

# BD Aspects Systèmes

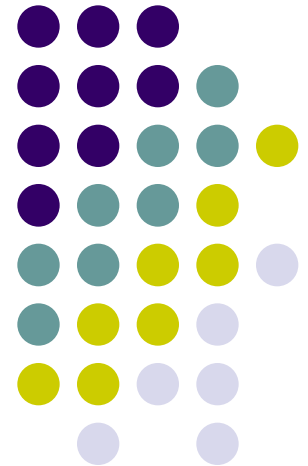
---

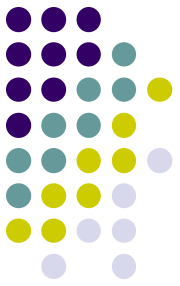
Licence 3  
ISTV – UPHF

Marie Thilliez

[marie.thilliez@uphf.fr](mailto:marie.thilliez@uphf.fr)

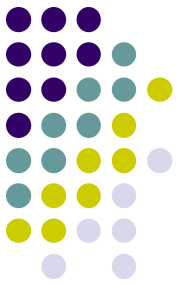
Clé Moodle : BDSYS2019





# Organisation du module

- Clé moodle : BDSYS2019
- Volume horaire :
  - 9h de cours
  - 12h de TD
  - 9h de TP

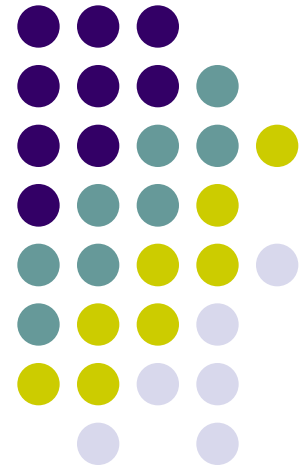


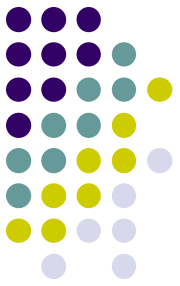
# Sommaire

1. Introduction et rappels
2. Gestion interne d'une base de données
3. Optimisation de requêtes
4. Fiabilité des données
5. BD Actives

# Introduction et rappels

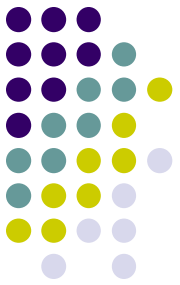
---





# Au sommaire

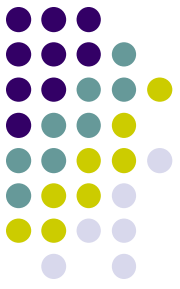
- Rappel des principes de base
- Les modèles de SGBD
- Le modèle relationnel et l'algèbre
- Conception



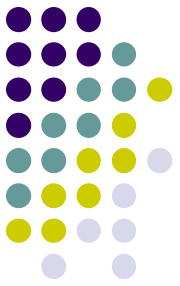
# Plan

- Définitions
- Nécessité des SGBDs
- Objectifs des SGBDs
- Notions de Base

# Qu'est ce qu'une Base de Données (BD) ?



- collection d'informations stockée de manière plus ou moins permanente dans un système informatique.
- Exemple :
  - les données décrivant l'activité d'une compagnie aérienne:
  - voyageurs, vols, avions, personnel, ...
  - réservations (un voyageur réserve pour un vol)

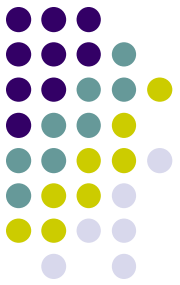


# Qu'est ce qu'un SGBD ?

- SGBD = Système de Gestion de Bases de Données
- un logiciel qui permet à une ou à plusieurs personnes :
  - de manipuler
  - de modifier les données d' une BD.
  - D' interroger d'une BD
  - De mettre à jour une BD

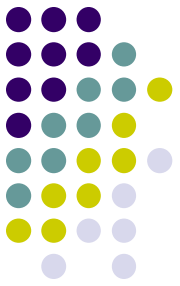


# Pourquoi un SGBD est il nécessaire ?



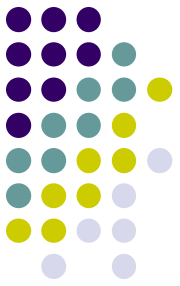
- Evolution du mode de développement des applications!
  - Jadis applications développées indépendamment...
    - multiplication des tâches (saisie,développement,) des supports
    - Incohérence possibles des données
    - Non portabilité des traitements
  - Dépendance à l'environnement
    - tout changement matériel ou logiciel a un impact sur les applications
  - Dépendance aux données
    - tout changement de la structure des données nécessite de modifier les programmes

# Pourquoi un SGBD est il nécessaire (2)?



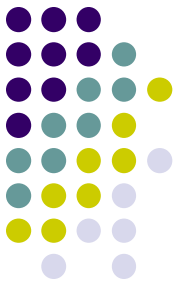
- ...vers un mode de développement des applications fédérant les données(données similaires et échangées)
- Pourquoi?
  - Les applications ne sont pas disjointes
  - Exemple :
    - comptabilité clients
    - comptabilité fournisseurs
    - gestion de stocks,...
  - Informations partagées
    - Clients, Articles, Fournisseurs et Personnels

# Pourquoi un SGBD est il nécessaire (3)?



- Solution : l'approche globale d'un Système d'Information
  - Description des données indépendante des traitements
  - Maintenance de la cohérence de données
  - Langage non procédural, interactif et structurel

# Objectifs initiaux des SGBDs

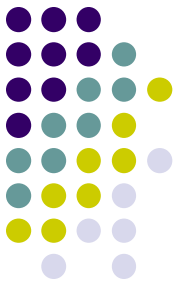


- Objectifs Initiaux

- Administration cohérente des données
- Indépendance physique des données
- Indépendance logique des données
- Contrôle de la redondance de données
- Manipulation des données par des non-informaticiens

# Objectif 1 :

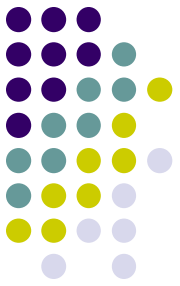
## Administration cohérente des données



- Etablissement d' une vue canonique des informations manipulées par l' entreprise
- Développement incrémental d'un modèle cohérent du système considéré
- Possibilité de modélisations particulières incluant des données privées
- Possibilité de concevoir des traitements nouveaux ou plus complexes (statistiques, aide à la décision, ...)

# Objectif 2 :

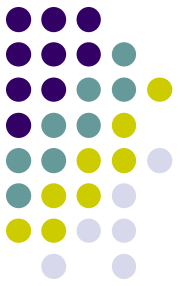
## Indépendance physique des données



- Objectif : la structure canonique des données doit être indépendant des structures de stockage en informatique.
  - Description logique / conceptuelle
    - Objets (données élémentaires) et classification d'objets
    - entité et association
    - règles de contrôle de données
  - Description internes de données
    - fichiers séquentiels, fichiers indexés (ex : B-Tree), fichiers hachés
- Résultats
  - Portabilité des applications
  - Indépendance au matériel

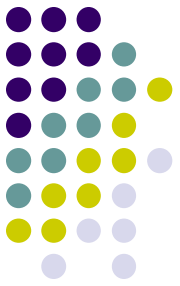
# Objectif 3:

## Indépendance logique des données



- Chaque application ou groupe d'applications désire utiliser ses propres descriptions conceptuelles de données en fonction de ses propres besoins.
- Chaque application établit une description des données quelle utilise :
  - les données écrites restent virtuelles
  - le SGBD se charge de leur faire correspondre des données réelles
- Exemple : Description conceptuelle intégrée
  - Véhicule(Immat, marque, type, couleur)
  - Personne(NumP, nom, prénom)
  - Propriété(#Immat, #NumP, date-achat)
    - Application 1: Personne(NumP, nom, Immat)
    - Application 2: Voiture(Immat, marque, date-vente)

## Objectif 3: Indépendance logique des données(suite)

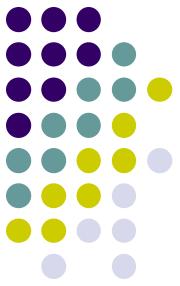


- Résultats :
  - Possibilité pour chaque application d'ignorer les besoins des autres applications tout en partageant la même BD
  - Possibilité d'intégrer des applications déjà écrites
  - Possibilité d'évolution de la BD sans réécriture des applications non concernées
  - Possibilité de limiter les conséquences du partage : données confidentielles



# Objectif 4 :

## Contrôle de la redondance de données

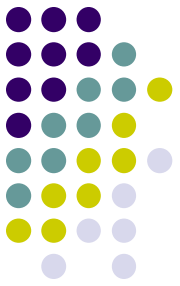


- **La redondance est contrôlée par le système pour ses propres besoins**
  - fiabilité
  - répartition géographique
  - performances en consultation
- **Résultats**
  - Diminution de tâches de saisie
  - Maintien de la cohérence des données partagées par plusieurs applications

# Objectif 5 : Manipulation des données par des non-informaticiens



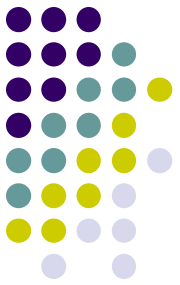
- Les non-informaticiens doivent pouvoir manipuler les données à partir de la seule connaissance du monde réel et de la modélisation qui en est faite.
- ↳ Langage descriptif proche du langage naturel.
- ↳ cependant, intégration possible de ce langage dans des langages de programmation plus généraux



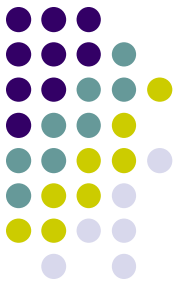
# Autres objectifs des SGBDs

- Conséquences : Nouvelles fonctions
  - optimisation de l'accès aux données
  - contrôle de l'intégrité sémantique des données
  - partage simultanée des données
  - sécurité des données
  - sûreté des données

# Nouvelles fonctionnalités



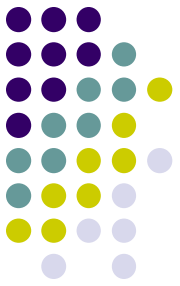
- Fonction 1 : Optimisation de l'accès aux données
  - gestion unique centralisée des chemins d'accès aux données par le SGBD
  - optimisation automatique des demandes d'accès des utilisateurs
  - économie de l'astuce des programmeurs et des spécialistes des fichiers, en sachant que l'astuce se chiffre en milliers heures d'écriture de logiciel.



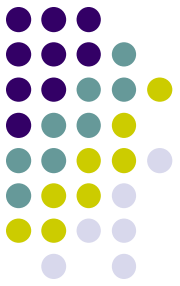
# Nouvelles fonctionnalités(2)

- Fonction 2 : Contrôle de l'intégrité sémantique des données
  - Contrôle sur les données élémentaire
    - contrôle de type
      - *NOM = alphabétique*
      - *SALAIRE X tel que  $1200 < X < 12000$*
  - Contrôle sur les relations entre les données
    - un électeur doit être inscrit sur une seule liste électoral
    - un vin bu par un buveur doit exister
    - prix de vente > prix d'achat

# Nouvelles Fonctionnalités (3)

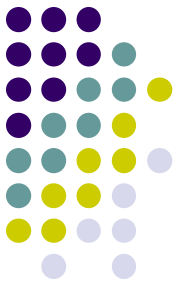


- Fonction 3 : Partage simultané des données  
(contrôle de la concurrence des accès)
  - les différentes applications qui partagent les mêmes données, veulent pouvoir s'ignorer et travailler simultanément



# Nouvelles Fonctionnalités (4)

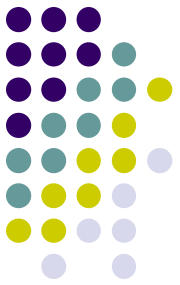
- Fonction 4 : Sécurité des données (confidentialité des données)
  - règles de partage des données
    - garantir la confidentialité de certaines informations
  - protection contre le vol et la corruption des données
  - Exemple
    - Compte Bancaire, Dossier médical, Procédé de fabrication



# Nouvelles Fonctionnalités (5)

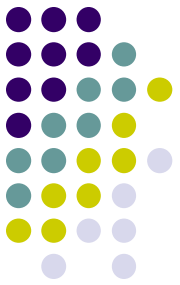
- Fonction 5 : Sûreté des données (résistance aux pannes)
  - intégrité des données en cas de panne
    - plan de sauvegarde
    - gestion de la redondance physique
  - Après une panne ou après l'abandon d'un traitement en cours, le SGBD doit rétablir les données dans un état cohérent et défini.





# Quelques Notions de Base

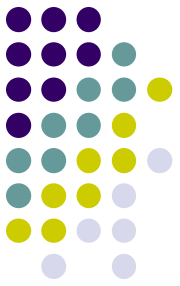
- Type d' Objets
  - ensemble des objets qui possèdent des caractéristiques similaires et manipulables par des opérations identiques.
    - entier = { 0, 1, 2, ... , 1000 }, vehicule = (immat, marque, type, couleur)
- Occurrence d' objets
  - Élément d'un ensemble d'objets
    - l'entier 486, le vehicule (AL 450 AB, Renault, Clio, Gris)
- Schéma d' une BD
  - Structure d'une BD
  - ensemble des types des objets de la base
    - Etudiant (NumEtud, nom, ville), Module(NumMod, titre), Inscription(#NumEtud, #NumMod, dateInscription)
- Instance de BD
  - occurrences d'objets appartenant aux types du schéma
    - 172 BDSYS 2019, 175 BDOR 2019



# Pour l'administration

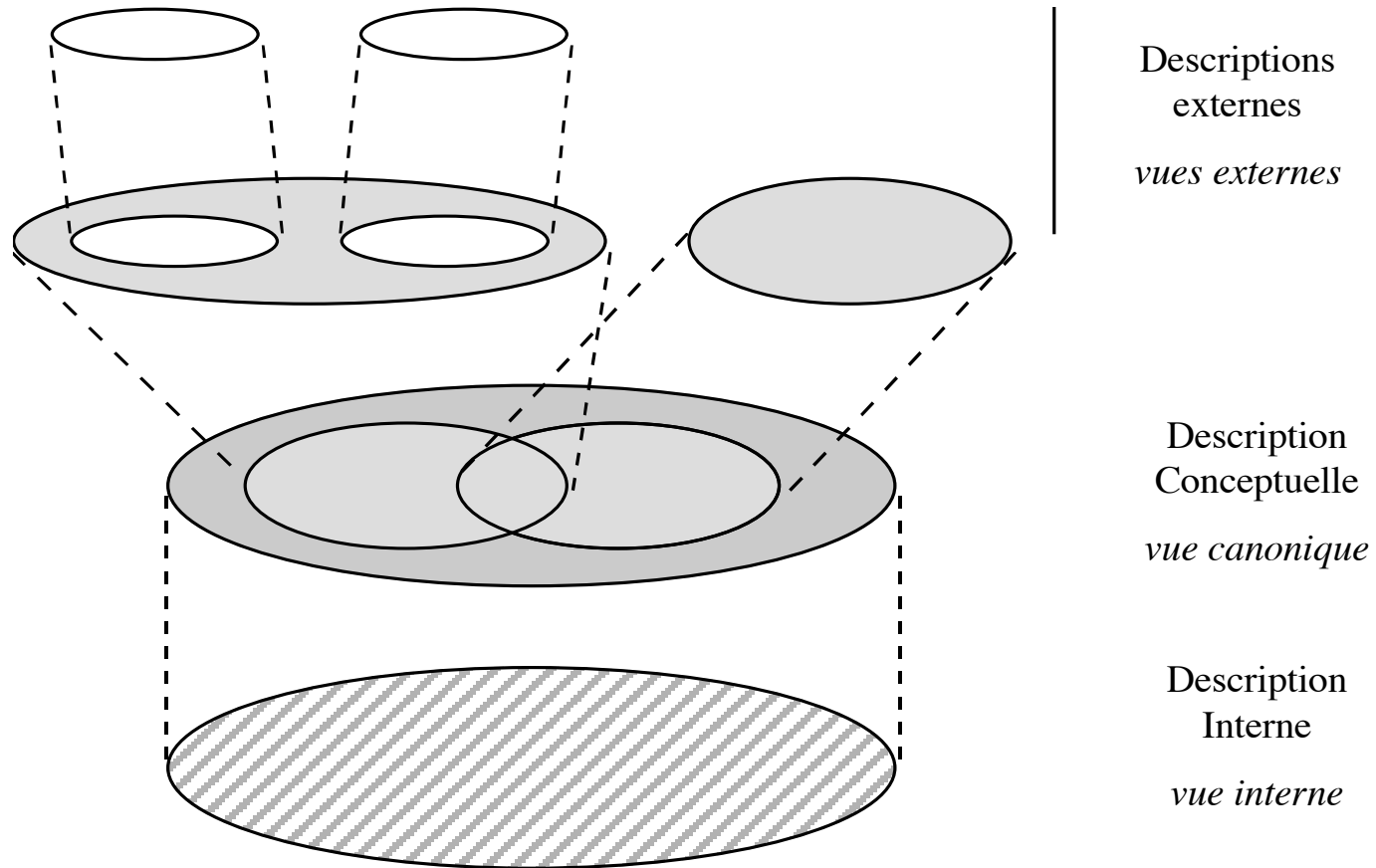
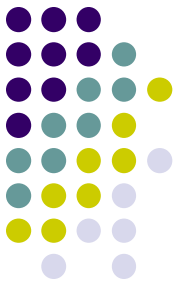
- **Groupe de personnes responsables**
  - de la définition des schémas des règles de correspondance
- **Administrateur Entreprise**
  - schéma conceptuel et règles de contrôle des données
- **Administrateur Bases de Données**
  - schéma interne et règles de correspondance conceptuel à interne
- **Administrateur Application**
  - schéma externe et règles de correspondance externe à conceptuel

# Vues

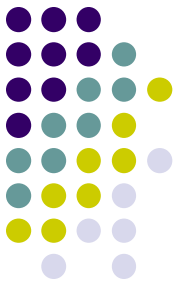


- **Vue canonique des données**
  - le schéma global fournit une modélisation globale des données (schéma conceptuel)
- **Vue externe des données**
  - une application particulière n'a qu'une vue partielle de la réalité correspondante (schémas externes)
- **Vue interne des données**
  - le schéma de stockage physique des données
- **Règles de correspondance**
  - règles de transformation des données correspondant à un niveau de schéma, en données conformes à un autre niveau.

# Les vues

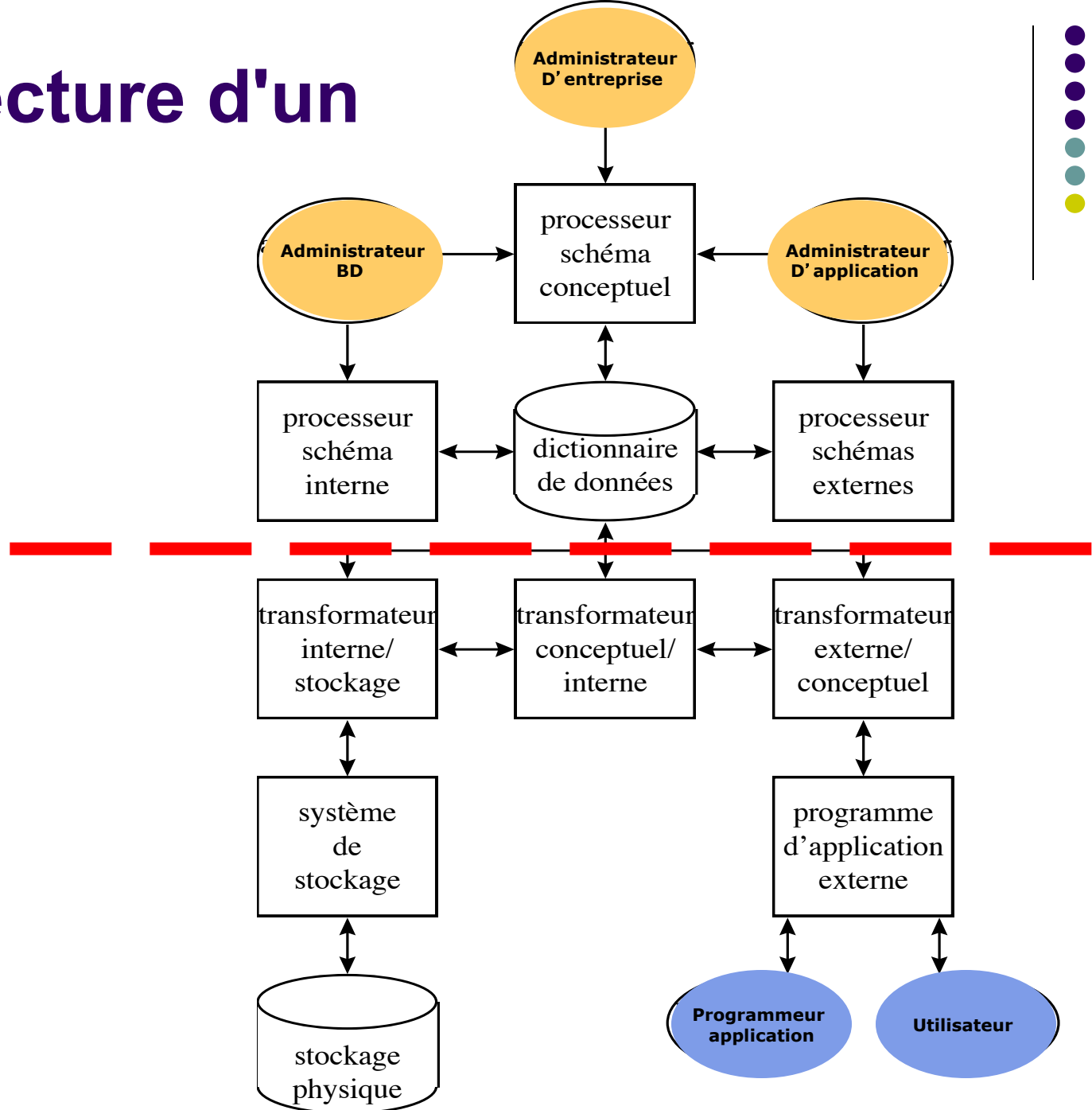
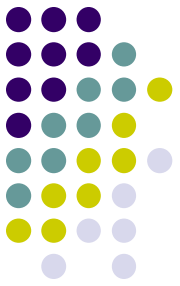


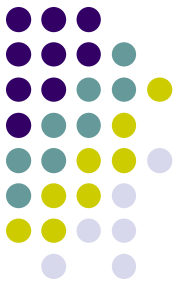
# la MétaBase = Dictionnaire des Données



- la base contient
  - les données
    - occurrences de types
  - les méta-données
    - schémas
    - règles de correspondance & règles de contrôle

# Architecture d'un SGBD



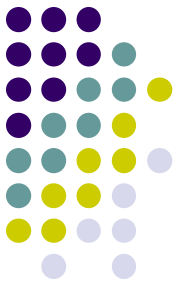


# Langages d'un SGBD

- LDD : Langage de Description de Données
  - **description du schéma et des règles**

```
create table etudiant(  
  NumEtud  number,  
  Nom      varchar2(20),  
  Ville    varchar2(15),  
  Primary key (NumEtud)  
);
```

```
create table inscription(  
  NumEtud      number,  
  Filiere      varchar2(15),  
  Annee        number,  
  Primary key (NumEtud, Filiere, Annee),  
  Foreign key (NumEtud) references Etudiant  
);
```



# Langages d'un SGBD

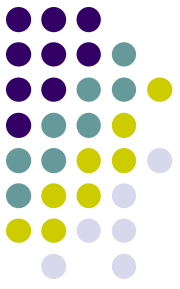
- LMD : Langage de Manipulation de Données
  - **interrogation et mise à jour de la BD**

```
select nom
from etudiant, inscription
where
Etudiant.NumEtud = inscription.NumEtud
and annee = 2017;
```

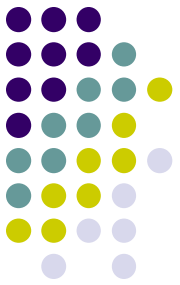
- LCD : Langage de Contrôle de Données
  - **Autorisation d'accès**
    - Grant, revoke



# Les différents types de bases de données

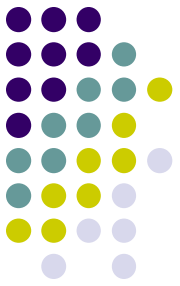


- Les bases de données objet
- Les bases de données XML (DB XML)
- Les bases de données NoSQL (Cassandra, MongoDB etc.)
- Le modèle relationnel (ou objet-relationnel)



# Théorème CAP

- Aucun système distribué ne peut fournir les 3 propriétés suivantes :
  - Consistency (ou cohérence)
  - Availability (ou disponibilité)
  - Partition tolerance (résistance au partitionnement)

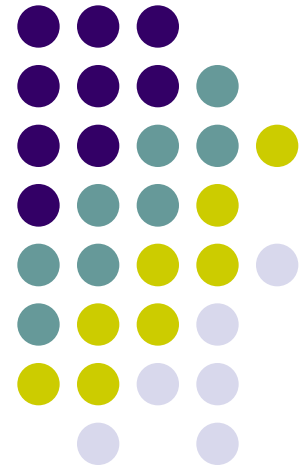


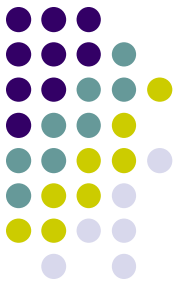
# CAP (brewer) : quel choix ?

- Système distribué :
  - A+P
  - C+P
- SGBD relationnel :
  - C+P
  - Propriété ACID au détriment des performances
- BD noSQL : souvent A+P

# Le modèle relationnel

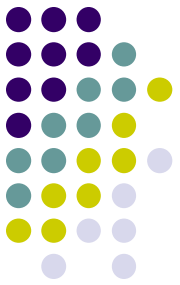
---





# Notions de base

- **Domaine**
  - **Relation**
  - **Attribut**
  - **Clé**
  - **Schéma d'une relation (intention vs Table : extension)**
  - **Clé étrangère (intégrité référentielle)**
- 
- **Algèbre relationnelle**
  - **Normalisation relationnelle**



# Algèbre relationnelle

- **Opérateurs Relationnels**

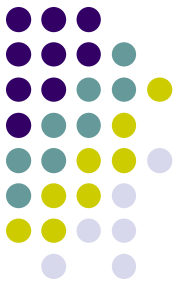
- **Opérateurs de Base**

- Union  $\cup$
    - Différence  $-$
    - Projection  $\Pi$
    - Restriction  $\sigma$
    - Produit Cartésien  $\times$

- **Opérateurs Additionnels**

- Intersection  $\cap$
    - Jointure  $\bowtie$
    - Produit  $\div$

- Toute requête est exprimée grâce à la composition d'opérateurs relationnels



# Exemple 1 de requête

- **Exemple de Base :**

Buveurs(NB, Nom, Prénom, Ville)

Abus(#NV, #NB, Quantité, Date)

Vins(NV, Cru, Millésime, Région, Couleur)

**Exemple 1 :**

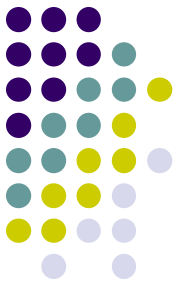
***"Donner le cru des vins rouges de millésime 2007"***

1. Temp =  $\sigma(\text{Vins} / \text{Couleur} = \text{"Rouge"} \wedge \text{Mill} = 2007)$

2. R =  $\Pi(\text{Temp} / \text{Vins.Cru})$

OU

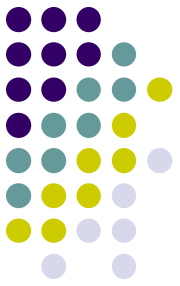
R =  $\Pi(\sigma(\text{Vins} / \text{Couleur} = \text{"Rouge"} \wedge \text{Mill} = 2007) / \text{Vins.Cru})$



# Exemple 2 de requête

- *"Donner les noms et prénoms des buveurs parisiens buvant du Mâcon 2007 en quantité supérieure à 10"*
- $R = \Pi(\quad$   
     $(\bowtie$   
         $(\bowtie$   
             $\sigma(\quad \text{Buveurs} / \text{Ville} = \text{'Paris'}),$   
             $\sigma(\quad \text{Abus} / \text{Qte} > 10) /$   
             $\text{Buveurs.NB} = \text{Abus.NB}),$   
             $\sigma(\quad \text{Vins} / \text{Cru} = \text{"Mâcon"} \wedge \text{Mill} = 2007),$   
             $\text{Abus.NV} = \text{Vins.NV}) /$   
     $\text{Buveurs.Nom, Buveurs.Prénom})$





# Autres exemples

- ***"Supprimer les vins blancs"***

$\text{Vins} = \text{Vins} - \sigma(\text{Vins} / \text{Vins.Couleur}=\text{BLANC})$

- ***"Insérer un vin rouge de Jurançon 2009 avec le numéro 105"***

$\text{Vins} = \text{Vins} \cup \{ ( 105, \text{"Jurançon"}, 2009, \text{ROUGE}) \}$

# Architecture d'un SGBD

- Introduction et rappels
- Disques et organisation
- Optimisation de requêtes

