



LE SON  
**SUN**  
UNIQUE

# Rapport de stage

SUN - Le Son Unique

Baptiste AUDEON

Tuteur de stage: François Picard

Enseignant référent : S.BENANAYA

2ème année BTS SIO

DU 08 janvier aux 23 février 2024

# Tables des matières

Tables des matières.....	2
Page de remerciements.....	3
Glossaire.....	4
Introduction.....	6
Présentation de l'Association.....	7
1. Présentation générale de la radio.....	7
2. L'équipe technique.....	7
Mes missions.....	8
1. Automatiser la création d'agenda.....	8
Conclusion.....	34
Bibliographie.....	35
Annexes:.....	36

## Page de remerciements

Avant d'entrer dans les détails de cette expérience professionnelle, il est important de débuter ce rapport de stage en exprimant ma gratitude envers ceux qui m'ont énormément appris durant cette période, ainsi qu'envers ceux qui ont généreusement contribué à faire de ce stage une expérience enrichissante.

Je tiens tout d'abord à remercier chaleureusement François Picard, mon maître de stage et son collègue Vincent Leveque, qui m'ont non seulement formé, mais m'ont également accompagné tout au long de cette expérience professionnelle avec une grande pédagogie.

Enfin, je souhaite exprimer ma gratitude envers l'ensemble de l'équipe de SUN, qui m'a prodigué des conseils précieux tout au long des cinq semaines de mon stage.

# Glossaire

Si un mot est suivi d'une astérisque\*, cela signifie qu'il est défini dans le glossaire.

**VM\*** (virtual machine) : machine virtuelle

**GRAM\*** (groupement des radios associatives de la métropole nantaise): A Nantes, le Gram regroupe des radios associatives de la métropole nantaise, notamment Alternantes, Eur@dio, Fidélité, Prun', Jet FM et Sun.

**DAB\*+<+(Digital Audio Broadcasting)**: en anglais. Le « + » correspond à une évolution de la norme mondiale de diffusion de la radio numérique terrestre (RNT). Le DAB\*+ est une technologie de modulation et de transmission numériques de la radio.

**Laravel\*** : Un framework PHP pour le développement web qui facilite des tâches courantes comme l'authentification, le routage, les sessions et le caching.

**Directus\*** : Un système de gestion de contenu (CMS) open-source qui vous permet de transformer n'importe quelle base de données SQL en API\* RESTful.

**N8N\*** : Un outil d'automatisation open-source qui permet de connecter diverses applications et systèmes via des workflows configurables.

**API\*** : Acronyme de Application Programming Interface, une interface logicielle qui permet à deux applications de communiquer entre elles.

**Flux RSS\***: c'est l'acronyme de « Rich Site Summary » (sommaire riche de site web), mais il est généralement traduit par « Really Simple Syndication » (syndication vraiment simple). Un flux RSS\* est un fichier texte qui contient les titres des derniers articles mis en ligne par un site web ainsi que les liens vers ceux-ci.

**Postman\*** : Un outil de développement API\* qui facilite la création, le test, et la documentation d'API\*s.

**Base de données MySQL\*** : Un système de gestion de base de données relationnelle open-source, largement utilisé pour stocker et gérer les données de manière structurée.

**Visual Studio Code\*** : Un éditeur de code source léger mais puissant qui supporte une multitude de langages de programmation et d'outils de développement.

**Termius\*** : Une application de terminal et de SSH qui facilite l'accès à distance et la gestion de serveurs.

**RustDesk\*** : Un logiciel open-source de bureau à distance pour le contrôle et l'accès à distance aux systèmes informatiques.

**LM Studio\*** : LM Studio\* est une application qui offre une solution simple pour télécharger et exécuter localement des Modèles de Langage de Grande Taille (LLM). Cette application intègre également une API\*.

**Text-Generation-WebUI\*** : Il s'agit d'une interface utilisateur graphique conçue pour faciliter l'utilisation de Modèles de Langage Étendus (LLM). Elle est dotée d'une API\* intégrée, rendant son utilisation encore plus pratique.

**DragUploader\*** Ce logiciel spécifique, développé par Vincent pour SUN, la radio nantaise, est dédié à la gestion de l'upload de fichiers audio, simplifiant ainsi les processus internes liés à la gestion de contenu.

**Node.js\*** : Un environnement d'exécution pour JavaScript construit sur le moteur V8 de Chrome. Node.js\* permet aux développeurs d'utiliser JavaScript pour écrire des scripts côté serveur.

**Scraping\*** : Le web scraping, est une technique d'extraction des données de sites web par l'utilisation d'un script ou d'un programme dans le but de les transformer et les réutiliser dans un autre contexte comme l'enrichissement de bases de données, le référencement ou l'exploration de données.

# Introduction

Durant la période du 08 janvier au 23 février 2024, j'ai eu l'opportunité d'effectuer un stage au sein de la radio SUN sun le son unique. Mon stage s'est déroulé à la technique constituée de François Picard et de Vincent Leveque. Ils m'ont confié la mission d'automatiser la création d'agenda d'évènement.

Dans un premier temps, nous présenterons la radio. Afin de mieux appréhender le contexte de mon stage, il est essentiel de comprendre l'importance de la gestion du système d'information au sein de la radio. Enfin, nous détaillerons les missions qui m'ont été confiées, ainsi que les connaissances et les compétences que j'ai pu développer au cours de cette expérience .

## Présentation de l'Association

### 1. Présentation générale de la radio

SUN, acronyme de « Son Unique à Nantes », était à l'origine une radio pirate fondée en 1985 par Pierre Boucard, Ludovic Avery et François Picard, qui est également mon maître de stage. Elle s'est transformée en radio associative en 1996.

Sa présence à Cholet marque un nouveau tournant pour l'identité de la radio puisque SUN « Le Son Unique à Nantes » devient SUN « Le Son Unique ».

Première radio associative des Pays de la Loire avec 18 000 auditeur·rice·s par jour. Elle couvre trois départements : la Loire-Atlantique, la Vendée et le Maine-et-Loire. Elle est disponible en FM à Nantes (93.0FM) et Cholet (87.7FM), et en DAB\*+ à Nantes, Angers, Saint-Nazaire, Pornic et La Roche-sur-Yon.

Depuis 1996, SUN s'attache à produire des contenus éditoriaux tournés vers le territoire, qui rendent compte de l'actualité, des événements, des initiatives citoyennes, associatives, entrepreneuriales... au sein d'un programme sans publicité. Cinq salarié·e·s encadrent chaque saison près de 90 bénévoles afin de donner vie à une soixantaine de programmes abordant des thématiques et sujets primordiaux : culture, jeunesse, égalité, environnement, inclusion ou encore lutte contre les discriminations. La ligne éditoriale se veut indépendante, apolitique, locale, accessible pour ceux·celles qui la font comme pour ceux·celles qui l'écoutent, et détachée de l'actualité à chaud. SUN s'engage pour la défense des médias indépendants et pour un travail sur l'information libre et éclairé.

## 2. L'équipe technique

Mon tuteur de stage, responsable technique de l'infrastructure chez SUN, possède une expertise étendue en gestion de l'infrastructure réseau et de la technique sonore. Par ailleurs, Vincent, développeur web et technicien multimédia, détient également une connaissance approfondie dans ces domaines, ce qui leur permet une collaboration efficace et la capacité de se remplacer mutuellement au besoin.

François en charge de l'infrastructure réseau, gère également la technique sonore. Son employeur donne de son temps à la gestion du flux audio pour les radios du GRAM\*, incluant Alternantes, Fidélité, Prun', et Jet FM. Il veillent au bon fonctionnement de l'équipement et du patrimoine informatique de SUN.

Vincent, quant à lui, est responsable du développement du site internet de SUN en utilisant le framework Laravel\* et du développement d'outils métiers comme DragUploader\*. Il s'occupe également de la prise et du montage vidéo.

# Mes missions

## 1. Automatiser la création d'agenda

### Qu'est-ce qu'un agenda

Les agendas des événements culturels locaux sont diffusés sur SUN, sous forme de flashes informatifs et en tant que brefs articles sur le site [lesonunique.com/mysun/agenda](http://lesonunique.com/mysun/agenda). Ces agendas visent à promouvoir les initiatives culturelles locales dans les zones de diffusion de la radio, notamment Nantes, Cholet, Saint-Nazaire, Angers et La Roche-sur-Yon. Ils offrent des informations sur les artistes, les lieux, les tarifs, et les modalités d'accès aux événements. Les événements sont souvent accompagnés d'une vidéo sur le site web.

### Les enjeux de ma mission

Durant mon stage, la préparation des agendas de Nantes était assurée par Taissia, qui est en service civique. Quant aux autres agendas, c'était Agathe, une autre stagiaire présente en même temps que moi, qui s'en occupait. Cette tâche répétitive a conduit l'équipe à envisager son automatisation, afin de réduire le temps consacré à cette activité. Leur travail consistait à rechercher des événements intéressants sur plusieurs sites web, à sélectionner une image, rédiger un texte avec les informations mentionnées plus haut, enregistrer les flashes radio, adapter le texte pour le site web, recadrer l'image pour le site via Photoshop, et trouver une vidéo si nécessaire.

Le but de ma mission était de centraliser toutes les données des sites habituellement utilisés sur Directus\*, afin de rassembler toutes les informations au même endroit. Grâce à

cela, nous pouvions trier les informations de multiples sites sur une seule plateforme. Il serait également possible de visualiser ces informations via un agenda et une carte interactifs. Nous pouvions éditer le texte, le titre et recadrer les images.

- Directus**
- City
- Modele Agenda
- Angers calend...
- Angers carte
- Angers liste
- Cholet calendr...
- Cholet carte
- Cholet liste
- La Roche-sur...
- La Roche-sur...
- La Roche-sur...
- Nantes calend...
- Nantes carte
- Nantes liste
- St Nazaire cale...
- St Nazaire carte
- St Nazaire liste

### Modele Agenda St Nazaire liste

Title	First Date	Link	Image
Chassol, Big Sun	8 février 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/chassol-big-sun">https://www.saintnazaire.fr/agenda/chassol-big-sun</a>	
Takomodé ! Un concert de chants et musiqu...	5 avril 2024 10:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/takomode-un-cc">https://www.saintnazaire.fr/agenda/takomode-un-cc</a>	
Concert The Forroscopic's	26 janvier 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/concert-the-forr">https://www.saintnazaire.fr/agenda/concert-the-forr</a>	
Ateliers Parents Bébé « éveil des sens » Déc...	11 avril 2024 10:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/ateliers-parents-">https://www.saintnazaire.fr/agenda/ateliers-parents-</a>	
Arlequin poli par l'amour, Marivaux, Thomas ...	13 mars 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/arlequin-poli-par">https://www.saintnazaire.fr/agenda/arlequin-poli-par</a>	
« En Terre », Cie NoMORPa	19 janvier 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/en-terre-cie-non">https://www.saintnazaire.fr/agenda/en-terre-cie-non</a>	
Ateliers Parents Bébé « éveil des sens » Déc...	4 avril 2024 10:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/ateliers-parents-">https://www.saintnazaire.fr/agenda/ateliers-parents-</a>	
Salti, Brigitte Seth, Roser Montlló Guberna	16 février 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/salti-brigitte-set">https://www.saintnazaire.fr/agenda/salti-brigitte-set</a>	
Stage de solo jazz roots, avec Julia Passot	17 février 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/stage-de-solo-ja">https://www.saintnazaire.fr/agenda/stage-de-solo-ja</a>	
Roman d'apprentissage, Valérie Mréjen, Mais...	17 avril 2024 10:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/roman-d-appren">https://www.saintnazaire.fr/agenda/roman-d-appren</a>	
Kaleidos Songs et les A'typiques en concert	19 janvier 2024 11:00:00 PM	<a href="https://www.saintnazaire.fr/agenda/kaleidos-songs-">https://www.saintnazaire.fr/agenda/kaleidos-songs-</a>	

- Modele Agenda
- Angers calend...
- Angers carte
- Angers liste
- Cholet calendr...
- Cholet carte
- Cholet liste
- La Roche-sur...
- La Roche-sur...
- La Roche-sur...
- Nantes calend...
- Nantes carte
- Nantes liste
- St Nazaire cale...
- St Nazaire carte
- St Nazaire liste

The map displays the coastline of the Loire River estuary and the surrounding land. Key locations labeled include: Morais Indivis de Grande Briere Mottière, île de Brécun, île de Bais, île de Fédrun, île de Ménac, île d'Aignac, Saint-Malo-de-Guersac, Montoir-de-Bretagne, Aéroport de Saint-Nazaire Montoir, Le Guézec, D 47, D 213, L'Immaculée, Prézégal-Berthauderie, Le Petit C, Rivot-Plaisance, La Baule-Escoublac, La Vézinière, Le Petit Gavy, Les Rochelles, Sainte-Marguerite, Saint-Marc, and Villés Martin. The map also shows the N 17 road and several bridges across the river.

ci dessus les différents formats d'affichages des données

L'intelligence artificielle va être utilisée pour trier les articles grâce à une liste de mots-clés générée automatiquement, et pour créer un brouillon de texte sur lequel l'opérateur pouvait s'appuyer.

À terme, l'objectif est que l'humain n'est plus qu'à sélectionner les événements les plus pertinents, vérifier et corriger les textes, titres et images recadré automatiquement, enregistrer le flash radio, puis appuyer sur bouton pour envoyer les articles, en brouillon sur le backend du site SUN lessonunique.com. Cette approche représente un gain de temps considérable.

## Démarche et Réflexions

Dans le cadre de ma mission, j'ai compris qu'il était primordial de maîtriser l'exécution manuelle d'une tâche avant de pouvoir l'automatiser. Pour cela, je me suis renseignée sur le processus de création manuelle d'agendas, que j'ai décrit auparavant.

J'ai ensuite cherché à comprendre et à débugger les prototypes N8N\* ainsi que l'API\* Directus\*, mis en place par le service technique avant mon arrivée. Le fait que ces éléments soient déjà en cours de développement a facilité ma compréhension rapide du fonctionnement de ces deux technologies.

## Outils Utilisés

Pour réaliser cette mission, j'ai utilisé plusieurs outils : N8N\*, Directus\*, Postman\*, une base de données MySQL, Visual Studio Code\*, l'API\* ChatGPT, Termius\*, RustDesk\*, LM Studio\* et Text-Generation-WebUI\*.

## Démarche et Réflexions

J'ai segmenté le processus de récupération des données en deux étapes distinctes afin d'améliorer la scalabilité du système.

### Workspace par Sites:

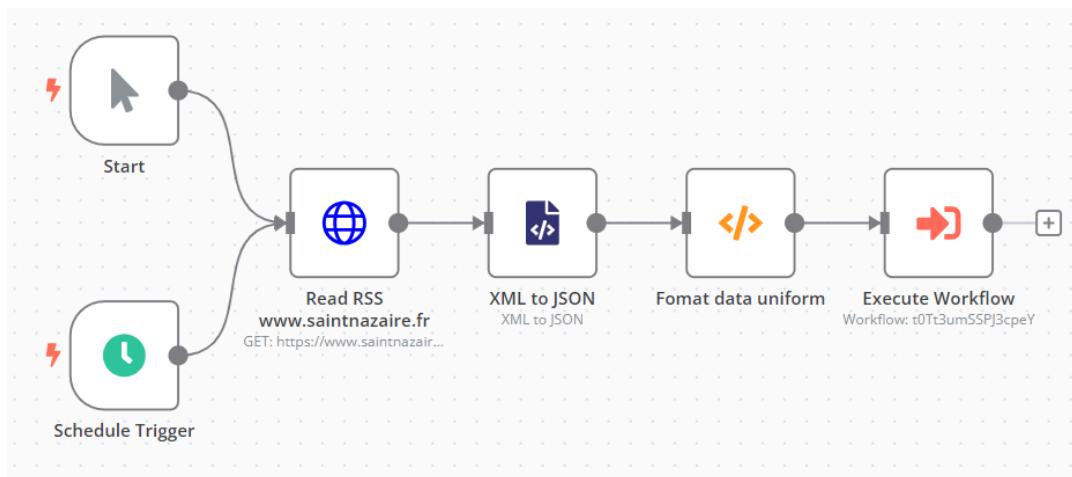
Pour la première phase du workflow, il est nécessaire d'établir un workspace distinct pour chaque site web afin de récupérer les données spécifiques à chacun. Cette approche permet une personnalisation et une adaptation aux différentes structures et formats de données présents sur chaque site.

### Workspace Directus\* :

La seconde workspace joue un rôle crucial en étant constante et uniforme, quelle que soit la source des données. Son objectif est d'insérer les données fournies par les premiers workspaces dans Directus\*. Cette standardisation assure une intégration fluide et efficace des données dans le système Directus\*, facilitant ainsi leur gestion et leur utilisation ultérieure.

Ces deux parties sont reliées par le nœud Execute Workflow, qui exécute et transmet les données au déclencheur du workflow. Il est crucial de le mettre en place une fois que le système est opérationnel, car la multiplication des espaces compliquerait significativement le débogage.

### Workspaces par Sites



Workflow qui récupère la data de Saint-Nazaire.fr

Le premier nœud récupère les données du flux RSS\* du site de Saint-Nazaire. Ensuite, le deuxième nœud convertit ces données du format XML au format JSON. Le dernier nœud, une boucle JavaScript, trie, crée et modifie les données pour les uniformiser.

```

items = $input.first().json.rss.channel.item;

return items.map((x => {

  if (typeof x.image === 'array') {
    x.image = x.image[0];
  }

  else if (typeof x.image !== 'string') {
    x.image = null;
  }

  if (x.image !== null) {
    x['nom_fichier_image'] = x.image.match(/\/([^\/]+)$/)[1];
  } else {
    x['nom_fichier_image'] = null
  }

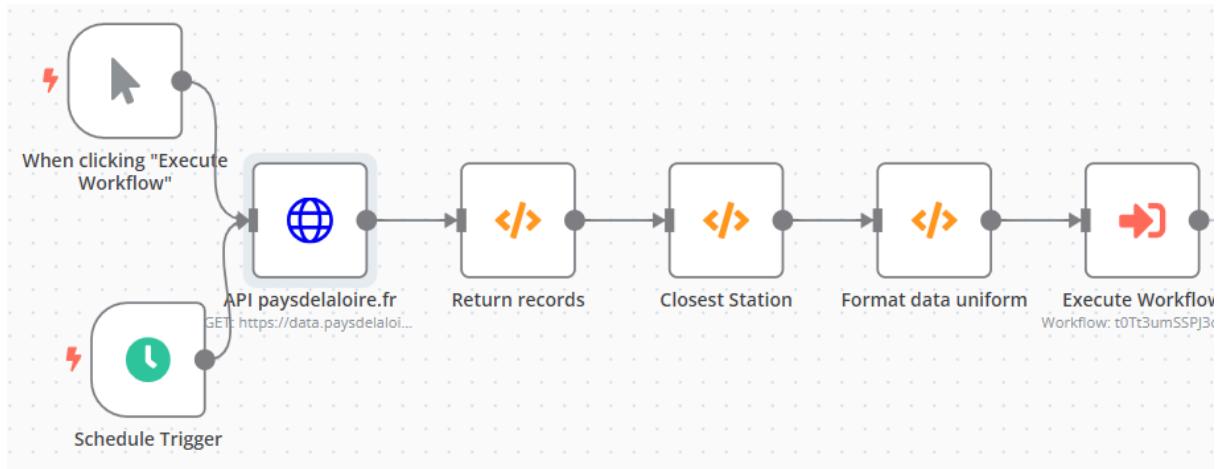
  x['title'] = x.title.replace('Saint-Nazaire | ', '')
  x['uid'] = 'saintnazairefr--' + x.link.replace('https://www.saintnazaire.fr/agenda/', '');
  x['source'] = "RSS SaintNazaire.fr";
  x['closest_station']= "SUN Saint-Nazaire";

  return
})

```

noeud Format data uniform

Cela garantit que les données en sortie soient toujours identiques, assurant ainsi leur compatibilité avec le deuxième workflow.



Workflow qui récupère la data paysdelaloire.fr

Ce workflow opère selon les mêmes principes que le premier. Cependant, au lieu d'utiliser le flux RSS\*, il se connecte via l'API\* du site. À la différence des sites précédents, celui-ci couvre l'ensemble de la région. J'ai donc récupéré les données spécifiques aux stations de radio et les ai triées par station.

```
const latitude = $input.item.json.fields.location_coordinates[0];
const longitude = $input.item.json.fields.location_coordinates[1];

$input.item.json.fields.closest_station = getClosestRadioStationFromUser(latitude, longitude);

return $input.item.json.fields;

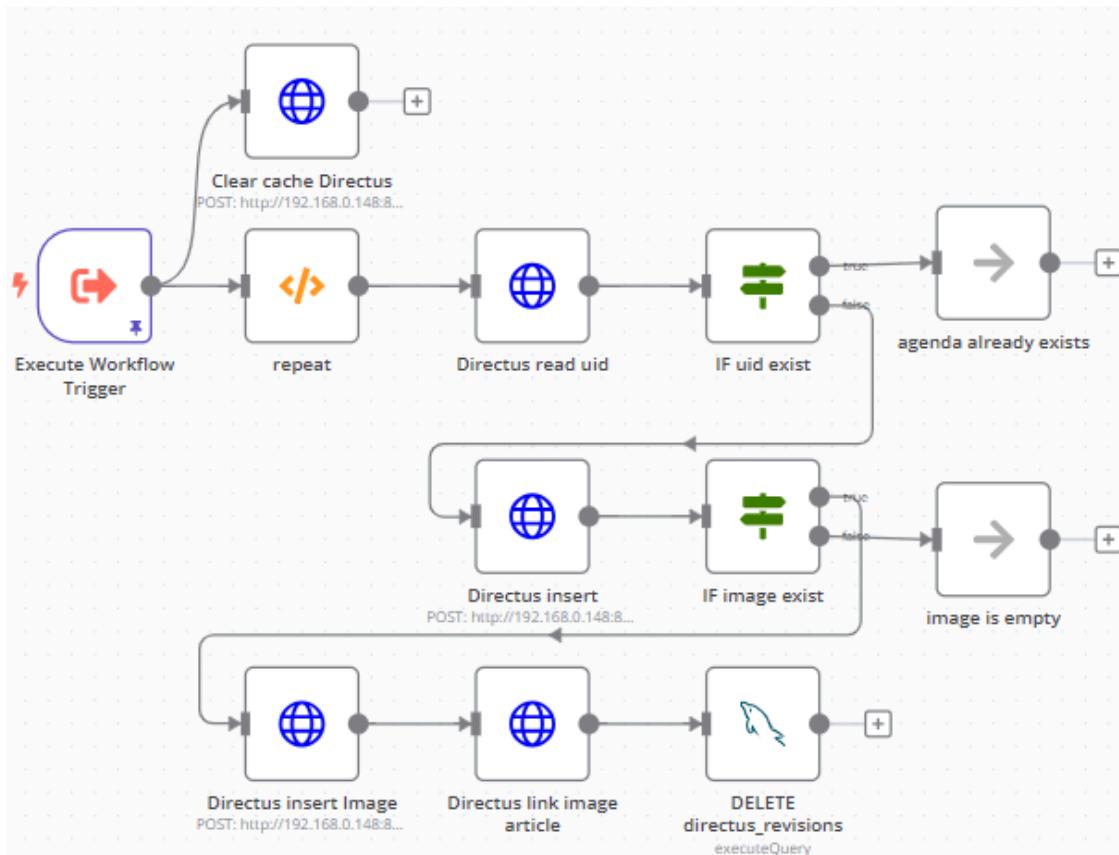
function getClosestRadioStationFromUser(browserLatitude, browserLongitude) {
  var locatableRadioStations = [
    { id: 2, name: 'SUN Nantes', latitude: 47.218372, longitude: -1.553621 },
    { id: 2, name: 'SUN Cholet', latitude: 47.057869, longitude: -0.87945 },
    { id: 2, name: 'SUN Saint-Nazaire', latitude: 47.27433, longitude: -2.21374 },
    { id: 2, name: 'SUN Angers', latitude: 47.473303, longitude: -0.5533221 },
    { id: 2, name: 'SUN La Roche-sur-Yon', latitude: 46.669848, longitude: -1.426689 }
  ];
  ...

}
```

#### noeud Closest Station

Dans la suite du processus, je formate les données pour que les JSON correspondent à ceux des workflows précédents.

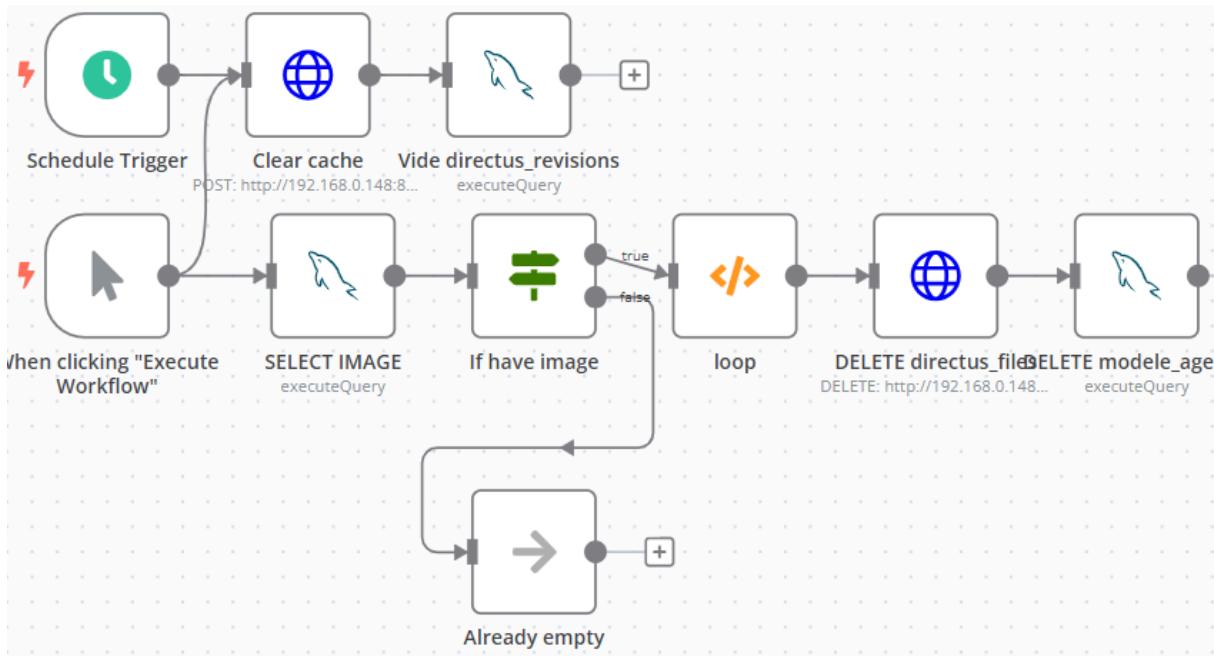
## Workspace Directus\*



Workflow insert donnée dans Directus\*

Ce workflow récupère les données des workflow précédents, les reproduit pour qu'elles soient accessibles dans ce nouveau workspace. Via l'API\* de Directus\*, je récupère l'identifiant unique pour vérifier si les articles existent déjà. Ensuite, j'insère les nouveaux articles dans Directus\*. Je tri les images absentes, je les insère également et les relie aux articles correspondants. Enfin, je supprime les données inutiles.

## Workflow de Suppression de Données



Workflow qui supprime les données correctement

J'ai développé un workflow qui supprime efficacement les données. Utilisant l'API\* pour les images et des requêtes SQL pour les articles, ce processus permet d'éliminer les données obsolètes, comme les événements passés.

## Utilisation de Directus\*

Les données récupérées et insérées dans Directus\* comprennent plusieurs éléments clés pour chaque événement. Voici la liste détaillée :

- Closest\_Station : Il s'agit de la station de radio la plus proche de l'événement.
- Description : Cette donnée représente la description officielle du site de l'événement.
- Description\_GPT : Un brouillon généré par Chat GPT pour chaque événement.
- First\_Date : La date de début de l'événement.
- Last\_Date : La date de fin de l'événement.
- Image : L'image originale associée à l'article de l'événement.
- Image\_600 : Une version recadrée de l'image originale, aux dimensions 600x844.
- Link : Le lien vers l'article original de l'événement.
- Location\_Name : L'adresse postale où se déroule l'événement.
- Mot\_ClefGPT : Les mots-clés générés pour classifier l'événement.
- Position : Les coordonnées géographiques de l'événement.
- Source : Le site source de l'événement.
- Title : Le titre de l'événement.

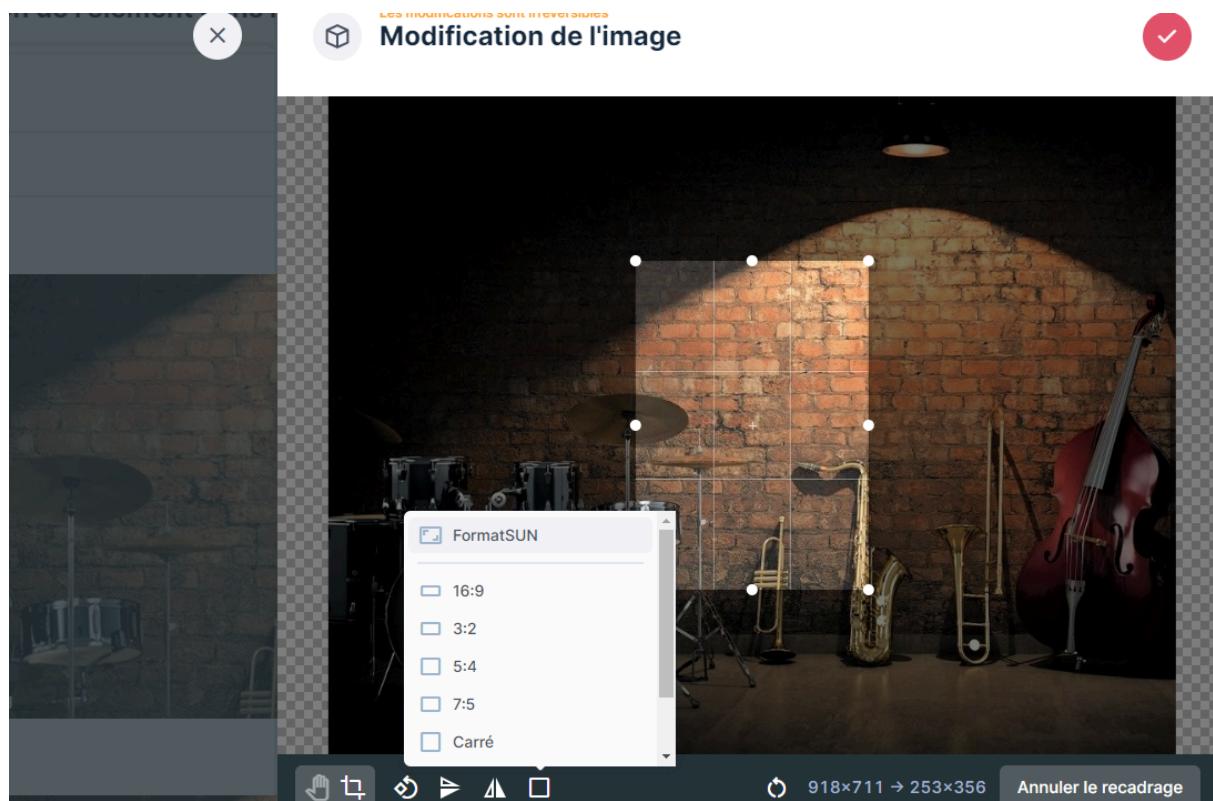
Dans les paramètres, j'ai généré un token pour me connecter à l'API\*. J'ai créé des marque-pages pour chaque station et générée les données correspondantes.

The screenshot shows the Directus interface for managing a 'Modele Agenda' item. The left sidebar lists various items under 'Directus' and 'Modele Agenda'. The main area displays the following fields:

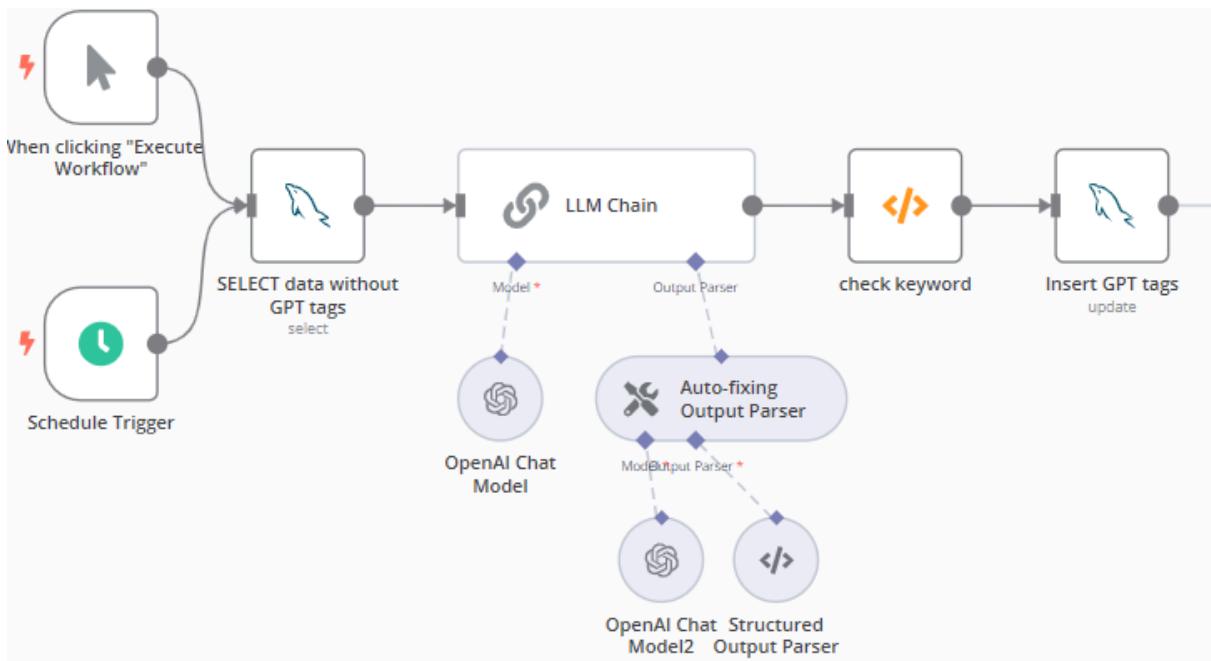
- Title:** Rencontre interprofessionnelle du spectacle vivant
- Link:** <https://www.saintnazaire.fr/agenda/rencontre-interprofessionnelle-du-spectacle-vivant-35422>
- link click:** (button)
- Description:** (Rich Text Editor)
  - Horaires : de 19h à 21h
  - Organisateur : Café-théâtre La P'tite scène des halles

interface directus

crée le Format SUN qui correspond aux ratio des images pour le site 844x600.



## Intégration de Chat GPT

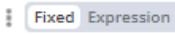


Workflow d'utilisation de chat GPT

Dans ce workflow, je commence par sélectionner certains articles dont le champ 'GPT' sont vides. Cette étape permet de limiter le nombre de requêtes simultanées. Par la suite, j'utilise une requête GPT pour trier ces articles, en leur attribuant des mots-clés. Ces mots-clés sont issus d'une liste que j'ai préalablement fournie à Chat GPT. En parallèle, je demande à Chat GPT de générer une description qui pourra servir de brouillon pour l'opérateur.

</> Structured Output Parser

Parameters Docs 

JSON Schema 

```

1 v {
2   "type": "object",
3   "properties": {
4     "id": {
5       "type": "integer",
6       "description": "id: Un identifiant unique pour l'entrée."
7     },
8     "mot_clefs": {
9       "type": "array",
10      "description": "Liste des mots clés identifiés.",
11      "items": {
12        "type": "string"
13      }
14    },
15    "DescriptionGPT": {
16      "type": "string",
17      "description": "La description détaillée fournie par le modèle GPT."
18    }
19  },
20  "required": ["id", "mot_clefs", "DescriptionGPT"]
21 }
```

The schema has to be defined in the [JSON Schema](#) format. Look at [this page](#) for examples.

## Format de sortie du prompt chat GPT

Les nœuds situés après la chaîne 'LLM' envoient une seconde requête pour formater le résultat obtenu. Cette étape garantit la structuration uniforme des résultats en format JSON, quel que soit l'article traité.

Le processus se poursuit avec un nœud de code qui vérifie la correspondance des cinq mots-clés renvoyés par Chat GPT avec la liste initiale. Cette vérification est cruciale car Chat GPT, en l'absence de correspondance exacte, peut générer des mots-clés alternatifs. Finalement, je transfère la description et les mots-clés validés dans la base de données.

## Démarche et Réflexions

Après avoir mis en place le système, il fonctionnait correctement. Toutefois, je me suis confronté à un problème majeur : l'utilisation de l'API\* du ChatGPT-4 engendrait des coûts significatifs. Pour mes tests, le coût s'élevait à 9 euros pour 100 requêtes, tandis que l'utilisation de ChatGPT-3.5 aurait coûté seulement 30 centimes. Ce coût élevé posait un problème de scalabilité, surtout en envisageant l'augmentation du nombre de sites pour enrichir notre contenu.

La solution que j'ai choisie consiste à utiliser un Large Language Model (LLM) en local, sur une machine virtuelle. Cette méthode entraîne certes un coût matériel important en termes de ressources mémoire et de capacité de la carte graphique, surtout pour obtenir des résultats probants. Cependant, ces ressources peuvent être réutilisées pour d'autres projets. De plus, si nous développons ultérieurement d'autres systèmes nécessitant de l'intelligence artificielle, cet investissement pourra être amorti.

L'association utilise déjà Chat GPT. Dans ce contexte, l'option de Mistral en local pourrait être envisagée, car elle a l'avantage de ne pas transmettre vos informations à OpenAI. Toutefois, il est important de noter que Chat GPT, est connecté à internet et est un LLM un peu plus puissant.

## Installation Text-Generation-WebUI\* sur une vm\* Linux Debian

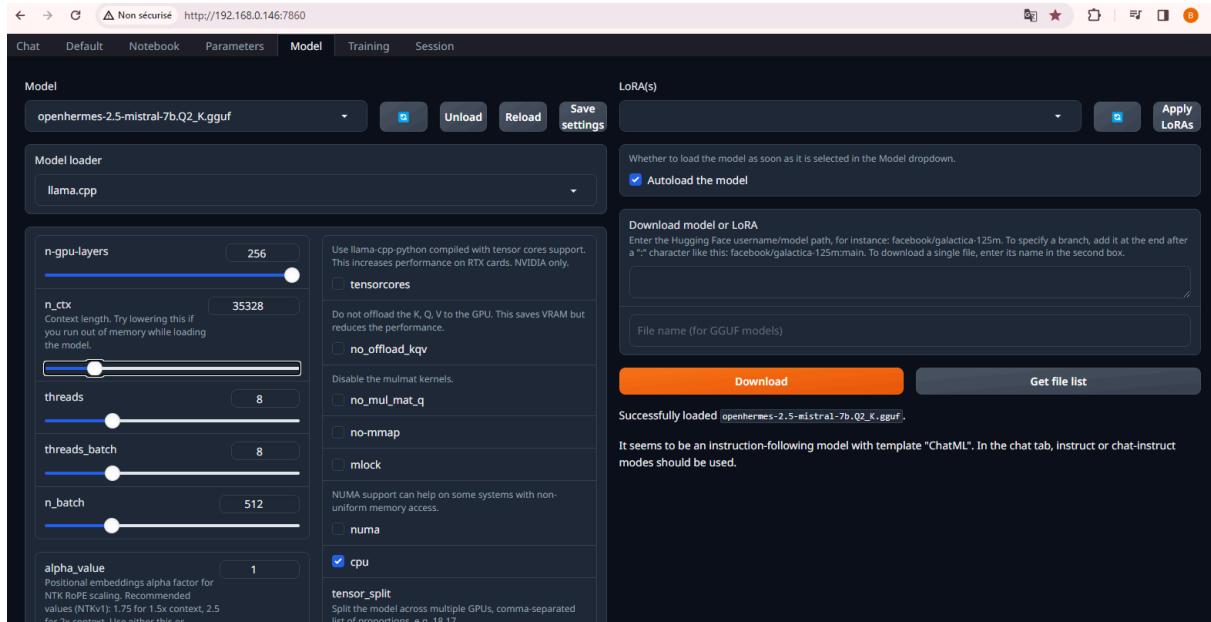
J'ai procédé à l'installation d'un LLM (Large Language Model) sur une machine virtuelle (vm\*) fonctionnant sous Linux Debian. Cette installation a été effectuée en ligne de commande, j'ai installé Git. Pour ce faire, avec Git j'ai d'abord cloné le dépôt depuis [https://github.com/oobabooga/Text-Generation-WebUI\\*](https://github.com/oobabooga/Text-Generation-WebUI*), un projet qui permet de lancer un LLM dans un navigateur et via une API\*.

```
su: Échec de l'authentification
membre@traitementLLM:~$ su root
Mot de passe :
root@traitementLLM:/home/membre# ./start_linux.sh history
bash: ./start_linux.sh: Aucun fichier ou dossier de ce type
root@traitementLLM:/home/membre# history
 1  nano /etc/network/interfaces
 2  nano /etc/ssh/sshd_config
 3  nano /etc/ssh/sshd_config
 4  systemctl reboot
 5  apt-get update
 6  apt-get upgrade
 7  systemctl poweroff
 8  apt-get update
 9  sudo apt-get install git
10  sudo -v
11  sudo -v
12  apt-get install git
13  git clone https://github.com/oobabooga/text-generation-w
ebui.git
14  git clone https://github.com/oobabooga/text-generation-w
ebui.git
15  gh repo clone oobabooga/text-generation-webui
16
17  git clone https://github.com/oobabooga/text-generation-webui.git
18  start_linux.sh
19  ls
20  cd
21  text-generation-webui
22  start_linux.sh
23  ls
24  ls
25  systemctl poweroff
26  ./start_linux.sh
27  apt get curl
28  apt-get install curl
29  ./start_linux.sh
30  share=True
31  launch/share=True
32  share=True
33  'share=True'
34  python app.py --share=True
35  ~/launch_app --share=True
36  ~/launch_app --share=True
37  ./launch_app --share=True
```

#### Historique des commandes linux

Avant de lancer l'installation, Suite à une erreur j'ai installé curl, un prérequis essentiel. Après cette vérification, j'ai procédé à l'installation en exécutant le script ./start\_linux.sh.

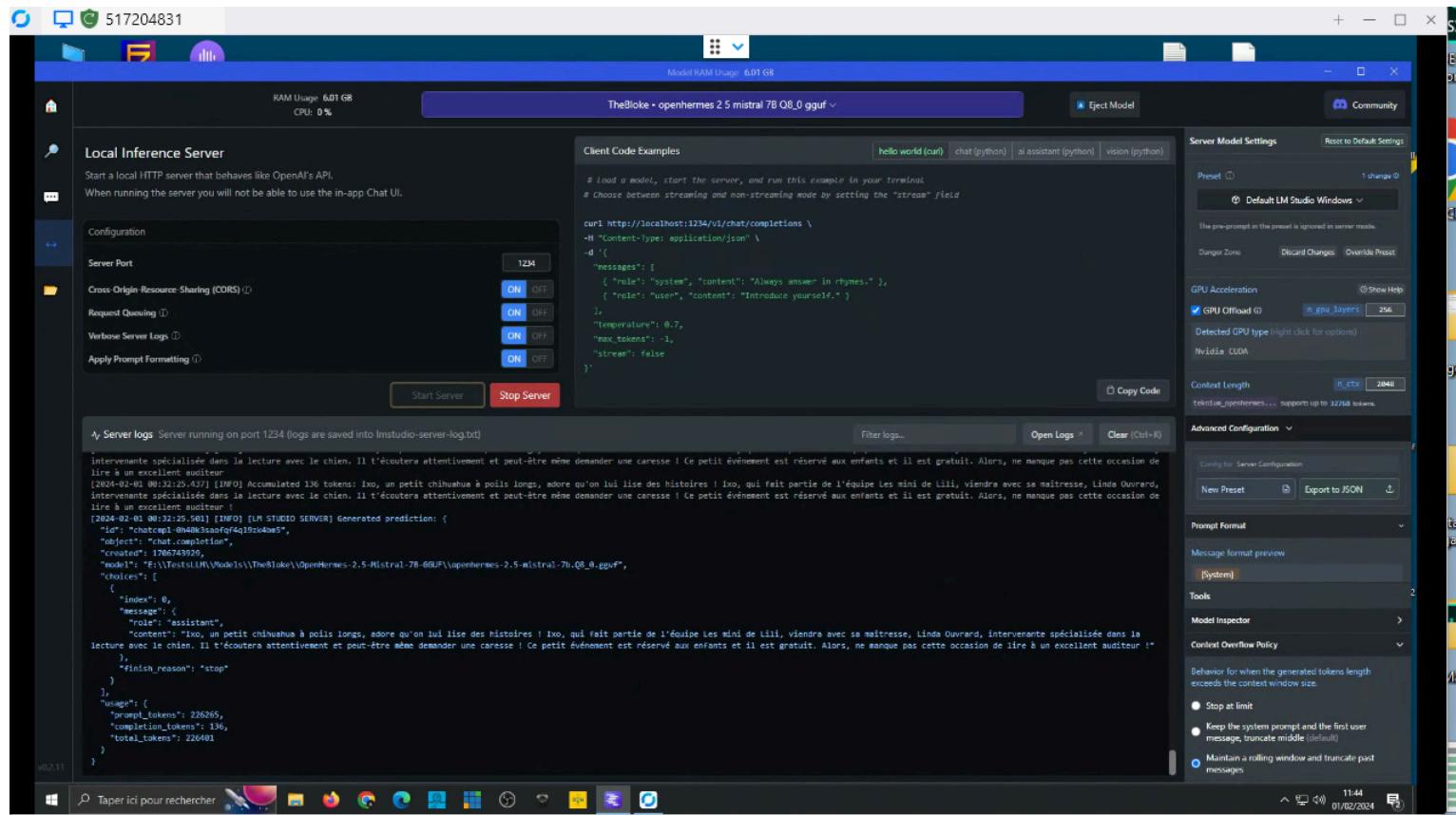
Une fois l'installation achevée, pour permettre l'accès à l'interface du LLM via le navigateur ou à l'API\* depuis un autre appareil que la vm\*, j'ai modifié les CMD\_FLAGS en décommentant les options --listen et --api.



ma Text-Generation-WebUI\*

## Installation de LM Studio\* sur une machine multimédia.

En attendant que l'équipe décide de l'investissement dans une nouvelle carte graphique, j'ai procédé à l'installation de LM Studio\* sur la machine multimédia équipée d'une carte graphique. Cependant, il est important de souligner que cette solution n'est pas viable sur le long terme. Il sera nécessaire de mettre en place la WebUI pour une meilleure intégration.



interface api de LM Studio\*

## Mise en place d'archivage.

J'ai mis en place un système d'archivage qui permet de supprimer les données de manière à éviter qu'elles ne réapparaissent à chaque synchronisation avec les flux RSS\* ou les API\*. Ce système fonctionne sur le principe du "soft delete". Pour cela, j'ai ajouté une colonne dans ma base de données avec des valeurs booléennes (true ou false) pour indiquer le statut d'archivage des données. J'ai également configuré l'option d'archivage dans Directus\*, permettant une gestion efficace des données sans les supprimer définitivement.

The screenshot shows the Directus admin interface. On the left, there's a sidebar with a tree view of datasets: 'Modele Agenda' is expanded, showing 'Angers calend...', 'Angers carte', 'Angers liste', 'Cholet calend...', 'Cholet carte', 'Cholet liste', 'La Roche-sur...', 'La Roche-sur...', 'La Roche-sur...', 'Nantes calend...', 'Nantes carte', 'Nantes liste', 'St Nazaire calend...', 'St Nazaire carte', 'St Nazaire liste', and 'Mots Clefs'. The 'Mots Clefs' item is selected. The main content area has a title 'Modele Agenda' with a back icon. It displays a table with 135 elements. The columns are 'Description' (with a dropdown arrow), 'Link', 'Image', and 'First Date'. Below the table is a search bar and a red '+' button. To the right of the table is a sidebar titled 'Information' with sections for 'Options d'affichage' and 'Archiver'. Under 'Archiver', there are three radio buttons: 'Afficher les éléments actifs' (selected, red circle), 'Afficher les éléments archivés' (blue circle), and 'Afficher tous les éléments' (blue circle). There are also icons for 'Ajouter' (add) and 'Supprimer' (delete).

mon système d'archivage sur directus

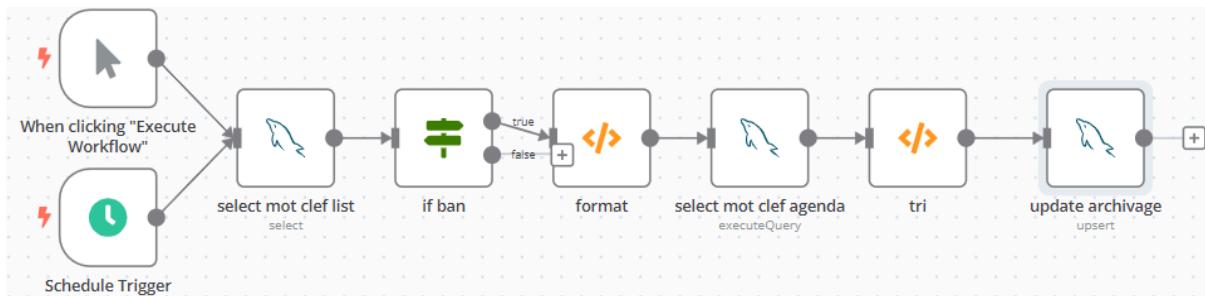
## Mot clefs editable

J'ai créé une nouvelle table intitulée "mot clef" dans Directus\*, avec une colonne "mots" pour faciliter l'ajout et l'édition des mots-clés directement depuis Directus\*.

The screenshot shows the 'Mots Clefs' dataset in Directus. The sidebar on the left shows 'Modele Agenda' expanded, with 'Mots Clefs' selected. The main content area has a title 'Mots Clefs' with a back icon. It displays a table with 16 rows. Each row has a checkbox in the first column and a tag name in the second column. The third column indicates the status: a grey 'X' for most rows, a red checkmark for 'Cirque', and a blue checkmark for 'Banni'. The rows are: Mots (Banni), Atelier, Brocante, Chant, Cinéma, Cirque (red checkmark), Classique, Commémoration, Concert, Conférence, Culture, Danse, and DJ Set.

table mots clefs sur directus

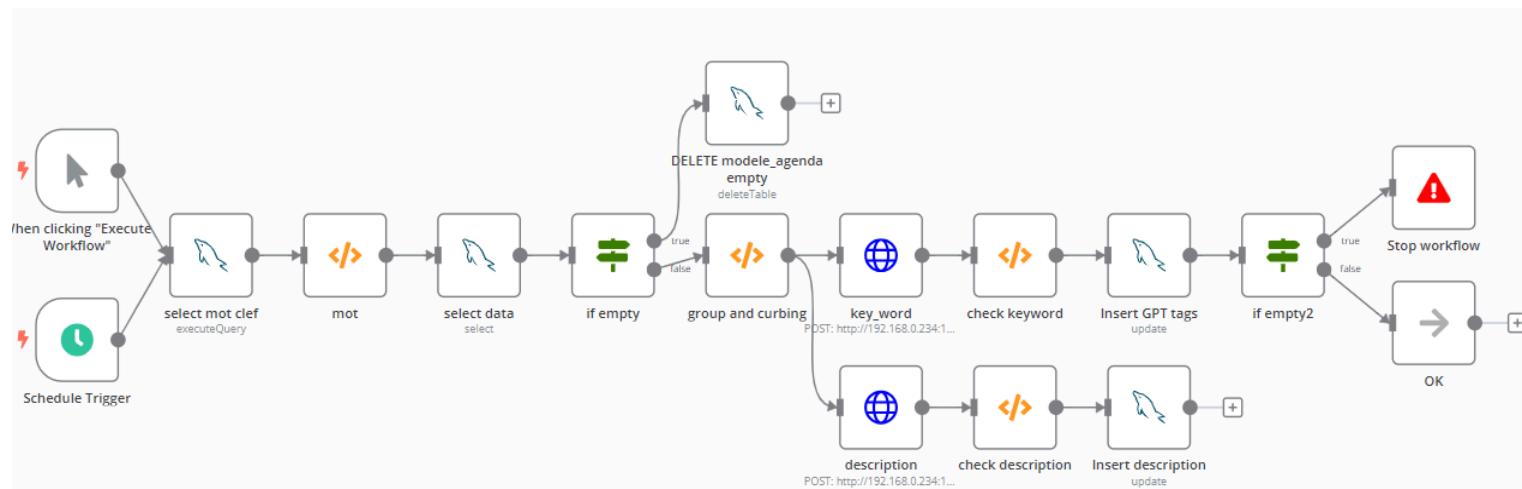
Dans cette nouvelle table, j'ai également mis en place un bouton poussoir doté d'une valeur binaire. Lorsqu'il est activé, ce bouton permet d'archiver automatiquement tous les articles associés à ce mot-clé spécifique dans la colonne "mots clefs GPT" de la base de données.



mon workflow de tri des agenda.

## Système en local.

La workspace en local est un peu différent car il n'y a pas les module gpt tout fait sur N8N\* et que il y a plus de restriction de requête ont ne paye plus les token. Il faut savoir que j'ai utilisé la version la moins quantifiée de mistral 7B. Ces techniques visent à diminuer la précision numérique des nombres flottants utilisés comme paramètres des modèles, qui sont des réseaux de neurones très profonds. La version la plus performante de LLM en local c'est Mixtral mais elle est vraiment trop gourmande en ressources.



La workspace qui génère les mots clefs et les descriptions avec mistral

Ce workspace est configuré pour récupérer la liste des mots-clés depuis Directus\*. J'organise ces données pour les réutiliser ultérieurement, en les formatant sous forme de chaînes de caractères et dans un tableau. Je sélectionne ensuite les agendas dont les

colonnes "description\_GPT" ou "mot\_clefGPT" sont vides afin de les compléter. Si la description d'un agenda est vide, je le supprime de la base de données. Pour constituer mon prompt, je combine la date de début avec ma description, qui est limitée à 1700 caractères. Cette limitation est nécessaire car un texte trop long empêcherait le bon fonctionnement de Mistral. Enfin, je réalise deux requêtes à mon API\* Mistral en local.

## Prompt 1:

```
{  
  "messages": [  
    {  
      "role": "system",  
      "content": "Sélectionne des mots-clés pertinents pour chaque événement à partir de cette liste : {{ $('format').item.json['mots_clefs_uniques"] }}. Choisis uniquement les mots-clés qui correspondent exactement à l'événement. Ne sélectionne pas de mots hors de la liste. Pour chaque événement traité, retourne un array de 5 mots-clés maximum, par exemple : [\"Atelier\", \"Brocante\", \"Chant\", \"Cinéma\", \"Cirque\"]"  
    },  
    {  
      "role": "user",  
      "content": "{{ JSON.stringify($json.descriptionJson) }}"  
    }  
  "model_name": "openhermes-2.5-mistral-7b.Q8_0.gguf",  
  "validation_rules": {  
    "check_json_format": true,  
    "max_keywords": 5,  
    "only_predefined_keywords": true  
  }  
}
```

prompt mistral qui génère les mots clefs

J'ai dû retravailler la requête pour extraire mes mots-clés de manière efficace. En effet, contrairement aux nœuds "output parser" dans ma version sur chat GPTI, le format de la réponse n'est pas toujours stable ni pertinent, ce qui nécessite de relancer l'espace de travail. Par la suite, j'ai créé un nœud de code qui filtre les mots ne figurant pas dans la liste, puis remplit la colonne "mots-clés" de l'agenda. Si cette colonne est vide, je stoppe le processus, car insérer du vide déclencherait inutilement une nouvelle exécution, entraînant une perte de temps et de ressource.

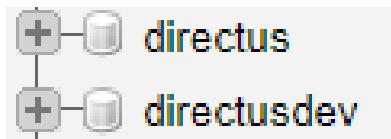
## Prompt 2:

```
{  
  "messages": [  
    {  
      "role": "system",  
      "content": "Sélectionne des mots-clés pertinents pour chaque événement à partir de cette liste : {{ $('format').item.json['mots_clefs_uniques"] }}. Choisis uniquement les mots-clés qui correspondent exactement à l'événement. Ne sélectionne pas de mots hors de la liste. Pour chaque événement traité, retourne un array de 5 mots-clés maximum, par exemple : [\"Atelier\", \"Brocante\", \"Chant\", \"Cinéma\", \"Cirque\"]"  
    },  
    {  
      "role": "user",  
      "content": "{{ JSON.stringify($json.descriptionJson) }}"  
    }  
  "model_name": "openhermes-2.5-mistral-7b.Q8_0.gguf",  
  "validation_rules": {  
    "check_json_format": true,  
    "max_keywords": 5,  
    "only_predefined_keywords": true  
  }  
}
```

prompt mistral qui génère la description

Si les mots-clés sont adéquats, je procède alors à la deuxième requête pour générer la description GPT, que j'insère ensuite dans la base de données.

## Environnement de production et d'un environnement de développement.



J'ai créé deux environnements de travail distincts afin de permettre à mes collègues de tester une version bêta de mon outil, tout en me donnant la possibilité de continuer à travailler sur le projet. Pour cela, j'ai dupliqué les espaces de travail, bases de données, ainsi que mon Directus\* sur Docker. J'ai également modifié les url, les titres et même la couleur de Directus\* pour bien différencier les deux environnements.

## Ajout de flux RSS\* et d'API\*

De la même manière que précédemment expliqué, j'ai ajouté des sites via l'API\* et le flux RSS\* afin de fournir suffisamment de données pour que les personnes en charge des agendas disposent d'assez de données pour nous faire part de leurs retours. L'intérêt de N8N\* réside dans sa capacité à fournir rapidement des retours, permettant d'évaluer les difficultés et la pertinence de notre outil. Cependant, à long terme, il n'est pas viable de rester sur N8N\* ; l'équipe technique envisage de se tourner vers Node.js\* pour le développement futur.

## Prise en compte des retours

Suite aux retours, j'ai réorganisé les données en les répartissant dans plusieurs tables par ville au sein de la base de données. Cette démarche vise à séparer et clarifier les informations. Cette organisation facilite la gestion des données puisque, actuellement, les agendas de chaque ville ne sont pas gérés par la même personne.

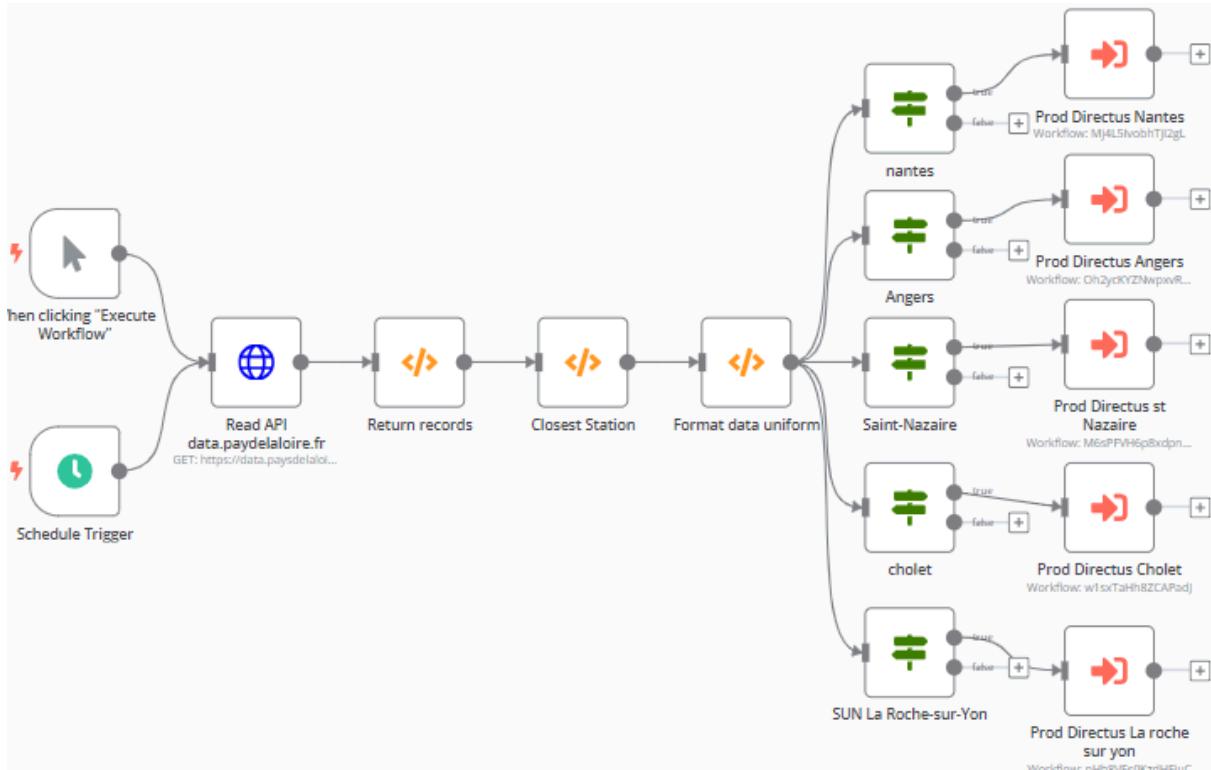
The screenshot shows the Directus application interface. On the left, there is a sidebar with a tree view of workspace structures. The main area displays a table of events with columns for Date, Titre, Mots Clefs, Source, and Image. A context menu is open on the right side of the table, listing options like 'Options d'affichage', 'Archiver', 'Actualisation automatique', 'Importer/Exporter', and 'Journal d'activité'.

Date	Titre	Mots Clefs	Source	Image
21 avril 2024 3:00:00 PM	Tom na Fazenda ...	Théâtre, Spectacle	JSON data.pays...	
22 mars 2024 8:00:00 PM	Plagier c'est copi...	Théâtre, Spectacle	JSON data.pays...	
15 mai 2024 6:00:00 PM	Bros - de Romeo ...	Théâtre, Spectacle	JSON data.pays...	
23 mars 2024 8:00:00 PM	Plagier c'est copi...	Théâtre, Spectacle	JSON data.pays...	
7 mars 2024 8:00:00 PM	Le mardi à Mono...	Théâtre, Réunions, Emergent	JSON data.pays...	
8 mars 2024 8:00:00 PM	Le mardi à Mono...	Théâtre, Réunions, Emergent	JSON data.pays...	
1 mars 2024 8:00:00 PM	Les belles-sœurs	Théâtre, Réunions	JSON data.pays...	
22 février 2024 7:30:00 PM	Comme un roma...	Théâtre, Musique, Histoire, Résidence, Litt...	JSON data.pays...	
22 février 2024 1:30:00 PM	Comme un roma...	Théâtre, Musique, Histoire, Résidence, Litt...	JSON data.pays...	
3 avril 2024 6:30:00 PM	Monte-Cristo - C...	Théâtre, Musique, Famille	JSON data.pays...	
5 avril 2024 6:30:00 PM	Monte-Cristo - C...	Théâtre, Musique, Famille	JSON data.pays...	
22 février 2024 6:00:00 PM	Sortie de Fabriqu...	Théâtre, Mémoire, Histoire, Littérature, Sp...	JSON data.pays...	
22 février 2024 6:00:00 PM	Sortie de Fabriqu...	Théâtre, Mémoire, Histoire, Littérature, Sp...	JSON data.pays...	

Directus\* avec les tables par ville et marque page

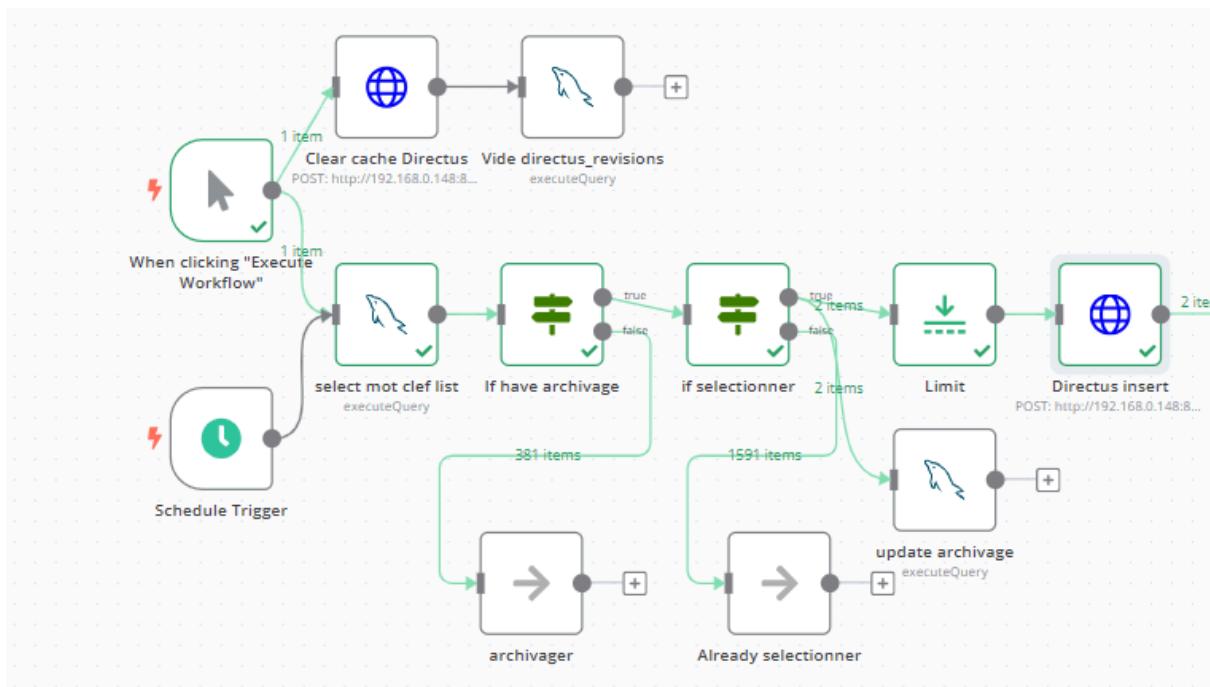
J'ai dupliqué tous les workspace directus, et j'ai modifié mes requêtes SQL afin de parcourir toutes les tables et rajouter du code pour aiguiller vers la bonne workspace .

<b>Prod Directus st Nazaire</b> Last updated 1 day ago   Created 26 January   directus api agenda	Inactive <input type="checkbox"/>	⋮
<b>Prod Directus Cholet</b> Last updated 1 day ago   Created 9 February   directus api agenda	Inactive <input type="checkbox"/>	⋮
<b>Prod Directus La roche sur yon</b> Last updated 1 day ago   Created 9 February   directus api agenda	Inactive <input type="checkbox"/>	⋮
<b>Prod Directus Angers</b> Last updated 1 day ago   Created 14 February   directus api agenda	Inactive <input type="checkbox"/>	⋮
<b>Prod Directus Nantes</b> Last updated 1 day ago   Created 14 February   directus api agenda	Inactive <input type="checkbox"/>	⋮



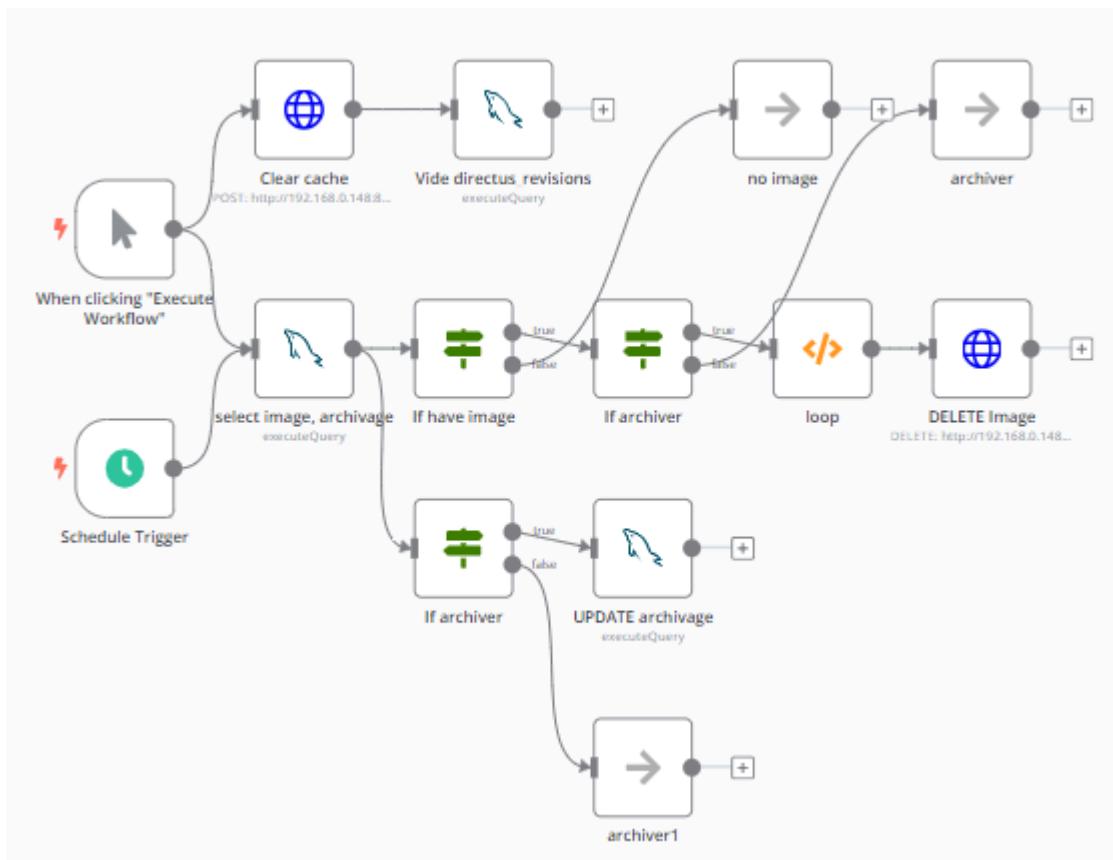
J'ai également revu la structure des marque-pages pour mettre en évidence les champs les plus pertinents : date, titre, mots-clés, source, image. La date étant un critère essentiel, j'ai exclu les données non datées et mis en place un système d'archivage pour les événements expirés.

Une table dédiée a été créée pour conserver les articles sélectionnés. Elle est dotée d'un champ spécifique permettant le transfert des données afin de les protéger contre la suppression automatique par l'archivage.



workspace qui transfère les données vers sélectionner

j'ai créé un workspace s'assurant que les données sont sélectionnées en vérifiant si le bouton a été en vérifiant la valeur binaire, duplique l'agenda vers la table choisie et archive l'article sélectionné dans sa propre table de ville.



workspace qui archive les évènement périmée

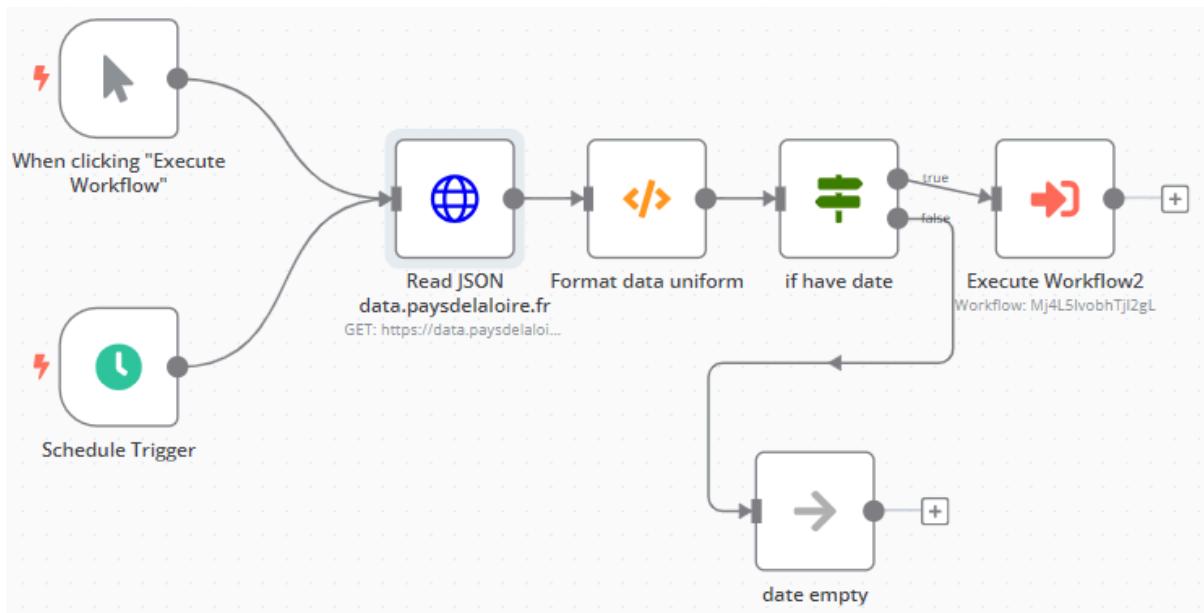
J'ai créé un workspace qui se charge d'archiver les données antérieures à la date du jour. Il vérifie également si l'agenda n'a pas déjà été archivé afin d'éviter la suppression des images des agendas sélectionnés.

## Bilan des technologies après six semaines de stage

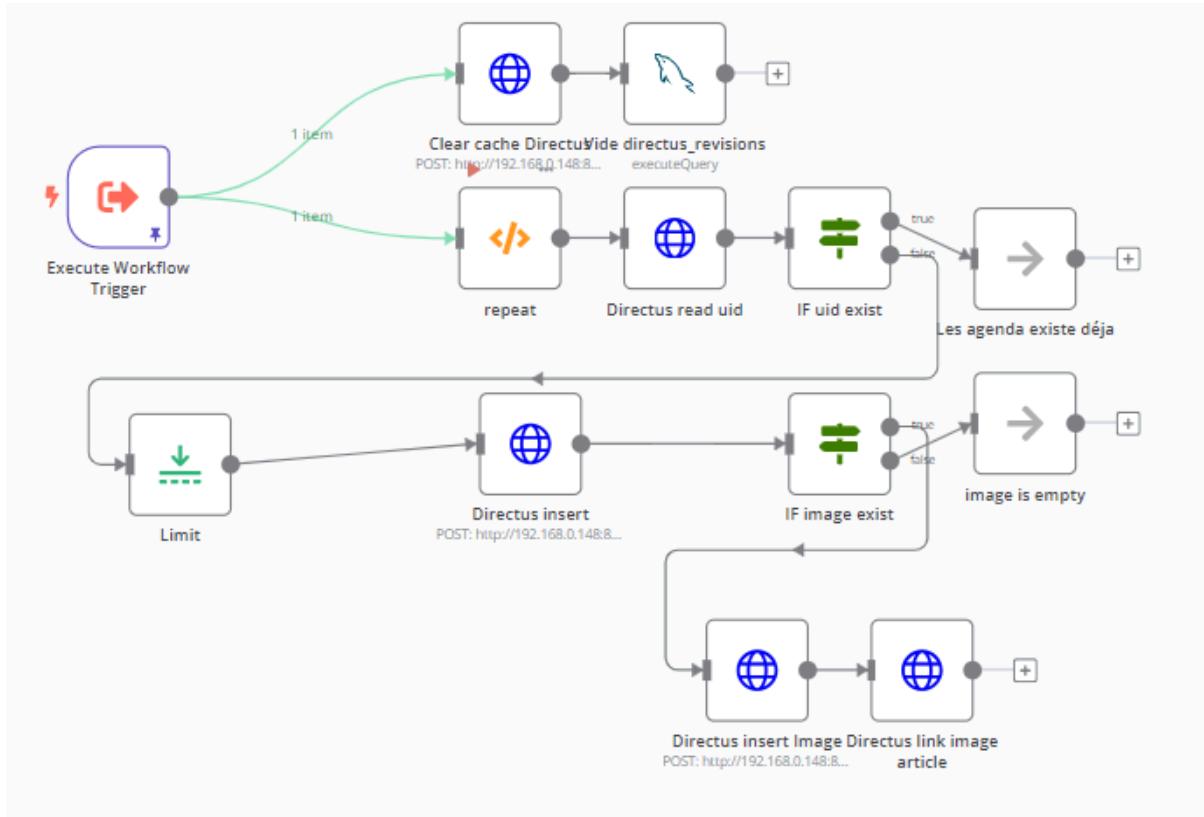
Après avoir utilisé n8n pendant plusieurs semaines, je considère cet outil comme intéressant pour réaliser des tâches simples ou pour tester des concepts. Néanmoins, à ce stade, mon travail s'apparente davantage à une première version du projet. J'ai complété la première version en utilisant n8n, car mon stage touchait à sa fin et je manquais de temps. L'utilisation de n8n s'est avérée chronophage. Opter pour la programmation, en Node.js\* ou Python par exemple, aurait certes nécessité plus de temps initialement, mais aurait rendu les modifications ultérieures significativement plus rapides, contrairement à l'ajustement des configurations sur n8n. En outre, la gestion des versions et le partage du code s'avèrent rapidement compliqués avec n8n, en raison de la multiplication des espaces de travail, ce qui peut conduire à une certaine confusion. On ne peut pas utiliser git pour le versionnage et la gestion d'environnement.

## Dernière modification

J'ai intégré une quantité importante de données issues d'un fichier JSON contenant un large éventail de données pertinentes (2000 événements) sur Nantes. Cette démarche m'a contraint à limiter le flux d'insertion à 250 événements à la fois afin d'éviter les plantages.



workspace Data [data.paysdeloire.fr/api/explore](https://data.paysdeloire.fr/api/explore)



workspace directus nantes dernière version

J'ai finalement transféré l'environnement de développement vers l'environnement de production en dupliquant la base de données de développement, puis en remplaçant celle de production, et en modifiant toutes les adresses des espaces de travail. J'ai également supprimé les espaces de travail en double, car la suite du projet ne se poursuivra pas sur N8N\*.

## Documentation et transmission des connaissances

J'ai élaboré une documentation qui décrit tous les systèmes et procédures que j'ai mis en place, lesquels sont toujours d'actualité. J'ai renommé les colonnes avec des noms plus appropriés. Cette documentation, jointe en annexe, présente la version finale de mon projet, tandis que ce rapport expose le code de manière chronologique, alors que toutes les parties du projet ont évolué de manière parallèle .

Lors d'une réunion avec Pierrick, bénévole et responsable chez Lonestone (<https://www.lonestone.io/>), nous avons évoqué la poursuite du projet. Grâce à son expertise, Pierrick a relevé des défis similaires à ceux auxquels j'ai été confronté, notamment en ce qui concerne scraping\* de sites web, une compétence qu'il a développée dans le passé pour des agences immobilières.

Nous avons échangé sur les choix techniques et étions, dans l'ensemble, en accord sur les décisions prises. Sauf concernant la duplication des bases de données. Cette approche, motivée par l'amélioration de l'expérience utilisateur et l'emploi de Directus\*, est reconnue comme étant loin d'être idéale en matière de gestion des bases de données. Cependant, elle est jugée acceptable dans notre contexte spécifique de première version fait pour avoir des retours expérimentation rapides avec les outils déjà mis en place.

Pierrick envisage de refondre la partie gérée par Directus\* en utilisant TypeScript et Node.js\*, et de remplacer Directus\* par un autre système de gestion de contenu (CMS), une solution low-code, ou une solution développée sur mesure. On a estimé que Directus\* et N8N\*, les outils recommandés par mon maître de stage, ne sont pas les plus adaptés à nos besoins. Il sera également nécessaire d'explorer des méthodes pour analyser les sites web qui ne disposent pas de flux RSS\* ou d'API\*. Enfin, SUN envisage d'utiliser l'intelligence artificielle pour générer la voix des capsules audio.

## Conclusion

Au cours de ce stage, de manière similaire à celui de ma première année, j'ai été affecté de façon autonome à une mission spécifique, bien que disposant d'une marge de manœuvre légèrement réduite du aux choix des technologies par mon maître de stage. J'aurais apprécié être impliqué dans des missions plus diversifiées et participer aux activités régulières de l'équipe technique. L'option de recourir au low code et l'utilisation de Directus\*, au fil du temps, se sont révélées être pesantes et chronophages. J'aurais préféré avoir la possibilité de programmer la solution moi-même. Bien que j'aie réalisé de nombreuses requêtes SQL, utilisé des API\* et écrit des fragments de code JavaScript du point de vue d'un étudiant, cela s'est avéré être assez limité et peu enrichissant.

Néanmoins, j'ai trouvé enrichissant d'approfondir mes connaissances sur l'utilisation des Modèles de Langage à Grande Échelle (LLM), tant avec l'API\* d'OpenAI et ChatGPT qu'avec des API\* installer avec des outils en local, notamment Mistral.

J'ai pu comprendre les enjeux liés à la création d'un outil métier, notamment l'importance de vérifier rapidement la faisabilité du projet, de créer une première version dans de brefs délais pour recueillir des retours, de se concentrer avant tout sur l'expérience utilisateur (car il est inutile d'avoir un outil fonctionnel mais inutilisé), et de choisir minutieusement les technologies les plus adaptées en prenant en compte de nombreux facteurs.

# Bibliographie

1. Wikipédia SUN:

<https://fr.wikipedia.org/wiki/SUN - Le Son Unique>

2. A propos SUN:

<https://lesonunique.com/page/a-propo>

3. Github installer LLM en local:

[https://github.com/oobabooga/Text-Generation-WebUI\\*](https://github.com/oobabooga/Text-Generation-WebUI*)

4. Documentation Directus\*:

<https://docs.directus.io/>

Annexes:

# Documentation automatisation du processus de création d'agenda.

1. Objectif de l'outil.	2
2. Les technologies utilisées	3
3. Directus	4
4. LLM en local / Mistral AI	8
5. N8N récolte de données	10
6. N8N outils	12
7. Prod / Dev	15
8. Adresse et Mot de Passe	15

## 1. Objectif de l'outil.

## Qu'est-ce qu'un agenda ?

Les agendas des événements culturels locaux sont diffusés sur SUN, sous forme de flashes informatifs et en tant que brefs articles sur le site [lesonunique.com/mysun/agenda](http://lesonunique.com/mysun/agenda). Ces agendas visent à promouvoir les initiatives culturelles locales dans les zones de diffusion de la radio, notamment Nantes, Cholet, Saint-Nazaire, Angers et La Roche-sur-Yon. Ils offrent des informations sur les artistes, les lieux, les tarifs, et les modalités d'accès aux événements. Les événements sont souvent accompagnés d'une vidéo sur le site web.

## Les enjeux.

Cette tâche répétitive a conduit l'équipe à envisager son automatisation, afin de réduire le temps consacré à cette activité. Leur travail consistait à rechercher des événements intéressants sur plusieurs sites web, à sélectionner une image, rédiger un texte avec les informations mentionnées plus haut, enregistrer les flashes radio, adapter le texte pour le site web, recadrer l'image pour le site via Photoshop, et trouver une vidéo si nécessaire.

Le but de ma mission était de centraliser toutes les données des sites habituellement utilisés sur directus, afin de rassembler toutes les informations au même endroit. Grâce à cela, nous pouvions trier les informations de multiples sites sur une seule plateforme, par titre d'article, date, source ou localisation. Il était également possible de visualiser ces informations via un agenda et une carte interactifs. Nous pouvions éditer le texte, le titre et recadrer les images.

L'intelligence artificielle va être utilisée pour trier les articles grâce à une liste de mots-clés générée automatiquement, et pour créer un brouillon de texte sur lequel l'opérateur pouvait s'appuyer.

À terme, l'objectif est que l'humain n'est plus qu'à sélectionner les événements les plus pertinents, vérifier et corriger les textes, titres et images recadré automatiquement, enregistrer le flash radio, puis appuyer sur bouton pour envoyer les articles, en brouillon sur le backend du site SUN [lesonunique.com](http://lesonunique.com). Cette approche représente un gain de temps considérable.

## 2. Les technologies utilisées

Dans le cadre de la mise en œuvre de la V1, nous avons employé n8n pour la récupération, le traitement et l'insertion des données, Directus comme CMS pour la visualisation, et Mistral AI pour la génération de descriptions.

Avant mon arrivée en stage, n8n et Directus étaient déjà en place. J'ai choisi de me tourner vers Mistral AI en local après avoir constaté que l'utilisation de ChatGPT s'avérait trop coûteuse en raison du prix des tokens.

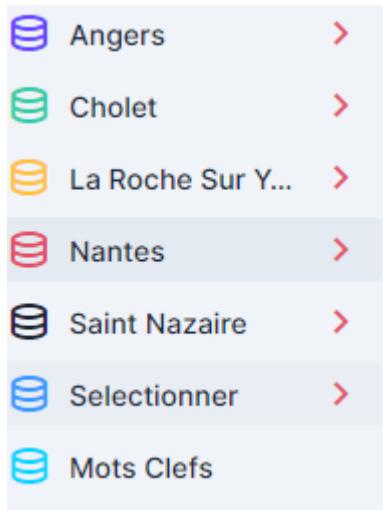
Concernant n8n, après plusieurs semaines d'utilisation, j'ai trouvé cet outil pratique pour exécuter des tâches simples ou pour tester des idées. Cependant, n8n montre rapidement ses limites. J'ai complété la première version en utilisant n8n, car mon stage touchait à sa fin et je manquais de temps .

L'utilisation de n8n s'est avérée chronophage. Opter pour la programmation, en Node.js ou Python par exemple, aurait certes nécessité plus de temps initialement, mais aurait rendu les modifications ultérieures significativement plus rapides, contrairement à l'ajustement des configurations sur n8n. En outre, la gestion des versions et le partage du code s'avèrent rapidement compliqués avec n8n, en raison de la multiplication des espaces de travail, ce qui peut conduire à une certaine confusion.

### 3. Directus

```
docker stop f3bf69414d52  
docker start f3bf69414d52  
docker ps -a
```

<http://192.168.0.148:8055/admin/login>



Sur Directus, j'ai configuré sept tables : une table dédiée à chacune des cinq stations de radio, une table pour l'édition des mots-clés, et une dernière table destinée au transfert des agendas sélectionnés

A screenshot of the Directus admin interface showing the 'Modification de l'élément dans Nantes' (Edit element in Nantes) screen. The left sidebar shows the same list of tables as the previous screenshot. The main area has tabs for 'Source' (containing 'JSON data.pays-de-la-loire'), 'Lien' (containing 'https://metropole.nantes.fr/infonantes/agenda/59807'), and 'Description' (containing rich text with a 'Visual aids' button). A note at the bottom of the description area says: 'Tous les jeudis, profitez de votre pause déjeuner pour suivre une visite, dessiner devant les œuvres, ou découvrir les trésors cachés conservés au cabinet d'arts graphiques du musée. Thème : Gare au gorille ! Le Gorille d'Emmanuel Frémiet a quitté le hall d'entrée pour rejoindre la Salle 12 au premier étage. Retrouvons-le dans son nouvel environnement.'

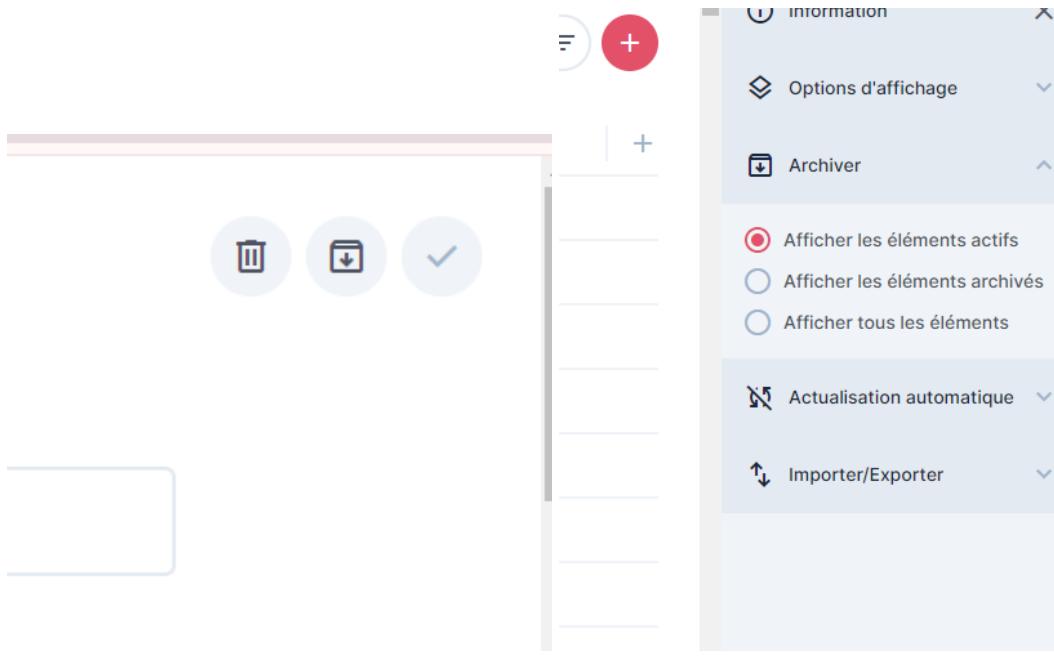
#	Nom	Type	Interclassement	Attributs	Null	Valeur par défaut	Commentaires	Extra	Action
1	<b>id</b> 🔗	int		UNSIGNED	Non	Aucun(e)		AUTO_INCREMENT	
2	<b>titre</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
3	<b>lien</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
4	<b>description</b>	text	utf8mb4_0900_ai_ci		Oui	NULL			
5	<b>description_IA</b>	text	utf8mb4_0900_ai_ci		Oui	NULL			
6	<b>mots_clefs_IA</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
7	<b>source</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
8	<b>adresse</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
9	<b>coordonnee_geographique</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
10	<b>station</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
11	<b>uid</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	NULL			
12	<b>archivage</b>	varchar(255)	utf8mb4_0900_ai_ci		Oui	FALSE			
13	<b>date</b>	datetime			Oui	NULL			
14	<b>image</b> 📸	char(36)	utf8mb4_0900_ai_ci		Oui	NULL			
15	<b>Selectionner</b>	tinyint(1)			Oui	0			

## Les tables Nantes, Angers, Cholet, La Roche Sur Yon, Saint Nazaire et Sélectionner.

Les 5 tables des villes sont constituées des mêmes colonnes.

La table "selectionner" est constituée des mêmes données sans la colonne "selectionner".

- Station : Il s'agit de la station de radio la plus proche de l'événement.
- Description : Cette donnée représente la description officielle du site de l'événement.
- Description\_IA : Un brouillon généré par Mistral pour chaque événement.
- Date : La date de début de l'événement.
- Image : L'image originale associée à l'article de l'événement.
- Lien: Le lien vers l'article original de l'événement.
- adresse: L'adresse postale où se déroule l'événement.
- Mots\_Clefs\_IA : Les mots-clés générés par Mistral pour classifier l'événement.
- Position : Les coordonnées géographiques de l'événement.
- Source : Le site source de l'événement.
- Titre : Le titre de l'événement.
- Archivage : Cette fonctionnalité utilise une valeur booléenne pour indiquer si un article est archivé ou non. (Action N8N)
- Sélection : Cette fonctionnalité repose sur une valeur booléenne associée à une case à cocher. Lorsqu'elle est activée, elle permet de transférer l'article vers la table de sélection.



L'archivage a été implémenté au moyen de ces boutons .

L'archivage joue un rôle clé dans ce projet. **En effet, supprimer directement les données entraînerait leur réapparition systématique tant qu'elles figurent dans le flux RSS ou l'API du site.**

la table de mot clef



## Contenu Mots Clefs

1-25 de 45 éléments



<input type="checkbox"/>	Mots	Banni	<a href="#">+</a>
<input type="checkbox"/>	Atelier	<a href="#">X</a>	
<input type="checkbox"/>	Brocante	<a href="#">X</a>	
<input type="checkbox"/>	Chant	<a href="#">X</a>	
<input type="checkbox"/>	Cinéma	<a href="#">X</a>	
<input type="checkbox"/>	Cirque	<a href="#">X</a>	
<input type="checkbox"/>	Classique	<a href="#">X</a>	
<input type="checkbox"/>	Commémoration	<a href="#">X</a>	
<input type="checkbox"/>	Concert	<a href="#">X</a>	
<input type="checkbox"/>	Conférence	<a href="#">X</a>	
<input type="checkbox"/>	Culture	<a href="#">X</a>	
<input type="checkbox"/>	Danse	<a href="#">X</a>	
<input type="checkbox"/>	DJ Set	<a href="#">X</a>	

#	Nom	Type	Interclassement	Attributs	Null	Valeur par défaut	Commentaires	Extra	Action
<input type="checkbox"/>	1	id	int	UNSIGNED	Non	Aucun(e)		AUTO_INCREMENT	<a href="#">Modifier</a> <a href="#">Supprimer</a> <a href="#">Plus</a>
<input type="checkbox"/>	2	mots	varchar(255)	utf8mb4_0900_ai_ci	Oui	NULL			<a href="#">Modifier</a> <a href="#">Supprimer</a> <a href="#">Plus</a>
<input type="checkbox"/>	3	banni	tinyint(1)		Oui	0			<a href="#">Modifier</a> <a href="#">Supprimer</a> <a href="#">Plus</a>



Mots Clefs

### Modification de l'élément dans Mots Clefs



#### Mots

 Atelier

#### Banni

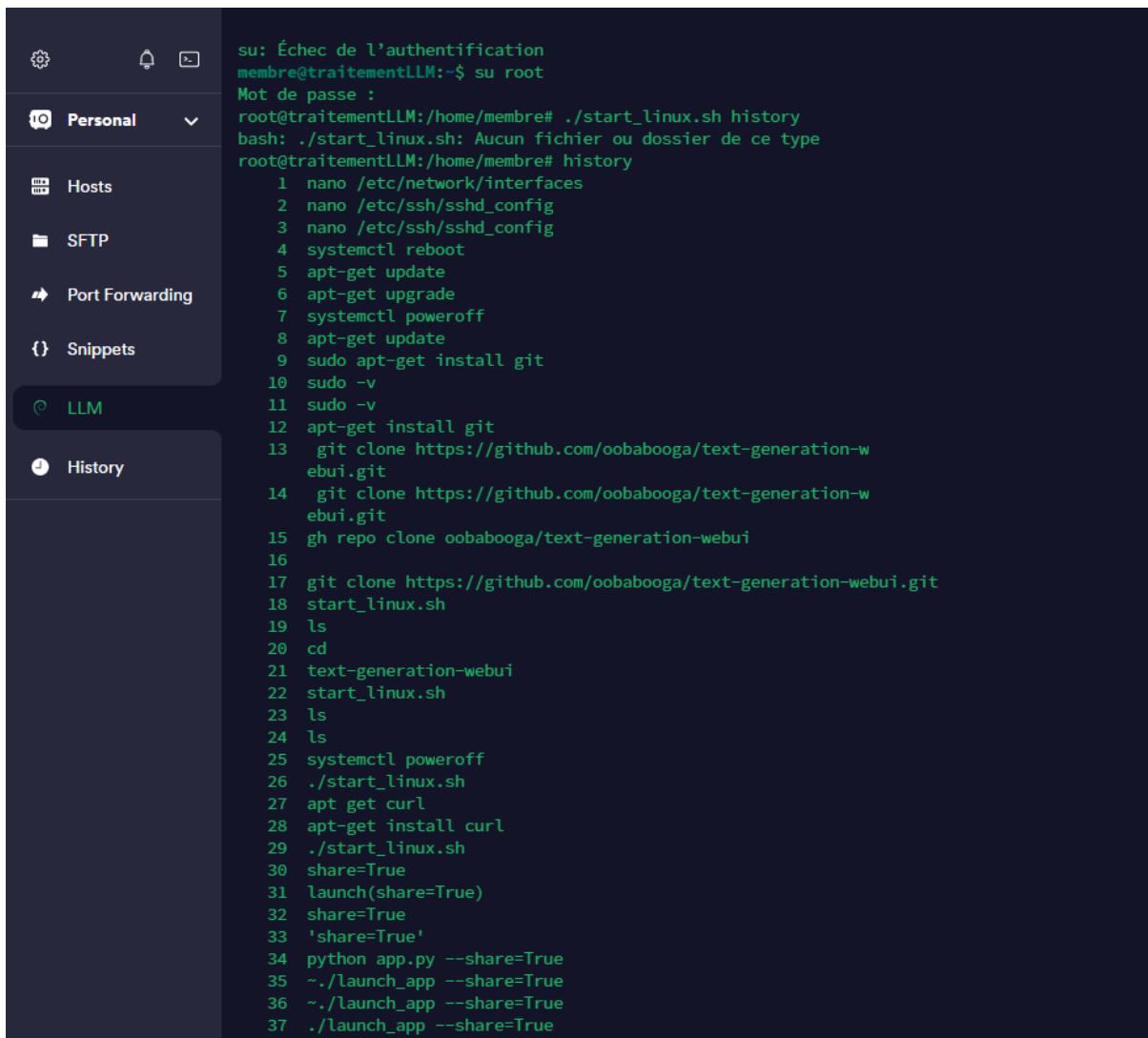
 Activé

La table des mots-clés offre la possibilité d'éditer la liste des mots-clés fournie à Mistral. J'ai également intégré une fonctionnalité permettant d'exclure les articles contenant des mots-clés bannis. (si bannis alors mis en archivage, attention si on réactive un mot, les anciens articles bannis restes archivés)

## 4. LLM en local / Mistral AI

### Installation text-generation-webui sur une VM Linux Debian

J'ai procédé à l'installation d'un LLM (Large Language Model) sur une **machine virtuelle (VM)** fonctionnant sous Linux Debian. Cette installation a été effectuée en ligne de commande, j'ai installé Git. Pour ce faire, avec Git j'ai d'abord cloné le dépôt depuis <https://github.com/oobabooga/text-generation-webui>, un projet qui permet de lancer un LLM dans un navigateur et via une API.



```
su: Échec de l'authentification
membre@traitementLLM:~$ su root
Mot de passe :
root@traitementLLM:/home/membre# ./start_linux.sh history
bash: ./start_linux.sh: Aucun fichier ou dossier de ce type
root@traitementLLM:/home/membre# history
 1  nano /etc/network/interfaces
 2  nano /etc/ssh/sshd_config
 3  nano /etc/ssh/sshd_config
 4  systemctl reboot
 5  apt-get update
 6  apt-get upgrade
 7  systemctl poweroff
 8  apt-get update
 9  sudo apt-get install git
10  sudo -v
11  sudo -v
12  apt-get install git
13  git clone https://github.com/oobabooga/text-generation-w
ebui.git
14  git clone https://github.com/oobabooga/text-generation-w
ebui.git
15  gh repo clone oobabooga/text-generation-webui
16
17  git clone https://github.com/oobabooga/text-generation-webui.git
18  start_linux.sh
19  ls
20  cd
21  text-generation-webui
22  start_linux.sh
23  ls
24  ls
25  systemctl poweroff
26  ./start_linux.sh
27  apt get curl
28  apt-get install curl
29  ./start_linux.sh
30  share=True
31  launch/share=True)
32  share=True
33  'share=True'
34  python app.py --share=True
35  ~./launch_app --share=True
36  ~./launch_app --share=True
37  ./launch_app --share=True
```

<https://gitcode.net/uflexnihao/text-generation-webui>

```
/etc/systemd/system#  txtwebui.service

[Unit]
Description=text generation webui
After=network-online.target

[Service]
Type=simple

User=root
Group=root
UMask=007

ExecStart=/home/membre/text-generation-webui/start_linux.sh --model
openhermes-2.5-mistral-7b.Q8_0.gguf

Restart=on-failure

# Configures the time to wait before service is stopped forcefully.
TimeoutStopSec=300

[Install]
WantedBy=multi-user.target

-----
systemctl enable txtwebui.service

kill -9 $(pidof python)
```

Avant de lancer l'installation, Suite à une erreur j'ai installé **curl**, un prérequis essentiel.  
Après cette vérification, j'ai procédé à l'installation en exécutant le script ./start\_linux.sh.

Une fois l'installation achevée, pour permettre l'accès à l'interface du LLM via le navigateur ou à l'API depuis un autre appareil que la **VM**, j'ai modifié les CMD\_FLAGS en décommentant les options --listen et --api.

## AJOUT CARTE GRAPHIQUE

<https://asded.gitlab.io/post/2023-07-01-pci-passthrough-proxmox-04/>

<https://medium.com/bigdatarepublic/gpu-machine-learning-on-linux-fa972dd77cf2>

Sur le serveur proxmox modifier le fichier grub : **/etc/default/grub**

mettre en commentaire : **#GRUB\_CMDLINE\_LINUX\_DEFAULT="quiet"**

Ajouter la ligne suivante : **GRUB\_CMDLINE\_LINUX\_DEFAULT="quiet intel\_iommu=on iommu=pt"**

Faire l'update de grub : **update-grub**

Sur la VM

Ajouter dans le fichier : **/etc/apt/sources.list**

**deb http://deb.debian.org/debian/ bullseye main non-free contrib**

**apt install nvidia-detect**

**apt install nvidia-driver**

Commande : **nvidia-smi**

NVIDIA-SMI has failed because it couldn't communicate with the NVIDIA driver. Make sure that the latest NVIDIA driver is installed and running.

<https://gist.github.com/espoirMur/65cec3d67e0a96e270860c9c276ab9fa>

Commande : **whereis nvidia**

nvidia: /usr/lib/x86\_64-linux-gnu/nvidia  
/usr/lib/nvidia /etc/nvidia  
/usr/share/nvidia  
/usr/src/nvidia-current-470.223.02/nvidia

<https://forum-debian.fr/viewtopic.php?t=2342>

nvidia-detect

```
apt-cache show nvidia-driver  
apt-cache show primus  
  
apt update && apt full-upgrade
```

On désinstalle LM studio

```
cd text-generation-webui/
```

```
rm -rf installer_files/
```

```
wget  
https://fr.download.nvidia.com/XFree86/Linux-x86\_64/535.154.05/NVIDIA-Linux-x86\_64-535.154.05.run
```

```
chmod 777 NVIDIA-Linux-x86_64-535.154.05.run
```

```
apt remove nvidia-driver
```

```
apt autoremove
```

```
./NVIDIA-Linux-x86_64-535.154.05.run
```

```
nvcc --version
```

```
apt install nvtop
```

```
https://github.com/Syllo/nvtop
```

nvtop équivalent htop pour carte graphique

```
https://forum-debian.fr/viewtopic.php?t=2342&start=100
```

nvidia-installer was forced to guess the X library path '/usr/lib' and X module path '/usr/lib/xorg/modules'; these paths were not queryable from the system. If X fails to find the NVIDIA X

driver module, please install the `pkg-config` utility and the X.Org SDK/development package for your distribution and reinstall the driver.

-----

WARNING: This NVIDIA driver package includes Vulkan components, but no Vulkan ICD loader was detected on this system. The NVIDIA Vulkan ICD will not function without the loader. Most distributions

package the Vulkan loader; try installing the "vulkan-loader", "vulkan-icd-loader", or "libvulkan1" package.

-----

```
/home/membre/text-generation-webui/installer_files/env/lib/python3.11/site-packages/torch  
/cuda/__init__.py:611: UserWarning: Can't initialize NVML
```

```
warnings.warn("Can't initialize NVML")
```

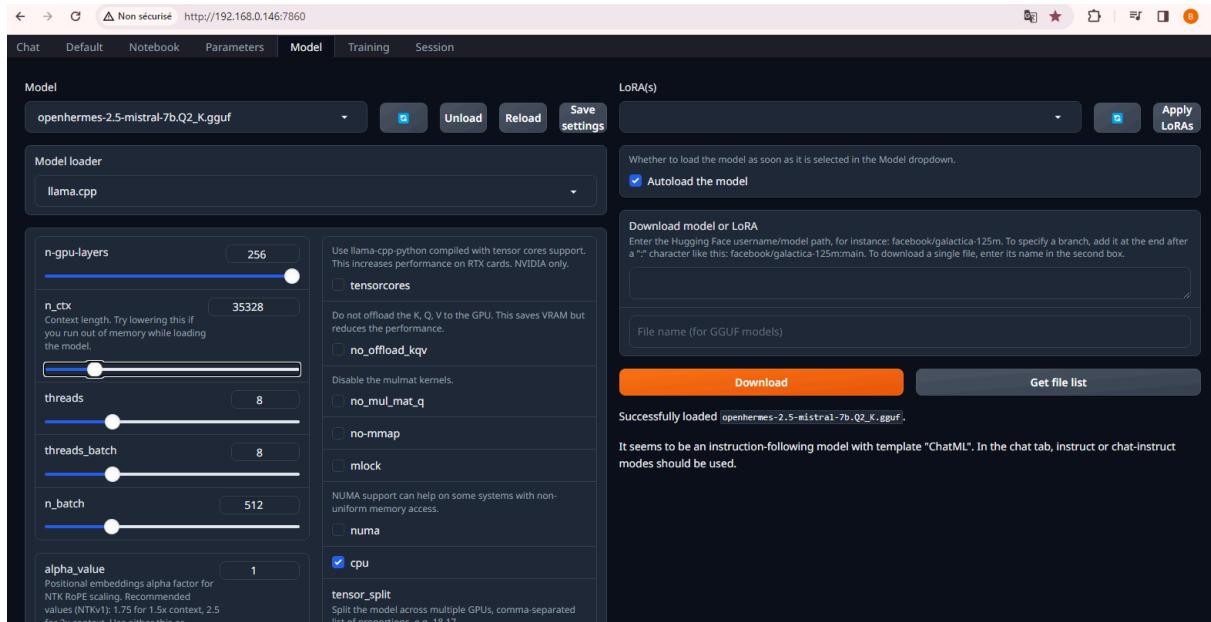
-----

WARNING: failed to allocate 4096.00 MB of pinned memory: no CUDA-capable device is detected

<https://docs.nvidia.com/cuda/cuda-installation-guide-linux/index.html#prepare-debian>

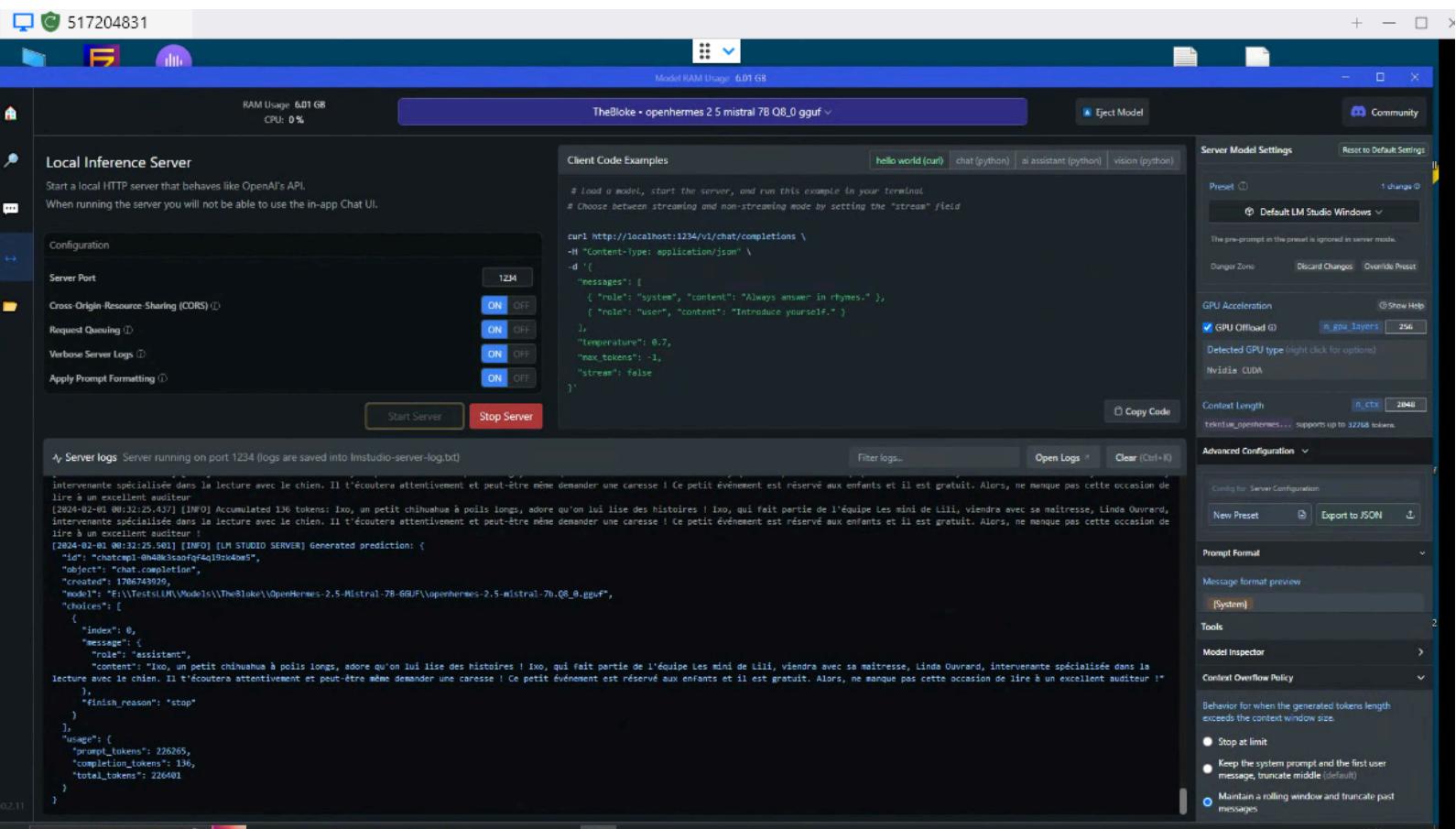
```
apt-get install linux-headers-$(uname -r)  
add-apt-repository contrib  
lsb_release -a  
wget  
https://developer.download.nvidia.com/compute/cuda/repos/debian11/x86_64/cuda-keyring_  
1.1-1_all.deb  
dpkg -i cuda-keyring_1.1-1_all.deb  
apt-get install ldmconfig  
su -  
dpkg -i cuda-keyring_1.1-1_all.deb  
cd /home/membre/
```

```
dpkg -i cuda-keyring_1.1-1_all.deb
apt-get update
apt-get -y install cuda
```



## Installation de LM studio sur une machine multimédia.

En attendant que l'équipe décide de l'investissement dans une nouvelle carte graphique, j'ai procédé à l'installation de LM Studio sur la machine multimédia équipée d'une carte graphique. Cependant, il est important de souligner que cette solution n'est pas viable sur le long terme. Il sera nécessaire de mettre en place la WebUI pour une meilleure intégration.



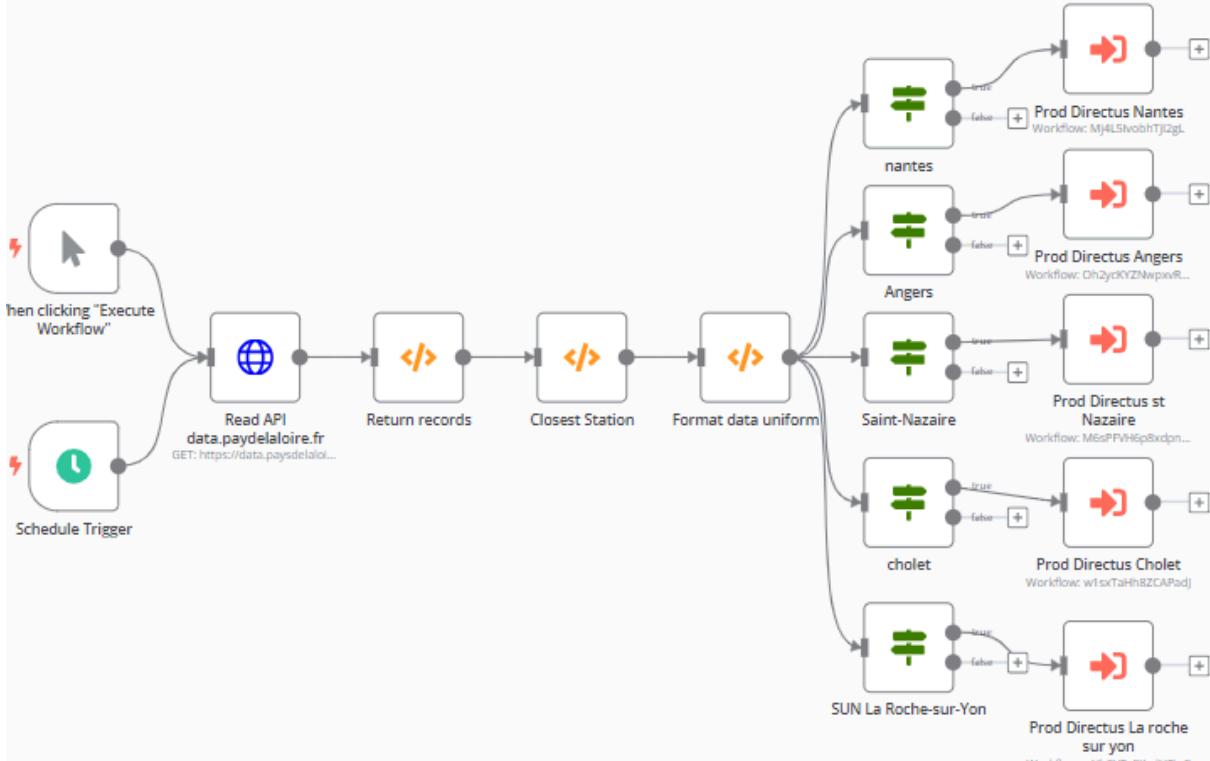
## 5. N8N récolte de données

Les espaces de travail dédiés à la collecte de données se divisent en deux catégories : les workspaces "Data", qui ont pour mission de récupérer des données via des flux RSS et des API et de les formater, et les workspaces "Directus", chargés d'insérer ces données uniformes dans Directus.

### Data

The screenshot shows three active workflows in the Data workspace:

- Prod Data www.saintnazaire.fr/liens-utiles/flux-rss**  
Last updated just now | Created 26 January directus api agenda
- Prod Data https://data.paysdelaloire.fr/api/explore**  
Last updated 4 minutes ago | Created 6 February directus api agenda
- Prod Data data.paysdelaloire.fr/api/records**  
Last updated 4 minutes ago | Created 26 January directus api agenda

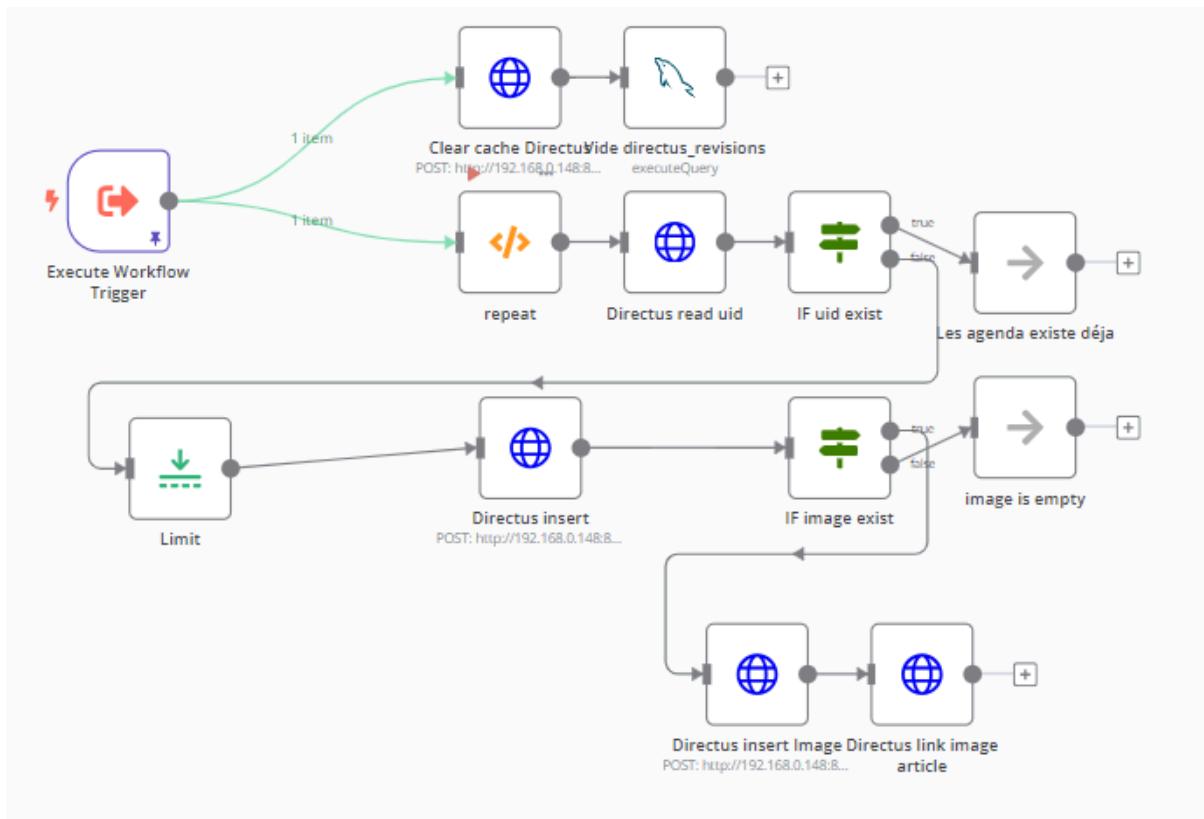


Comme mentionné précédemment, ces espaces sont constitués d'une requête qui récupère les données pour ensuite les uniformiser dans un même format. Pour le

site [data.paysdelaloire.fr/api/records](http://data.paysdelaloire.fr/api/records), en raison de la diversité des stations concernées, nous avons développé un système capable de rediriger les données vers la station appropriée en fonction de sa position géographique.

## Directus

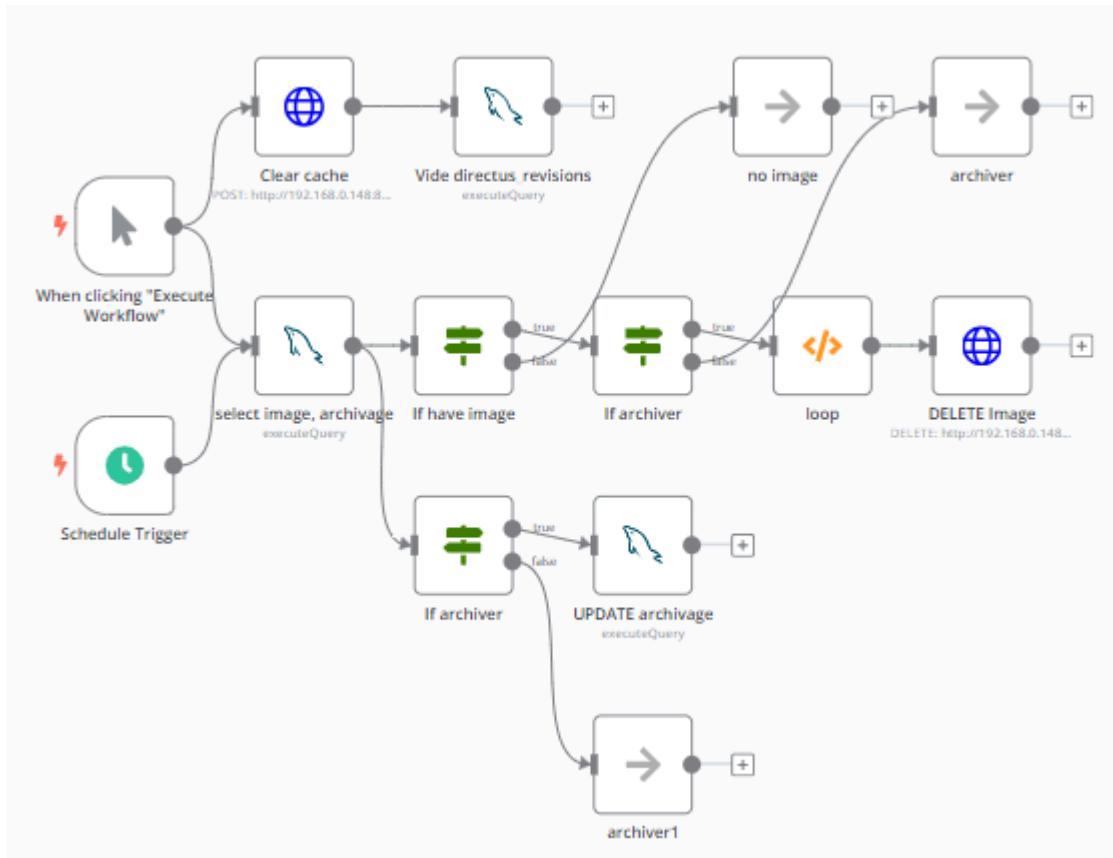
Prod Directus st Nazaire Last updated 1 day ago   Created 26 January   directus   api agenda	Inactive	⋮
Prod Directus Cholet Last updated 1 day ago   Created 9 February   directus   api agenda	Inactive	⋮
Prod Directus La roche sur yon Last updated 1 day ago   Created 9 February   directus   api agenda	Inactive	⋮
Prod Directus Angers Last updated 1 day ago   Created 14 February   directus   api agenda	Inactive	⋮
Prod Directus Nantes Last updated 1 day ago   Created 14 February   directus   api agenda	Inactive	⋮



Les données formatées sont transférées depuis l'espace de travail "Data" vers l'espace "Directus" correspondant à la ville ciblée par les agendas.

Chaque instance de Directus est configurée pour insérer les données dans la table spécifique à la ville concernée.

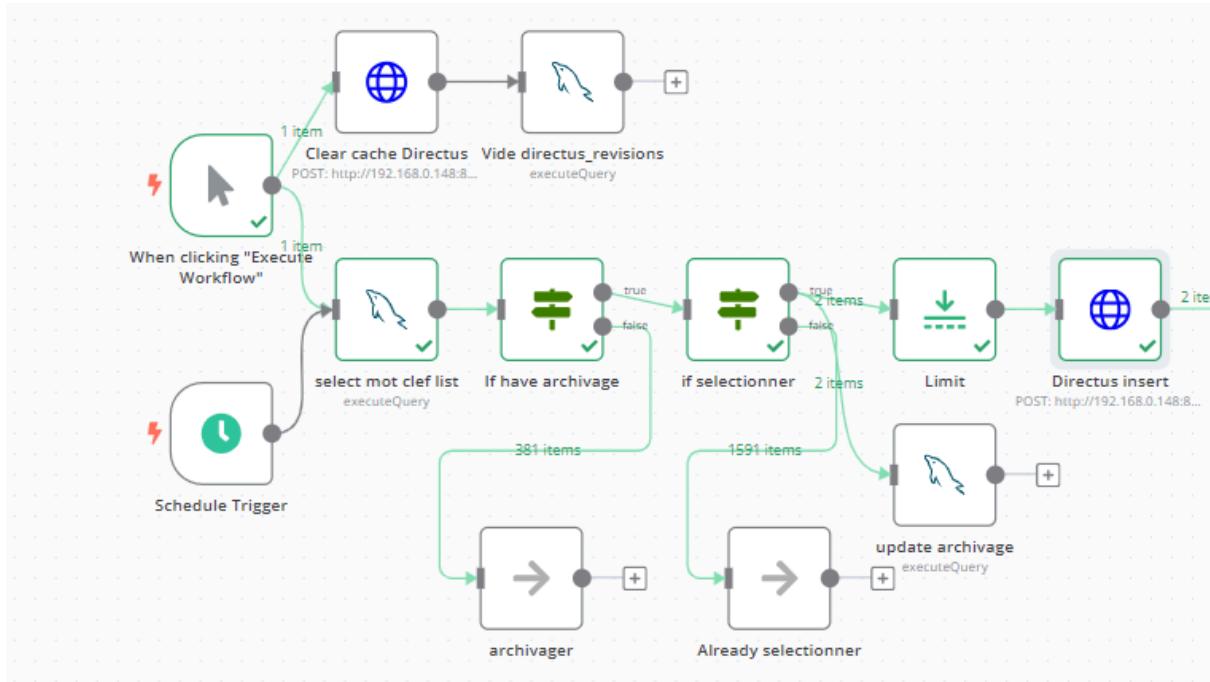
## 6. N8N outils



### Archivage d'agendas périmés

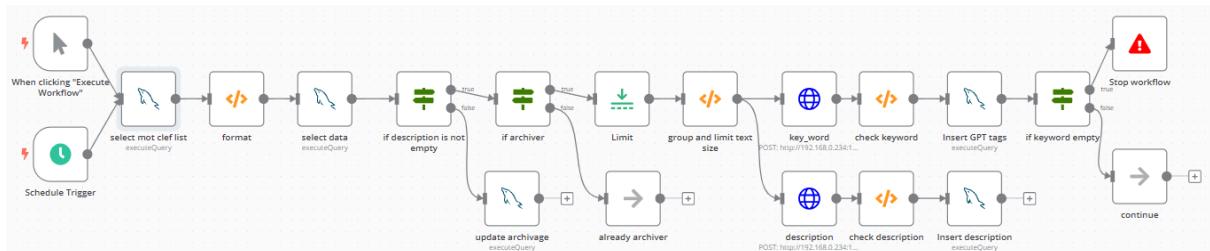
Ce workspace se charge d'archiver les données antérieures à la date du jour. **Il vérifie également si l'agenda n'a pas déjà été archivé afin d'éviter la suppression des images des agendas sélectionnés.**

Prod push to sélectionner



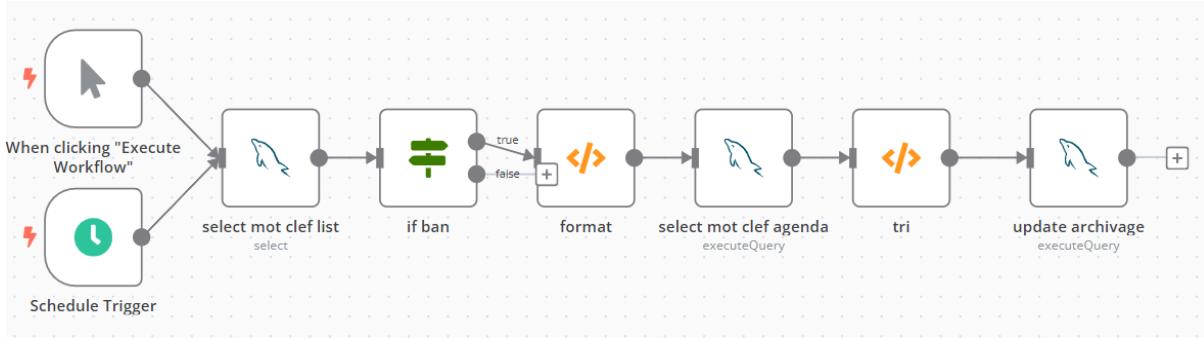
Ce workspace s'assure que les données sont sélectionnées, duplique l'agenda vers la table choisie et archive l'article sélectionné dans sa propre table de ville.

## Prod LM studio Tags from Description



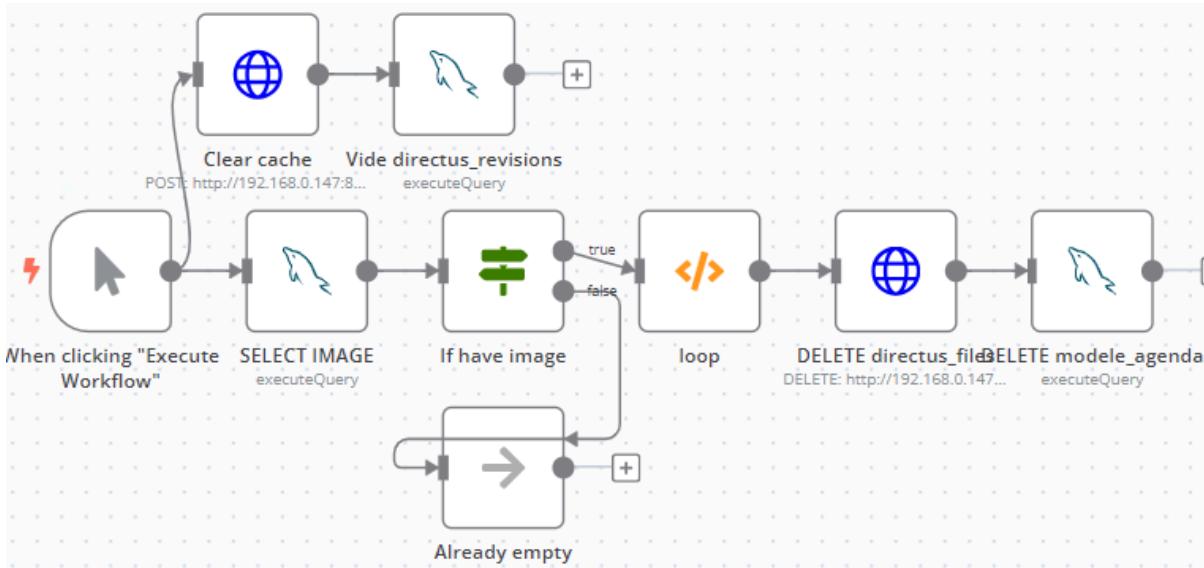
Ce workspace crée des mots-clés à partir d'une liste préétablie sur Directus et génère une description pour l'article choisi. L'API de génération de texte fonctionne de manière similaire à celle de LM Studio.

## Filtrage des agendas



Ce workspace archive les articles contenant des mots-clés banni afin de filtrer les sujets non pertinents.

## Dev Supprime les agenda via l'API



Ce workspace est dédié à la suppression propre des données. Il n'est pas encore configuré, car le moment optimal pour nettoyer la base de données n'a pas encore été déterminé.

(Dans l'idée il faut supprimer les données quand elles ne sont plus dans le rss ou l'API.)

## Autres

**Data lekiosquenantais.fr** : Il s'agit d'un workspace de données mis de côté en raison du manque d'informations du flux rss.

**Tags GPT à partir de la description** : C'est le workspace que j'avais utilisé lors de mes tests avec Chat GPT.

**Développement de WebUI pour les tags à partir de la description** : C'est le workspace utilisé lors des tests de l'interface WebUI sur une machine Linux qui n'était pas suffisamment puissante.

## 7. Prod / Dev

Environnement de production et d'un environnement de développement.

J'ai mis en place deux environnements de travail distincts pour permettre à mes collègues de tester une version bêta de l'outil tout en continuant le développement du projet.

J'ai supprimé les workspaces ayant le préfixe "Dev" qui étaient en double, car l'objectif désormais est de reprogrammer les workspaces dans un langage de programmation(python, node js). Cela permettra d'aligner les fonctionnalités du Directus de développement sur celles du Directus de production.

## 8. Adresse et Mot de Passe

(supprimé pour le rapport )

Fait à Nantes le 23 / 02 / 2024

François PICARD

Responsable technique

