

Jeudi 09 Janvier 2025

Nom : TIOMELA ZANGUE Jorel

Matricule : 21U2144

Filière : INF L3

Projet INF 351 : Mise en œuvre d'un Système d'Aide à la Décision

Rapport R1 : Conception, implémentation et alimentation de la base de données opérationnelle

1- Analyse et conception

L'analyse et la conception de ce problème a été fait en cours et le MLD qui en découle est le suivant :

- ville(id_ville, nom_ville, #id_pays)
- pays(id_pays, nom_pays)
- affiliation(id_affiliation, nom_affiliation, #id_ville)
- auteur(id_auteur, nom_auteur, #id_affiliation)
- article(id_article, nom_article)
- rediger(#id_article, #id_auteur, annee)

2- Implémentation

Pour l'implémentation de la base de données, nous avons utilisé le SGBD MySql et les requêtes que nous avons utilisés pour créer les tables sont les suivantes :

--Créer la BD

```
CREATE DATABASE bd_articles;
```

-- Table Pays

```
CREATE TABLE pays (  
  id_pays INT PRIMARY KEY AUTO_INCREMENT,  
  nom_pays VARCHAR(255) NOT NULL  
);
```

-- Table Ville

```
CREATE TABLE ville (  
  id_ville INT PRIMARY KEY AUTO_INCREMENT,  
  nom_ville VARCHAR(255) NOT NULL,  
  id_pays INT NOT NULL,  
  FOREIGN KEY (id_pays) REFERENCES pays(id_pays)  
);
```

```

-- Table Affiliation
CREATE TABLE affiliation (
    id_affiliation INT PRIMARY KEY AUTO_INCREMENT,
    nom_affiliation VARCHAR(255) NOT NULL,
    id_ville INT NOT NULL,
    FOREIGN KEY (id_ville) REFERENCES ville(id_ville)
);

-- Table Auteur
CREATE TABLE auteur (
    id_auteur INT PRIMARY KEY AUTO_INCREMENT,
    nom_auteur VARCHAR(255) NOT NULL,
    id_affiliation INT NOT NULL,
    FOREIGN KEY (id_affiliation) REFERENCES affiliation(id_affiliation)
);

-- Table Article
CREATE TABLE article (
    id_article INT PRIMARY KEY AUTO_INCREMENT,
    titre_article VARCHAR(255) NOT NULL
);

-- Table Rediger
CREATE TABLE rediger (
    id_article INT NOT NULL,
    id_auteur INT NOT NULL,
    annee YEAR NOT NULL,
    PRIMARY KEY (id_article, id_auteur),
    FOREIGN KEY (id_article) REFERENCES article(id_article),
    FOREIGN KEY (id_auteur) REFERENCES auteur(id_auteur)
);

```

3- Alimentation

Une fois la base de données créée, nous avons effectué plusieurs opérations ETL sur les fichiers afin d'avoir un contenu raisonnable dans la base de données. Les opérations effectuées sont les suivantes :

- **Extraction et Transformation** : Nous avons écrit un script python appelé [script_ET](#) qui récupère essentiellement les informations telles que nomArticle, nomAuteur, nomAffiliation, nomVille, nomPays dans tous les fichiers contenus dans chaque dossier, pour chaque dossier un nouveau fichier csv est généré contenant tous les infos bien structurés des articles de chaque année.
- **Stockage** : Une fois que toutes les données sont transformées sous un nouveau format bien structuré, nous avons écrit un nouveau script python qui lit toutes les informations qui sont structurées dans les fichiers csv qui ont été générés, ensuite charge ces informations dans la base de données tout en tenant compte des clés primaires et clés étrangères de chaque table.

et en se rassurant qu'il n'y ait pas de données en double dans la base de données.