

Fiche de TD/TP N°11- BDD NoSQL : Orienté Documents avec MongoDB

Exercice 1 : Modélisation (*Relationnel vers Orienté Documents*)

On souhaite réaliser une base de données orientée documents pour gérer des papiers scientifiques et leurs auteurs, étant données les informations suivantes :

- Un papier est décrit par les attributs titre, année et un ensemble de références.
- Les références sont d'autres papiers.
- Un auteur est décrit par les attributs nom, prénom et une affiliation.
- Une affiliation est composée du nom de l'établissement de l'auteur, la ville et le pays.
- Un auteur peut écrire plusieurs papiers et un papier peut être écrit par un ou plusieurs auteurs tout en prenant en considération leur ordre.

1. Proposer un modèle orienté documents normalisé.
2. Proposer un modèle orienté documents purement dénormalisé en favorisant les *papiers*. Quel est le principal défaut de cette solution ? Est-il toujours possible d'avoir une solution ne reposant que sur l'imbrication ?
3. Proposer un modèle orienté documents basé sur les références.
4. Sachant que l'objectif de l'application est de visualiser une liste des *auteurs* avec leurs *papiers*, et d'accéder aux détails des *papiers* uniquement lorsque l'on clique sur son titre. Proposer une solution adaptée à ce problème mobilisant référence et imbrication ?

Exercice2 : Manipulation des données sous mongoDB

1. Lancer un container MongoDB avec docker en utilisant docker-compose
 - a. `docker-compose up -d`
2. Installer *pymongo* afin de se connecter au cluster mongodb
 - a. `pip install pymongo`
3. Télécharger le notebook « **demo_Pymongo_esi-sba_2023** » et essayer de pratiquer les différents blocks.
4. Importation du dataset:
 - a. Décompresser les fichier **dblp.zip**
 - b. `docker cp dblp.json mongo-esi:/dblp.json`
 - c. `docker exec mongo-esi mongoimport -d dblp -c publis /dblp.json`
5. Répondre aux requêtes MongoDB suivantes :
 - Donner le titre des publications dont le type est « Book ».
 - Donner le titre des publications ayant «publisher » et « editor » renseignés.
 - Donner le titre et l'année des publications qui sont publiées entre 2005 et 2007.
 - Donner le nombre des publications dont le titre contient « Ontology » .



- Donner le titre des publications publiées par l'auteur «Mimoun Malki ».
- Donner le titre des publications dont «Mimoun Malki » est premier auteur.
- Liste distincte des auteurs.
- Retourner le titre et l'année des publications de type article, et trier sur année dans le l'ordre décroissant.
- Donner le nombre publications par type.
- Pour chaque auteur le nombre des publications. Trier par ordre décroissant.
- Donner le nombre de pages moyen par type. (Utiliser l'opérateur *\$subtract*)