



كلية العلوم
والتقنيات - مراكش
FACULTÉ DES SCIENCES
ET TECHNIQUES - MARRAKECH

Faculté des Sciences et Techniques
Marrakech



Module Bases de Données

Modélisation des données

Prof. Sara Qassimi

LE CYCLE D'ABSTRACTION		
Niveaux	DONNEES	TRAITEMENTS
CONCEPTUEL	M C D	M C T
QUOI	<i>Modèle conceptuel des données</i> Signification des informations sans contraintes techniques, organisationnelles ou économiques. Modèle entité – association	<i>Modèle conceptuel des traitements</i> Activité du domaine sans préciser les ressources et leur organisation
ORGA-NISATIONNEL	M O D	M O T
QUI, OU, QUAND	<i>Modèle organisationnel des données</i> Signification des informations avec contraintes organisationnelles et économiques. (Répartition et quantification des données ; droit des utilisateurs)	<i>Modèle organisationnel des traitements</i> Fonctionnement du domaine avec les ressources utilisées et leur organisation (répartition des traitements sur les postes de travail)
LOGIQUE	M L D	M L T
COMMENT	<i>Modèle logique des données</i> Description des données tenant compte de leurs conditions d'utilisation (contraintes d'intégrité, historique, techniques de mémorisation). Modèle relationnel	<i>Modèle logique des traitements</i> Fonctionnement du domaine avec les ressources et leur organisation informatique.
PHYSIQUE	M P D	M P T
COMMENT	<i>Modèle physique des données</i> Description de la (ou des) base(s) de données dans la syntaxe du Système de Gestion des données (SG.Fichiers ou SG Base de Données) Optimisation des traitements (indexation, dénormalisation, triggers).	<i>Modèle physique des traitements</i> Architecture technique des programmes

La modélisation

- Les deux modèles dominant Merise sont :

- le modèle relationnel : MR (qui correspond aux SGBD-R) et le modèle entité-association : MEA (qui est indépendant du type de SGBD utilisé). Ces deux modèles correspondent à 2 langages différents.
- Les schémas **entité-relation** et les diagrammes de classe UML peuvent être utilisés comme autres langages à peu près équivalents au **MEA**.

- La méthode MERISE distingue entre 3 types de modèles selon des critères méthodologiques :

- le modèle conceptuel des données : **MCD**, le modèle logique des données : **MLD** et le modèle physique des données : **MPD**.

- L'usage tend à rendre équivalents:

- **MCD** et **MEA**,
- **MLD** et **MR**,
- **MPD** et **SQL**.

La « jungle » des modèles !			
Méthode	MCD	MLD	MPD
Langage	MEA, schéma E-R, UML	MR	SQL

- Le **MCD** est du niveau de l'analyse fonctionnelle et est adapté à la maîtrise d'ouvrage (MOA). Le MLD est du niveau de l'analyse organique et est adapté à la maîtrise d'œuvre (MOE).

Les différents modèles et leurs relations

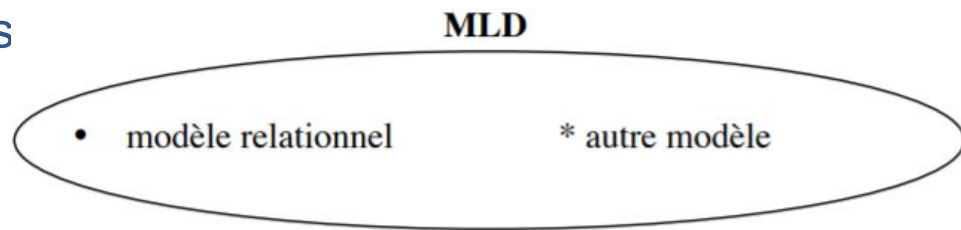
MCD

- 
- modèle Entité-Association * autre modèle

- **MCD et Modèle Entité-Association**

- Le **MCD**, c'est l'ensemble des modèles qui intègrent les contraintes conceptuelles définies par Merise.
- Parmi ces modèles, le plus couramment utilisé est le modèle **Entité-Association**.
- Le **MCD** est donc une abstraction (un modèle abstrait), tandis que le modèle **Entité-Association** est un modèle concret. C'est une instance possible du **MCD**. Une autre instance possible d'un **MCD** peut être réalisée avec le formalisme des diagrammes de classes UML.
- Toutefois, quand on parle du **MCD**, le plus souvent, on parle du modèle concret réalisé pour intégrer les contraintes conceptuelles définies par Merise :
donc on parle d'un modèle **Entité-Association**.

Les différents modèles et leurs relations



● **MLD et modèle relationnel**

- La notion de **MLD** correspond à l'ensemble des modèles qui intègrent les contraintes organisationnelles et logiques définies par Merise.
- Parmi ces modèles, le plus couramment utilisé est le **modèle relationnel**.
- La notion de **MLD** est donc une abstraction (un modèle abstrait), tandis que le **modèle relationnel** est un modèle concret.
- Toutefois, quand on parle du **MLD**, le plus souvent, on parle du modèle concret réalisé pour intégrer les contraintes organisationnelles et logiques définies par Merise :
donc on parle d'un **modèle relationnel**.

MCD : Modèle Conceptuel des données

- Le MCD modélise les données (les informations) sans tenir compte des contraintes
 - techniques, • organisationnelles, • économiques.
- Concrètement, cela veut dire qu'au niveau du MCD :
 - Aucune contrainte technique ne sera prise en compte : on n'utilise pas nécessairement un SGBD relationnel.
 - Toutes les données de l'entreprise seront prises en compte (pas de répartition selon les sites ou les applications : pas de contraintes organisationnelles).

Modèle Conceptuel des données- Schéma entités-associations

- L'élaboration du MCD passe par les étapes suivantes :
 - la mise en place de **règles de gestion** ;
 - l'élaboration du **dictionnaire des données** ;
 - la recherche des **dépendances fonctionnelles** entre ces données ;
 - l'élaboration du MCD
 - (création des **entités** puis des **associations** puis ajout des **cardinalités**).

Modèle Conceptuel des données- Schéma entités-associations

Recueillir les besoins des futurs utilisateurs de votre application. Et à partir de ces besoins, il faut établir les règles de gestion des données à conserver:

Règles de gestion

Exemple d'un développeur qui doit informatiser le SI d'une bibliothèque. On lui fixe les règles de gestion suivantes :

- pour chaque livre, on doit connaître le titre, l'année de parution, un résumé et le type (roman, poésie, science-fiction...);
- un livre peut être rédigé par un ou plusieurs auteurs dont on connaît le nom, le prénom, la date de naissance et le pays d'origine ;
- chaque exemplaire d'un livre est identifié par une référence composée de lettres et de chiffres et ne peut être paru que dans une et une seule édition ;
- un inscrit est identifié par un numéro et on doit mémoriser son nom, prénom, adresse, téléphone et adresse e-mail ;
- un inscrit peut faire zéro, un ou plusieurs emprunts qui concernent chacun un et un seul exemplaire. Pour chaque emprunt, on connaît la date de l'emprunt et la date de retour .

Modèle Conceptuel des données- Schéma entités-associations

Le dictionnaire des données

document qui regroupe toutes les données (et qui figureront donc dans le MCD).

le **code mnémonique** : il s'agit d'un libellé désignant une donnée (par exemple «*titre_1*» pour le titre d'un livre) ;

la **désignation** : il s'agit d'une mention décrivant ce à quoi la donnée correspond (par exemple «*titre du livre*») ;

le **type de donnée** :

Alphabétique : lorsque la donnée est uniquement composée de caractères alphabétiques (de 'A' à 'Z' et de 'a' à 'z'),

Numérique : lorsque la donnée est composée uniquement de nombres (entiers ou réels),

Alphanumérique : lorsque la donnée peut être composée à la fois de caractères alphabétiques et numériques,

Date : lorsque la donnée est une date (au format AAAA-MM-JJ),

Booléen : Vrai ou Faux ;

la **taille** : elle s'exprime en nombre de caractères ou de chiffres.

Dans le cas d'une date au format AAAA-JJ-MM, on compte également le nombre de caractères, soit 10 caractères. Pour ce qui est du type booléen, nul besoin de préciser la taille ;

● Exemple de Dictionnaire des données

Code mnémonique	Désignation	Type	Taille	Remarque
id_i	Identifiant numérique d'un inscrit	N		
nom_i	Nom d'un inscrit	A	30	
prenom_i	Prénom d'un inscrit	A	30	
rue_i	Rue où habite un inscrit	AN	50	
ville_i	Ville où habite un inscrit	A	50	
cp_i	Code postal d'un inscrit	AN	5	
tel_i	Numéro de téléphone fixe d'un inscrit	AN	15	
tel_port_i	Numéro de téléphone portable d'un inscrit	AN	15	
email_i	Adresse e-mail d'un inscrit	AN	100	
date_naissance_i	Date de naissance d'un inscrit	Date	10	Au format AAAA-JJ-MM
id_l	Identifiant numérique d'un livre	N		
titre_l	Titre d'un livre	AN	50	
annee_l	Année de parution d'un livre	N	4	
resume_l	Résumé d'un livre	AN	1000	
ref_e	Code de référence d'un exemplaire d'un livre	AN	15	Cette référence servira également d'identifiant dans ce système
id_t	Identifiant numérique d'un type de livre	N		
libelle_t	Libellé d'un type de livre	AN	30	

Modèle Conceptuel des données- Schéma entités-associations

Les Dépendance Fonctionnelle

Soit deux propriétés (ou données) P1 et P2. On dit que P1 et P2 sont reliées par une **dépendance fonctionnelle** (DF) si et seulement si une **occurrence** (ou valeur) de P1 permet de connaître une et une seule occurrence de P2.

Cette dépendance est représentée comme ceci :

$P1 \rightarrow P2$

On dit que P1 est la **source** de la DF et que P2 en est le **but**.

Par ailleurs, plusieurs données peuvent être source comme plusieurs données peuvent être but d'une DF.

Exemples :

$P1, P2 \rightarrow P3$

$P1 \rightarrow P2, P3$

$P1, P2 \rightarrow P3, P4, P5$

En reprenant les données du dictionnaire précédent, on peut établir les DF suivantes :

$id_em \rightarrow date_em, delais_em, id_i, ref_e$

$id_i \rightarrow nom_i, prenom_i, rue_i, ville_i, cp_i, tel_i, tel_port_i, email_i, date_naissance_i$

Modèle Conceptuel des données- Schéma entités-associations

Les entités

Un objet, une chose concrète ou abstraite qui peut être reconnue distinctement et qui est caractérisée par son unicité.

Chaque **entité** est unique et est décrite par un ensemble de propriétés encore appelées attributs ou caractéristiques.

Une des **propriétés** de l'entité est **l'identifiant**:

Cette propriété doit posséder des occurrences uniques et doit être source des dépendances fonctionnelles avec toutes les autres propriétés de l'entité.

Ex: On utilise une donnée de type entier qui s'incrémente pour chaque occurrence, ou encore un code unique spécifique du contexte.

Le formalisme d'une entité est le suivant :

Nom de l'entité
<u>identifiant</u>
propriété1
propriété2
...

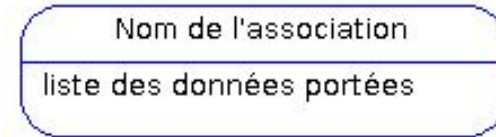
Modèle Conceptuel des données- Schéma entités-associations

Les associations

Une **association** définit un **lien sémantique** entre une ou plusieurs **entités**.

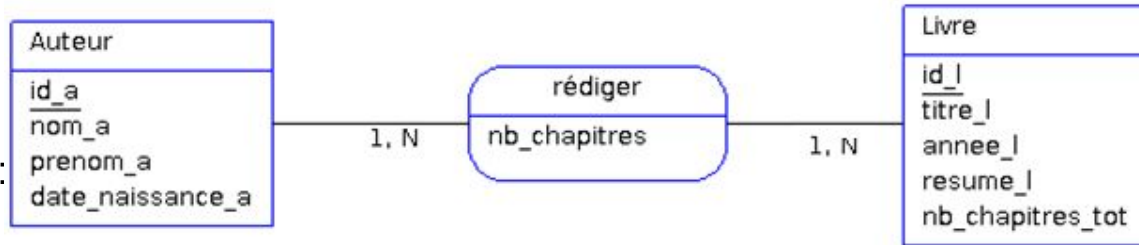
la définition de liens entre entités permet de traduire une partie des règles de gestion qui n'ont pas été satisfaites par la simple définition des entités.

Le formalisme d'une association est le suivant :



Les cardinalités

Une cardinalité est définie comme ceci :
minimum, maximum



le nombre de fois minimal et maximal d'interventions d'une entité dans une association

Les cardinalités les plus répandues sont les suivantes : **0,N ; 1,N ; 0,1 ; 1,1.**

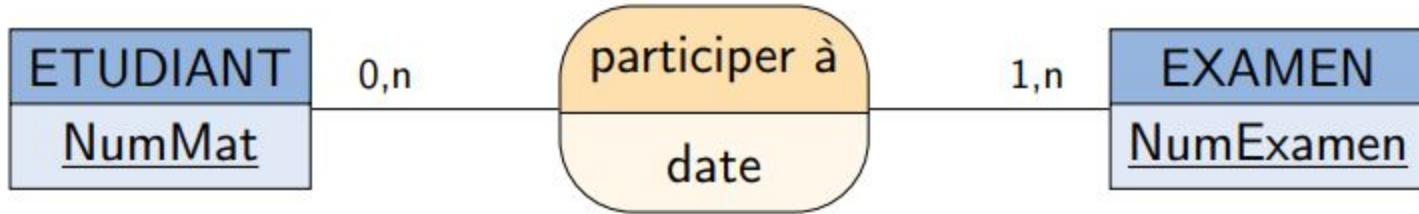
Un auteur rédige au moins un ou plusieurs livres.

Modèle Conceptuel des données- Schéma entités-associations

Les associations

Il existe deux types d'association en fonction des cardinalités entre les entités concernées. Il est donc important de déterminer les cardinalités des associations lors de la conception d'une base de données car cela a un impact sur la structure des

données.



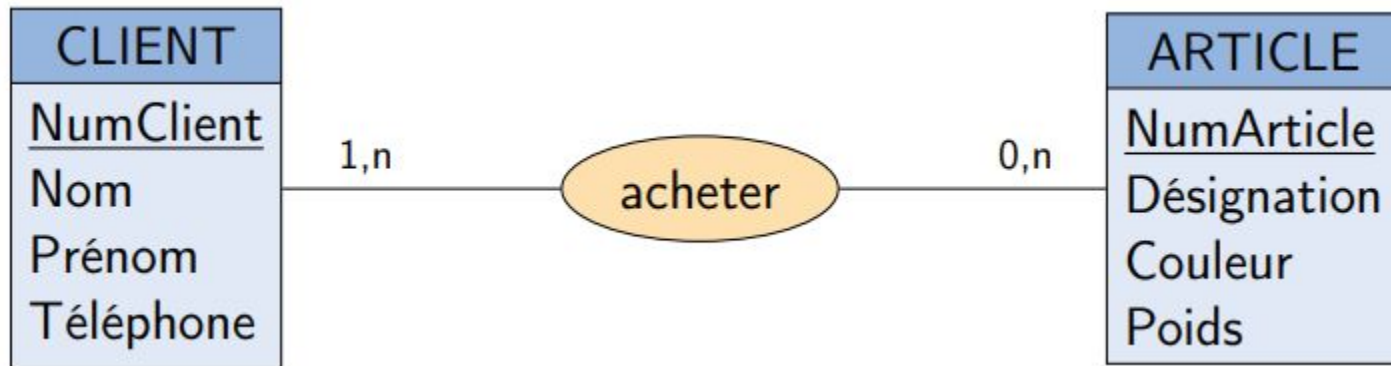
L'association participer explicite le lien entre les entités ETUDIANT et EXAMEN. Au minimum aucun étudiant participe à un examen et au maximum n étudiants participent à un examen, ce qui explique la cardinalité 0, n. À un examen participe au minimum un étudiant et au maximum n, d'où la cardinalité 1, n.

Modèle Conceptuel des données- Schéma entités-associations

Association non hiérarchique

Une association est dite non hiérarchique lorsque les cardinalités maximales sont à n pour les deux

entités.



Dans la relation non hiérarchique, un client peut acheter 1 à n articles et un article peut être achetée par 0 ou n client.

Modèle Conceptuel des données- Schéma entités-associations

Association hiérarchique

Une association hiérarchique est une association où pour une entité, la cardinalité maximale est 1 et de l'autre entité, la cardinalité maximale est à n.

L'entité ayant le maximum à 1 est appelée entité «fils». L'entité ayant un max à n est appelée entité «père», car «un fils n'a qu'un seul père, et un père peut avoir plusieurs fils ». Connaissant l'occurrence de l'entité «fils », il est possible de retrouver l'occurrence de l'entité «père».

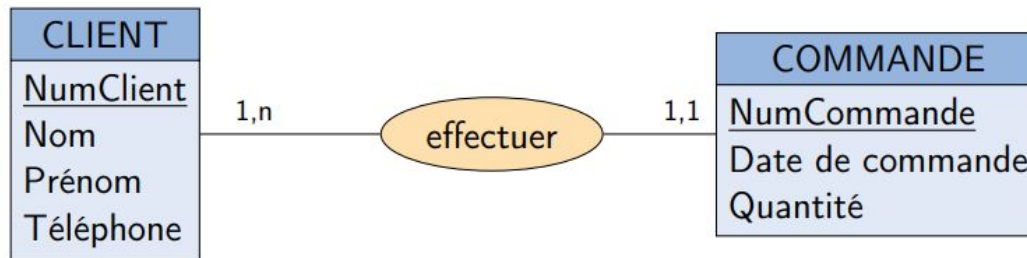
Association faible, forte

Si la cardinalité côté fils d'une association hiérarchique est (0,1), alors l'association est dite faible. Si la cardinalité du côté fils est (1,1) alors l'association est dite forte. Une association hiérarchique forte est aussi appelée Contrainte d'Intégrité Fonctionnelle (CIF). Il existe ainsi systématiquement un lien pour l'ensemble des occurrences de l'association.

Modèle Conceptuel des données- Schéma entités-associations

Association faible, forte

Dans la relation, toute commande est systématiquement rattachée à un client. Il existe forcément une commande pour un client car il est considéré comme client à partir du moment où il commande.

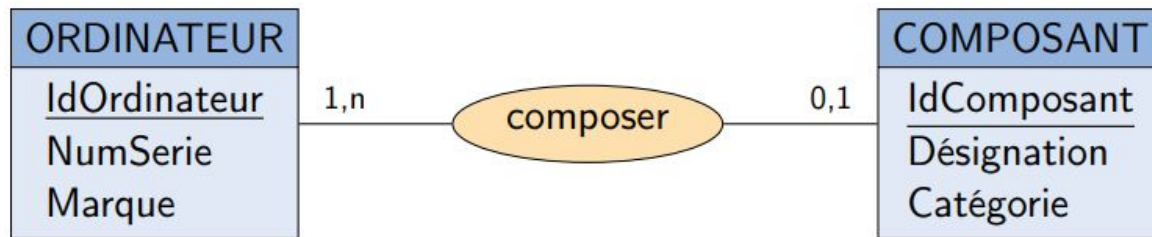


Exemple d'association hiérarchique.

Modèle Conceptuel des données- Schéma entités-associations

Association faible, forte

Cet exemple montre une relation pour une base de données pour un assembleur d'ordinateur. Un ordinateur est composé de 1 à N composant. Un composant peut, par contre, n'être utilisé dans aucun ordinateur.



Exemple d'association hiérarchique faible.

Modèle Conceptuel des données- Schéma entités-associations

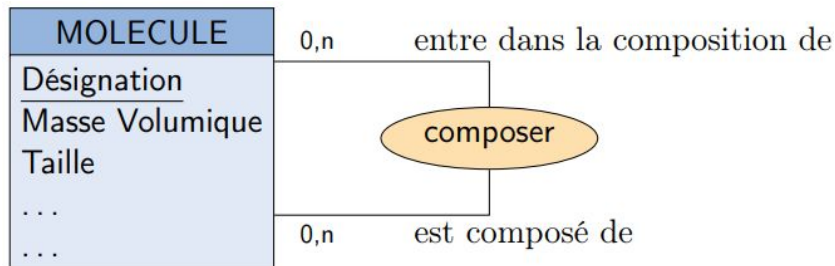
Association réflexive

Une association réflexive est une association hiérarchique reliant des occurrences de la même entité. Pour lire une association réflexive, il faut connaître le rôle attribué à chaque branche de l'association.

Association réflexive non hiérarchique

Une association réflexive non hiérarchique est une association reliant des occurrences de la même entité, pour lesquelles il n'existe pas de considération hiérarchique, c'est à dire l'association est non hiérarchique.

Dans l'association réflexive non hiérarchique suivante;
une molécule peut être composée de plusieurs molécules.



Modèle Conceptuel des données- Schéma entités-associations

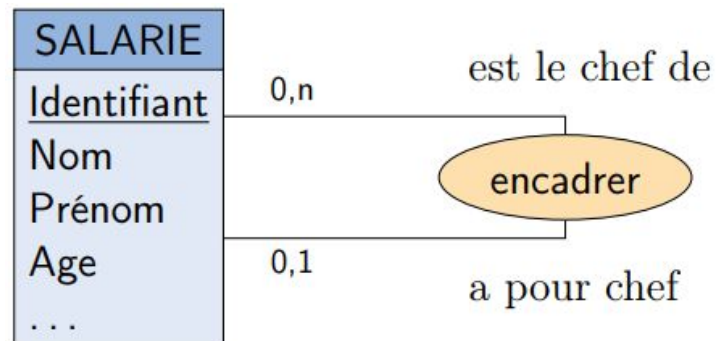
Association réflexive

Une association réflexive est une association hiérarchique reliant des occurrences de la même entité. Pour lire une association réflexive, il faut connaître le rôle attribué à chaque branche de l'association.

Association réflexive hiérarchique

Une association réflexive hiérarchique est une association hiérarchique reliant des occurrences de la même entité.

Dans l'association réflexive hiérarchique suivante;
un salarié a pour supérieur direct un autre salarié.



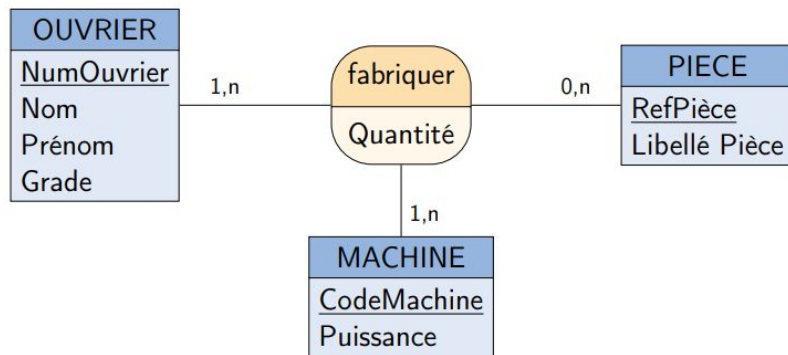
Modèle Conceptuel des données- Schéma entités-associations

Associations de dimension 3 ou plus (ternaires ou plus)

Une association peut relier 3 entités (ou plus) autour d'une même association. Une telle association n'a que pour but d'alléger le MCD mais n'est jamais obligatoire. Il faut donc faire très attention car il est très rare de faire une association ternaire (de dimension 3) sans se tromper.

Cette association ternaire n'est possible que si les trois entités sont indissociables. C'est-à-dire qu'il n'est pas possible d'utiliser cette relation pour définir "quel ouvrier travaille sur une machine ? ". En effet, dans cet exemple, tant qu'il n'y a de fabrication d'une pièce par un ouvrier (matérialisée à gauche par une relation et à droite par une entité), il n'est pas possible de faire le lien entre machine et ouvrier.

Pour connaître la pièce fabriquée par un ouvrier, il faut savoir sur quelle machine il a travaillé.

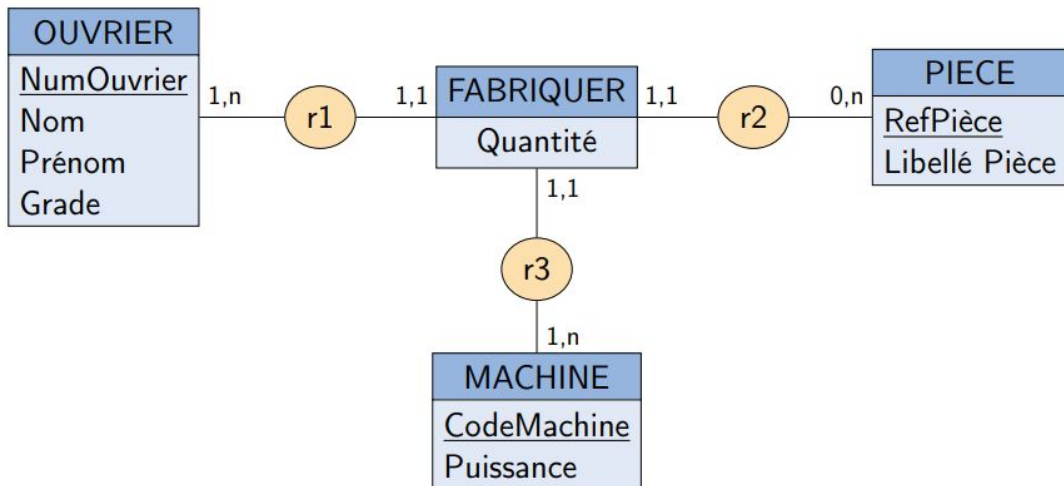


Exemple de relation ternaire qui sont équivalentes

Modèle Conceptuel des données- Schéma entités-associations

Associations de dimension 3 ou plus (ternaires ou plus)

Une association peut relier 3 entités (ou plus) autour d'une même association. Une telle association n'a que pour but d'alléger le MCD mais n'est jamais obligatoire. Il faut donc faire très attention car il est très rare de faire une association ternaire (de dimension 3) sans se tromper.

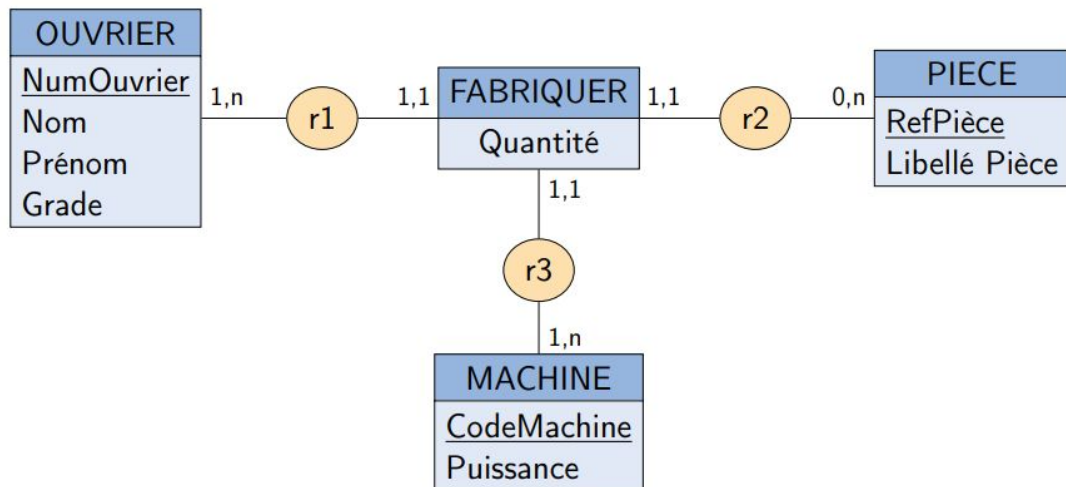


Exemple de relation ternaire qui sont équivalentes

Modèle Conceptuel des données- Schéma entités-associations

Associations de dimension 3 ou plus (ternaires ou plus)

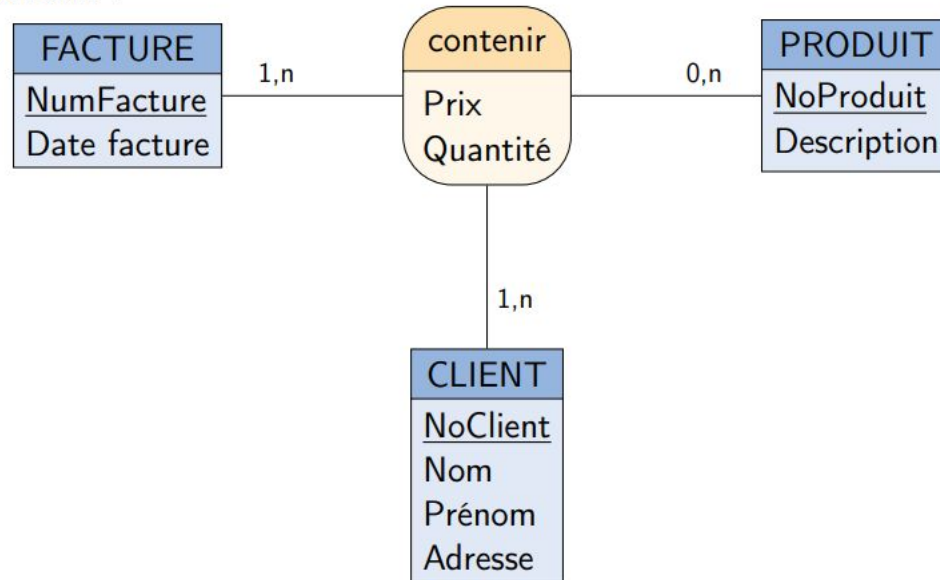
Une association peut relier 3 entités (ou plus) autour d'une même association. Une telle association n'a que pour but d'alléger le MCD mais n'est jamais obligatoire. Il faut donc faire très attention car il est très rare de faire une association ternaire (de dimension 3) sans se tromper.



Exemple de relation ternaire qui sont équivalentes

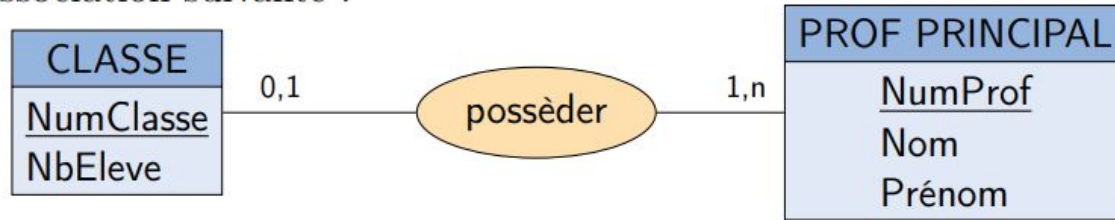
Exercice 1:

Est ce que cette association ternaire est juste ? Si oui expliquez pourquoi, sinon donner l'association qui est correcte :



Exercice 2:

Caractériser l'association suivante :



LE CYCLE D'ABSTRACTION		
Niveaux	DONNEES	TRAITEMENTS
CONCEPTUEL	M C D	M C T
QUOI	<i>Modèle conceptuel des données</i> Signification des informations sans contraintes techniques, organisationnelles ou économiques. Modèle entité – association	<i>Modèle conceptuel des traitements</i> Activité du domaine sans préciser les ressources et leur organisation
ORGA-NISATIONNEL	M O D	M O T
QUI, OU, QUAND	<i>Modèle organisationnel des données</i> Signification des informations avec contraintes organisationnelles et économiques. (Répartition et quantification des données ; droit des utilisateurs)	<i>Modèle organisationnel des traitements</i> Fonctionnement du domaine avec les ressources utilisées et leur organisation (répartition des traitements sur les postes de travail)
LOGIQUE	M L D	M L T
COMMENT	<i>Modèle logique des données</i> Description des données tenant compte de leurs conditions d'utilisation (contraintes d'intégrité, historique, techniques de mémorisation). Modèle relationnel	<i>Modèle logique des traitements</i> Fonctionnement du domaine avec les ressources et leur organisation informatique.
PHYSIQUE	M P D	M P T
COMMENT	<i>Modèle physique des données</i> Description de la (ou des) base(s) de données dans la syntaxe du Système de Gestion des données (SG.Fichiers ou SG Base de Données) Optimisation des traitements (indexation, dénormalisation, triggers).	<i>Modèle physique des traitements</i> Architecture technique des programmes

MOD: Modèle Organisationnel des Données

- Le modèle organisationnel des données permet d'apporter un certain nombre de précisions par rapport au modèle conceptuel des données. Ces précisions portent sur les 4 aspects suivants :
 - que faut-il informatiser ?
 - quantification et historicité des informations à mémoriser
 - répartition des données informatisées par unités organisationnelles
 - sécurisation des données

Problématique

Modélisation **conceptuelle** des données :

- **signification des informations** utilisées dans un **domaine d'activité**
- **sans tenir compte de contraintes** organisationnelles, économiques ou techniques

Modélisation **organisationnelle** des données : la **mémorisation**

- **choix des informations** à mémoriser "**informatiquement**"
- **volume** des informations (quantification)
- **durée de vie** et **historisation** des informations
- **répartition organisationnelle** des données informatisées dans l'organisation (l'entreprise) selon des **unités organisationnelles** (UOs)
- **sécurité des données** : modalités de partages/droits d'accès aux ²⁸niveaux des UOs et au niveau des acteurs de ces UOs ne prend pas compte **l'optimisation** de la **technologie de mémorisation informatique** retenue.

MOD: Modèle Organisationnel des Données

Données informatisables

- il s'agit simplement de faire la différence entre ce qui est informatisable et ce qui doit rester manuel.
- De fait, on peut être amené à supprimer des entités qui ne seraient pas informatisables.
- On peut aussi introduire de nouvelles données (lien entre un numéro, donnée informatisable, et le même numéro pouvant être lu sur une fiche papier).

MOD: Modèle Organisationnel des Données

Quantification

- L'objectif principal est de donner une estimation du "volume" du système d'information.
- Plusieurs examens sont à effectuer :
 - taille des propriétés : pour chaque entité, il convient de déterminer la nature des données.
- Dans une première approche, on peut classer ces données en trois catégories :
 - chaîne de caractères (n caractères),
 - numérique entier (n chiffres),
 - numérique fractionnaire (n chiffres pour la partie entière, m chiffres pour la partie décimale)
- nombre d'occurrences des entités : une association du type 1,n ne donne pas avec précision la cardinalité maximale ; celle-ci a pourtant une borne réaliste ;
- il conviendra pourtant d'en donner une estimation qui peut être la cardinalité maximale réelle, la cardinalité maximale à 95%, la cardinalité modale, la cardinalité moyenne.

Quantification MOD

- **Objectif :**

- **déterminer** en première approximation **le volume de données à mémoriser**

- **Pour cela :**

- **taille des propriétés**
- **nombre d'occurrences des entités et relations :**
 - mémoire immédiate/mémoire à long terme
 - durée de vie des propriétés
 - quantification des cardinalités
- **évaluation du volume global du MOD**

Prise en compte de la durée de vie des propriétés

• MOD sans durée de vie :

- Pour l'entité CLIENT, le chiffre d'affaire depuis le début de l'année est égale à la somme, sur toutes ses commandes passées depuis le début de l'année, du montant total de la commande (propriété éventuellement calculée elle aussi).

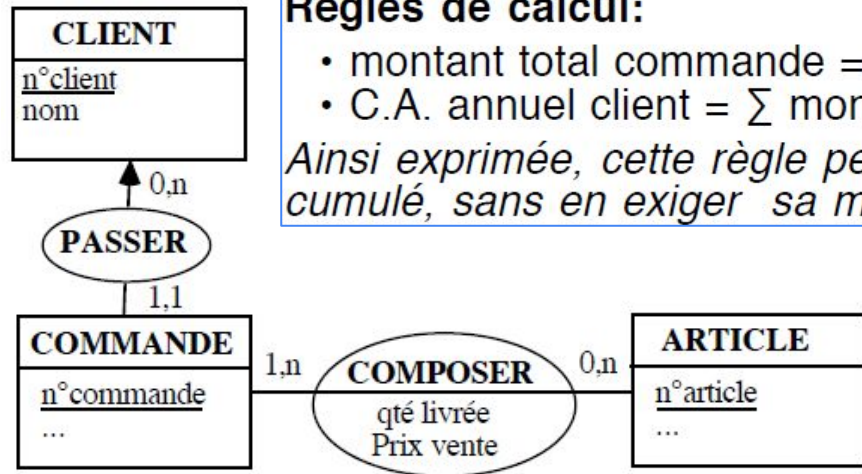
Durées de vie:

- CLIENT, ARTICLE : permanent
- COMMANDE, PASSER, COMPOSER : 3 mois

Règles de calcul:

- montant total commande = $\sum \text{qté livrée} \times \text{Prix vente}$
- C.A. annuel client = $\sum \text{montant total commande}$

Ainsi exprimée, cette règle permet toujours, sémantiquement de calculer ce C.A cumulé, sans en exiger sa mémorisation.



Prise en compte de la durée de vie des propriétés

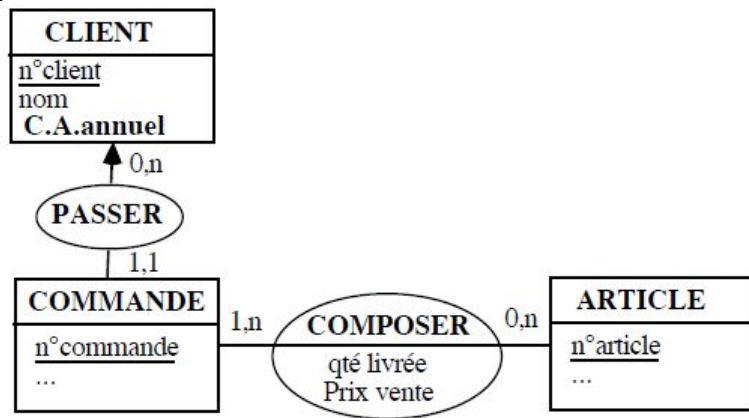
- **MOD avec durée de vie :**

Si la durée de vie des commandes est inférieure à un an (3 mois par exemple), il sera impossible de reconstituer ce cumul au-delà de la durée de vie des commandes

- chaque fois qu'une information est calculée à partir d'autres information **de durée de vie inférieure** à la sienne :

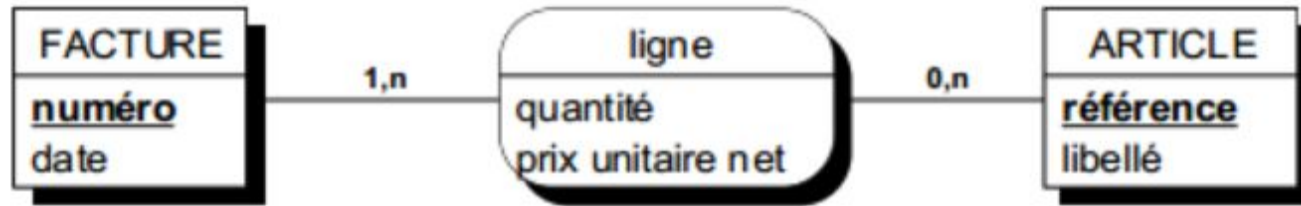
-> considérer cette information calculée comme une **information primaire**, et la modéliser comme une propriété d'une entité/relation.

- affecter cette propriété à une entité ou relation déjà existant
- créer une entité-type ou relation-type nouveau
- modifier la collection d'une relation déjà existante
- modifications **ne remettant pas en cause la signification du MCD initial**
- possible d'élaborer **MOD de la mémoire long terme** (informations à archiver)



Quantification des cardinalités

- Le nombre d'occurrences des entités est généralement assez aisé à estimer ; celui des relations est beaucoup plus difficile, compte tenu de la nature même des relations : des associations entre objets.
- Dans la plupart des cas, ce dénombrement s'effectuera à partir du nombre d'occurrences d'une entité de sa collection et de la quantification de la cardinalité de la « patte ».
- Dans le modèle conceptuel de données, les cardinalités maximales multiples sont spécifiées jusqu'à présent par n. Il faut, au niveau du MOD, évaluer cette multiplicité.



La valeur de la cardinalité maximale n de Facture par rapport à ligne doit être précisée.³⁴
La valorisation des cardinalités peut s'exprimer par :

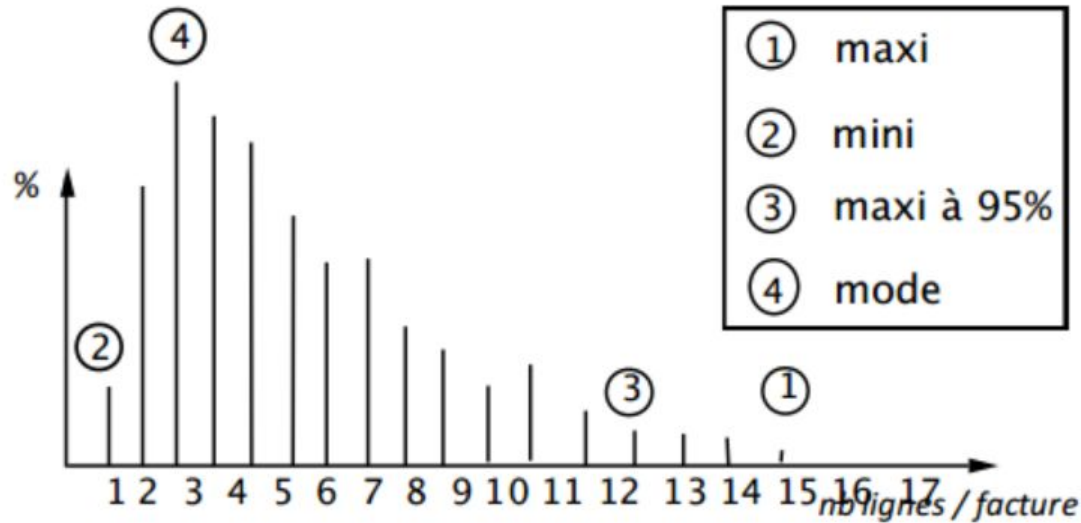
- la cardinalité maximale ;
- la cardinalité maximale à 95 % ;
- la cardinalité modale ;
- la cardinalité moyenne.

Quantification des cardinalités

- Le choix de telle ou telle valeur se fera selon l'intérêt dans le processus d'optimisation ; dans tous les cas, au niveau du MOD, on définira la **cardinalité moyenne**, en particulier pour calculer le nombre d'occurrences des relations.
- La quantification de la cardinalité mini s'exprime par le taux de participation.
 - En effet, au niveau conceptuel, la cardinalité mini peut prendre les valeurs 0 (optionnelle) ou 1 (obligatoire).
 - Le taux de participation ne se détermine que pour les cardinalités mini à 0 et mesure le pourcentage d'occurrences de l'entité qui participent aux occurrences de la relation.
 - Il prend donc une valeur décimale comprise entre $[0, 1[$.
- La **quantification de la cardinalité moyenne** n'est pas toujours facile.
- Soit on fixe arbitrairement cette cardinalité, soit à partir de statistiques sur les populations d'entités concernées par la relation, il est possible de la déterminer de façon plus précise.

Quantification des cardinalités

- Par exemple, pour l'exemple précédent, on pourrait avoir la distribution des lignes de facture par facture illustrée par:



Notons que la cardinalité moyenne d'une cardinalité (1,1) est toujours 1 et que la cardinalité moyenne d'une cardinalité (0,1) est égale au taux de participation.

Quantification des cardinalités

- Pour notre exemple, on peut supposer que la loi de répartition des cardinalités est approximativement triangulaire

Étant donné une telle répartition triangulaire, on peut calculer la cardinalité moyenne par la formule suivante :

$$\text{cardinalité moyenne} = [(m + 2M + N) / 4] \times P$$

avec :

m = valeur mini (0 exclu)

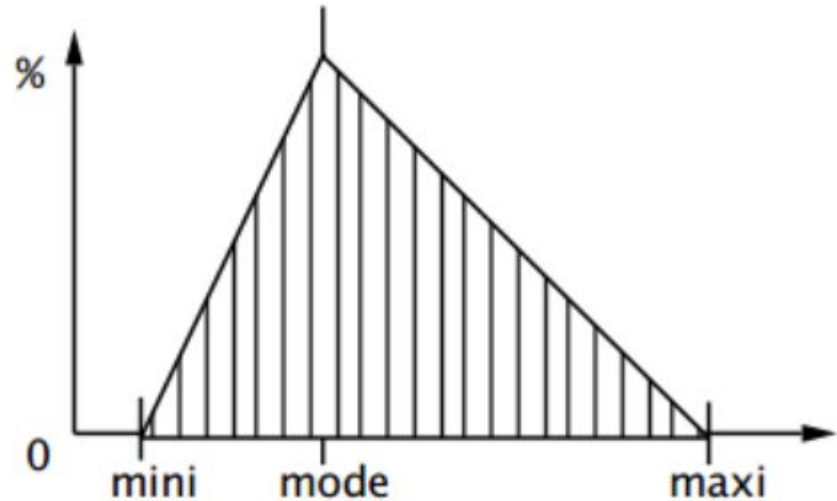
M = valeur modale

N = valeur maxi

P = taux de participation

Par exemple, si M = 3 ; N = 17 ; m = 1 ; P = 1 ; on a :

$$\text{cardinalité moyenne} = ((1 + 2 \times 3 + 17) / 4) \times 1 = 6$$

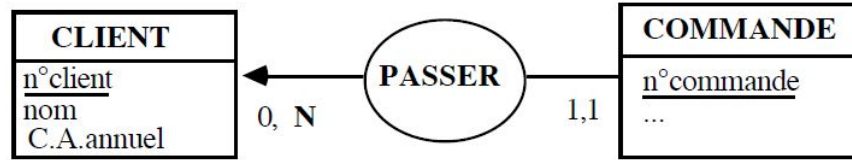


Loi de répartition triangulaire.

Cette loi de distribution triangulaire permet d'estimer les moyennes sur des répartitions dissymétriques (assez fréquentes dans la réalité) et donne d'assez bons résultats pratiques.

Quantification des cardinalités

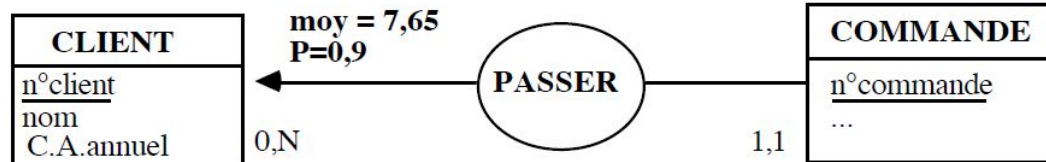
MCD :



- la valeur de la **cardinalité maximale N** doit être précisée :
 - cardinalité maxi
 - cardinalité maxi à 95 %
 - cardinalité modale
 - cardinalité moyenne** <- la plupart du temps retenue
- on peut aussi définir un **taux de participation** :

$$\begin{aligned}\text{taux participation} &= \frac{\text{nb clients ayant } \geq 1 \text{ commande}}{\text{nb total de clients}} \\ &= 1 - \% \text{ clients sans commande}\end{aligned}$$

MOD :

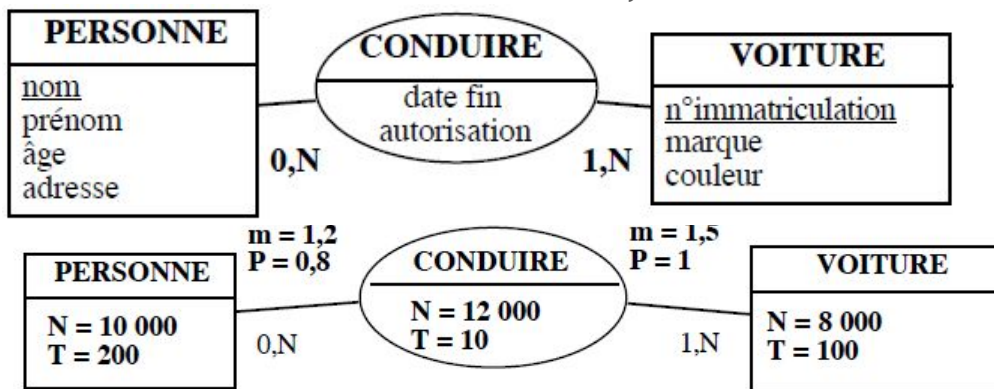


Evaluation du volume global du MOD

- Sachant que :
- on estime les effectifs à 10 000 personnes et 8000 voitures
- une personne conduit généralement qu'une voiture et 3 au maximum
- 80% des personnes conduisent une voiture
- **Card. Moy. de PERSONNE dans CONDUIRE** = $((1+2 \times 1+3)/4) \times 0,8 = 1,2$ (loi triangulaire)
=> nb. d'occurrences de CONDUIRE = 10 000 x 1,2 = 12 000 occurrences
=> **Card. Moy de VOITURE dans CONDUIRE** = 12 000 / 8000 = 1,5

- Supposons tailles suivantes :

- PERSONNE : 200 caractères
- VOITURE : 100 caractères
- CONDUIRE : 10 caractères
- On a alors :



-> estimation du volume brut :

$$10\,000 \times 200 + 8\,000 \times 100 + 12\,000 \times 10 = 2\,920\,000 \text{ caractères}$$

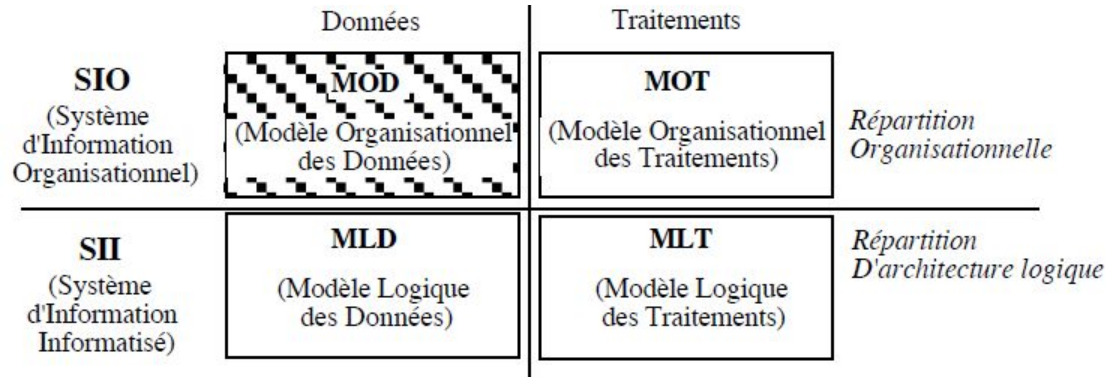
soit 3 millions de caractères environs

MOD: Modèle Organisationnel des Données

Répartition des données informatisables

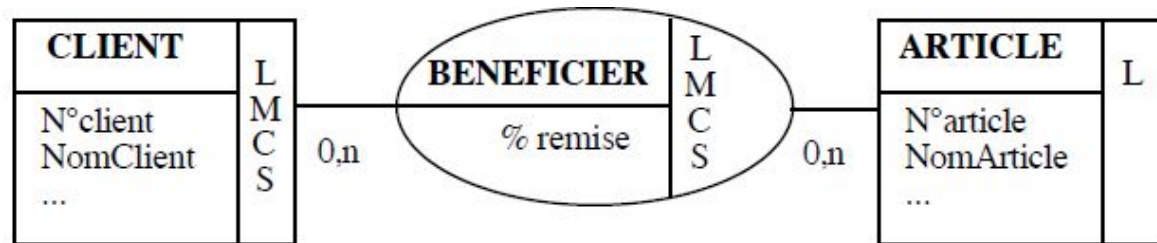
- On peut concevoir qu'une entreprise n'est pas localisée dans un seul bureau, voire sur un seul site géographique.
- Par suite les données seront nécessairement réparties et il peut être intéressant de considérer des MOD "locaux" ce qui aboutit au concept d'unité organisationnelle ayant certains droits d'action sur les données:
 - Lecture Modification Création Suppression
- Des unités locales pourront avoir sur certaines entités des droits partiels.

MOD Locaux: Répartition Organisationnelle des données



- pour un type d'unités organisationnelles (UOs) donné : préciser le **MOD local associé** (entités ou relations accessibles par ces UOs)
- pour chaque entité/relation du MOD local : **nature de l'accès autorisé pour l'UO** :

- lecture (L),
- modification (M),
- création (C),
- suppression (S)

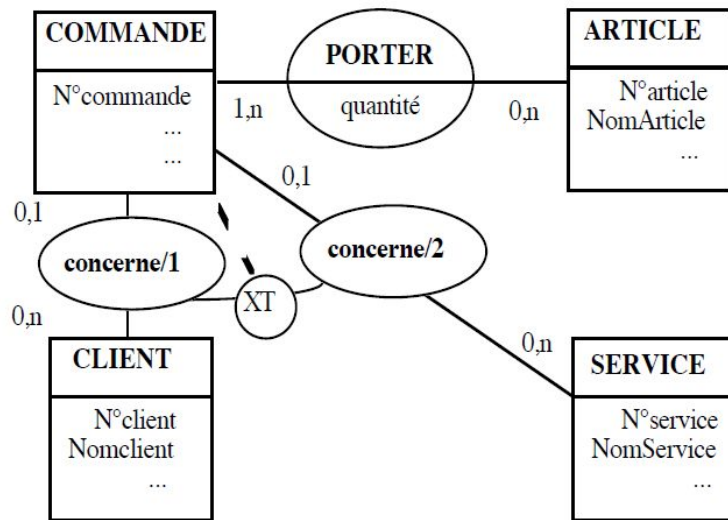


Répartition Organisationnelle des données

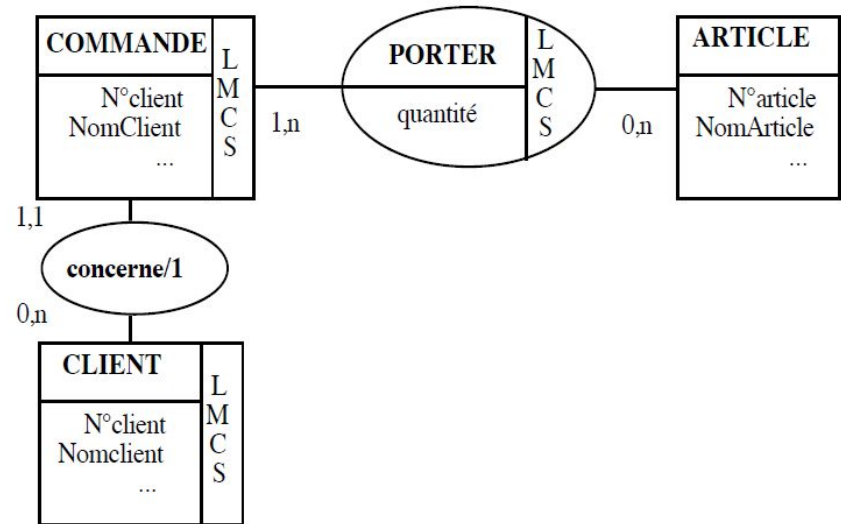
Le MOD local d'une UO peut différer du MOD global au niveau :

- des propriétés,
- des entités,
- des relations

Exemple :



MOD Global



MOD Local
de l'unité organisationnelle
Point de vente Client

MOD: Modèle Organisationnel des Données

Sécurité des données

- Ce point rejoint le précédent en ce qui concerne les unités organisationnelles.
- Mais il existe des données qui sont communes à plusieurs unités organisationnelles ;
- ces données peuvent être classées en deux catégories : données privées, données partagées.

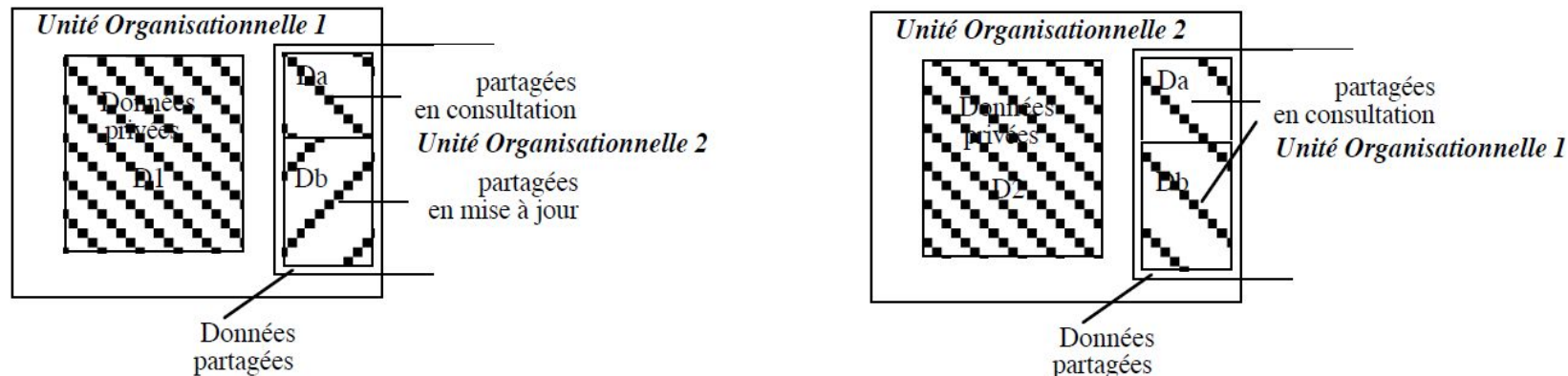
Sécurité intra unité organisationnelle :

-> accessibilité à un type d'utilisateur :

[illegible]

Sécurité inter-unités organisationnelles :

-> données privées, partagées ou protégées :



MOD / U.O.1:

individu/ relation	U.O. 1	U.O.2	U.O.3	U.O.4
Ind.1	LC	L	LCM	...		
Ind.2	LC	L	LCM	...		
Ind.3	LCM	L	LCM			
Ind.4	LCM	L	LCM			
....						
Rel. 1	CM	L	...			
Rel. 2	LS	L	...			
Rel. 3	..					
Rel. 4	...					
....						
...						