Base de données et structures de données

CM1 : Base de données et SGBDR

Mickaël Martin Nevot

V1.0.2

Cette œuvre de Mickaël Martin Nevot est mise à disposition sous licence Creative Commons Attribution - Utilisation non commerciale - Partage dans les mêmes conditions.

BY-NC-SA

Base de données et structures de données

I. Prés.
II. BD et SGBD
III. DF et FN
IV. Merise
V. LDD
VI. LMD
VII. LCT
VIII. Droits
IX. LDSP
X. SQL avancé

 (cc) BY-NC-SA

Base de données (BD) ?

- **Ensemble de données modélisant un univers (domaine)**
- Outil de stockage structuré et de consultation correspondant
- **Système de gestion de base de données (SGBD)** :
  - Logiciel de manipulation et de stockage de BD
- Administrateur de BD : est responsable (en permanence) du bon fonctionnement des serveurs de bases de données

BdD, BDD : base de données
DB : *database*
ABD : administrateur de base de données
DBA : *database administrator*

SGBD : système de gestion de base de données
DBMS : *database-management system*

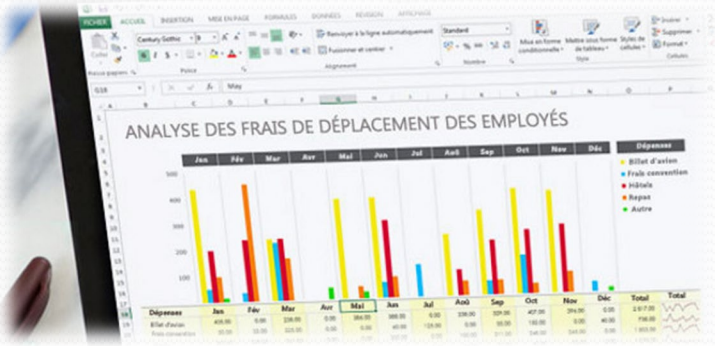
Banque de données : collection de bases de données

 BY-NC-SA

Avant les bases de données

- Utilisation de fichiers non informatisés

- Tableur (Microsoft Excel)



- Autres simples fichiers informatisés

 BY-NC-SA

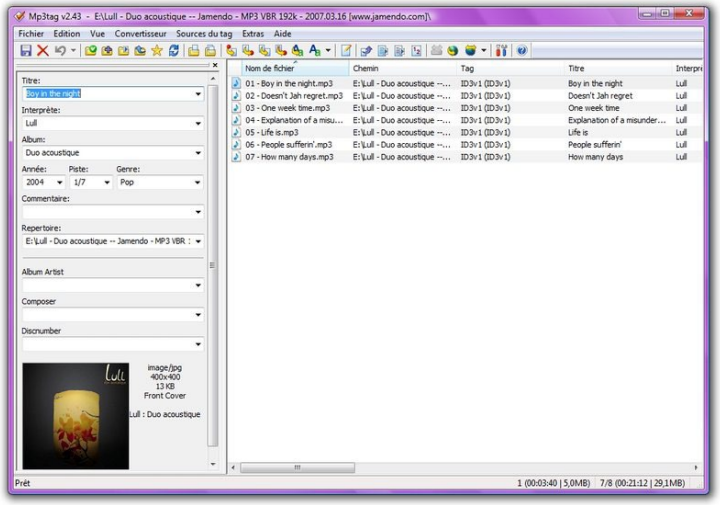
Avant les bases de données

- Limites des approches traditionnelles :
  - Complexité (exponentielle)
  - Redondance (duplication des données)
  - Coûts importants de développement et de maintenance
  - Manque de flexibilité
  - Manque de sécurité
- Concepts fondamentaux (partagés avec une BD) :
  - Organisation des données sur disque (ou autres stockages)
  - Procédé de récupération des données

BY-NC-SA

Intérêts d'une BD ?

- Stocker de (très) **gros volumes de données** durablement
- **Protéger** les données (avaries, piratage)
- **Simplifier l'accès aux données**
- Diminuer (drastiquement) les **temps de recherche**
- Prise en compte de liens **(dépendances)** entre les données
- Plusieurs **utilisateurs simultanés**



 (cc) BY-NC-SA

Historique

- Modèles de données :
  - 1960 : modèle hiérarchique
  - 1970 : modèle en réseau
  - 1980 : **modèle relationnel**
    - 1995 : PostgreSQL
  - 1990 : modèle objet
- **SQL** (norme) :
  - 1986 : SQL-86
  - 1989 : SQL-1
  - 1992 : **SQL2**
  - 1999 : SQL3

CODASYL : organisme à l'origine des BD

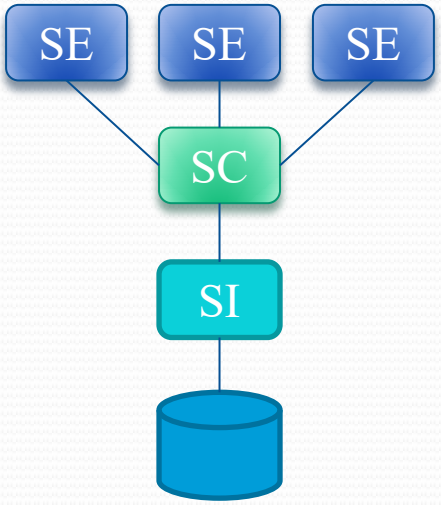
Inventé par E.F.CODD

SQL *(structured query language)* : langage de requête structurée, 1974

BY-NC-SA

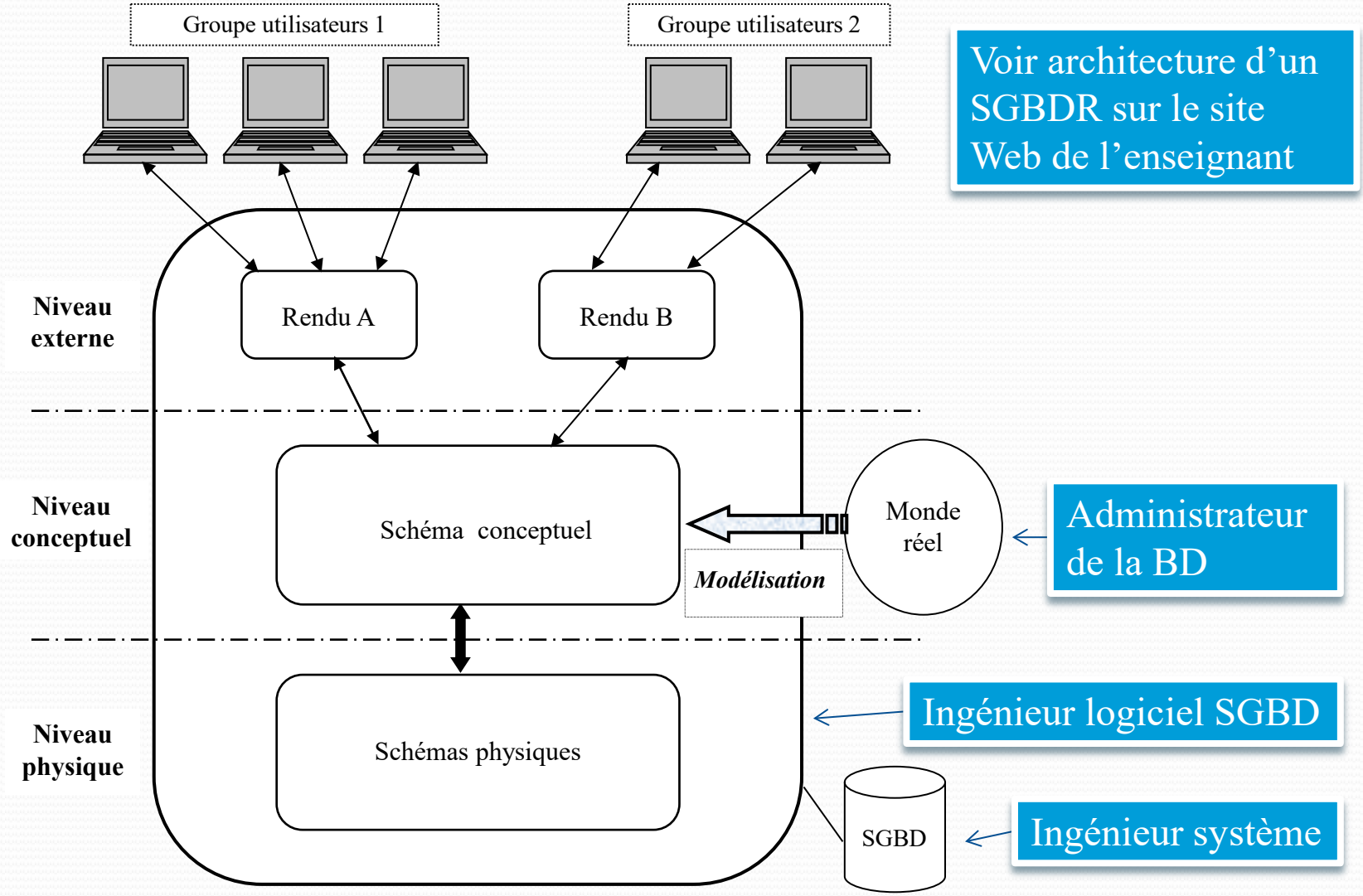
Architecture ANSI-SPARC

- Schéma externe (SE) :
  - Comment l'utilisateur voit les données
- Schéma **conceptuel** (SC) :
  - Structure logique des données
- Schéma interne (SI) :
  - Structure **physique** de stockage des données



(cc) BY-NC-SA

Architecture d'un SGBD



 (cc) BY-NC-SA

Fonctionnalités d'un SGBD

- **Description des données** : redondance minimale
- **Intégrité de la base** : cohérence avec la réalité transcrite
- **Indépendance des données** : indépendance applicative
- **Sécurité de fonctionnement** : journalisation
- **Administration et contrôle** : privilèges, optimisation
- **Partage des données** : transactions (sérialisées)
- **Souplesse d'accès aux données** : langages déclaratifs

Un **langage déclaratif** suit un paradigme produisant des résultats contextuellement indépendants

 (cc) BY-NC-SA

Schéma conceptuel

- Représentation de concepts sémantiquement liés entre eux :
  - Exprimer un besoin
  - Traduire le monde réel (modélisation)

Représentation abstraite d'un système qui facilite l'étude et la compréhension du système et permet de le simuler. Vue subjective, décomposée mais pertinente de la réalité. Représentation d'un système dans un autre monde que celui du système

- Représentation « graphique » décrivant une base de données :
  - Modèle hiérarchique :
    - Défauts : redondance, dissymétrie, parcours long
  - Modèle réseaux sémantiques (ou de réseau) :
    - Défaut : trop faible indépendance des données
  - Modèle objet :
    - Défaut : importante quantité de ressources nécessaire
  - **Modèle entité / association (relationnel)**

(cc) BY-NC-SA

Modèle relationnel

- Théorie mathématique : algèbre relationnelle (ensembliste)
- Tables **(relations)** avec colonnes **(attributs)** et lignes **(tuples)**

Ou champs, rubriques

Ou enregistrements, fiches

	Attributs					
R = (	A_1	A_2	…	A_n	)	Intention
						Extension
N-uplet ou tuple	D_1	D_2	…	D_n	}	
						↓
	Domaine (D_i)				→	iR - Instance de R

Domaine : ensemble nommé de valeurs non vide que peut prendre un attribut

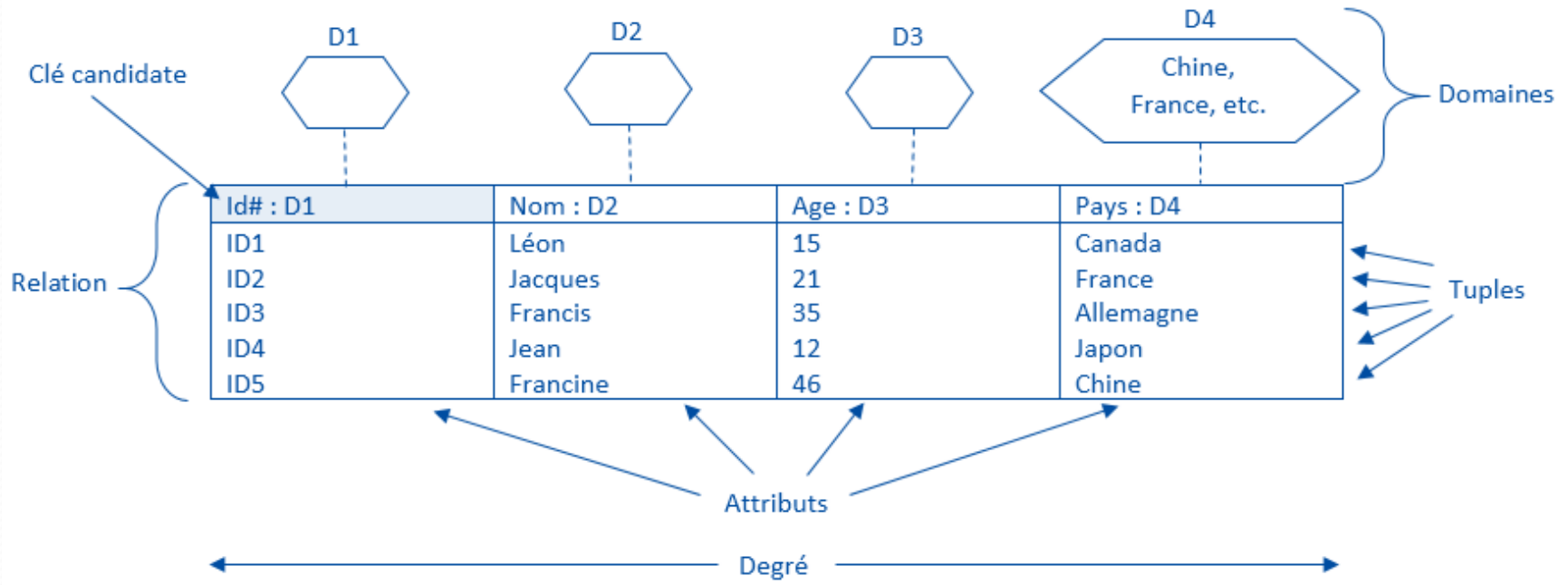
Intention : structure d'une base de données

Extension : données d'une base de données

(cc) BY-NC-SA

Modèle relationnel

- **Degré** : nombre d'attributs d'une relation
- **Cardinalité** : nombre de tuples d'une relation
- **Clef candidate** (potentielle) : ensemble des données permettant d'indexer chaque ligne de manière différenciée



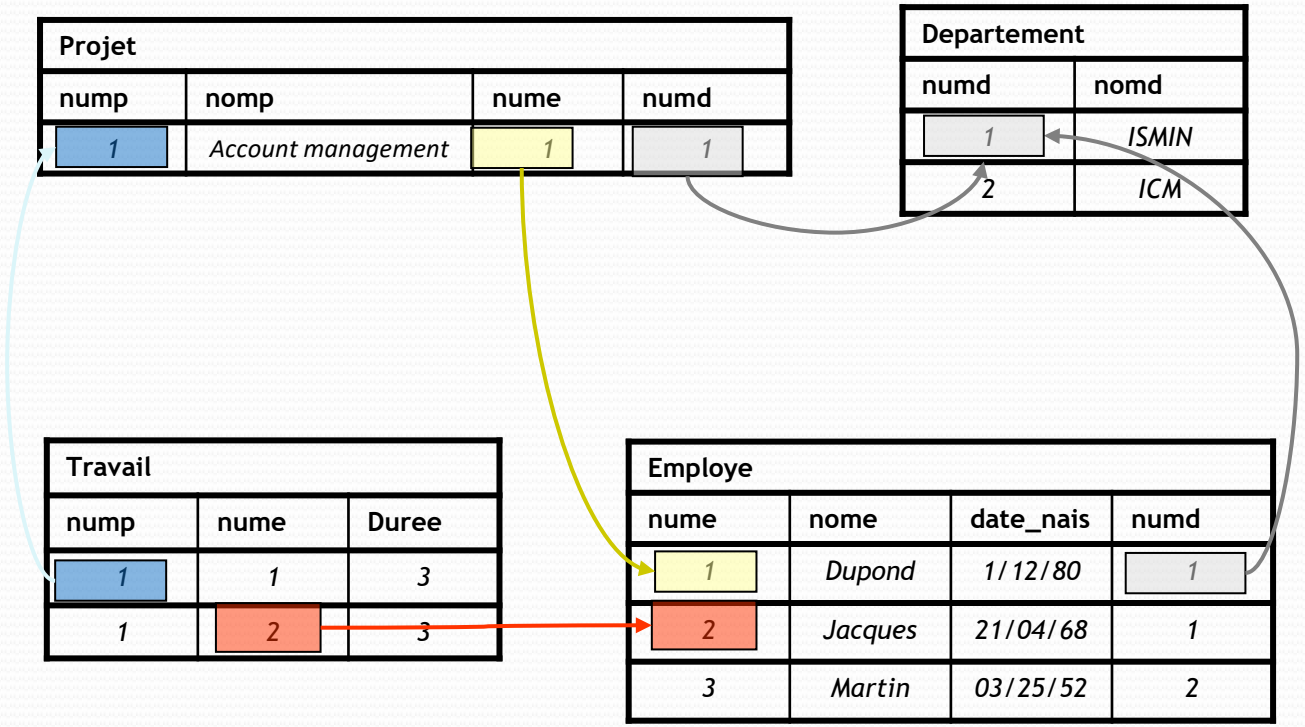
Id# : D1	Nom : D2	Age : D3	Pays : D4
ID1	Léon	15	Canada
ID2	Jacques	21	France
ID3	Francis	35	Allemagne
ID4	Jean	12	Japon
ID5	Francine	46	Chine

(cc) BY-NC-SA

Modèle relationnel

Voir les 12 règles de Codd sur le site Web de l'enseignant

- Navigation grâce à SQL
- Associé à la théorie de la normalisation
- **SGBDR** : SGBD relationnel



(cc) BY-NC-SA

Clefs

- **Clef primaire** : Obligatoire
  - **Une seule par relation** (clef candidate retenue comme primaire)
  - **Simple** (un seul attribut) ou **composée** (plusieurs attributs)
  - **Unique** et **non nulle**
- **Clef étrangère** :
  - Clef primaire d'une autre relation de la BD

Employe

nume	nome	daten	numd
1	*Dupond*	*1/12/80*	1
2	*Jacques*	*21/04/68*	1
3	*Martin*	*03/25/52*	2

Departement

numd	nomd
1	*ISMIN*
2	*ICM*

BY-NC-SA

Principaux SGBDR

 BY-NC-SA

Pourquoi PostgreSQL ?

- Un **SGBDR** très utilisé
- Préinstallé/proposé par beaucoup d'offres d'hébergements
- Facilité de déploiement et de prise en main
- Libre et gratuit
- S'intègre très bien avec Linux, Apache HTTP Server, PHP…

Apache HTTP Server : Serveur HTTP libre et très répandu

- Inconvénients de PostgreSQL :
  - Ne respecte pas complètement la norme SQL
  - Sauvegardes et réplications peu évoluées

(cc) BY-NC-SA

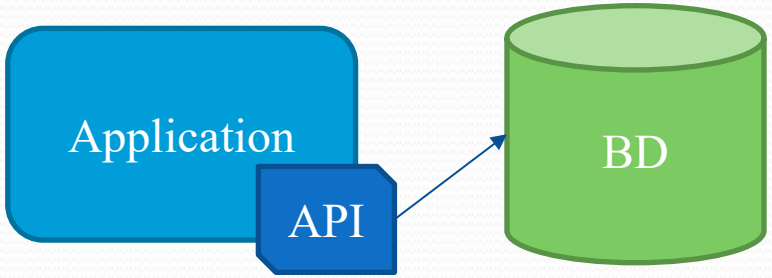
Utilisation d'un SGBDR

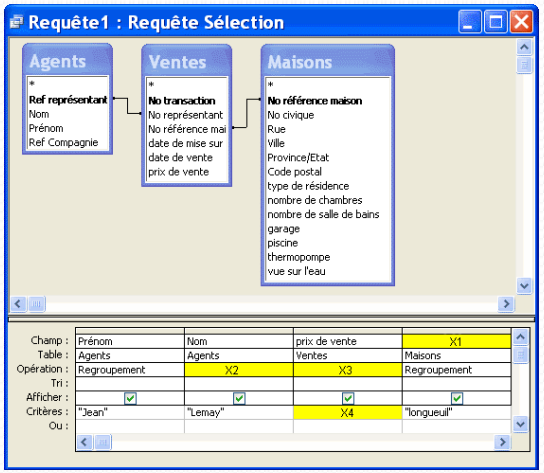
- **L3G** (langage de troisième génération) :
  - Procédural, séquentiel (C, Java, PHP)
  - Inversion de commandes avec une **API**

Préserve l'indépendance physique/logique

- **L4G** (Langage de quatrième génération) :
  - SGBDR intégré (Microsoft Access)
  - Générateur d'application

Rapide et simple d'utilisation pour de petites applications

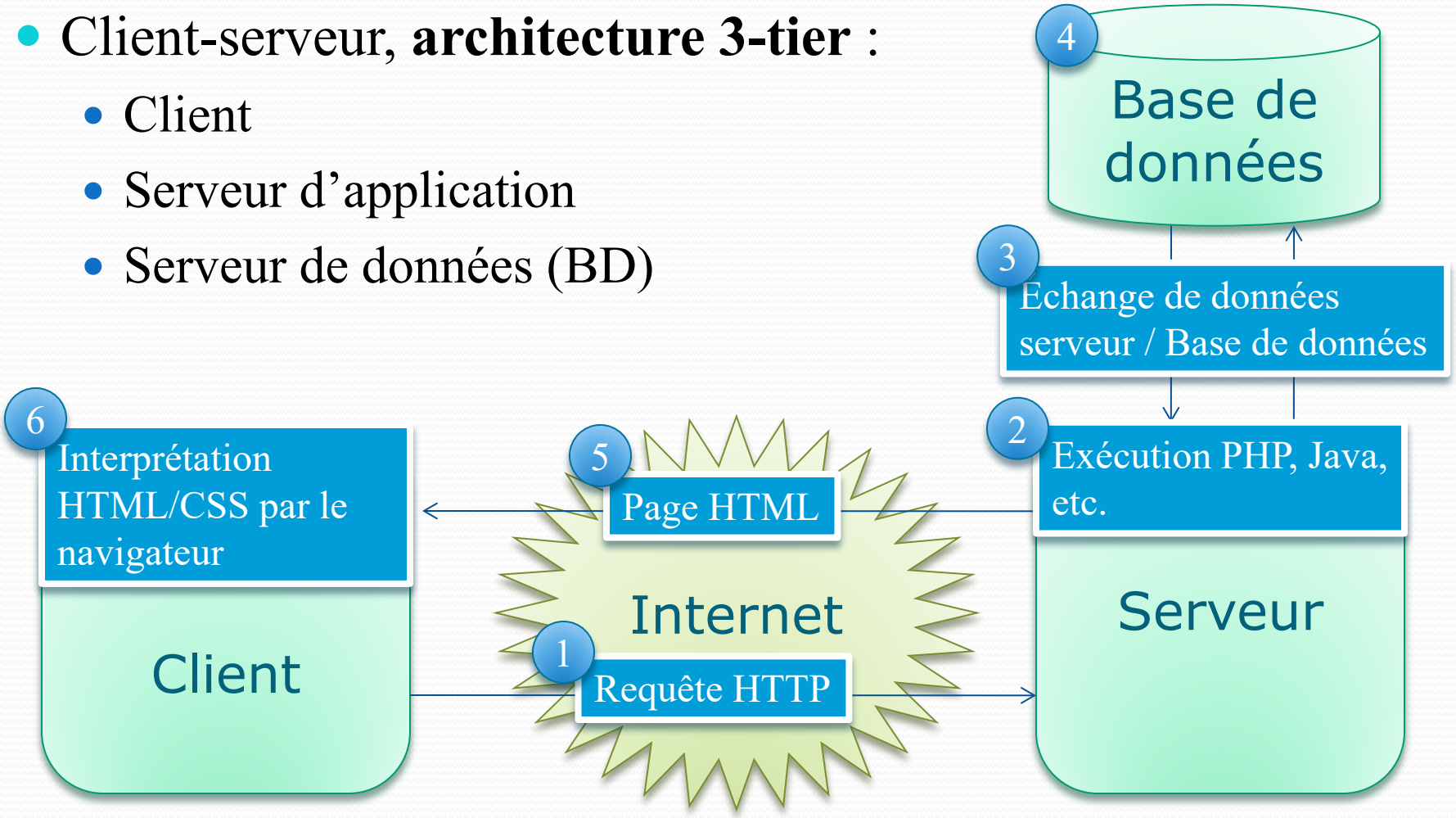




(cc) BY-NC-SA

Utilisation d'un SGBDR

- Client-serveur, **architecture 3-tier** :
  - Client
  - Serveur d'application
  - Serveur de données (BD)



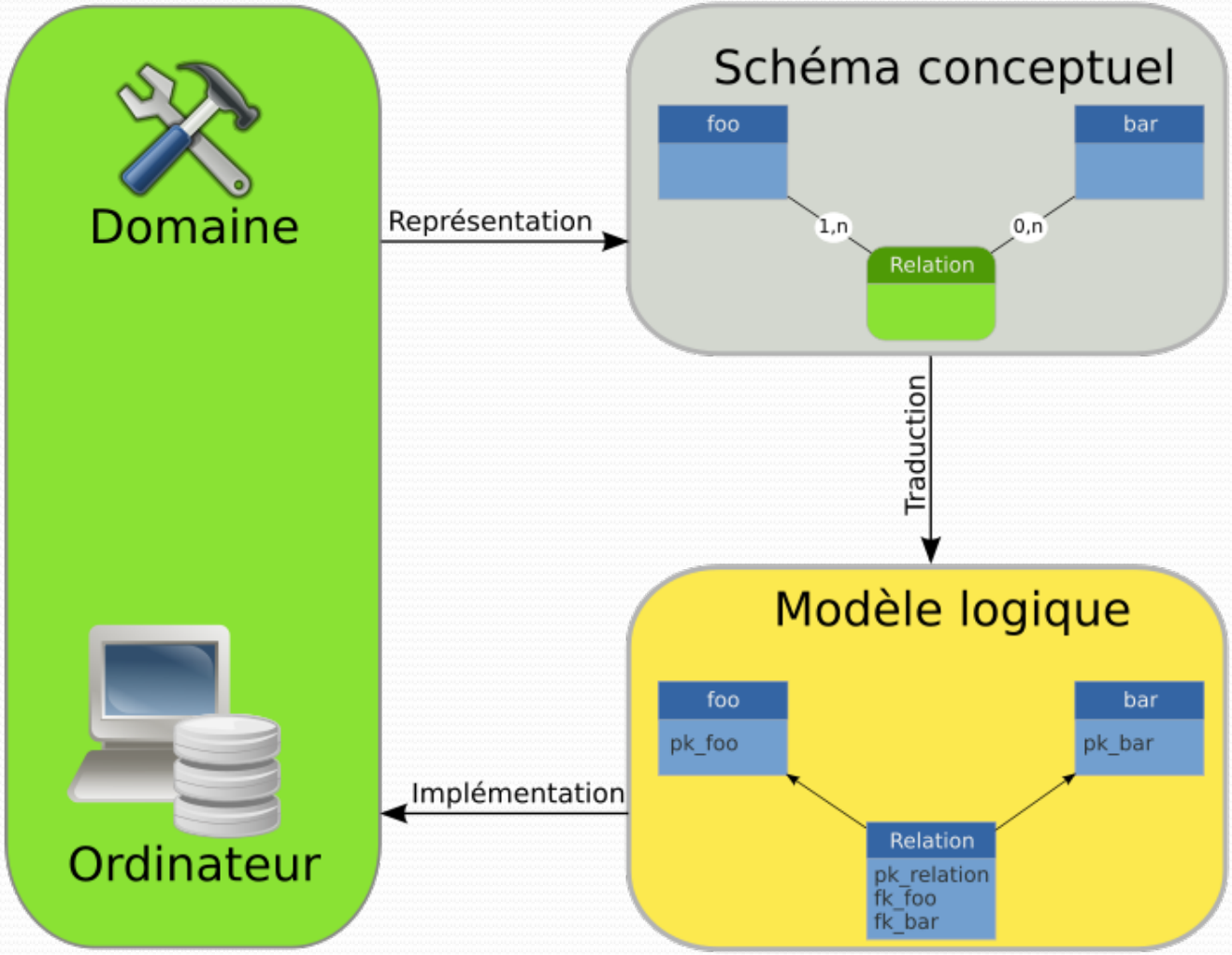
(cc) BY-NC-SA

Cycle de vie d'une BD

Conception	Merise
Implantation	Langage de définition de données (LDD)
Utilisation	Langage de manipulation de données (LMD)
Maintenance	Corrective, adaptative, perfective

(cc) BY-NC-SA

Conception d'une BD



Crédits

Auteur

Mickaël Martin Nevot
mmartin.nevot@gmail.com

- Laurent Carmignac

Carte de visite électronique

Relecteurs

Cours en ligne sur : www.mickael-martin-nevot.com