

---

# CULTUR'ADVISOR - DÉVELOPPEMENT D'UNE PLATEFORME

---

MASTER 2 IMIS  
ANNÉE 2022-2023



PRÉSENTÉ PAR : BARODINE ANAËL, MBELEK-NOUGA  
PAUL-DUBIEN, TESSIER ADRIEN, VERVERKE XAVIER  
ENCADRÉ PAR : TANDJAOUI ROSA ET HAKIM AOUDIA



## Remerciements

Tout d'abord, nous tenons à remercier Rosa Tandjaoui, notre mentor, et Matthieu Exbrayat notre encadrant universitaire de nous avoir accompagnés et guidés pour ce projet.

Nous voulons aussi remercier toutes les personnes qui ont contribué au succès de ce projet et qui nous ont aidés lors de la rédaction de ce rapport.

## Table des matières

<b>1</b>	<b>contexte du projet</b>	<b>4</b>
<b>2</b>	<b>organisation de l'équipe</b>	<b>5</b>
2.1	Moyens de communication . . . . .	5
2.2	Répartitions des tâches . . . . .	5
<b>3</b>	<b>Projet</b>	<b>6</b>
3.1	Mis en place du projet . . . . .	6
3.2	Création du serveur Flask . . . . .	6
3.3	Récupération des données . . . . .	6
3.4	Normalisation des données . . . . .	7
3.5	Mise sous forme de base de données . . . . .	8
3.6	Requêtes . . . . .	8
<b>4</b>	<b>Conclusion</b>	<b>8</b>

## 1 contexte du projet

À la croisée de l'agenda culturel, du média, du magazine et du réseau social, Cultur'Advisor est un espace de productions, de consultations et d'échanges autour de la culture.

En France, il n'existe pas de base de données regroupant l'ensemble des produits culturels du territoire. Il existe bien de telles bases mais elles sont restreintes à de plus petites zones géographiques (département, région) ou à certains événements précis (festivals...). Tout cela rend plus difficile pour le public de les découvrir.

L'objectif de l'entreprise est donc de faciliter l'accès à la culture en aidant les acteurs à trouver de nouveaux publics, améliorer leur visibilité, augmenter leur fréquentation et pour le public, à trouver l'expérience culturelle qui leur ressemble, leur correspond car encore aujourd'hui, beaucoup de produits ou services culturels ne trouvent pas leurs clientèles et les clients finaux ne sont souvent pas informés de l'offre, ni même conscient de son intérêt.

Pour cela le projet consiste en la réalisation d'une application afin de permettre aux utilisateurs de découvrir plus facilement cette culture en France.

## **2 organisation de l'équipe**

### **2.1 Moyens de communication**

Après avoir été attribué à ce projet, nous avons rencontré notre mentor aux dill Bootcamp à Bourge et nous avons fait une première réunion en sa compagnie afin de découvrir plus en détails ce que nous devrions faire pour ce projet.

À la suite de cette réunion, nous avons créé un dépôt git sur GitHub afin que nous puissions bien travailler ensemble sur ce projet. Ensuite, pour pouvoir discuter entre étudiants de ce groupe et connaître l'avancement de chacun, nous avons utilisé Discord pour pouvoir partager nos ressources et nos rendus.

Afin que notre mentor puisse connaître l'avancement de notre projet, nous avons utilisé Whatsapp et fait plusieurs réunions avec elle.

### **2.2 Répartitions des tâches**

Afin de mener ce projet à bien, nous avons commencé par se répartir les différentes tâches du projet.

Nous avons, avec notre mentor défini ces différentes tâches qui sont de récupérer l'ensemble des données disponible sur les API, les rassembler et les uniformiser, en faire une base de données, puis pouvoir faire des requêtes sur ces données en fonction des utilisateurs.

Pour commencer, Anaël s'est occupé de créer le serveur Flask ainsi que de regarder comment récupérer les données sur les API.

Ensuite, Adrien s'est occupé de récupérer le reste des données ainsi que de normaliser ces données. Et enfin, Xavier s'est occupé de créer une base de données à partir des données normalisées.

Pour finir, il nous restera à faire des requêtes sur cette base de données.

## 3 Projet

### 3.1 Mis en place du projet

Premièrement, nous avons d’abord commencé par choisir ensemble un langage dans lequel nous pourrions réaliser ce projet et nous avons choisi Python qui nous permettait de créer un serveur Flask et de bien manipuler les différentes données que nous avons. Flask permet de mettre en place un serveur web. Il est chargé de recevoir des requêtes HTTP et éventuellement d’y répondre. Il suffit de déterminer les URL, appelés routes, qui réceptionnent les requêtes et les traitent.



### 3.2 Création du serveur Flask

La création d’un serveur Flask est simple. Pour cela, il faut importer la bibliothèque Flask et appeler le constructeur du même nom. Ensuite il faut rédiger les routes et le traitement des requêtes. Cela se fait en écrivant une méthode qui sera le traitement de la requête à laquelle on ajoute une annotation `@app.route` qui prend en paramètre la route d’où doit venir la requête.

```
@app.route('/')
```

FIGURE 1 – Annotation définissant une route

### 3.3 Récupération des données

Une fois le serveur Flask mis en place, il faut maintenant pouvoir récupérer les différentes données disponibles. Pour cela, nous avons eu par notre mentor une liste d’API contenant différents produits culturels à récupérer et nous avons commencé par analyser ces différentes bases de données. Ensuite, nous avons récupéré chaque colonne de ces bases de données puis nous en avons fait un tableau comme nous montre les 2 images ci-dessous qui correspondent à 2 bases de données que nous avons récupérées.

	source	titre	niveau_scolaire_enfant	contenue_adapte
0	culturecheznous	Le rap français et américain racontés dans le ...	NaN	NaN
1	culturecheznous	Bibliofil : journées confinées, lectures parta...	NaN	NaN
2	culturecheznous	Mission Carcas	107 Primaire Collège Lycée	Audiodescription (AD);Jeu utilisant du HTML;Na...
3	culturecheznous	Jouer avec le patrimoine du Cher	NaN	NaN
4	culturecheznous	Archives thématiques	NaN	NaN
...	...	...	...	...
1144	culturecheznous	Delacroix et le duel romantique	NaN	NaN
1145	culturecheznous	Les Vizim'	Primaire Collège	NaN
1146	culturecheznous	Les curieuses histoires du Muséum	NaN	NaN
1147	culturecheznous	Les Tribunes du Muséum : Agissons pour la gran...	NaN	NaN
1148	culturecheznous	Les playlists des Transmusicales	NaN	NaN

[1149 rows x 19 columns]

FIGURE 2 – Dataframe contenant les données de culture chez nous

	source	titre	site_internet	identifiant_CNM	identifiant_agence_A	sous_categorie_art_visuel
0	festivals-global-festivals_-pl	Tous courts	https://festivaltouscourts.com	NaN	NaN	NaN
1	festivals-global-festivals_-pl	Les enchanteurs	https://www.festival-lesenchanteurs.com/	NaN	NaN	NaN
2	festivals-global-festivals_-pl	Corsica polar	http://polarcorse.kazeo.com/	NaN	NaN	NaN
3	festivals-global-festivals_-pl	Festival international de la BD	http://festivalbdajaccio.corsica/	NaN	NaN	NaN
4	festivals-global-festivals_-pl	Les estivales	https://www.allauch.com/	NaN	NaN	NaN
...	...	...	...	...	...	...
7277	festivals-global-festivals_-pl	P'tits mots P'tits mômes	https://www.sortiravizille.com/2022/03/19/fest...	NaN	NaN	NaN
7278	festivals-global-festivals_-pl	Le chant des sucs	https://sites.google.com/site/lechantdessucs/	NaN	NaN	NaN
7279	festivals-global-festivals_-pl	Fana Manga	https://www.rom-game.fr/agenda/4633-Fana+Manga...	NaN	NaN	NaN
7280	festivals-global-festivals_-pl	Festival Rire	NaN	NaN	NaN	NaN
7281	festivals-global-festivals_-pl	Projet Images	NaN	NaN	NaN	Performance, vidéos d'artistes

[7282 rows x 31 columns]

FIGURE 3 – Dataframe contenant les données des festivals en France

Nous avons donc récupéré chaque colonnes une à une de ces bases de données ce qui nous donne en sortie 5 tableaux avec respectivement 19, 22, 31, 35 et 17 colonnes. Ensuite, avant de regrouper tous ces tableaux en un, nous avons déjà fait une normalisation à l'intérieur de ces tableaux car par exemple dans un tableau, nous avons 2 colonnes différentes qui représentaient le lien d'un site internet (lien1 et lien2) et nous l'avons regrouper en une seule colonne sous le nom de (site-internet) comme le montre l'image ci dessous.

```
# Lien1
if "lien1" in record["fields"]:
    resource.fields["site_internet"] = record["fields"]["lien1"]

# Lien2
if "lien2" in record["fields"]:
    resource.fields["site_internet"] = record["fields"]["lien2"]
```

FIGURE 4 – Exemple de deux colonnes d'un tableaux rassemblé en une

### 3.4 Normalisation des données

Après avoir récupéré ces données et les avoir mis dans des tableaux différents, nous devons donc les rassembler et les normaliser.

Pour cela, nous avons utilisé un Dataframe afin de mettre tous les tableaux récoltés dans un seul. Pour tous les tableaux que l'on a, nous avons regardé les colonnes qui contenaient les mêmes données

mais qui avaient des noms différents puis on a changé ces noms de colonnes afin que chaque tableau de données qui contenait les mêmes données ait le même nom de colonnes afin qu'il soit normalisé. On a donc pu réduire le nombre de colonnes dans ce dernier Dataframe ci-dessous pour passer de 124 colonnes à 90.

	source	titre	description	nb_ecrans	nb_salles_theatre	jauge_theatre
0	culturechezvous	Le rap français et américain racontés dans le ...	Chroniques d'albums, interviews, mixtapes et a...	NaN	NaN	NaN
1	culturechezvous	Bibliofil : journées confinées, lectures parta...	<p>Amateurs de théâtre et de littérature, lise...	NaN	NaN	NaN
2	culturechezvous	Mission Carcas	<p>Un jeu conçu spécialement pour les enfants ...	NaN	NaN	NaN
3	culturechezvous	Jouer avec le patrimoine du Cher	<p>Depuis le début du confinement, une activit...	NaN	NaN	NaN
4	culturechezvous	Archives thématiques	Naviguez parmi les archives en ligne : iconogr...	NaN	NaN	NaN
...	...	...	...	...	...	...
9995	base-des-lieux-et-des-equipements-culturels	Château de Crévy	NaN	NaN	NaN	NaN
9996	base-des-lieux-et-des-equipements-culturels	Eglise Notre-Dame de Lorette	NaN	NaN	NaN	NaN
9997	base-des-lieux-et-des-equipements-culturels	Cimetière	NaN	NaN	NaN	NaN
9998	base-des-lieux-et-des-equipements-culturels	Camp Protohistorique dit du Castel-Finans	NaN	NaN	NaN	NaN
9999	base-des-lieux-et-des-equipements-culturels	Camp Protohistorique de Kastel-Ker-Nevé	NaN	NaN	NaN	NaN

FIGURE 5 – Dataframe contenant toutes les données après la normalisation

Une fois qu'on a construit ce Dataframe avec toutes les données normalisées, nous l'avons transformé en fichier CSV Pour pouvoir ensuite récupérer ce CSV et en faire une base de données. Nous avons aussi remarqué qu'en normalisant tous ces tableaux en un seul Dataframe, Nous avons beaucoup de valeurs Null car par exemple sur le Dataframe ci dessus nous voyons que le tableau culturechezvous contient une description mais le tableau base-des-lieux-et-des-equipements-culturels n'en contient pas ce qui pourrait poser d'éventuelle problèmes lors des requêtes que nous ferons par la suite.

### 3.5 Mise sous forme de base de données

Une fois les données récupérées et stockées sous CSV, il a fallu le parcourir ligne par ligne pour remplir la base de données. Pour réaliser cela nous avons utilisé une base de données MySQL. Malgré le fait que les données soient normalisées, un problème subsiste, les champs sont de longueur variable. Par exemple, le champ description est parfois absent dans certaines lignes mais une autre contient une portion entière de code HTML. Pour permettre l'insertion de ces variables à taille indéterminée nous les stockons sous le type Text.

### 3.6 Requêtes

Le stockage sous forme de base de données nous permettra de réaliser des requêtes sur les données. Nous pourrons donc obtenir des informations en fonction de paramètres.

## 4 Conclusion

Pour conclure, lors de ce projet, nous avons répondu au besoin initial qui était de rassembler toutes les données culturelles afin de les mettre en valeurs pour le public. Malgré quelques soucis lors de la lecture des données disponible sur les différentes API et la mise en place de la base de données, nous avons pu surmonter ces erreurs en nous entraïdant. Nous avons pris un peu de retard au début de ce projet à cause d'une mauvaise compréhension du travail réalisé ainsi que l'absence d'un des membres de notre équipe mais nous avons quand même pu continuer à mener ce projet à bien et qui continuera d'être amélioré jusqu'à la fin.