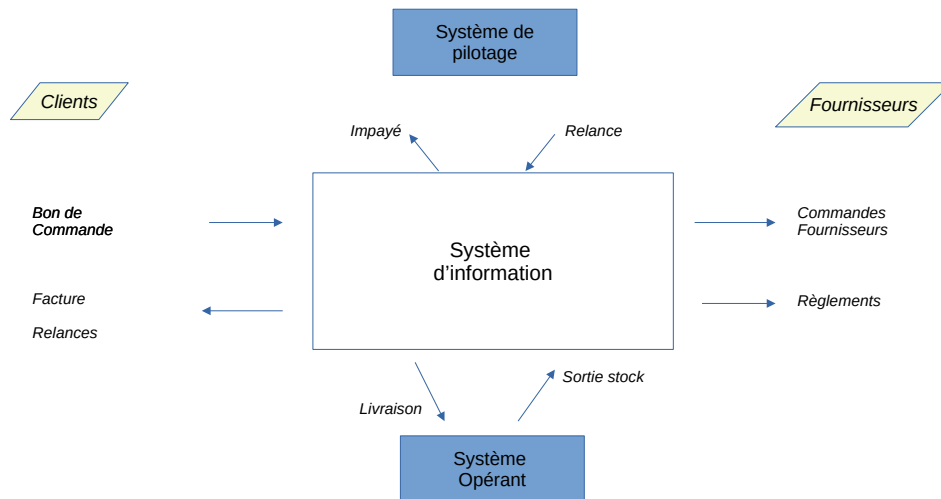


FIGURE 1.5: Exemple d'système d'information

Pour garantir un stockage durable des informations collectées dans le système d'information, celles-ci sont structurées et sauvegardées dans des fichiers gérés par des Système de gestion de fichiers (SGF) ou dans des bases de données gérées par des Système de gestion de bases de données (SGBD). Pour être exploitées par ces systèmes, les informations doivent subir des traitements : en particulier, une transformation car l'ordinateur ne sait stocker que des données et inversement, on doit être capable de reconstituer ces informations à partir des données stockées.

Nous présentons les limites de l'utilisation de l'approche fichiers dans un système d'information.

## 2 Limites des fichiers

- Isolement et séparation des données : les données sont contenues dans des fichiers isolés et séparés les uns des autres
- Des données sont dupliquées à plusieurs endroits dans des fichiers,
- Les structures de données sont définies dans des fichiers. Tout changement dans ces structures impliquent des modifications dans des programmes d'application,



- Format des fichiers différents : cette multitude de format nécessite un effort supplémentaire pour gérer des données stockées dans plusieurs fichiers.

Dans une base de données la description des données se fait à part dans un dictionnaire des données lui assurant une indépendance programme-données : *on peut modifier les structures de données sans affecter les programmes d'application.*

Aujourd'hui les bases de données sont largement utilisées en particulier dans : les banques, les réservations sur le Web etc. Nous introduisons l'approche base de données.

### 3 Bases de données

Une base de données (BD) est une collection de données inter reliées et stockées ensemble avec aussi peu de redondance que possible pour servir une ou plusieurs applications.

Un système de gestion de bases de données (SGBD) est l'ensemble de programmes assurant le stockage, la mise à jour et la recherche de données dans une base de données. Il se présente sous la forme d'un ensemble de modules intégrés écrits en langage d'assemblage ou en langage évolué. Il se greffe sur le système d'exploitation de la machine hôte.

#### Le rôle des SGBD

Un SGBD assure principalement les fonctions suivantes :

1. décrire les données qu'on souhaite stocker dans la base de données (dictionnaire des données accessible aux utilisateurs),
2. décrire les contraintes d'intégrité sur ces données,
3. consulter ces données,
4. mettre à jour ces données,
5. prévoir des procédures de reprise en cas d'incident,
6. résoudre des problèmes d'accès multiples aux données : confidentialité des données, cohérence des données stockées, contrôle de concurrence, récupération des données après panne.

On peut citer quelques inconvénients concernant les SGBD :

1. Complexité : les fonctionnalités des SGBD font de ceux-ci des systèmes complexes,
2. Taille : Cette complexité fait que ces systèmes occupent un espace disque important et qu'ils sont gourmands en mémoire vive,

3. Coût : le coût des SGBD peut être élevé selon l'environnement et les fonctionnalités offertes. A cela on ajoute le coût d'une maintenance récurrente,
4. Panne : En cas de panne, du fait que toutes les données sont centralisées dans une base de données, une panne pourrait entraîner l'arrêt de tout le système.

## 4 Architecture ANSI / SPARC

ANSI/SPARC est une norme qui spécifie l'architecture d'une base de données en trois niveaux :

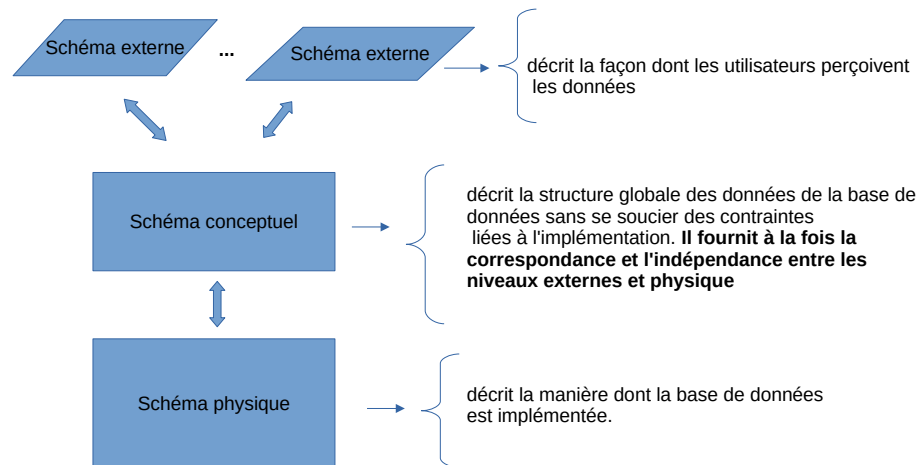


FIGURE 1.6: Système d'information

### Avantage :

1. indépendance physique des données : donne la possibilité de modifier l'implémentation physique des données sans modifier le schéma conceptuel,
2. indépendance logique des données : donne la possibilité de modifier le schéma conceptuel de données sans modifier les vues externes ni les programmes applicatifs,
3. seules les correspondances externe/conceptuel et conceptuel/physique sont susceptibles de changement.

## 5 Conception de bases données

La conception de bases de données est le résultat du processus appelé modélisation de données :

1. La modélisation permet de représenter, par le biais d'un modèle de données (semi-) formel, la réalité comme elle est comprise par les concepteurs de la base de données
2. Le résultat obtenu à la fin du processus de modélisation est la définition d'un schéma appelé Schéma Conceptuel de Données (SCD)

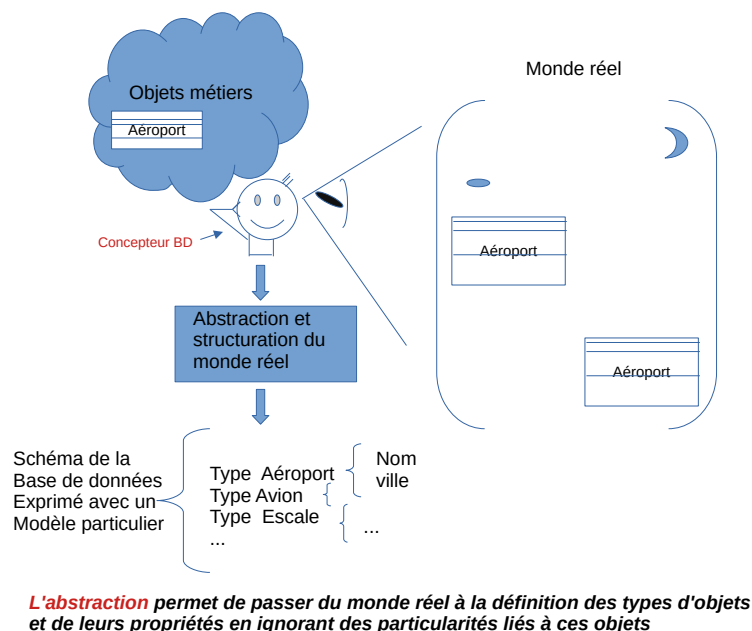


FIGURE 1.7: Système d'information

**Model vs Schéma** La modélisation est une étape importante avant toute implémentation d'une base de données. Les risques d'omettre cette étape sont nombreux : la base de données ne répond pas aux objectifs fixés au départ, le coût de la maintenance est plus élevé et les performances ne sont pas au rendez-vous. Par analogie on peut imaginer qu'un modèle représente la langue française et un schéma est un texte décrit en français. Un modèle de données est un langage permettant de décrire un domaine ou une réalité alors qu'un schéma est la description du domaine avec ce langage. La qualité du schéma (description fidèle à la réalité) dépend finalement de la richesse du modèle utilisé pour décrire le schéma.

La conception d'une base de données passe par plusieurs étapes (ou niveaux) de modélisation.

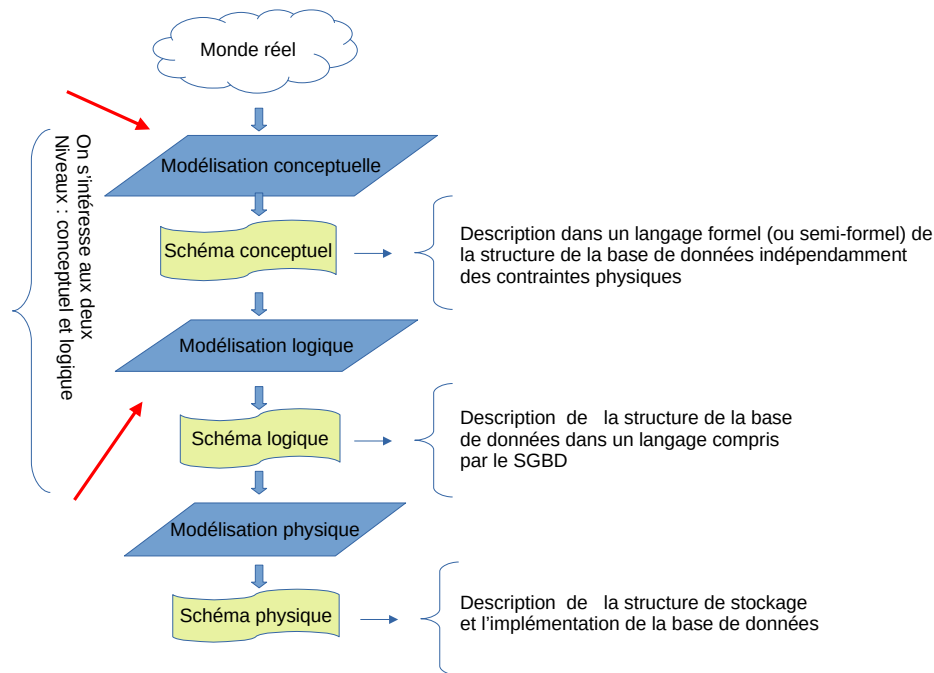


FIGURE 1.8: Système d'information

### Plusieurs niveaux de modélisation

- Niveau 1 : Modélisation conceptuelle avec l'obtention du schéma conceptuel de données (SCD)
- Niveau 2 : Modélisation logique avec l'obtention du schéma logique de données (SLD)
- Niveau 3 : Modélisation physique avec la description du schéma physique (cette partie ne sera pas abordée dans ce cours)

## 6 Modèles et systèmes

Différents modèles et types de SGBD sont apparus dans le passé avant le modèle relationnel.

1. SGBD hiérarchiques (modèle de données hiérarchique) :
2. S.G.B.D. Réseaux (modèle de données réseau)

**SGBD hiérarchiques** Les données sont exprimées avec le modèle de données hiérarchique :

1. permet de représenter des entités (ensemble d'objets) et des relations de type père - fils,
2. l'ensemble de ces entités constitue une arborescence,
3. les SGBD hiérarchiques offrent des primitives pour naviguer dans de telles structures,
4. les utilisateurs doivent raisonner au niveau interne (entités, pointeurs etc.) et non au niveau conceptuel.

**S.G.B.D. Réseaux** Les données sont exprimées avec le modèle de données réseau :

1. permet de représenter des entités (ou un ensemble d'objets) et des relations de type père - fils,
2. les SGBD réseaux offrent des primitives pour naviguer dans de telles structures,
3. à la différence des SGBD hiérarchiques, ils autorisent d'avoir plusieurs entités « père » reliées à une entité fils .

**Aujourd'hui : SGBD Relationnel** Les données sont décrites avec le modèle de données relationnel :

1. à la différence des deux modèles présentés il permet de voir une base de données comme un ensemble de tables,
2. le langage de manipulation des SGBD relationnels est le SQL ,
3. SQL est un langage non procedural.