

REPUBLIC OF CAMEROON  
Peace-Work-Fatherland  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR

\*\*\*\*\*

UNIVERSITE DE YAOUNDE I

\*\*\*\*\*

FACULTE DES SCIENCES

\*\*\*\*\*

DEPARTEMENT D'INFORMATIQUE

\*\*\*\*\*



REPUBLIC OF CAMEROON  
Peace-Work-Fatherland  
MINISTRY OF HIGHER EDUCATION

\*\*\*\*\*

UNIVERSITY OF YAOUNDE I

\*\*\*\*\*

FACULTY OF SCIENCES

\*\*\*\*\*

DEPARTMENT OF COMPUTER SCIENCE

\*\*\*\*\*

**INF 351 : RAPPORT DU TP2 (CONCEPTION ET ALIMENTATION D'UN  
ENTREPOT DE DONNEES)**

**SUPERVISEUR : Dr. Valéry MONTHE**

Année Académique : 2023 - 2024

NOMS	MATRICULES	PARTICIPATIONS (En %)
NOUBISSIE KAMGA WILFRIED	20U2671	100
NDEUNA NGANA OUSMAN SINCLAIR	21T2433	100
TCHOUDJI YEMEFACK JAURES FERRINO	20U2870	100
FOGUE GUEMGNE NELLY SORELLE	20U2993	100
NYOBE LUCRECE GEORGINA	20U2704	100



## Contenu

Définition des besoins _____	1
Conception de l'entrepôt _____	2
Schema de l'architecture de notre systeme _____	6
Description de l'architecture_____	8



## Définition des besoins

Pour le cadre de ce TP, nous avons décidé de répondre à 08 besoins qui sont :

- Collecter les données des publications, y compris les informations sur les institutions liés à chaque publication d'article ;
- Récupérer les publication par auteur afin de déterminer le nombre de publication par individu ;
- Regrouper les publications par affiliation pour connaître le nombre de publications associées à chaque institution ;
- Identifier l'auteur ayant publié le plus grand nombre d'articles aux USA ;
- Déterminer l'établissement qui a publié le plus grand nombre d'articles ;
- Analyser les collaborations entre auteurs pour identifier l'article ayant la plus grande collaboration d'auteurs
- Regrouper les publications par année afin de connaître le nombre de publications par année ;
- Identifier l'auteur, son établissement et son pays ayant publié le plus grand nombre d'articles, ainsi que ce nombre .



## Conception de l'entrepôt

### DEFINITION DE DIFFERENTS SUJETS D'ANALYSE

- Nombre de publication par institution
- Nombre de publication par auteur
- Nombre d'auteur par affiliation
- L'auteur ayant publié le plus grand nombre d'article aux USA
- Etablissement ayant le plus d'article publié
- Le nombre d'article publié par année
- L'article ayant la plus grande collaboration d'auteurs
- L'auteur, son établissement, son pays, ayant publié le plus grand nombre d'article, ainsi que son nombre d'article publié



## DICTIONNAIRE DES MESURES

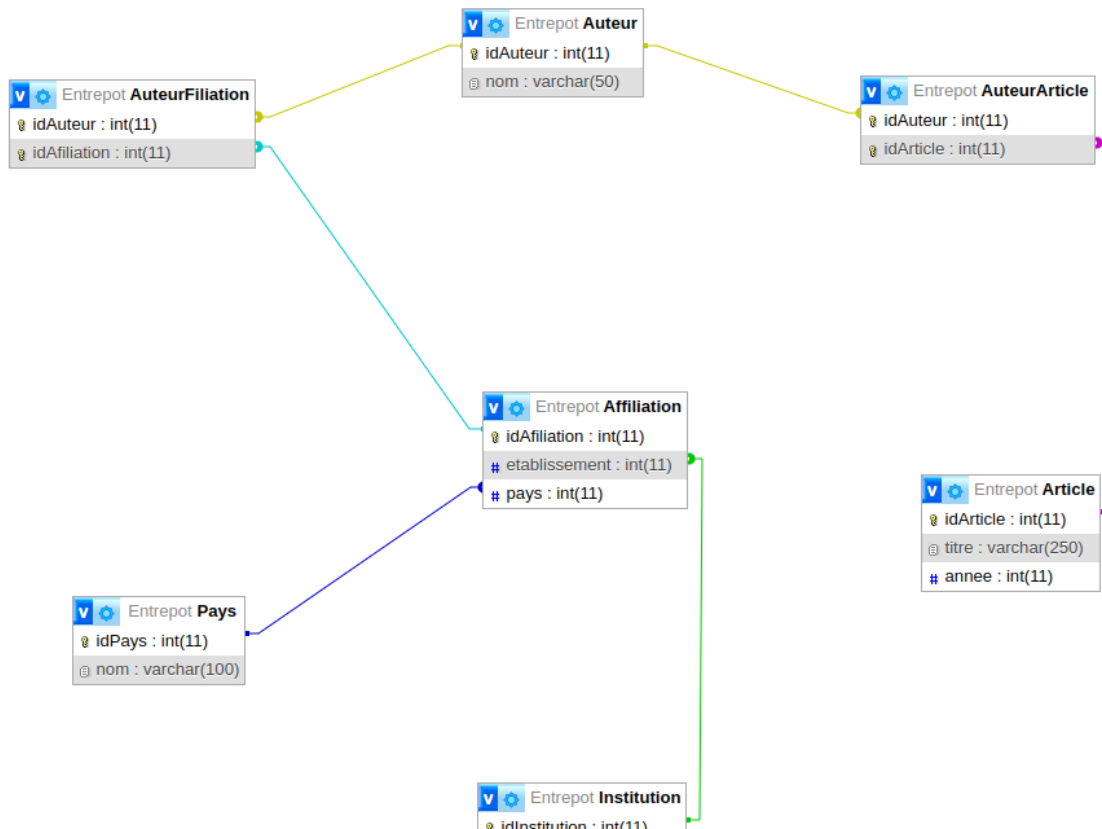
Code	Description	Type	Formule d'extraction
<b>Nbre_Pub_Inst</b>	Nombre de publication par institution	Entier	<pre>SELECT Institution.nom AS institution, COUNT(*) AS nombre_publications FROM Affiliation JOIN Institution ON Affiliation.etablissement = Institution.idInstitution JOIN AuteurFiliation ON Affiliation.idAfiliation = AuteurFiliation.idAfiliation JOIN AuteurArticle ON AuteurFiliation.idAuteur = AuteurArticle.idAuteur GROUP BY Institution.nom;</pre>
<b>Nbre_Pub_Aut</b>	Nombre de publication par auteur	Entier	<pre>SELECT Auteur.nom AS auteur, COUNT(*) AS nombre_publications FROM Auteur JOIN AuteurFiliation ON Auteur.idAuteur = AuteurFiliation.idAuteur JOIN AuteurArticle ON AuteurFiliation.idAuteur = AuteurArticle.idAuteur GROUP BY Auteur.nom;</pre>
<b>Nbre_Aut_Aff</b>	Nombre d'auteur par affiliation	Entier	<pre>SELECT Affiliation.idAfiliation, COUNT(AuteurFiliation.idAuteur) AS nombre_auteurs FROM Affiliation LEFT JOIN AuteurFiliation ON Affiliation.idAfiliation = AuteurFiliation.idAfiliation GROUP BY Affiliation.idAfiliation;</pre>
<b>Max_Aut_Art</b>	L'auteur ayant publié le plus grand nombre d'article aux USA	String	<pre>SELECT Auteur.nom AS auteur, COUNT(AuteurArticle.idArticle) AS nombre_articles FROM Auteur JOIN AuteurFiliation ON Auteur.idAuteur = AuteurFiliation.idAuteur JOIN Affiliation ON AuteurFiliation.idAfiliation = Affiliation.idAfiliation JOIN Pays ON Affiliation.pays = Pays.idPays JOIN AuteurArticle ON Auteur.idAuteur = AuteurArticle.idAuteur WHERE Pays.nom = ' USA' GROUP BY Auteur.idAuteur ORDER BY nombre_articles DESC LIMIT 1;</pre>
<b>Max_Eta_Art</b>	Etablissement ayant le plus d'article publié	String	<pre>SELECT Institution.nom AS etablissement, COUNT(AuteurArticle.idArticle) AS nombre_articles FROM Institution JOIN Affiliation ON Institution.idInstitution = Affiliation.etablissement JOIN AuteurFiliation ON Affiliation.idAfiliation = AuteurFiliation.idAfiliation JOIN AuteurArticle ON AuteurFiliation.idAuteur = AuteurArticle.idAuteur GROUP BY Institution.idInstitution ORDER BY nombre_articles DESC LIMIT 1;</pre>
<b>Nbre_Art_An</b>	Le nombre d'article publié par année	Entier	<pre>SELECT annee, COUNT(idArticle) AS nombre_articles FROM Article GROUP BY annee ORDER BY annee;</pre>
<b>Max_Art_Coll</b>	L'article ayant la plus grande collaboration d'auteurs	String	<pre>SELECT Article.titre AS titre_article, COUNT(Auteur.idAuteur) AS nombre_auteurs FROM Article JOIN AuteurArticle ON Article.idArticle = AuteurArticle.idArticle JOIN Auteur ON AuteurArticle.idAuteur = Auteur.idAuteur GROUP BY Article.idArticle ORDER BY nombre_auteurs DESC LIMIT 1;</pre>



Aut_Eta_Pays	L'auteur, son établissement, son pays, ayant publié le plus grand nombre d'article, ainsi que son nombre d'article publié	String	<pre>SELECT Auteur.nom AS auteur, Institution.nom AS etablisement, Pays.nom AS pays, COUNT(AuteurArticle.idArticle) AS nombre_articles FROM Auteur JOIN AuteurFiliation ON Auteur.idAuteur = AuteurFiliation.idAuteur JOIN Affiliation ON AuteurFiliation.idAfiliation = Affiliation.idAfiliation JOIN Institution ON Affiliation.etablisement = Institution.idInstitution JOIN Pays ON Affiliation.pays = Pays.idPays JOIN AuteurArticle ON Auteur.idAuteur = AuteurArticle.idAuteur GROUP BY Auteur.idAuteur, Institution.idInstitution, Pays.idPays ORDER BY nombre_articles DESC LIMIT 1;</pre>

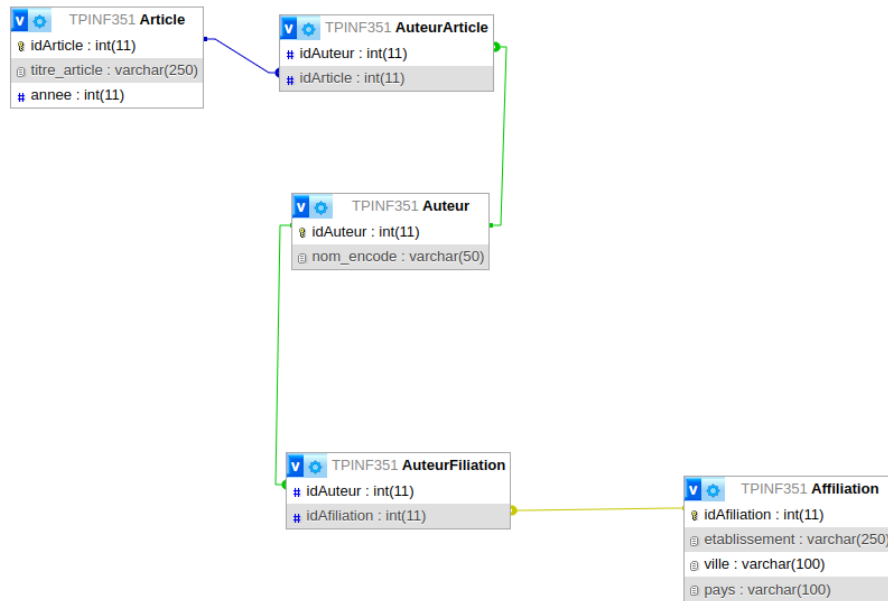
SCHEMA DE L'ENTREPOT DE DONNEES

Nous utiliserons un model en flocon de neige pour réduire les redondances dans les données à manipuler

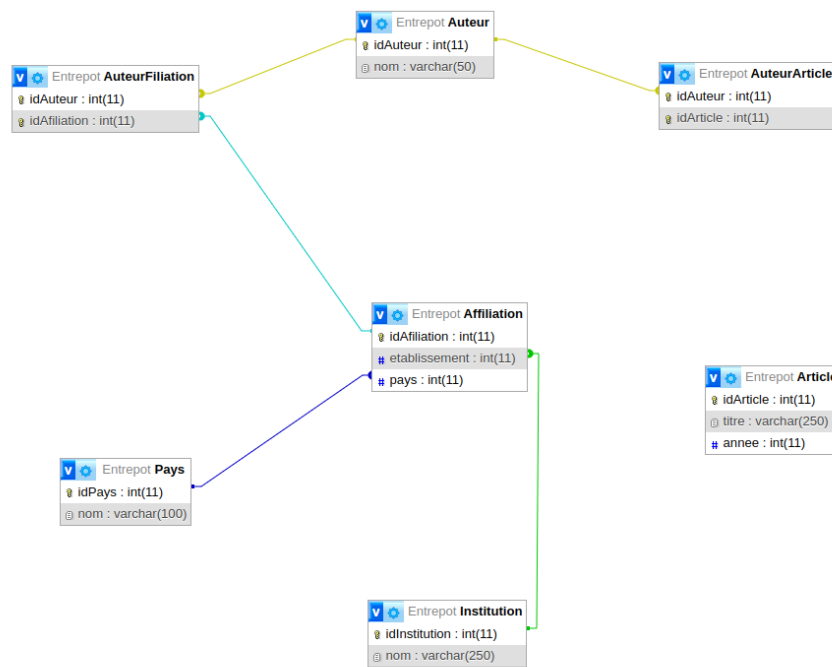




## Schema de l'architecture de notre systeme



## BASE DE DONNEES SOURCE



## ENTREPOT



Les différences entre votre base de données de base et votre entrepôt de données sont les suivantes :

- L'ajout des tables "Institution" et "Pays" dans votre entrepôt de données.
- Colonnes des tables : Dans notre base de données de base, nous avons la colonne "titre\_article" dans la table "Article" et la colonne "nom\_encode" dans la table "Auteur" ; Dans notre entrepôt de données, ces colonnes ont été renommées respectivement en "titre" et "nom". De plus, la colonne "etablissement" dans la table "Affiliation" a été modifiée de type VARCHAR à INTEGER pour référencer la clé primaire de la table "Institution", et la colonne "pays" a été modifiée de type VARCHAR à INTEGER pour référencer la clé primaire de la table "Pays".



## Description de l'architecture

Notre entrepôt de données suit une architecture de base de données relationnelle. Il est composé de plusieurs tables qui sont reliées entre elles par des clés étrangères pour établir des relations. Nous avons créé de nouvelles tables qui sont la table « Pays » et « Institution » pour établir une hiérarchie avec la table « Affiliation » afin d'éviter le plus que possible les redondances et permettre des analyses par palier sur les dimensions hiérarchisées.