



**Base de données universelle
ou
universalité des bases de
données**

Qu'est ce qu'une base de données ?? (SGBD)

- C'est un logiciel capable de gestion le stockage de données (SGBD) selon un schéma d'organisation définit à l'avance.

Les contraintes en terme de capacité sont faibles mais elles demeurent.

- Il est aussi capable de restituer toutes ou PARTIE de ces données aux travers de requêtes.
- Il gère la sécurité des modifications (ajout, modification, suppression)
- Il gère les autorisations des utilisateurs

Qu'est ce n'est pas une base de données ?

- Il n'est pas capable d'analyser les données qu'il contient (sauf somme, moyenne), de fournir des visualisations graphiques

Base de données universelle ? ou !

- Si on parle de base de données universelle; cela sous-entends qu'il pourrait exister qu'UNE seule base de données sur terre. Une base qui pourrait gérer les bourses du monde, le trafic aérien, le fonctionnement des sous-marins nucléaires, ...

Base de données universelle ? ou !

- Cela semble risible, mais pas impossible...
- Pourtant cela pose plein de problème :
 - Taille gigantesque
 - Problème de sécurité
 - Problème des gestion de mise à jour à l'échelle planétaire

Base de données universelle ? ou !

- Mais les personnes ne sont concernées que par une toute petite part de ces informations. Hors, toutes les informations seront traitées à chaque fois, d'où un temps de traitement considérable complètement inutile...
- D'où, l'utilisation de bases de données spécifiques à chaque besoin.

Base de données universelle ? ou !

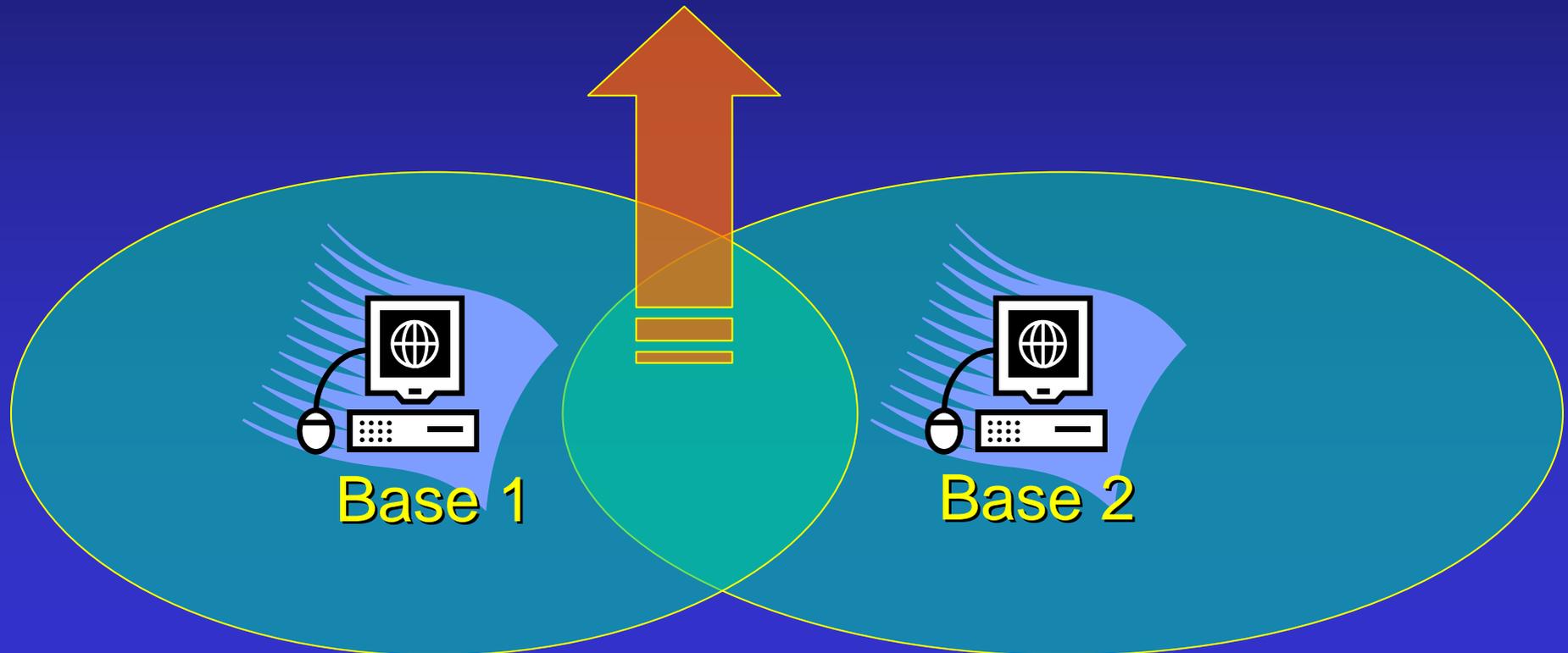
- Il en est de même avec les données "environnementales".
Comment concilier des données aussi diverses ?
- Possible, mais difficile...
- La conception serait très très longue.
- L'utilisation aussi...
- Là aussi, il est plus efficace de créer des bases de données spécifiques

Base de données universelle ? ou !

- Là aussi, il est plus efficace de créer des bases de données spécifiques
- Mais, pour permettre l'interopérabilité des différentes bases, il "suffit" des champs communs (avec des codes communs également) permettant ainsi une universalité des bases de données.

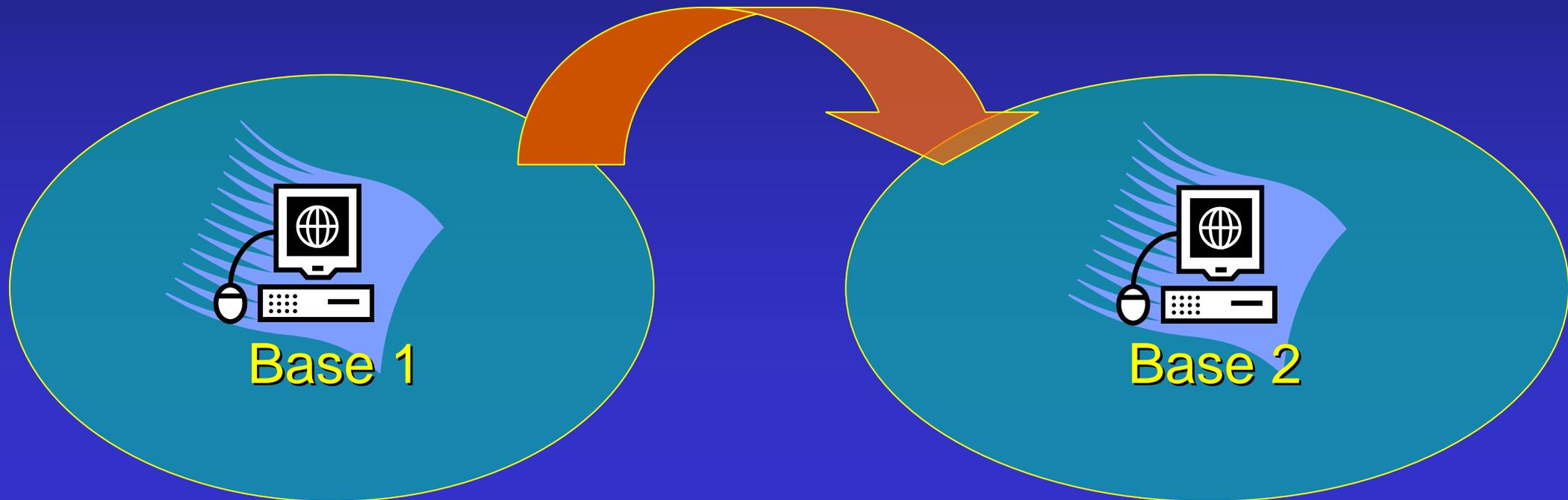
Universalité des bases de données

Permet de "récupérer" toutes les données des deux bases liées par un code commun



Universalité des bases de données

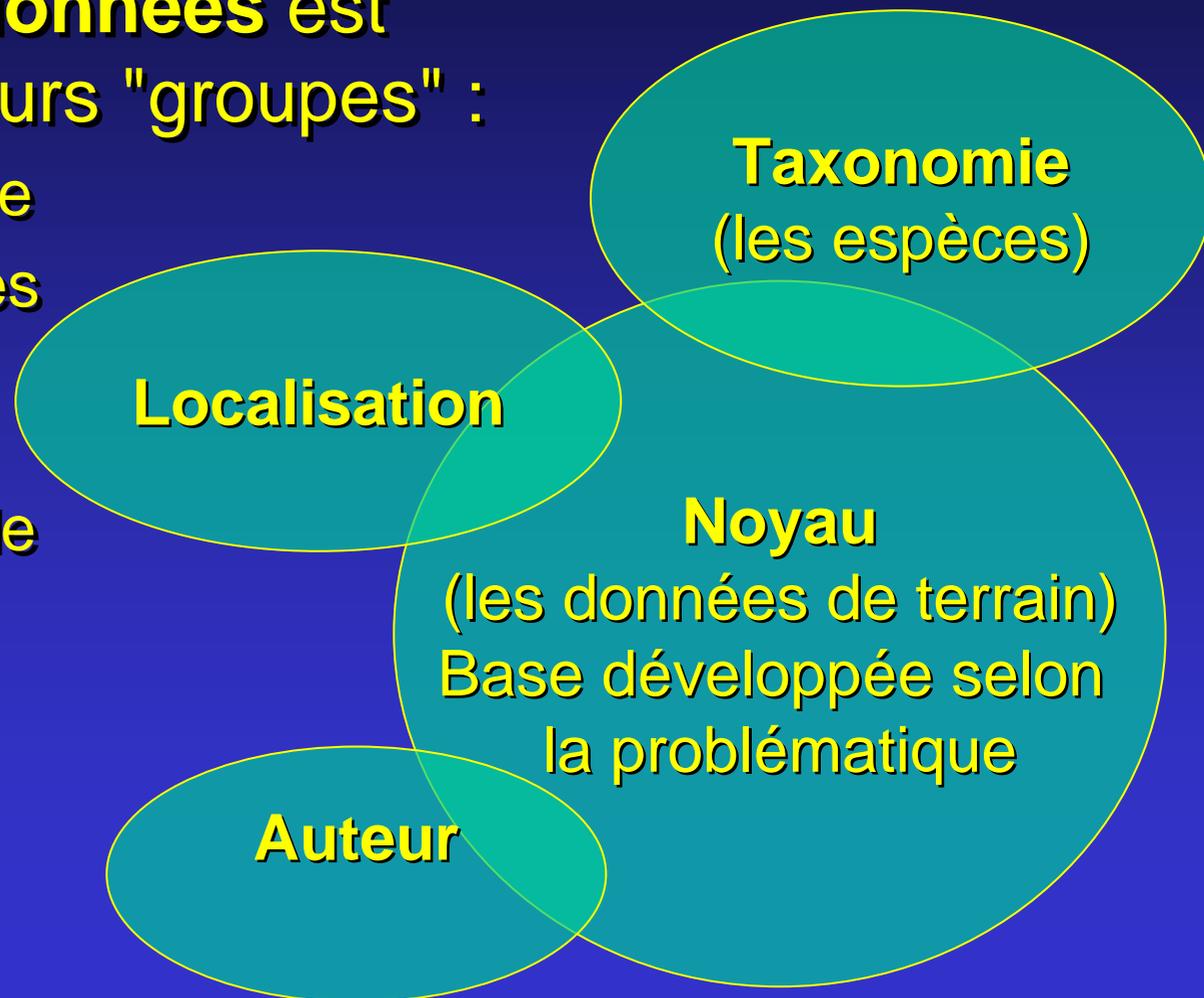
Permet de "transférer" des données d'une base vers une autre.



Universalité des bases de données

- Chaque **base de données** est constitué de plusieurs "groupes" :
 - Le noyau de la base
 - Les tables connexes

c'est là que seront situées les codes de connectivité.



Universalité des bases de données

- Et maintenant, commençent d'autres problèmes !
- Il y a tellement de types de code disponibles.
Alors comment rendre les choses compatibles ?

- Prenons l'exemple du "groupe" taxonomie...

Universalité des bases de données

Taxon	Français	Anglais	FNAT	EURING	MNHN	
Passer domesticus	Moineau domestique	House sparrow	PASDOM	13205		
Turdus merula	Merle noir		TURMER	11502		

- Quels champs faut-il utiliser comme clé primaire ? (i.e. le champs stocké dans le noyau)

Universalité des bases de données

id	Taxon	Français	Anglais	FNAT	EURING	MNHN	
1	Passer domesticus	Moineau domestique	House sparrow	PASDOM	13205		
2	Turdus merula	Merle noir		TURMER	11502		

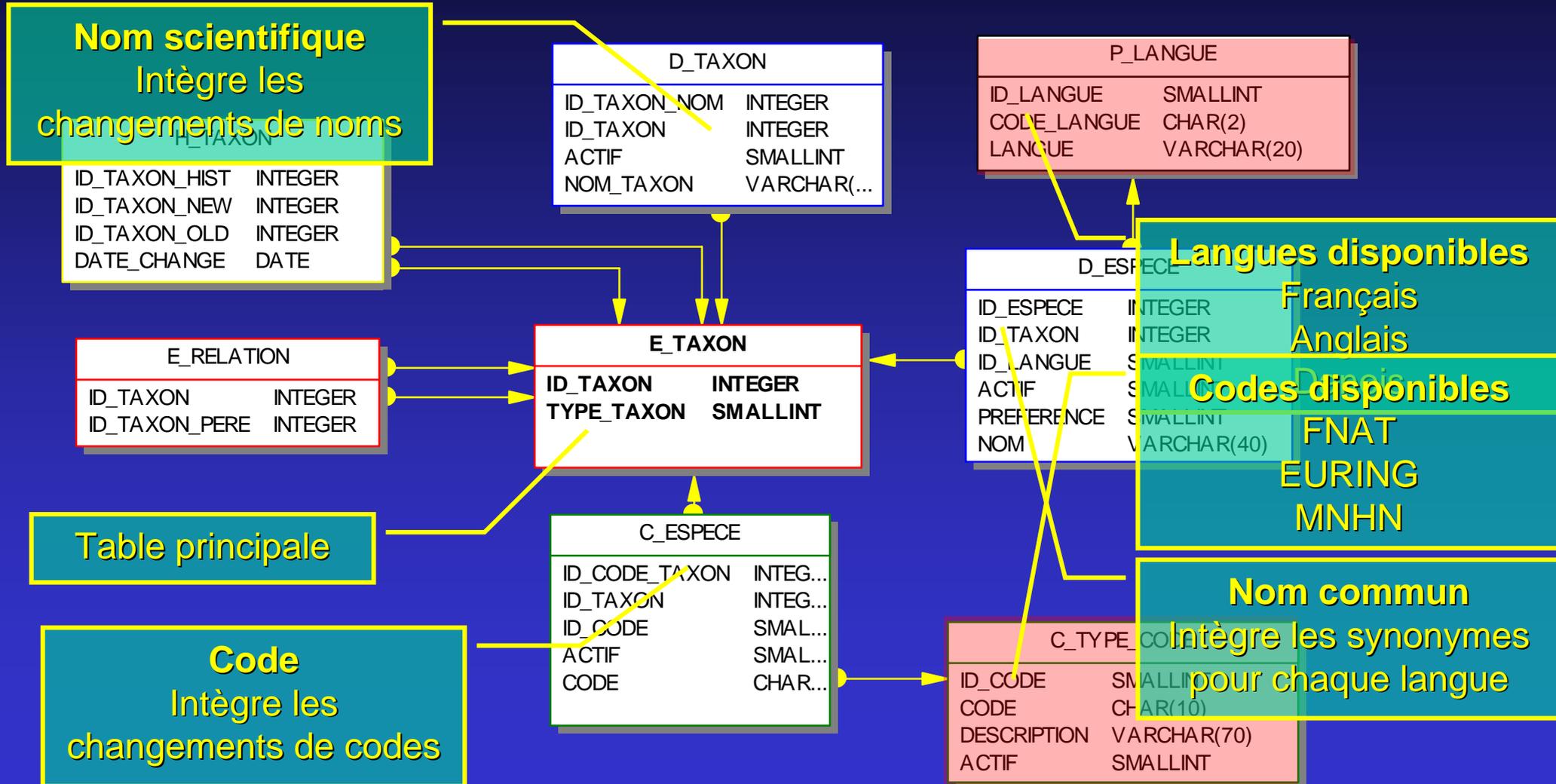
- Quels champs faut-il utiliser comme clé primaire ? (i.e. le champs stocké dans le noyau)
- Aucun, il vaut mieux utiliser un identifiant unique généré par la base. Les autres champs peuvent être modifiés.

Universalité des bases de données

id	Taxon	Français	Anglais	FNAT	EURING	MNHN	
1	Passer domesticus	Moineau domestique	House sparrow	PASDOM	13205		
2	Turdus merula	Merle noir		TURMER	11502		

- Avec ce système, le nombre de type de code est limité à 3. Mais, que faire s'il faut transférer des données à une autre base qui n'utilise aucun de ces codes ? ou langue ou non scientifique ???
- Je propose de disjoncter ce système en 3 groupes.

Universalité des bases de données : Modèle taxonomique multi-codes et multi-langues



Universalité des bases de données

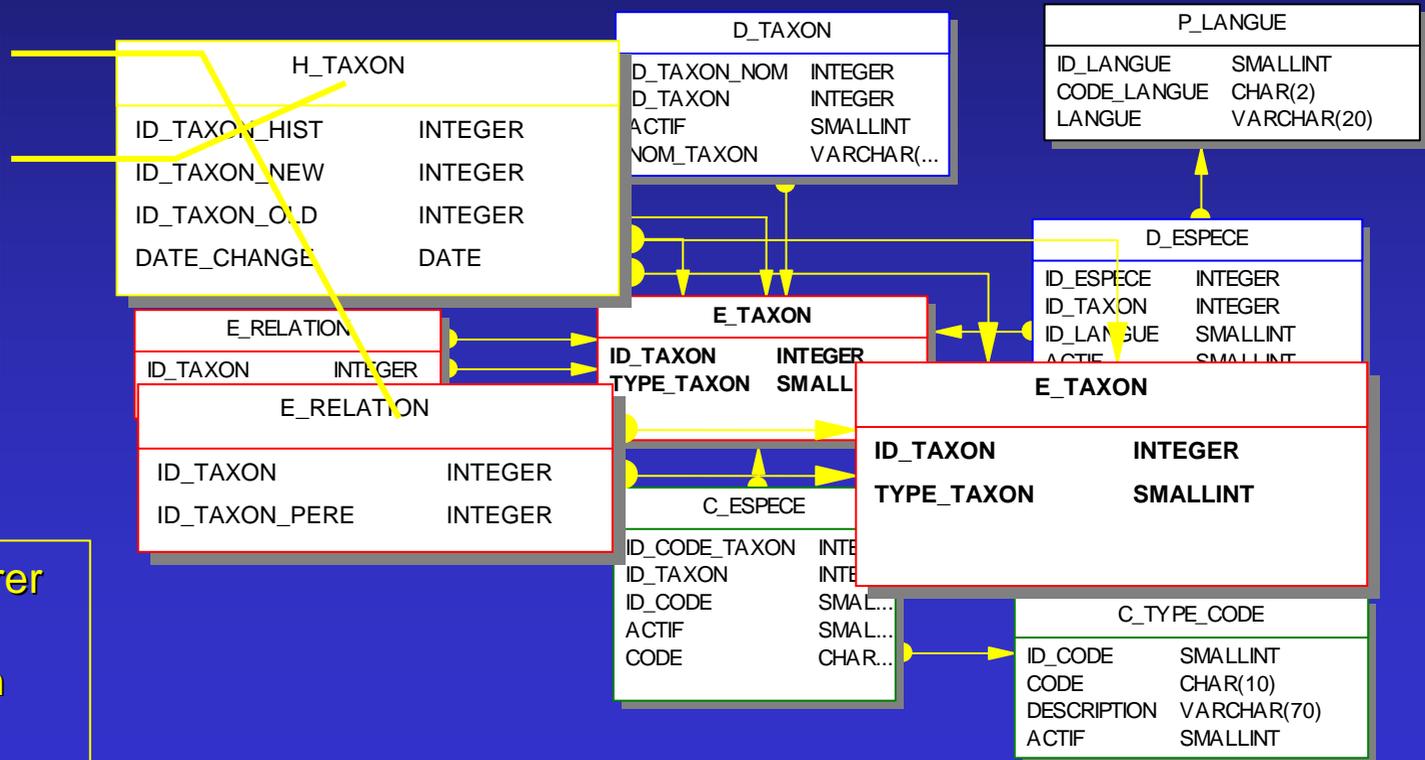
- Avec ce type de structure, il n'y a plus de limitations dans les échanges entre bases de données.
- Evidemment, il reste à remplir les tables...
- On peut adapter cette structure aux localisations ou à tout autre groupe.

Les petits plus ...

- Vous avez pu observer 2 tables supplémentaires.
- A quoi servent-elles ?

Hiérarchisation des taxons
 Pour ne pas être limité qu'au niveau de l'espèce
 Intègre les regroupements ou séparations de taxons

Intéressant pour pouvoir intégrer des données anciennes. C'est aussi l'objet de la conservation des anciens codes et anciens noms



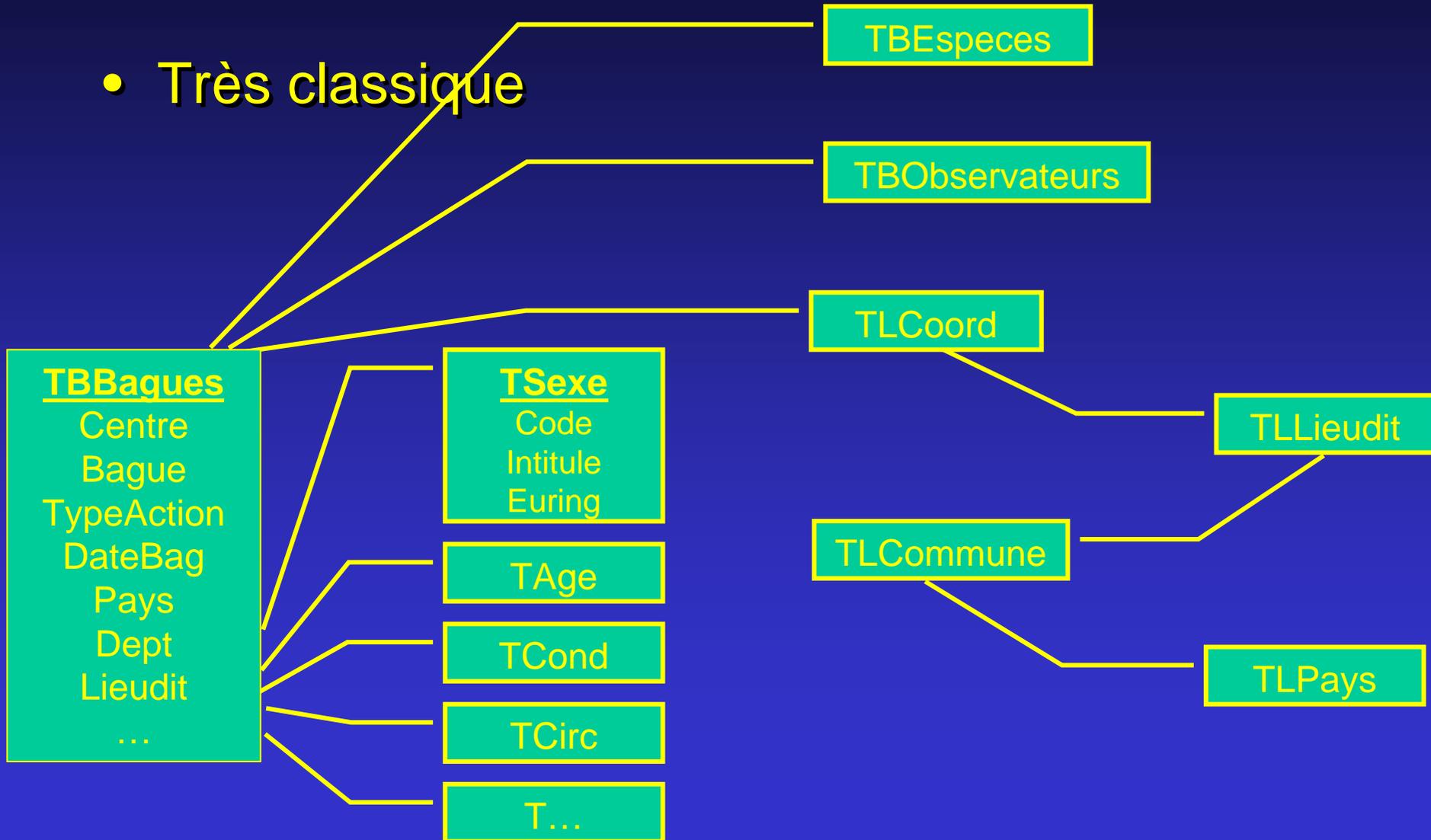
Les petits plus ...

- Les éléments hiérarchisés (arborescence) sont assez classiques :
 - Taxonomie : Ordre, Classe, Famille, Genre, Espèce, ...
 - Localisation : Pays, régions, département, commune, ...
- Comment modéliser ce besoin ?
- Plusieurs solutions
 - Tables liées
 - Représentation intervallaire
 - Notification des relations

Les petits plus ...

Arborescence : tables liées

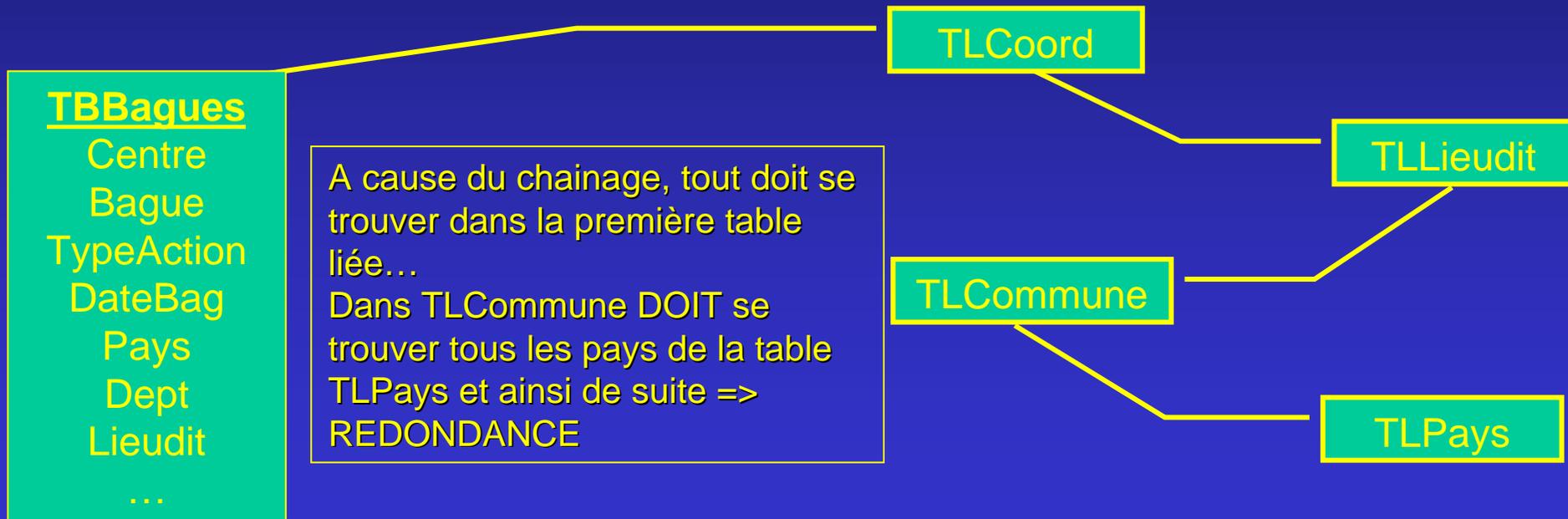
- Très classique



Les petits plus ...

Arborescence : tables liées

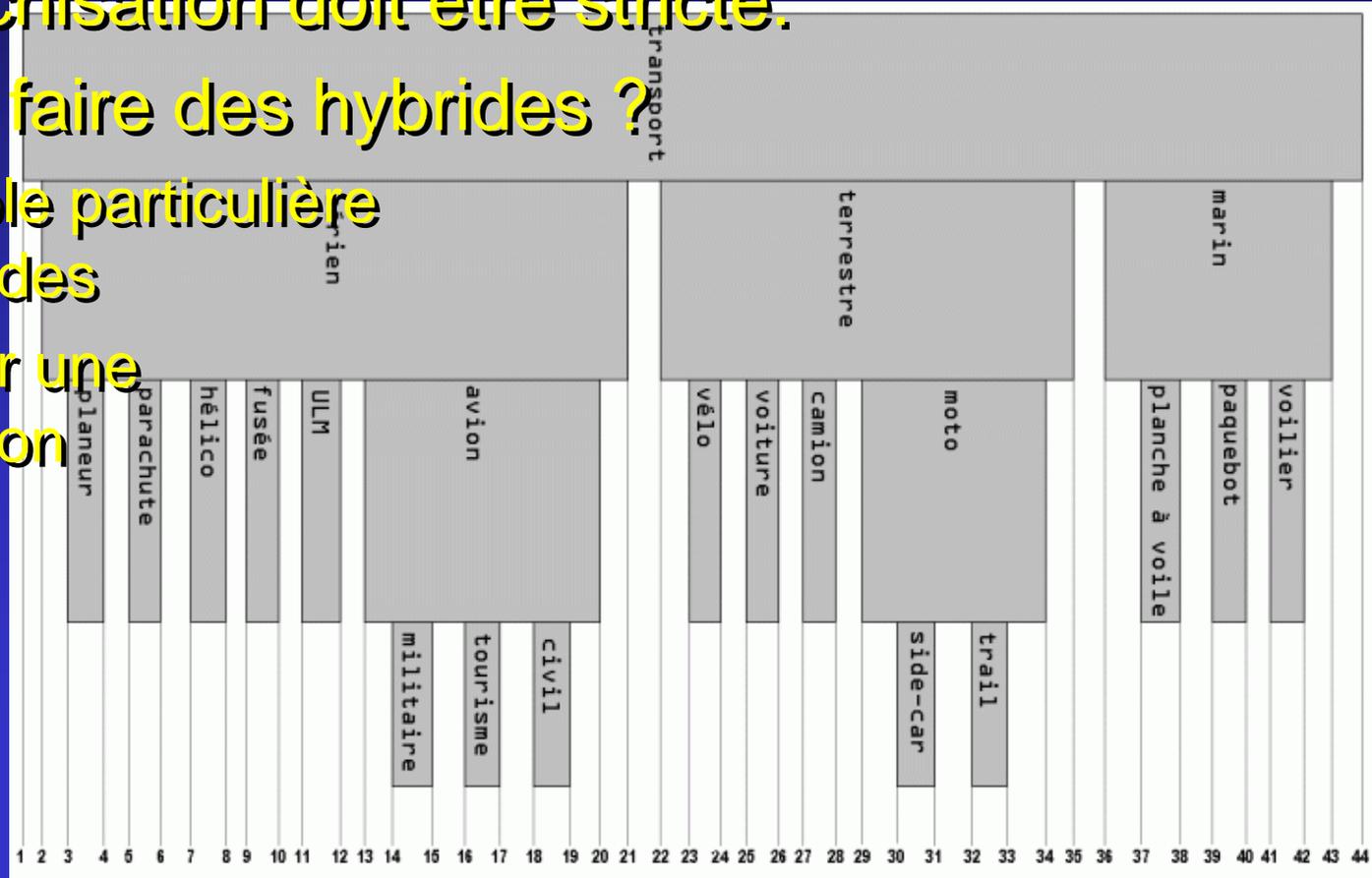
- Très classique
- Problème !



Les petits plus ...

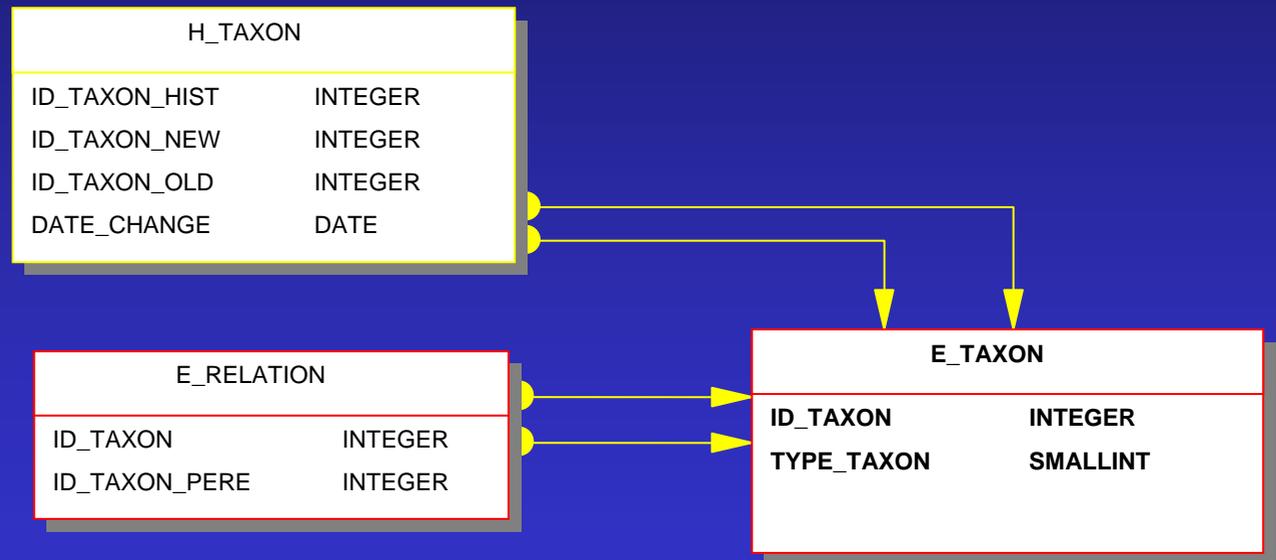
Arborescence : Représentation intervallaire

- Subtil, et très rapide dans les requêtes (bornes)
- Mais la hiérarchisation doit être stricte.
- Oui, mais que faire des hybrides ?
 - Créer une table particulière pour les hybrides
 - Ou passer par une table de relation



Les petits plus ...

Arborescence : Notification des relations



- Pas de redondance d'informations
- Atomicité (dehorter olivier / dehorter et olivier)

- Entites (table)
- Attribut (colonne)
- Relation
- Cle
- Pour être **valide**, un modèle relationnel doit être **normalisé**.
- Le but de la normalisation est de supprimer tout risque de redondance d'informations et donc d'éviter tout problème lié aux anomalies de stockage des informations.

Il s'agit de décomposer les relations en sous-tables de telle sorte que les différents concepts à gérer soient isolés les uns des autres

- Base de synthese issus de synthese ou d'analyse des autres bases

Protocole STOC point d'écoute

- Carré (aléatoire)
- 10 point de 5mn (choisis) en tout ou rien
- 2 passages par an

STOC-EPS

Tableau de données

Id	Date	Carre	Pass	Point	PASDOM	TURMER	COLPAL	PARMAJ	STUVUL	...
1	15/04/07	1	1	1	10	1	0	0	0	
2	15/04/07	1	1	2	2	0	0	0	0	
3	15/04/07	1	1	3	0	3	4	1	2	
4	15/04/07	2	1	1	0	0	0	0	0	
5	15/04/07	2	1	2	0	5	15	2	1	
6	15/04/07	2	1	3	0	6	6	10	0	
7	18/05/07	1	2	1	12	2	1	1	0	
8	18/05/07	1	2	2	0	0	0	0	0	
9	18/05/07	1	2	3	0	0	0	0	0	
10	18/05/07	2	2	1	1	2	5	1	3	
11	18/05/07	2	2	2	5	12	5	20	2	
12	18/05/07	2	2	3	2	0	0	0	1	
13	16/04/08	1	1	1	20	2	0	1	5	
14	16/04/08	1	1	2	5	0	5	2	0	

STOC-EPS

Tableau de données

Id	Date	Carre	Pass	Point	PASDOM	TURMER	COLPAL	PARMAJ	STUVUL	...
1	15/04/07	1	1	1	10	1	0	0	0	
2	15/04/07	1	1	2	2	0	0	0	0	
3	15/04/07	1	1	3	0	3	4	1	2	
4	15/04/07	2	1	1	0	0	0	0	0	
5	15/04/07	2	1	2	0	5	15	2	1	
6	15/04/07	2	1	3	0	6	6	10	0	
7	18/05/07	1	2	1	12	2	1	1	0	
8	18/05/07	1	2	2	0	0	0	0	0	
9	18/05/07	1	2	3	0	0	0	0	0	
10	18/05/07	2	2	1	1	2	5	1	3	
11	18/05/07	2	2	2	5	12	5	20	2	
12	18/05/07	2	2	3	2	0	0	0	1	
13	16/04/08	1	1	1	20	2	0	1	5	
14	16/04/08	1	1	2	5	0	5	2	0	

STOC-EPS

Tableau de données

Id	Date	Carre	Pass	Point	PASDOM	TURMER	COLPAL	PARMAJ	STUVUL	...
1	15/04/07	1	1	1	10	1	0	0	0	
2	15/04/07	1	1	2	2	0	0	0	0	
3	15/04/07	1	1	3	0	3	4	1	2	
4	15/04/07	2	1	1	0	0	0	0	0	
5	15/04/07	2	1	2	0	5	15	2	1	
6	15/04/07	2	1	3	0	6	6	10	0	
7	18/05/07	1	2	1	12	2	1	1	0	
8	18/05/07	1	2	2	0	0	0	0	0	
9	18/05/07	1	2	3	0	0	0	0	0	
10	18/05/07	2	2	1	1	2	5	1	3	
11	18/05/07	2	2	2	5	12	5	20	2	
12	18/05/07	2	2	3	2	0	0	0	1	
13	16/04/08	1	1	1	20	2	0	1	5	
14	16/04/08	1	1	2	5	0	5	2	0	

- Espece
- localisation

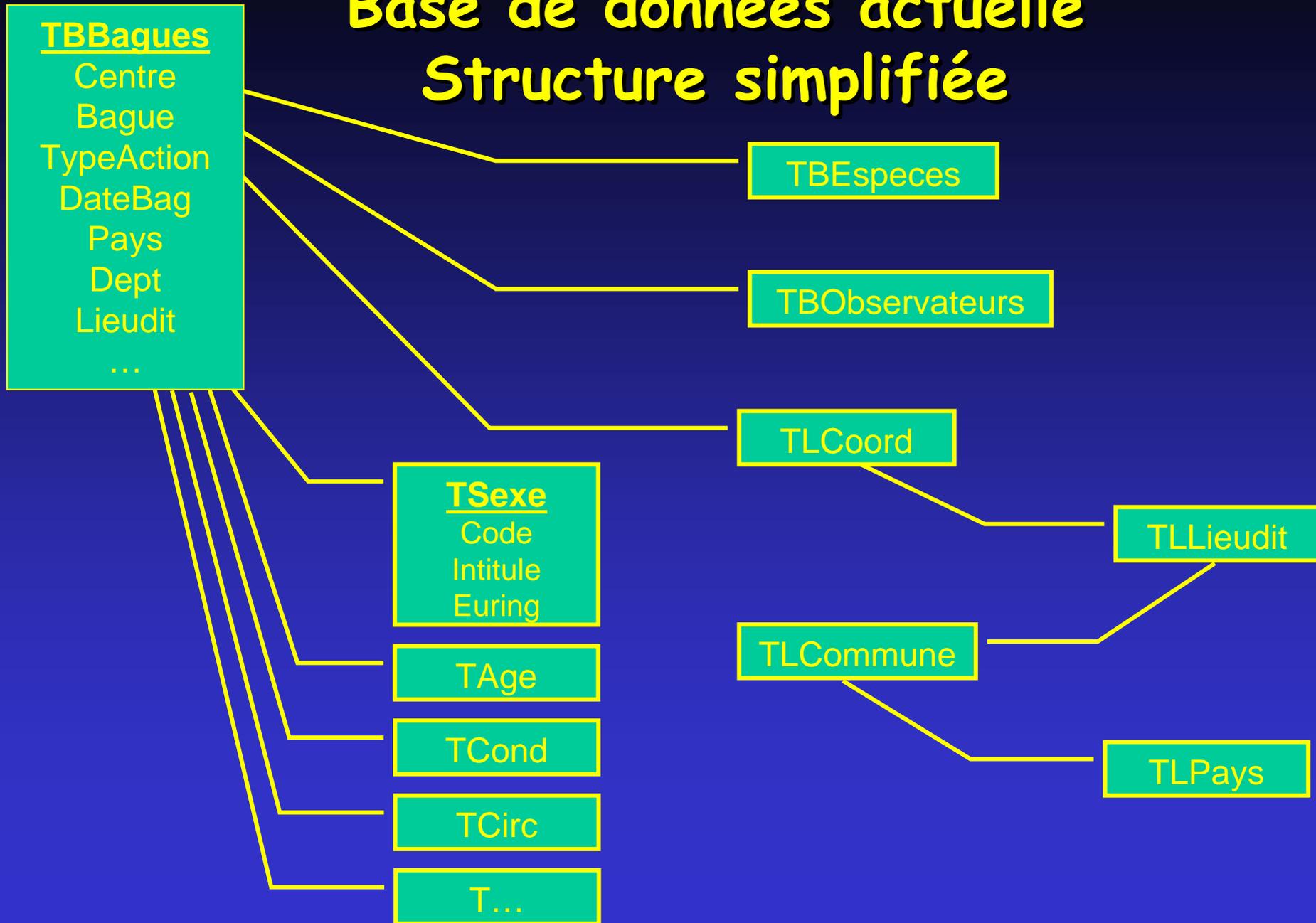


Base de données actuelle

Structure

Base de données actuelle

Structure simplifiée



Nouvelle base

- 150 tables (entités)