



Dossier d'initialisation du projet d'interopérabilité

AlpesTransport

Mohammed ROUABAH
Ilyes ZEGHDALLOU
William MAILLARD

01/01/2023

1 Réunion du Lundi 31/01

Ordre du jour :

- Mise en place des idées utiles.
- Création d'un répertoire GitLab.
- Mise à jour du diagramme de Gantt.

2 Tour de table

Un répertoire GitLab a été créé, les différentes tâches pour la première phase de développement ont été réparties.

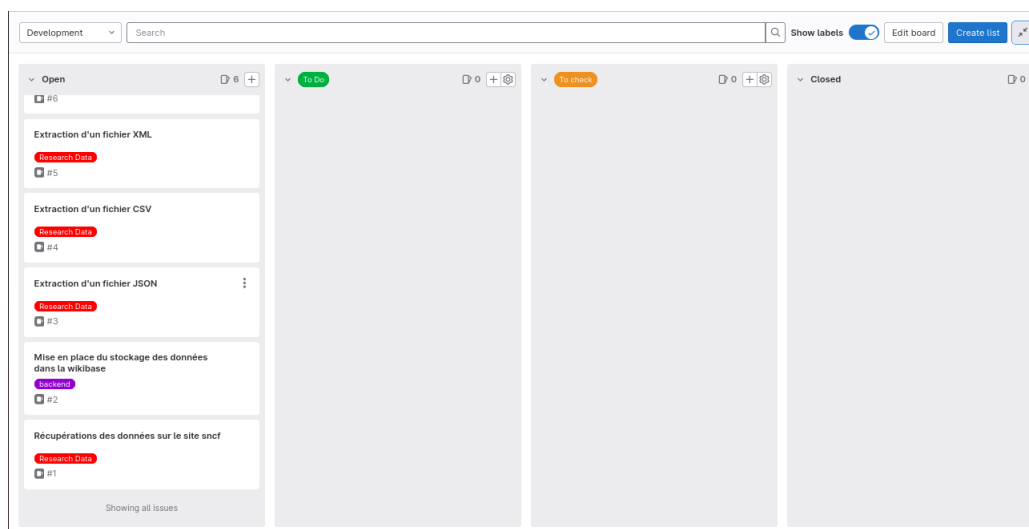


FIGURE 1 – Gestion de tâche

Un groupe discord a aussi été ouvert, avec la création de différentes sections pour optimiser le temps et partager les informations nécessaires au développement du projet.

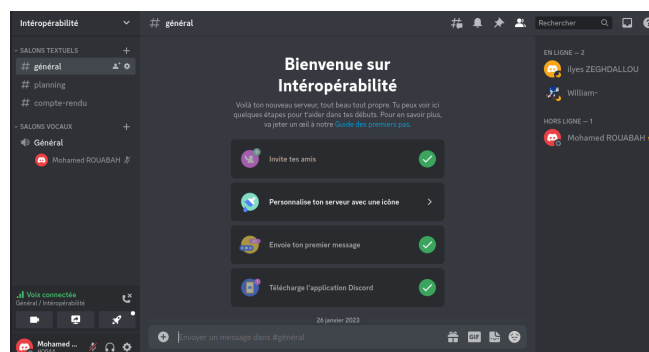


FIGURE 2 – Groupe discord

3 Planning previsionnel

AlpesTransport

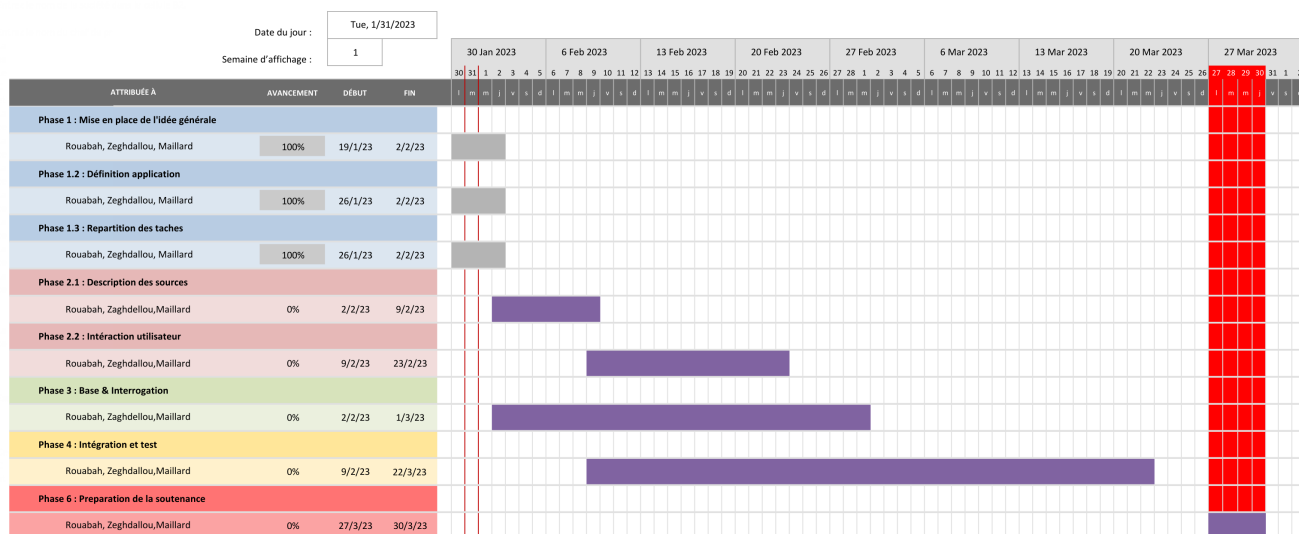


FIGURE 3 – Diagramme de Gantt

Cette semaine nous avons effectué la recherche des différentes sources et types de fichiers (XML,JSON,CSV,GTFS) qui seront a utiliser pour collecter les données. Nous avons aussi établis le sujet du projet et les différentes données que nous allons collecter sur les transports et réaliser un planning et un *taskboard* pour l'organisation du projet.

4 Le projet

Le projet a pour objectif de promouvoir les activités locales en rassemblant des informations locales, sur un domaine choisis, dans une WIKIBASE mise à jour régulièrement. Nous avons choisi de nous concentrer sur les informations relatives aux transport en commun.

Ainsi notre wikibase permettra de centraliser les informations tel que les noms des arrêts, les horaires et les tarifs pour différentes entités de transport en commun (stas, sncf, bus région Auvergne-Rhône-Alpes) afin d'aider un particulier à se déplacer dans la région.

Les données insérées dans notre wikibase seront collectées à partir de 3 sources, qui fournissent chacune des données sous un format différent (csv, xml/json, gtfs/txt, API wikidata) :

1. [SNCF](#)
2. [STAS](#)
3. [RegionRhôneAlpes](#)

Chacune de ces sources sera associé à un *parseur* permettant d'extraire les informations des documents collectés, et de les insérer dans la WIKIBASE en respectant le schéma préétablit.

En effet, la WIKIBASE sera alimenté en respectant un schéma définit par nos soins permettant de mettre en exergue les relations entre les entités créées.

Ces relations seront ensuite utilisées pour rechercher les informations requêtées par l'utilisateur via une application. Cette application permettra entre autres de consulter les transports en communs passant à un lieu donnée, les différents tarifs, et les horaires, afin de permettre à l'utilisateur d'organiser au mieux son déplacement dans la région Auvergne-Rhône-Alpes.

5 Déroulement du projet

Le projet se déroule sur 7 semaines et se compose de 3 principales parties :

- Collecte et extraction de données
- Modélisation et remplissage d'une WIKIBASE
- Interface utilisateur

5.1 Collecte et extraction de données

Pour chaque source de données (4) nous allons programmer un *parseur* pour extraire les données qui nous intéressent. Pour cela il va falloir identifier les données utiles pour chaque source et analyser la structure des documents afin de les extraire. Les données collectés seront en lien avec :

- horaires
- tarifs
- lieux des arrêts
- mode de transport

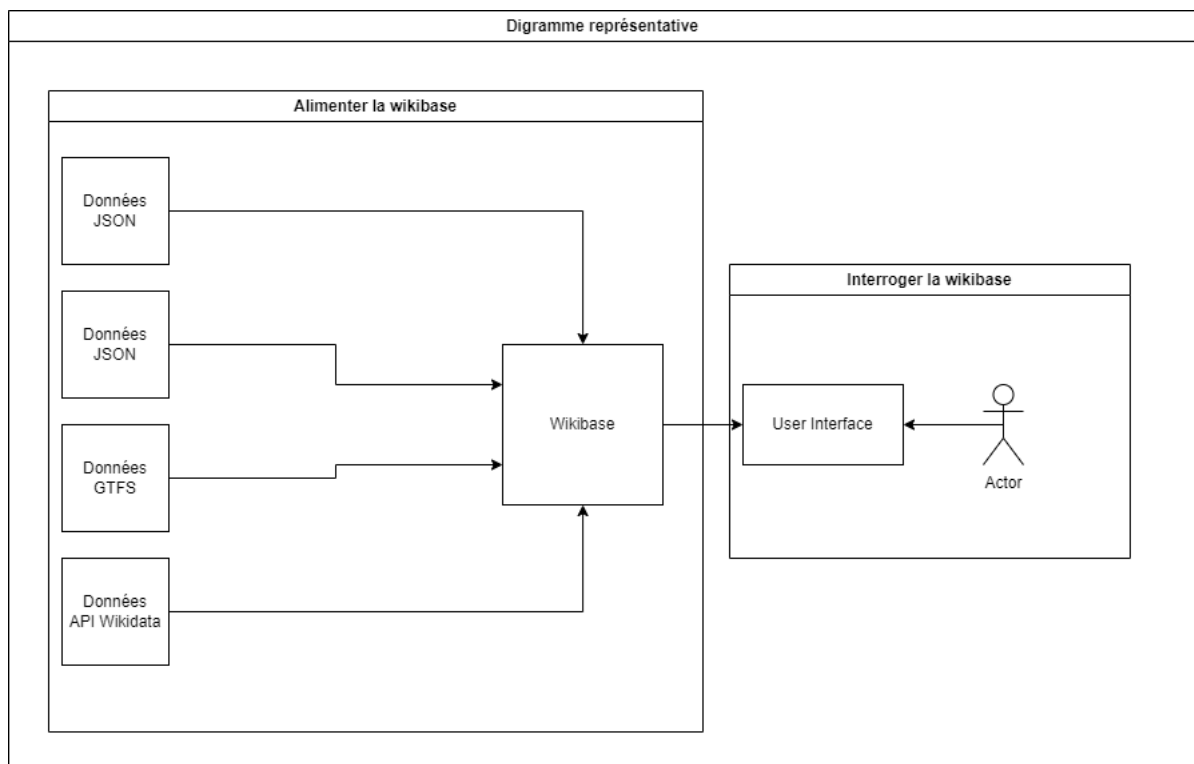


FIGURE 4 – Architecture du projet

5.2 Modélisation et remplissage d'une wikibase

La WIKIBASE nécessite la représentation des données sous forme de graphes, il nous faut donc créer un modèle de données à partir des entités et leurs relations, extraites de nos différentes sources d'informations.

Nous avons pour l'instant identifié les entités suivantes :

- entreprise de transport (stas, sncf, region Auvergne-Rhône-Alpes)
- ligne de transport (bus, train)
- les lieux des arrêts et leurs horaires

Voici un exemple de ce à quoi va ressembler notre modèle de wikibase :

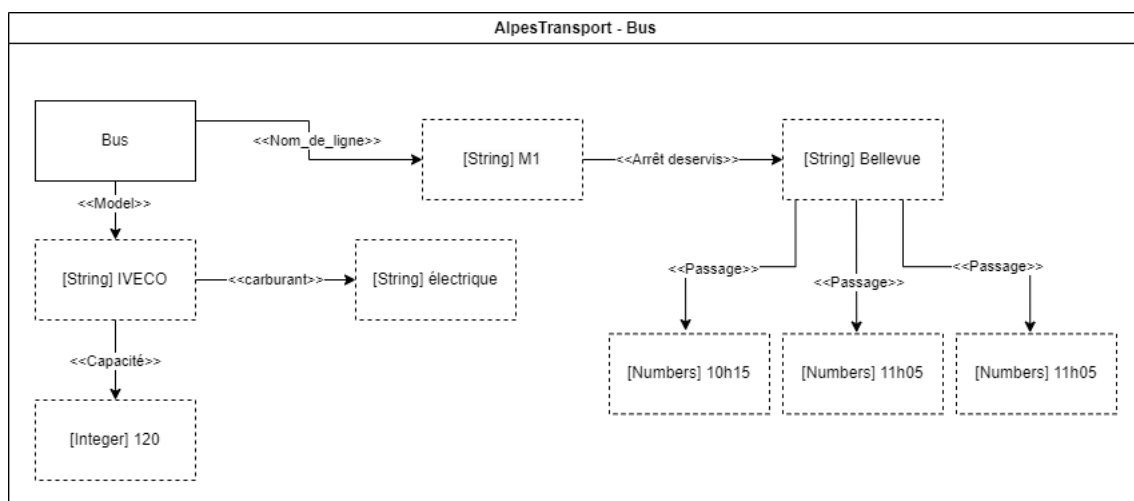


FIGURE 5 – Exemple de modèle de données pour la wikibase

5.3 Interface utilisateur

L'objectif final pour l'utilisateur est de lui permettre d'obtenir des informations sur les transports en communs. Pour cela nous allons développer une application qui permettra de faire des requêtes à notre WIKIBASE afin d'en extraire les informations demandées par un utilisateur. Les informations demandées pourront concerner :

- consulter les horaires des transports en communs à un arrêt
- les tarifs des transports pour un déplacement
- les différentes options de transport pour aller d'un point A à un point B

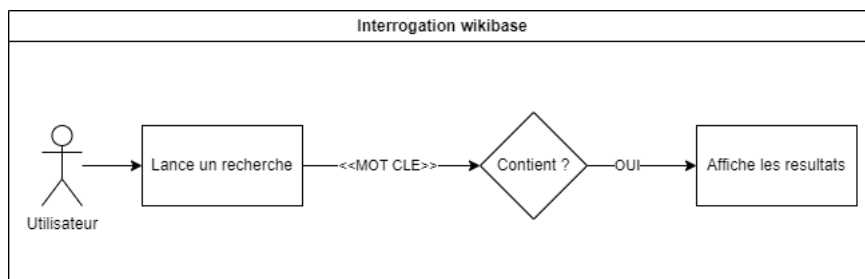


FIGURE 6 – Recherche d'information par un utilisateur

5.4 Parallélisation des tâches

Le projet contient 3 parties distinctes, qui vont être développées en parallèles. Dans un premier temps nous avons identifié des sources d'information, ce qui va nous permettre de collecter les données. En même temps que cela nous pouvons d'ores et déjà commencer le design du modèle de notre WIKIBASE, et l'UX-design de l'application client.

Nous allons ensuite pouvoir passer à l'insertion des données dans la WIKIBASE et la finalisation de son modèle de représentation des données. En même temps que cela on peut commencer à réfléchir aux requêtes utilisateurs qui seront envoyées par l'application.

Et enfin, nous réaliseront l'extraction et la constitution des réponses aux requêtes utilisateurs ainsi que des tests pour vérifier le bon fonctionnement de notre application.

6 Prochaine réunion et objectifs de la semaine

Les objectifs de la semaine prochaine seront d'analyser les différents documents récoltés depuis les sources énoncées précédemment (4) et d'en extraire leurs structures. Cela permettra ensuite de repérer les données qui nous intéressent et de réfléchir à un moyen de les extraire.

Suivant l'avancement de cette première tâche, nous pourrions envisager de commencer à programmer un *parseur*, pour chaque document, permettant d'en extraire les informations pertinentes pour le projet.

Ce jeudi nous allons nous répartir un format de document chacun pour effectuer cela d'ici la semaine prochaine, et partager nos idées sur le modèle de la *wikibase* pour stocker nos données.