

MSIS Chapitre 1

Introduction aux bases de données et aux systèmes d'information

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

Définition

Un **système d'information (SI)** est un ensemble organisé de ressources et de dispositifs permettant de collecter, stocker, traiter et diffuser les informations nécessaires au fonctionnement d'une organisation (administration, entreprise, institution ...).

⚠ **Attention** : le système d'information ne doit pas être confondu avec le système informatique. Le système informatique est un sous-ensemble du système d'information.

Exemple : entreprise de transports publics

- désire se doter d'un système d'information (SI) pour la gestion de son réseau,
- celui-ci comprend des lignes, des véhicules ainsi que des chauffeurs,
- le SI doit prendre en compte des faits comme :
 - *le chauffeur « Dupont » est en congé le lundi 30 octobre*
 - *le 31 octobre, il assure la ligne 2 avec le véhicule No 56....*



Michel Léonard, Fondements de la Modélisation de Systèmes d'information, Université de Genève, 2008

Entreprise aimerait que le SI lui permette de...

Trouver des réponses à des questions concernant son activité :

- *Qui a assuré la ligne 18 le 15 mars entre 16h et 18h?*
- *Est-ce que le véhicule No45 est en marche aujourd'hui?*
- *Un véhicule doit-il toujours assurer la même ligne?*

Faire face à des situations critiques :

- *Dupont est malade aujourd'hui ; comment peut-on le remplacer ?*
- *Le tram de la ligne 12 est bloqué par un accident ; comment aménager une solution de secours ?*

Quelle solution?

Entreprise aimerait que le SI lui permette de...

Gérer des événements uniques et/ou récurrents :

- *Il faut enlever le véhicule 124 de la circulation.*
- *Une nouvelle ligne de tram sera prête pour la mise en circulation début septembre.*

S'assurer que ses règles de gestion sont respectées :

- *Un chauffeur ne peut conduire un véhicule que s'il possède un permis de conduire valable pour ce type de véhicule.*
- *Un chauffeur peut conduire plusieurs types de véhicules.*

Quelle solution?

Quelle solution pour cette entreprise ?



Comment stocker, mettre à jour et trouver les informations nécessaires pour assurer les activités de cette entreprise ?

Exemple : gestion du planning de différentes lignes de transport

On veut savoir pour chaque jour :

- *Quel véhicule est affecté à quelle ligne et quel chauffeur le conduit.*
- *Le numéro de téléphone et l'adresse du chauffeur pour pouvoir le/la contacter à tout moment.*
- *Les terminus 1 et 2 de chaque ligne.*
- *Le type et la capacité de chaque véhicule.*



Solution 1 : stockage dans un fichier

Un tableau où chaque ligne représente une affectation d'un chauffeur et d'un véhicule sur une ligne à un jour donné:

Date	Ligne	Terminus1	Terminus2	Chauffeur	Adresse Chauffeur	Tel Chauffeur	Véhicule	Type Véhicule	Capacité Véhicule
15.07	2	Bernex	Cressy	P. Dupont	Nyon	076 5553	56	Bus	90
15.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
15.07	15	Nations	Palettes	M. Jaques	Morges	076 2323	12	Tram	360
15.07	18	CERN	Palettes	T. Martin	Carouge	079 4422	4	Tram	400
16.07	15	Nation	Patelles	P. Dupont	Nyon	076 5553	4	Tram	400
16.07	2	Bernex	Cressy	T. Martin	Carouge	079 4422	86	Bus	90
16.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
...									

Y a-t-il des données redondantes ?

Problèmes de stockage dans un fichier



Redondance de données

Ex: répétition de l'adresse et du téléphone de chauffeurs

répétition de terminus des lignes

répétition de la capacité des véhicules

Date	Ligne	Terminus1	Terminus2	Chauffeur	Adresse Chauffeur	Tel Chauffeur	Véhicule	Type Véhicule	Capacité Véhicule
15.07	2	Bernex	Cressy	P. Dupont	Nyon	076 5553	56	Bus	90
15.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
15.07	15	Nations	Palettes	M. Jaques	Morges	076 2323	12	Tram	360
15.07	18	CERN	Palettes	T. Martin	Carouge	079 4422	4	Tram	400
16.07	15	Nation	Patelles	P. Dupont	Nyon	076 5553	4	Tram	400
16.07	2	Bernex	Cressy	T. Martin	Carouge	079 4422	86	Bus	90
16.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
...									

Problèmes de stockage dans un fichier



La taille du fichier

Consultation de l'information – lecture et recherche dans le fichier

- Facile si accès par une seule clé
 - ex: *recherche par le nom de chauffeur :*
 - trouver l'adresse de **Dupont**
 - trouver les véhicules qui ont été conduits par **Dupont**
- Plus difficile si accès par plusieurs clés simultanément.
 - ex: *recherche par le type de véhicule et la date :*
 - trouver les chauffeurs qui ont conduit un **tram** le **15.07**



Problèmes de stockage dans un fichier

Recherche avec deux clés

Recherche avec une clé

Date	Ligne	Terminus1	Terminus2	Chauffeur	Adresse Chauffeur	Tel Chauffeur	Véhicule	Type Véhicule	Capacité Véhicule
15.07	2	Bernex	Cressy	P. Dupont	Nyon	076 5553	56	Bus	90
15.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
15.07	15	Nations	Palettes	M. Jaques	Morges	076 2323	12	Tram	360
15.07	18	CERN	Palettes	T. Martin	Carouge	079 4422	4	Tram	400
16.07	15	Nation	Patelles	P. Dupont	Nyon	076 5553	4	Tram	400
16.07	2	Bernex	Cressy	T. Martin	Carouge	079 4422	86	Bus	90
16.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
...									

Problèmes de stockage dans un fichier



Modification de l'information – écriture dans le fichier – problèmes de cohérence de l'information

- Insertion de nouvelles informations, ex: *insertion d'une nouvelle affectation*

16.07	3	Gardiol	Gos	T. Martin	Morges	076 2323	33	Trolleybus	120
-------	---	---------	-----	-----------	--------	----------	----	------------	-----

Il faut s'assurer que T. Martin n'est pas déjà affecté à une autre ligne pour ce jour-là et que son adresse et téléphone sont corrects.

- Modification de données redondantes

ex: *modification du numéro de téléphone d'un chauffeur* – il faut le faire pour toutes les affectations de ce chauffeur

- Suppression de données

ex: *suppression de l'unique affectation pour le chauffeur M. Jaques* – on perd également son adresse et son téléphone.



Problèmes de stockage dans un fichier

Suppression de l'affectation
de M. Jaques le 15.07.

Modification du numéro de
téléphone de P. Dupont

Date	Ligne	Terminus1	Terminus2	Chauffeur	Adresse Chauffeur	Tel Chauffeur	Véhicule	Type Véhicule	Capacité Véhicule
15.07	2	Bernex	Cressy	P. Dupont	Nyon	076 5553	56	Bus	90
15.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
15.07	15	Nations	Palettes	M. Jaques	Morges	076 2323	12	Tram	360
15.07	18	CERN	Palettes	T. Martin	Carouge	079 4422	4	Tram	400
16.07	15	Nation	Patelles	P. Dupont	Nyon	076 5553	4	Tram	400
16.07	2	Bernex	Cressy	T. Martin	Carouge	079 4422	86	Bus	90
16.07	A	Rives	Meinier	A. Petit	Genève	077 1212	25	Trolleybus	140
...									

Solution 2 : stockage dans plusieurs fichiers

Structuration de données dans plusieurs fichiers (tableaux)

Fichier des chauffeurs

Chauffeur	Téléphone	Adresse
P. Dupont	076 5553	Nyon
A. Petit	077 1212	Genève
T. Martin	079 4422	Carouge
M. Jaques	076 2323	Morges

Fichier des véhicules

Véhicule	Type	Capacité
4	Tram	400
56	Bus	90
25	Trolleybus	140
86	Bus	90
12	Tram	360

Fichier des affectations journalières

Date	Ligne	Chauffeur	Véhicule
15.07	2	P. Dupont	56
15.07	A	A. Petit	25
15.07	15	M. Jaques	12
15.07	18	T. Martin	4
16.07	15	P. Dupont	4
16.07	2	T. Martin	86
16.07	A	A. Petit	25

Fichier des lignes

Ligne	Terminus 1	Terminus 2
2	Bernex	Cressy
3	Gardiol	Gos
A	Rives	Meinier
15	Nation	Patelles
18	Palettes	CERN



Stockages dans plusieurs fichiers

Avantages

- Pas de redondance de données
- Pas de problèmes de mise à jour – cohérence des fichiers peut être assurée

Nouveau défis – pour trouver l'information complète il faut chercher dans plusieurs fichiers

ex.: *pour trouver l'information complète sur une affectation il faut combiner les informations contenues dans tous les 4 fichiers (tableaux)*

Difficile à programmer pour un système à base de fichiers, mais...

Devient facile avec l'approche Base de données !

De l'approche fichiers vers l'approche base de données

- Avant de proposer une solution informatique il faut étudier la structure de données de manière approfondie.
- S'appuyer sur des concepts et une méthode de structuration de données – **approche base de données**.

Quelle solution pour l'entreprise de transports publics ?



Système de gestion de fichiers ?

- redondance de données
- difficulté de mettre à jour les données
- accès aux données limité
- dépendance entre données et traitements
- manque de vision globale des données



Technologie des bases de données ?

- aucune redondance de données
- mise à jour de données facile et continue
- données facilement accessibles
- indépendance entre données et traitements
- vision globale des données



Qu'est-ce qu'une base de données ?

Définition

Une **base de données** (BD) est un ensemble structuré de données enregistrées sur des **supports informatiques** pour satisfaire simultanément **plusieurs personnes** de façon sélective et en temps opportun.

(Adiba, Delobel 1982)

Une BD sert à...

- stocker de gros **volumes** de données
- **partager** des informations entre plusieurs personnes, plusieurs entreprises
- gérer l'**accès** à ces informations
- assurer la **cohérence** et **non-redondance** des informations

Une BD doit traduire la connaissance du domaine / du métier ...

- Les faits élémentaires

le chauffeur « Dupont » est affecté à la ligne 2 le 24 juin 2023

- Les propriétés et des règles

un chauffeur a un seul permis de conduire, mais est compétent pour plusieurs types de véhicule

- Les événements

de véhicule 124 est supprimé de la circulation

... en un ensemble de données structurées

Quels sont les domaines d'application des BD ?

Partout où l'information tacite n'est pas suffisante et l'information écrite et structurée est nécessaire pour exercer une activité / un métier

- Gestion d'entreprise
- Finances
- Industrie
- Ingénierie
- Commerce
- Transport
- Administration publique
- Droit
- Médecine
- Éducation
- Bibliothèque
- Édition
- Culture
- Écologie
- Services
- Restauration
- Agriculture
- ...

Utilisation des BD a tout son sens quand...

- Les applications sont **centrées sur les données**
- La **complexité de la structuration des données** l'emporte sur la complexité des algorithmes de traitement de données (contre-exemples : traitement d'images, cryptographie, traduction automatique)
- Le **volume des données** l'emporte sur le nombre de traitements
- Les données doivent être stockées de manière **persistante**
- Il est possible de décrire les données avec un **modèle de données**

Approche base de données

- Modélisation conceptuelle du domaine à l'aide d'un **modèle/schéma conceptuel de données** (concepts)
- **Règles de validation** du schéma conceptuel (méthode de normalisation)
- Réalisation informatique: avec un **Système de Gestion de Base de Données (SGBD)**
- Programmation des **accès aux données** avec un langage spécifique

Caractéristiques de l'approche BD

Abstraction des données

La représentation conceptuelle des données n'est pas liée à la manière dont sont stockées les données physiquement.

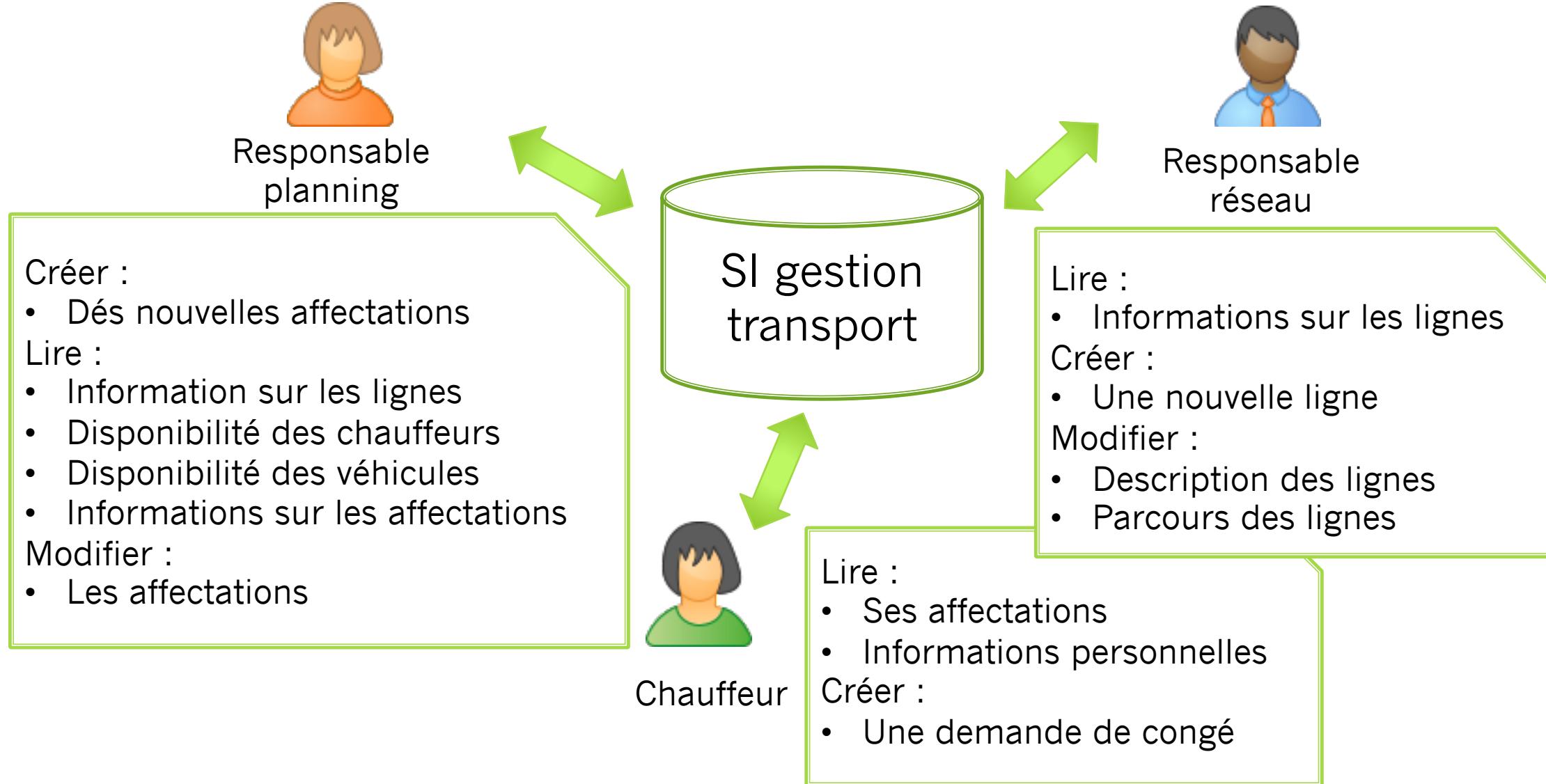
Indépendance données / programmes

Les programmes d'accès aux données sont écrits indépendamment du stockage physique de données -> un changement de données et /ou de la structure des données n'invalide pas les programmes sur ces données.

Vues multiples des données

Plusieurs vues différentes sur les mêmes données pour plusieurs types d'utilisateurs différents.

Utilisation d'une BD / SI



Les défis de l'approche BD

Stockage des données

- Quel sont les données (ou type de données) qui devrait être stockées dans la base?
- Sous quelle forme?
- Comment structurer les données pour éviter la redondance?
- Comment organiser les données pour faciliter leur accès?

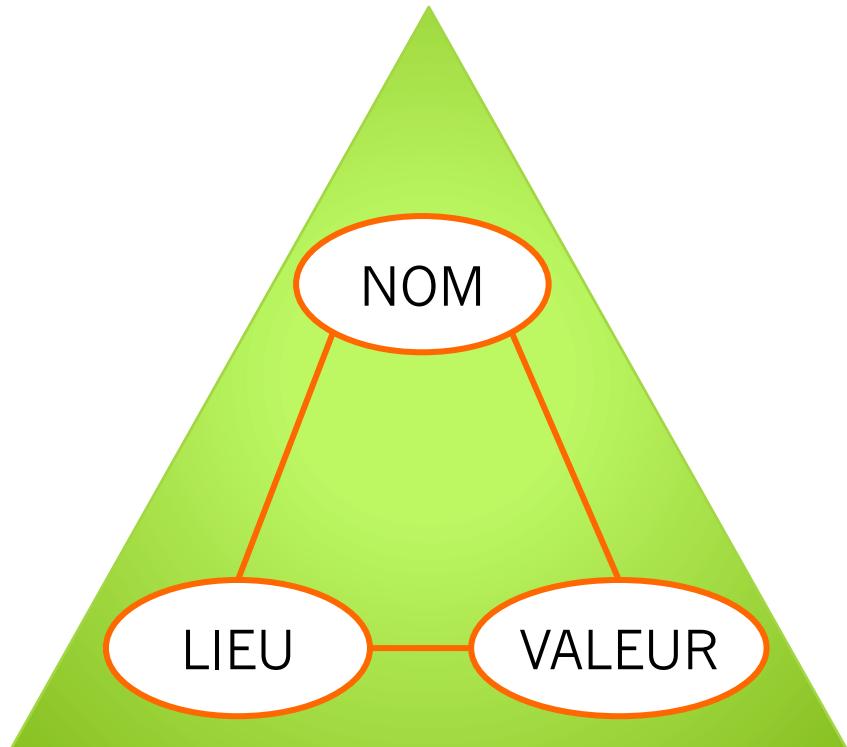
Gestion des accès et des mises à jour des données

- Qui a droit de lire, créer, modifier et supprimer les données?

Traitements des données

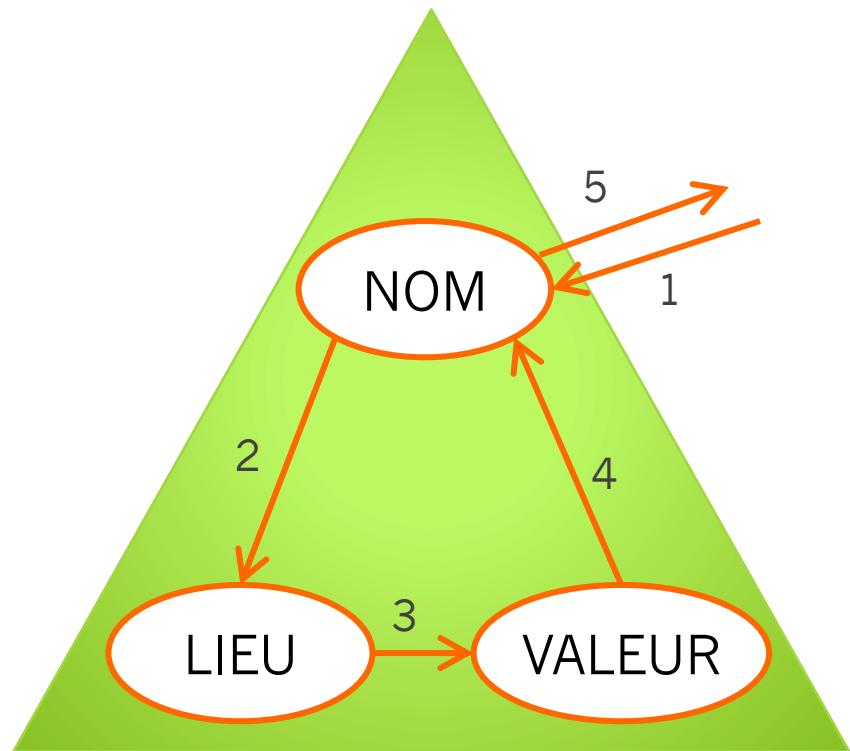
- Applications/programmes permettant de transformer les données en information nécessaire à la gestion d'une organisation

Définition des données dans un SGBD



- Une donnée a un **nom**.
Une donnée ne peut être atteinte qu'au travers de son nom, autant par les développeurs que par les utilisateurs.
- Une donnée prend un **lieu** de stockage et y reste.
- Une donnée représente une **valeur** pour l'utilisateur.

Définition des données dans un SGBD



Le SGBD

1. à partir du nom reconnaît le domaine de la donnée :
 - le nombre d'octets
 - le mode d'interprétation (chaînes de caractères, entiers, réels, dates...)
2. calcule le lieu de stockage de la donnée
3. retrouve l'ensemble des octets de la donnée
4. interprète les octets en termes compréhensibles aux utilisateurs
5. et leur fournit le résultat

Un modèle de données est nécessaire pour gérer l'espace de noms!

Bases de données relationnelles

Le modèle relationnel – un ensemble de concepts pour décrire le contenu structuré des bases de données relationnelles sous forme tabulaire.

- Les données sont structurées en relations – **tables**.
- Une relation regroupe les données de même type.

Table des chauffeurs, table des véhicules, tables des lignes, table des affectations journalières des chauffeurs à des lignes

- Les lignes de la table sont appelées **tuples** ou **enregistrements**.
- Les colonnes sont appelées **attributs**.

The diagram shows a table titled "Chauffeur" (Driver) with four columns: Nom (Name), Prénom (First Name), Téléphone (Phone), and Adresse (Address). The table contains four tuples (rows): Dupont/Pierre (076 5553, Nyon), Petit/Annaëlle (077 1212, Genève), Martin/Thomas (079 4422, Carouge), and an ellipsis (...). A callout bubble labeled "Attribut" (Attribute) points to the "Téléphone" column. Another callout bubble labeled "Tuple/enregistrement" (Tuple/Record) points to the second tuple (Petit/Annaëlle).

Nom	Prénom	Téléphone	Adresse
Dupont	Pierre	076 5553	Nyon
Petit	Annaëlle	077 1212	Genève
Martin	Thomas	079 4422	Carouge
...			

Structuration des données

Les données structurées sous forme de tables interconnectées

Chauffeur

Nom	Prénom	Téléphone	Adresse
Dupont	Pierre	076 5553	Nyon
Petit	Annaëlle	077 1212	Genève
Martin	Thomas	079 4422	Carouge
...			

Véhicule

NoVéhic	Type	Capacité
4	Tram	400
56	Bus	90
25	Trolleybus	140
86	Bus	90
...		

Ligne

NoLigne	Terminus1	Terminus2
2	Bernex	Cressy
A	Rives	Meinier
15	Nation	Patelles
18	Palettes	CERN
...		

Affectation

Nom	NoLigne	NoVéhic	Date
Dupont	2	56	15.07.23
Dupont	15	4	16.08.23
Martin	2	86	16.08.23
...			

SQL – Structured Query Language

Pour la création et manipulation des BD relationnelles

Langage de définition de données

Création, modification et suppression des tables

Langage de manipulation de données

Création, modification et suppression de données dans les tables

Langage de contrôle de données

Définition des droits d'accès

Exemple de définition de données en SQL

```
CREATE TABLE Chauffeur  
(  
    Nom    CHAR(50),  
    Prénom CHAR(50),  
    Téléphone CAHR(8),  
    Adresse CHAR(100),  
    DateNaissance DATE  
);
```

Exemples de création, de modification et d'interrogation de la BD

Création de données

```
INSERT INTO Chauffeur
```

```
VALUES ('Aubry', 'Christine', '075 3456', 'Lancy', '12.03.1990');
```

Interrogation

```
SELECT Nom, DateNaissance  
FROM Chauffeur  
WHERE Adresse='Genève';
```

Modification

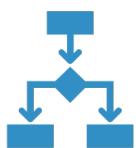
```
UPDATE Chauffeur  
SET Adresse = 'Genève'  
WHERE Nom = 'Aubry';
```

Qualifications requises pour l'usage de la technologie des bases de données



Maîtrise des technologies d'information, surtout des systèmes de gestion de bases de données (SGBD)

Conceptualisation des schémas



de données

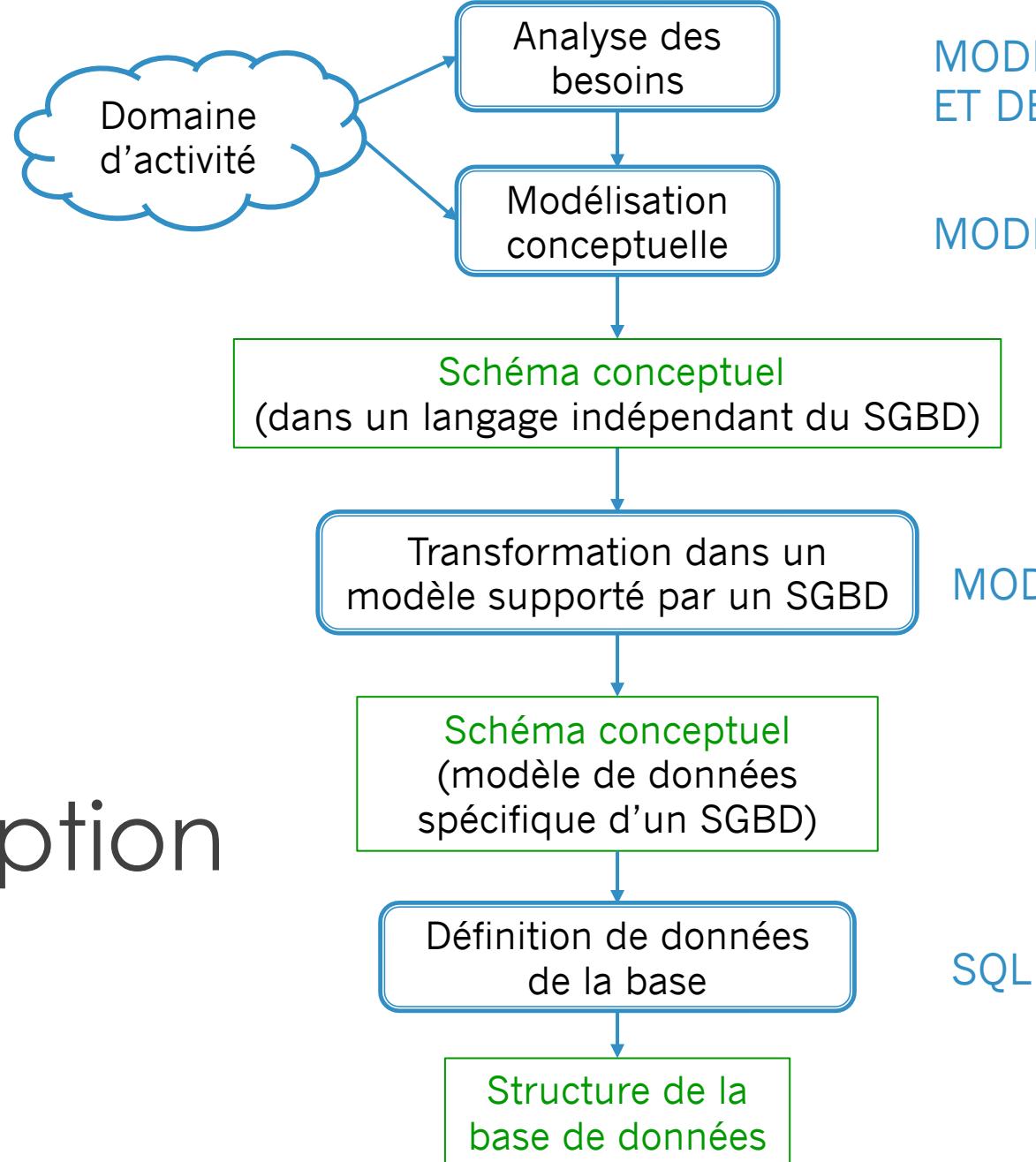
des traitements de données

des interfaces avec les acteurs de l'entreprise



Organisation des activités humaines sous-jacentes à l'utilisation intensive du SI dans le cadre de différents processus organisationnels

Les étapes de la conception d'une BD



Langage de modélisation

MODÈLES DES OBJECTIFS
ET DES CAS D'UTILISATION

MODÈLE DES CLASSES

MODÈLE RELATIONNEL

SQL