

# Réalisez une application de recommandation de contenu

Hourdin Charlène - Février 2023



# Agenda



**Présentation du projet**



**Pré-traitement des  
données**



**Système de  
recommandation**



**Azure Fonction, Git et  
déploiement de l'API**



## **PRÉSENTATION DU PROJET**

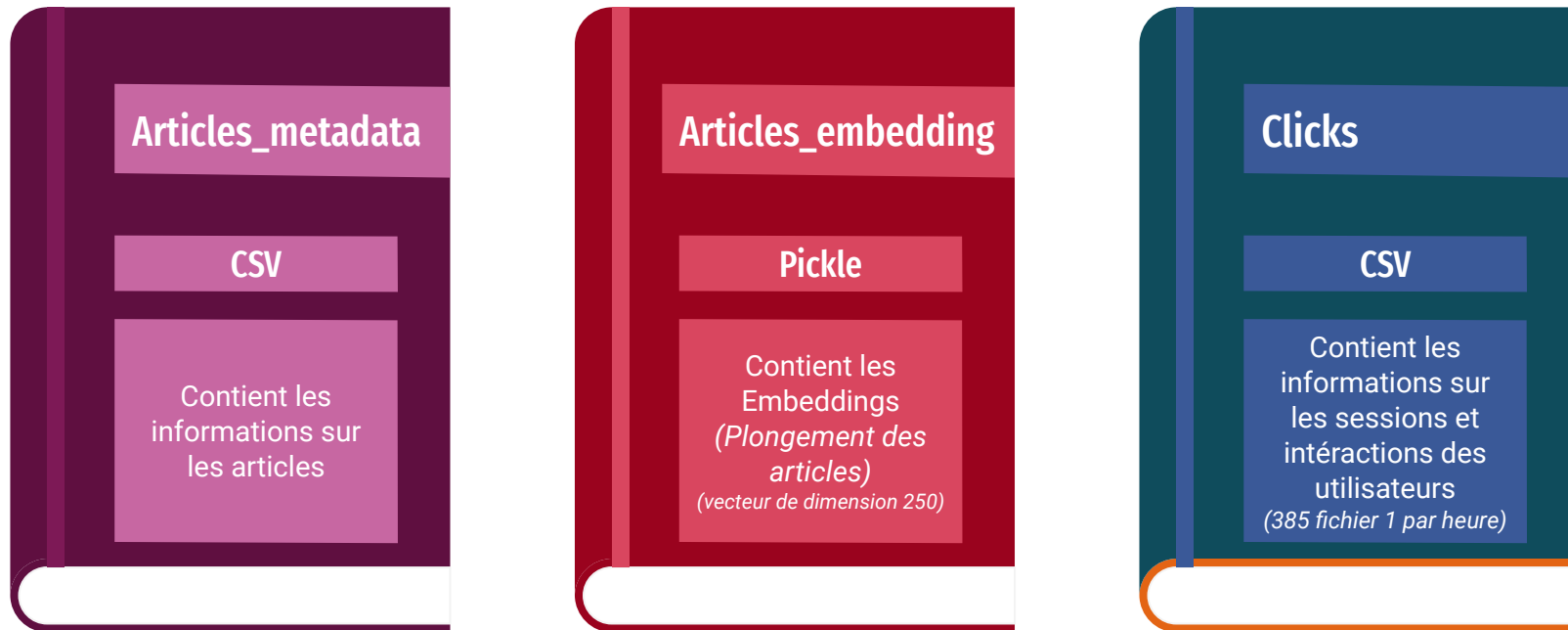
# Appel à projet

01	<b>Le projet</b>	Tester une solution de recommandation d'articles et de livres
02	<b>L'objectif</b>	Réaliser un MVP sous la forme d'une application
03	<b>La mission</b>	<b>Mettre en production une API avec le système de recommandation serverless</b> : Le modèle envoie un ID utilisateur et retourne les recommandations d'articles associées.
04	<b>La méthode</b>	Livrer un prototype fonctionnel du modèle



## **PRÉTRAITEMENT DES DONNÉES**

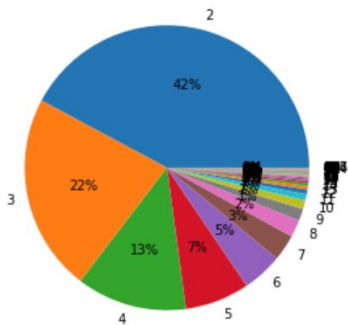
# Jeu de données



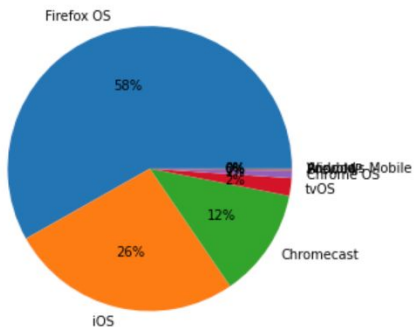
Concaténation des 385 fichiers **clicks** pour en créer un seul  
Regroupement des dataset **articles\_metadata** et **clicks** pour créer notre jeu de données

# EXPLORATION DES DONNÉES

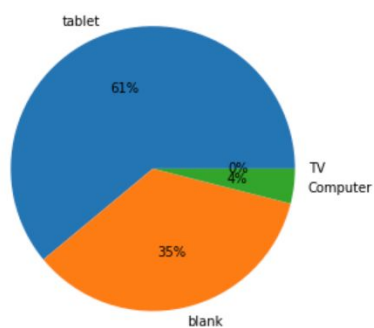
### Distribution du nb d'article lus par sessions



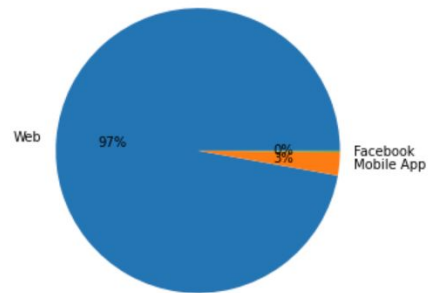
Systeme d'exploitation



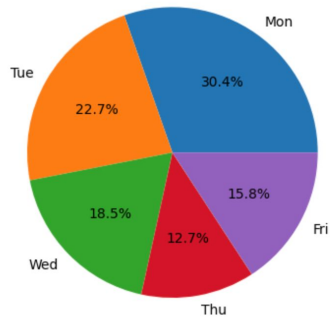
Type de terminal utilisées pour la lecture



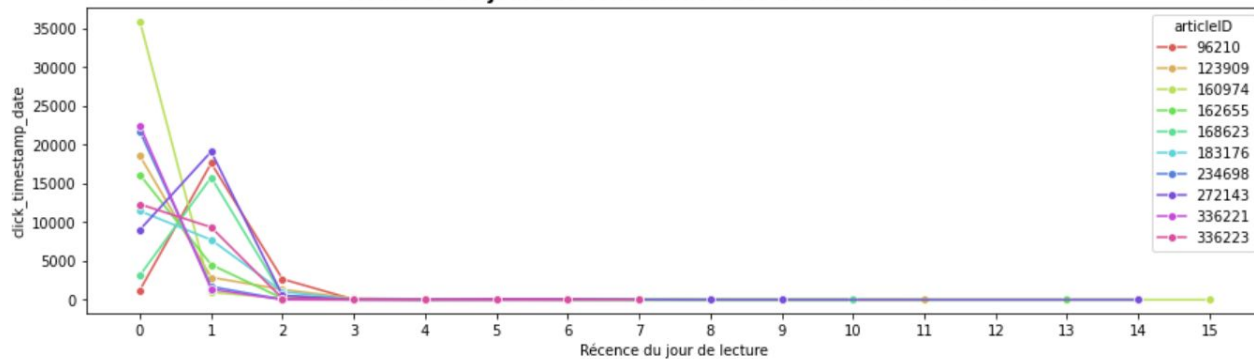
Lecture des articles par plateforme de connexion



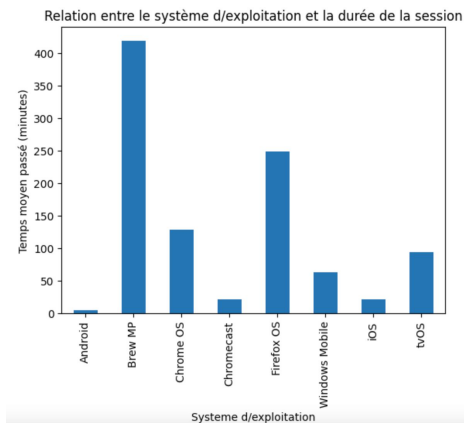
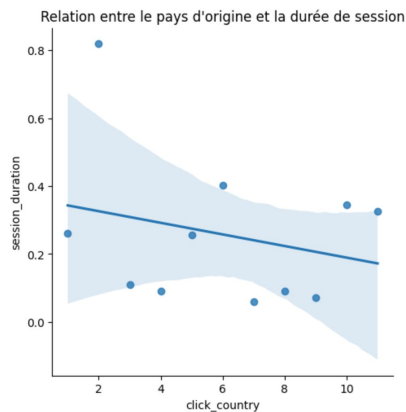
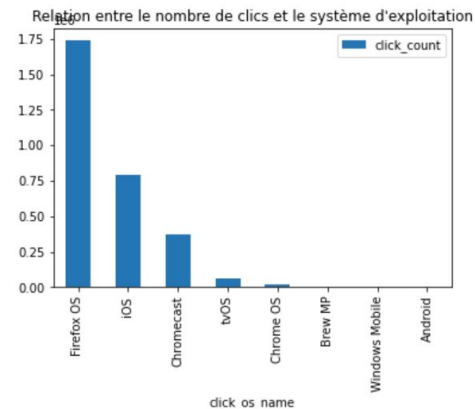
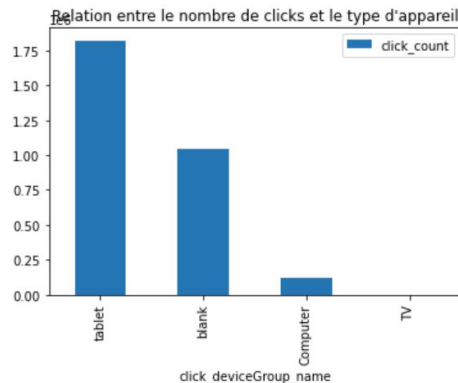
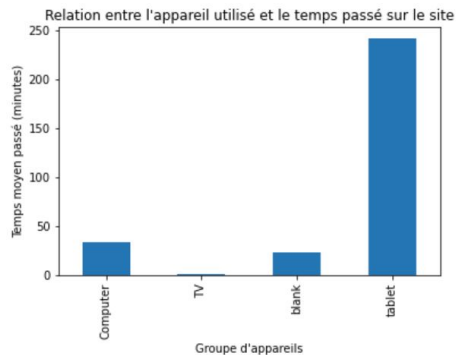
### Distribution du nombre d'articles par jour



### Nombres de jours de consultation des 10 meilleurs articles



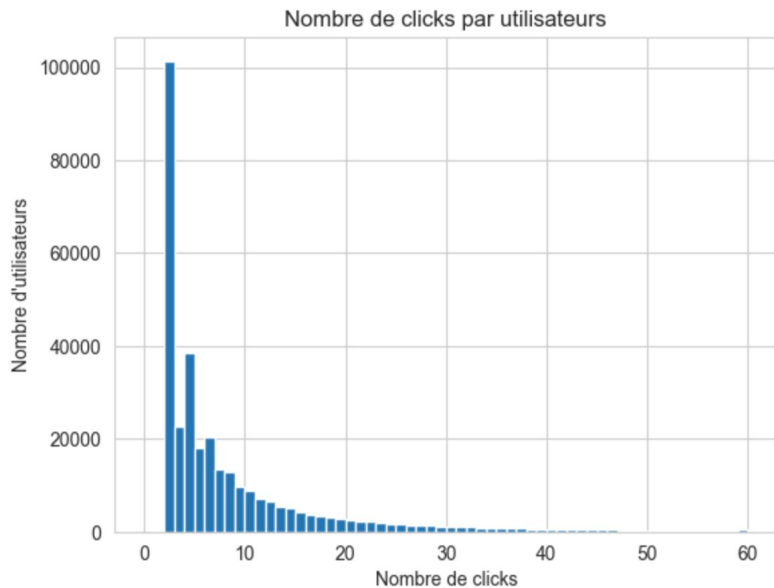
# Analyse des corrélations



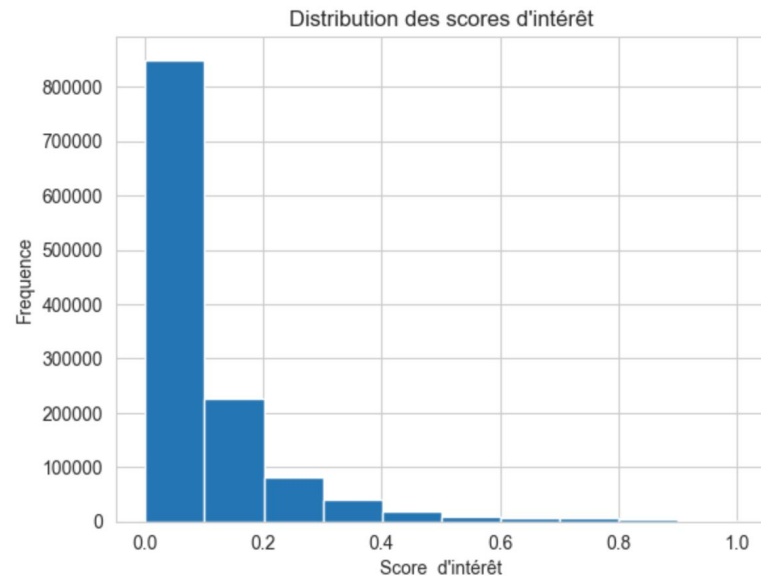


# PRÉTRAITEMENT DES DONNÉES

Suppression des utilisateur ayant moins de 10 clicks



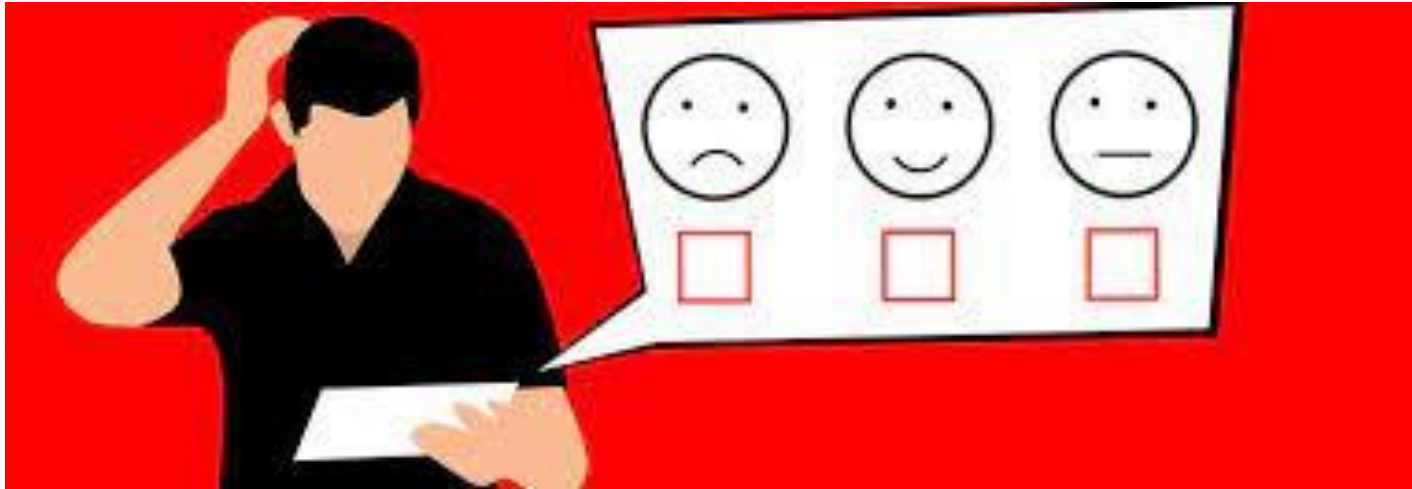
Création d'un score d'intérêt





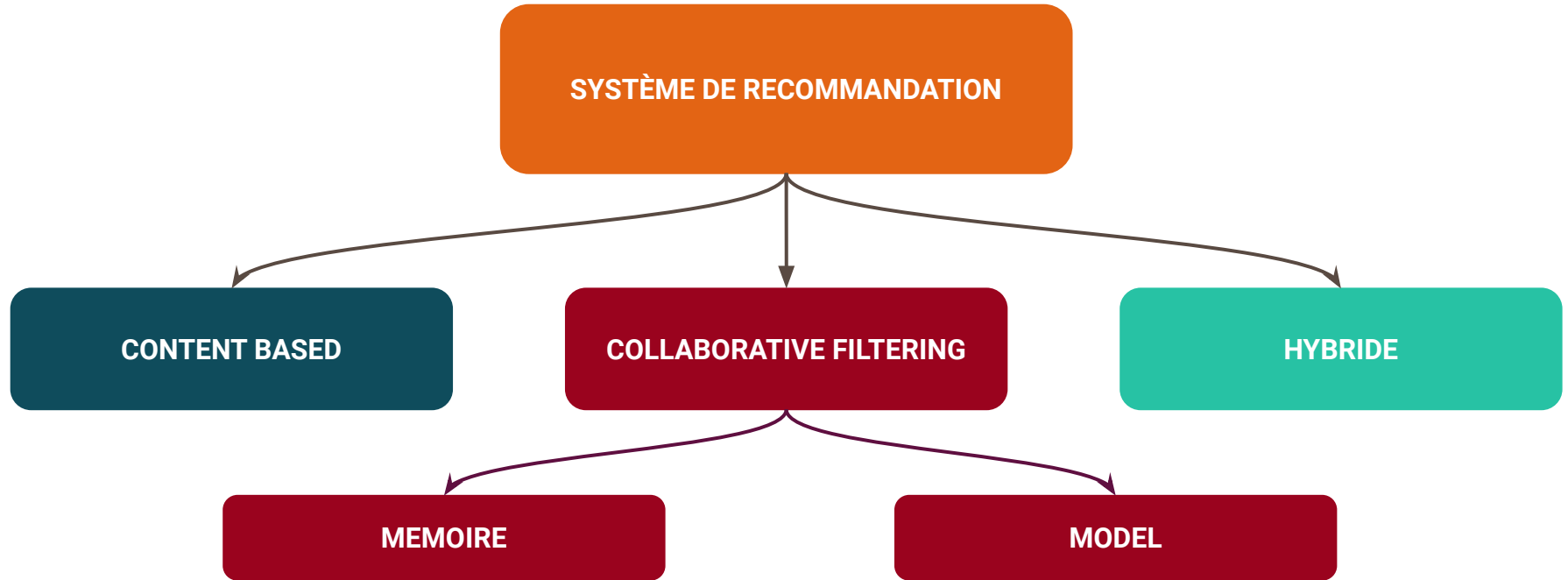
# **SYSTÈME DE RECOMMANDATION**

# SYSTÈME DE RECOMMANDATION

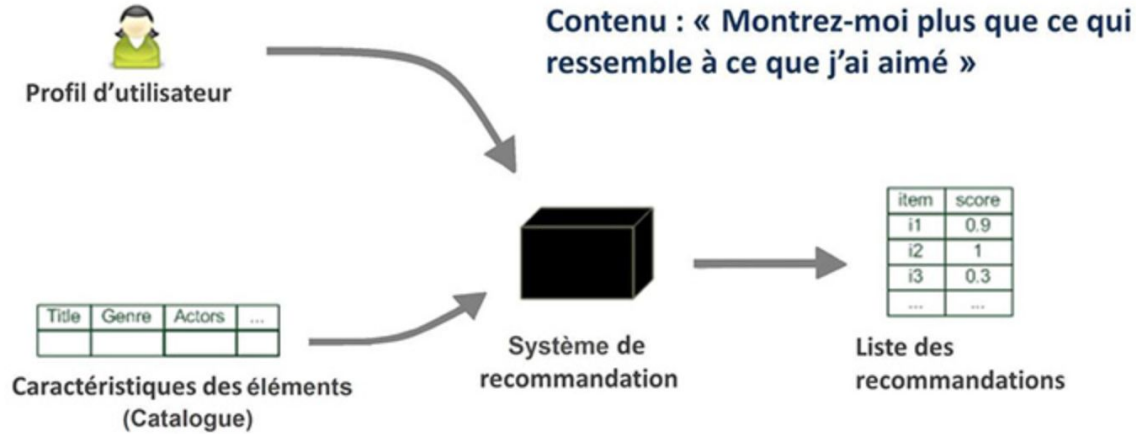


**Les systèmes de recommandation** analysent les préférences et les comportements d'un utilisateur pour lui recommander des articles ou produits similaires, en utilisant différentes techniques

# TYPE DE SYSTÈME DE RECOMMANDATION

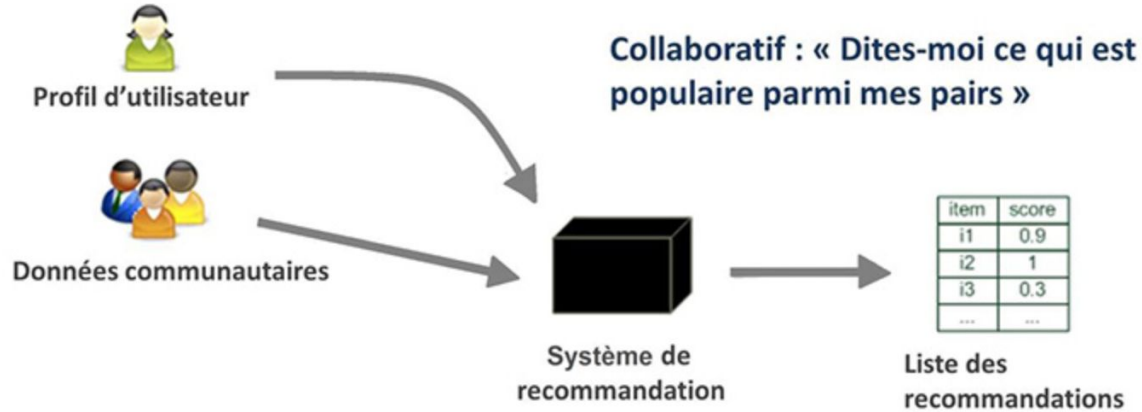


# Modèles basés sur le contenu (Content-Based)



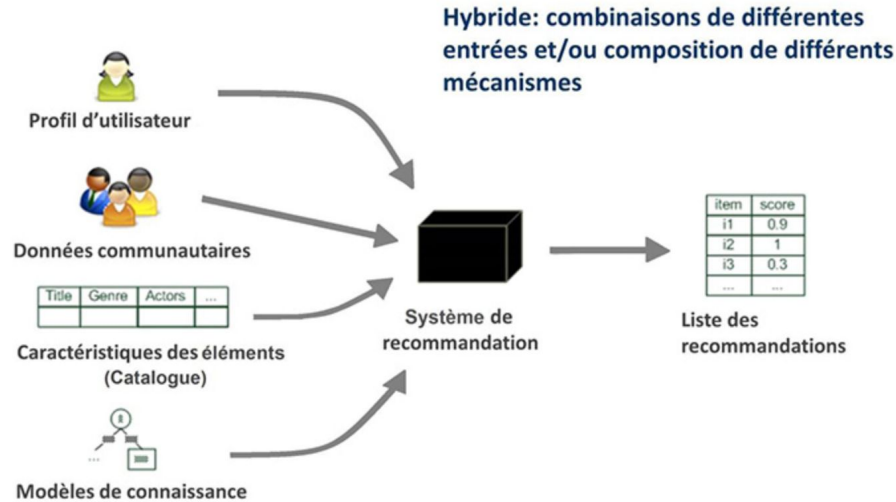
**Le système de recommandation basé sur le contenu (content-based) fonctionne en analysant les caractéristiques et les attributs des articles préférés par un utilisateur, puis en recommandant des articles similaires en termes de contenu.**

# Filtrage collaboratif (Collaborative Filtering)



***Le système de recommandation basé sur le filtrage collaboratif** fonctionne en analysant les préférences d'un utilisateur et en trouvant d'autres utilisateurs ayant des préférences similaires, pour recommander des articles appréciés par ces utilisateurs similaires à l'utilisateur en question.*

# Systèmes de recommandation hybrides (Hybrid Recommender Systems)



**Le système de recommandation hybride** utilise une combinaison des approches basées sur le contenu et le filtrage collaboratif pour fournir des recommandations personnalisées et précises en utilisant les avantages de chaque méthode.

# Avantages et inconvénient



Recevoir des recommandations pertinentes dès les premières interactions sans données des autres utilisateurs.

## Filtrage collaboratif

Utilise les comportements passés des utilisateurs pour recommander des éléments similaires, sans besoin d'informations sur les articles ou les utilisateurs.

## Hybride

Combine Content-based et filtrage collaboratif pour des recommandations précises et personnalisées.



Nécessite beaucoup d'articles et d'infos pour être efficace en utilisant les caractéristiques pour des recommandations personnalisées.

Difficultés à recommander des éléments impopulaires ou nouveaux avec peu d'interactions utilisateurs-articles.

Plus complexe et coûteux

la combinaison de différentes approches peut rendre le système plus difficile à comprendre et à interpréter,



# Choix du modèle (*Implicit*)

## Sans matrice d'embedding

	model	precision	map	ndcg	train_time
0	AlternatingLeastSquares	0.00628	0.00349	0.00470	9.86
1	BayesianPersonalizedRanking	0.00024	0.00020	0.00023	6.09
2	LogisticMatrixFactorization	0.01726	0.00887	0.01231	0.52

## Avec matrice d'embedding

	model	precision	map	ndcg	train_time
0	AlternatingLeastSquares	0.00518	0.00284	0.00385	9.23
1	BayesianPersonalizedRanking	0.00033	0.00019	0.00024	6.11
2	LogisticMatrixFactorization	0.02411	0.01220	0.01696	0.53

## Résultat :

ID Utilisateur = 44 :

	article_id	score
0	336223	6.59
1	96074	6.18
2	285719	4.81
3	96470	4.70
4	235230	3.95

Le score est calculé en utilisant une **fonction sigmoïde**, cette fonction transforme la prédiction en un score compris entre 0 et 1, qui représente la **probabilité** que l'utilisateur **aime l'élément**.



## **AZURE FONCTION, GIT ET DÉPLOIEMENT DE L'API**

# Serveless



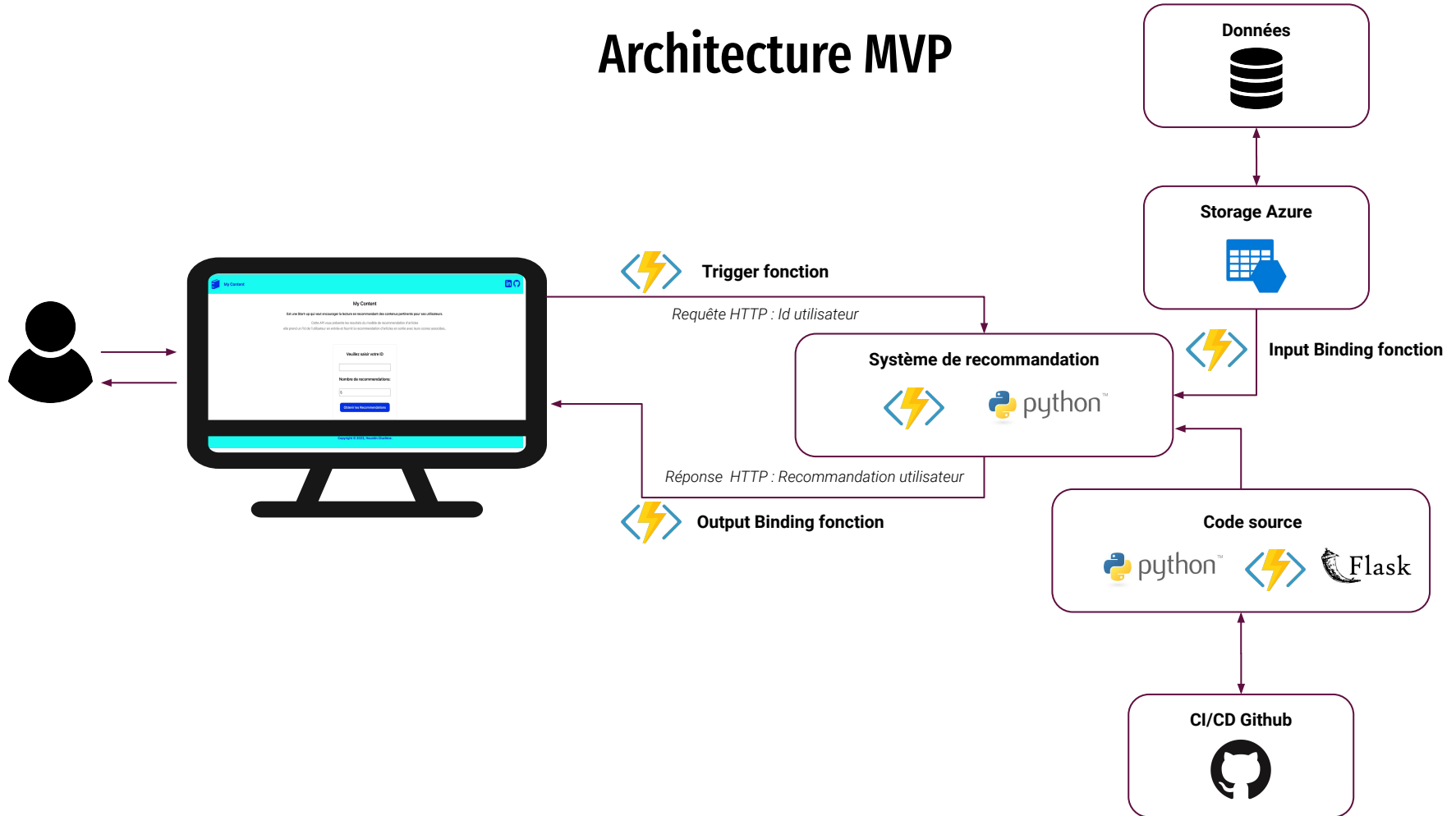
**Le serverless ("sans serveur")** est une méthode de développement qui permet l'exécution de code dans un environnement de cloud computing sans gérer les serveurs physiques, laissant ainsi les développeurs se concentrer sur le code plutôt que sur l'infrastructure.

# Azure functions

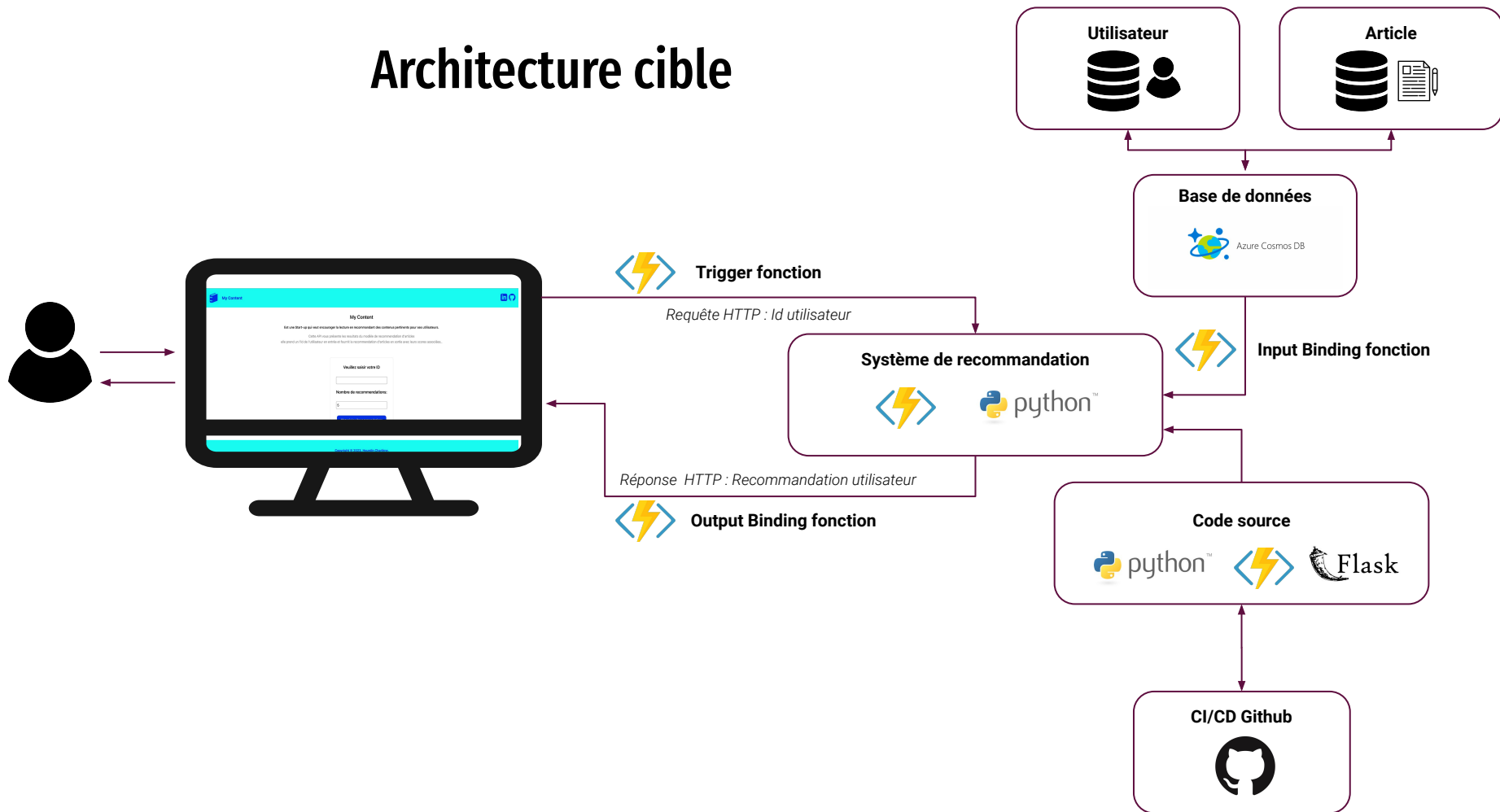


**Une Azure Function** est un service cloud serverless permettant aux développeurs de coder des fonctions qui permettent l'exécution de code en réponse à des événements spécifiques.

# Architecture MVP



# Architecture cible



# Déploiement du modèle



## Flask

Flask est un **micro-framework** web en Python pour créer rapidement des applications web légères et simples.



## GIT

Git est un système de **contrôle de version** open source, permet ainsi de garder une trace de chaque version de votre projet



## Azure

Azure est une plateforme de **cloud computing** de Microsoft qui offre une gamme de services pour créer, déployer des applications dans le cloud.

# Conclusion



## Résultat

- ✓ Nous avons développé une première version de notre système de recommandation sous forme **d'Azure Functions** avec un **déploiement continue CI/CD**
- ✓ Les scripts sont stocké dans un **dossier Github**
- ✓ L'API est **fonctionnelle**

## Axe d'amélioration

- ✓ Optimiser les **hyperparamètres**
- ✓ Implémenter l'**architecture cible**
- ✓ Tester une approche de type **réseaux de neurones** *(Les modèles de **NLP** peuvent être utilisés pour analyser le contenu textuel des articles et des commentaires d'utilisateurs pour recommander des articles similaires ou complémentaires.)*



# MERCI

Avez-vous des questions?



 **git** <https://github.com/charlenehourdin/P9-mycontent>

 **Azure** <https://mycontent-reco.azurewebsites.net>