

# Projet Fil Rouge SRI 2020-2021

## Moteur de recherche

### Participants

Mathias Bossaerts  
Baptiste Pomarelle  
Théo Trafny  
Martial Bailly  
Davy Duvivier

Client :	Julien Pinquier
Responsable Projet Fil Rouge :	Julien Vanderstraeten
Date de rendu :	17 / 11 / 2020
Version du document :	0.2

# 1. Introduction

Ce dossier de spécification a pour objectif de lister les fonctionnalités du “projet fil rouge 2020-2021” dans le cadre de la première année de la formation “Systèmes Robotiques et Interactifs” à UPSSITECH (voir figure 1).

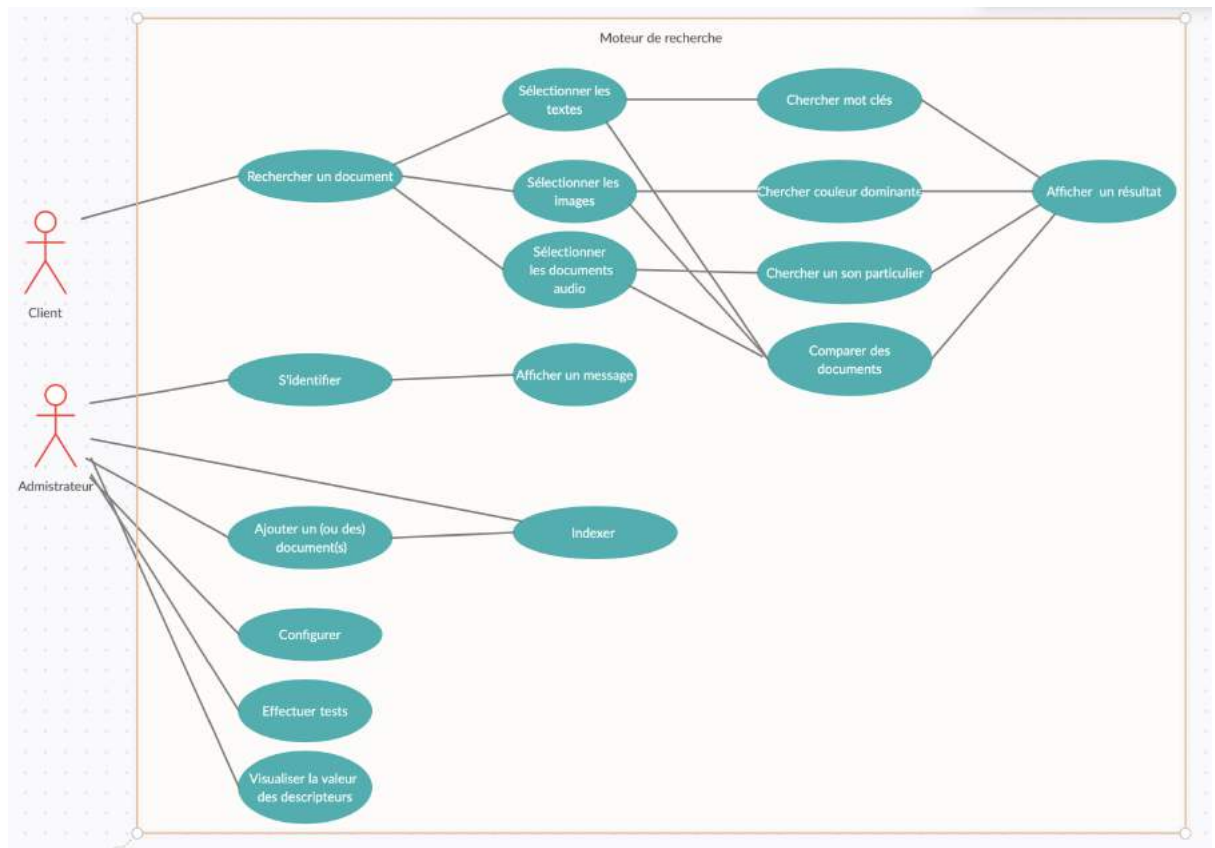


Figure 1 : Diagramme Use Case

Le projet fil rouge porte sur la conception et la réalisation d'un moteur de recherche. A travers ce document, nous allons développer la spécification de ce projet. Un moteur de recherche peut se séparer en 2 parties :

- Indexation des documents

Un administrateur fournit un document, ou un ensemble de documents, au moteur de recherche, celui-ci va devoir le(s) traiter afin de pouvoir le(s) stocker et ainsi y accéder rapidement par la suite.

- Recherche des documents

Permet à l'utilisateur de rechercher une liste de documents correspondant aux critères fournis par ce dernier lors de sa recherche.

Deux rôles seront disponibles dans le moteur de recherche :

- Administrateur
  - Ajouter un (ou des) document(s)
  - Configurer chaque processus d'indexation
  - Lancer un processus d'indexation
  - Visualiser la valeurs des descripteurs
  - Effectuer des tests
- Utilisateur
  - Saisie d'une requête selon des critères (cf Tableau 1 pour les type de documents gérés) :
    - Recherche de document texte à partir de mots clés
    - Recherche de document audio à partir d'un son particulier
    - Recherche d'un document image à partir d'une couleur dominante
    - Recherche de documents similaires à partir d'un autre document
  - Visualiser les résultats

Une discussion concernant la rigidité au niveau de la sécurité a été faite. La sécurité au niveau de la connexion d'un administrateur n'a donc pas besoin d'être sophistiquée, mais ne sera pas pour autant laissée sur le côté.

Notre moteur de recherche gérera trois types de documents (voir Tableau 1).

*Tableau 1 : type de documents gérés et extensions*

Type de document	Extension des fichiers pour la visualisation	Extension de fichiers pour le traitement
Texte	.xml	.xml
Audio	.wav	.bin
Image	.jpg, .bmp	.txt

## 2. Indexation automatique (Mode administrateur)

**Fonction :** indexation

**Entrées :**

- Chemin vers un dossier. Ce chemin représente un dossier qui contient les fichiers à indexer (cf. tableau 1 pour les types de fichiers gérés).

**Sortie :** Un ensemble de descripteurs associé à chaque document indexé.

**Description :** La fonctionnalité d'indexation va permettre à l'administrateur (cf. accès administrateur) de générer automatiquement des descripteurs de fichiers à travers un processus d'indexation. Ce processus sera configuré à l'aide d'un ensemble de paramètres (cf. ci-dessous) et générera un ensemble de descripteurs associés à un ensemble de fichiers demandés.

**Fonction :** tests (optionnelle)

**Entrées :**

- Chemin vers un dossier. Ce chemin représente un dossier qui contient les fichiers à indexer (cf. tableau 1 pour les types de fichiers gérés).
- Un ensemble de descripteurs valides associés aux fichiers à indexer (réalisés manuellement ou fournis)

**Sortie :** Une comparaison entre les descripteurs obtenus par indexation et ceux validés.

**Description :** L'administrateur va lancer une indexation automatique sur un dossier contenant des fichiers destinés aux tests, qui auront donc leurs descripteurs validés manuellement ou fournis par avance. Les descripteurs obtenus seront comparés avec ceux fournis. L'administrateur pourra visualiser les différences et déterminer les éventuelles erreurs.

Voir en annexe 2 : [Diagramme de séquence - Indexation](#)

Voir en annexe 6 : [Scénario - Indexation](#)

### 3. Moteur de recherche (Mode utilisateur)

La fonction moteur de recherche permet à un utilisateur de rechercher, à partir d'un fichier initial ou d'un critère propre à chaque type de documents (*audio* ou *image* ou *texte*) que l'on peut traiter, les documents qui ressemblent ou qui sont pertinents vis-à-vis des paramètres de recherche.

**Fonction :** recherche

**Entrées :** Critère de recherche OU document

- **Un critère de recherche.** Exemple : un mot pour la recherche de documents texte, une couleur dominante pour la recherche d'une image, recherche d'un son dans un document audio.
- **Un document.** Recherche les documents similaires au document fourni en fonction d'un seuil de similarité.

**Sortie :** Document(s) correspondant à la recherche, c'est-à-dire de même type que la recherche.

**Description :** La fonctionnalité de recherche permet de lancer une recherche en fonction d'un critère de recherche choisi par l'utilisateur ou en fonction d'un document.

**Fonction :** comparaison

**Entrées :** 2 descripteur de même type

**Sortie :** Trois cas possibles

- Identiques : les deux descripteurs sont strictement les mêmes
- Similaires : les deux descripteurs de même type ont plus de x% de caractéristiques en commun. X étant le seuil de similarité.
- Non similaires : les deux ne sont pas de même type ou ont moins de x% de caractéristiques en commun.

**Description :** La fonctionnalité de comparaison permet de comparer deux descripteurs (cf. indexation automatique) de même type (audio ou image ou texte) en fonction d'un seuil de similarité.

Voir en annexe 1 : [Diagramme de séquence - Moteur de recherche](#)

Voir en annexe 3 : [Scénario - Moteur de recherche Texte](#)

Voir en annexe 4 : [Scénario - Moteur de recherche Audio](#)

Voir en annexe 5 : [Scénario - Moteur de recherche Image](#)

## 4. Répartition du travail

Tableau 2 : répartition du travail

Modules	Spécifications (Image, Audio et Texte)	Tous
Module Image	Indexation automatique	Mathias Bossaerts + Davy Duvivier
	Moteur de recherche	
Module Audio	Indexation automatique	Théo Trafny + Baptiste Pomarelle
	Moteur de recherche	
Module Texte	Indexation automatique	Baptiste Pomarelle + Martial Bailly
	Moteur de recherche	
Programme principal (couche supérieure)	/	Baptiste Pomarelle
Organisation / Communication	/	Mathias Bossaerts

### Précisions :

- Le programme principal est la partie du logiciel qui lie les différents modules décrits ci-dessus. Il comporte le menu et la configuration des paramètres.
- Organisation / Communication regroupe toutes les tâches non techniques :
  - Communication avec le client
  - Répartition des tâches
  - Gestion du board GitLab
  - Gestion du répertoire Git

## 5. Planning

Nous vous proposons de consulter notre planning pour ce projet à travers un diagramme de Gantt.

La présentation du planning se présente en plusieurs figures. Tout d'abord, un résumé des données présentées dans le diagramme : dates du projet, ressources, tâches. Ces éléments nous permettent d'avoir une vue d'ensemble sur le projet (voir figure 2).

Nous avons ensuite les ressources humaines disponibles pour le projet ainsi que leur éventuel rôle (voir figure 3).

Puis la liste des tâches et leurs sous-tâches à effectuer avec leurs dates de début et fin (voir figure 4).

Enfin, le diagramme de Gantt en novembre qui permet d'avoir une vision globale de la répartition des tâches dans le temps. Le diagramme est découpé en 3 figures pour faciliter la lecture. (voir figure 5, 6 et 7).

Projet Fil Rouge		14 nov. 2020
Chef de projet	Mathias Bossaerts	
Dates du projet	6 nov. 2020 - 4 janv. 2021	
Avancée	0%	
Tâches	13	
Ressources	5	
Projet informatique se déroulant tout au long de la première année du cycle ingénieur		

Figure 2 : Entête du diagramme

### Ressources

Nom	Rôle par défaut
Mathias Bossaerts	Chef de projet
Martial Bailly	Non défini
Davy Duvivier	Non défini
Théo Trafny	Non défini
Baptiste Pomarelle	Non défini

Figure 3 : Ressources humaines disponibles

## Tâches

Nom	Date de début	Date de fin
Projet Fil Rouge	06/11/2020	03/01/2021
Création Dossier spécification	06/11/2020	13/11/2020
Développement moteur de recherche	15/11/2020	03/01/2021
Développement module image	15/11/2020	03/01/2021
Indexation image	15/11/2020	19/12/2020
Sous-moteur de recherche image	19/12/2020	03/01/2021
Développement module texte	15/11/2020	03/01/2021
Indexation texte	15/11/2020	19/12/2020
Sous-moteur de recherche texte	19/12/2020	03/01/2021
Développement module audio	15/11/2020	03/01/2021
Indexation audio	15/11/2020	19/12/2020
Sous-moteur de recherche audio	19/12/2020	03/01/2021
Intégration du programme	12/12/2020	03/01/2021

Figure 4 : Liste des tâches à réaliser avec les dates associées

## Diagramme de Gantt

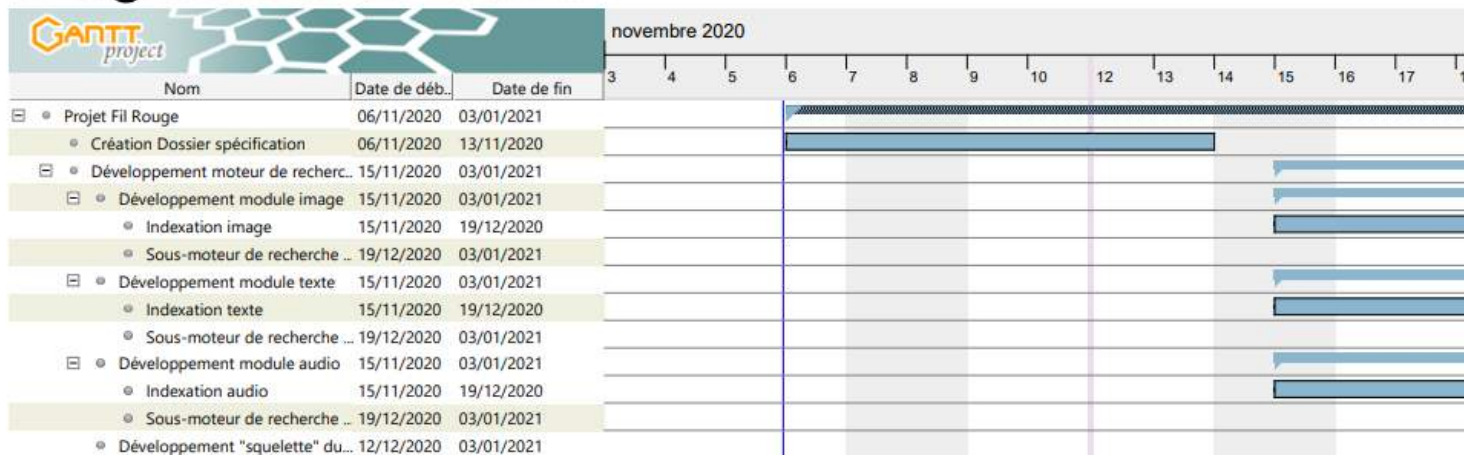


Figure 5 : Diagramme de Gantt du projet partie 1



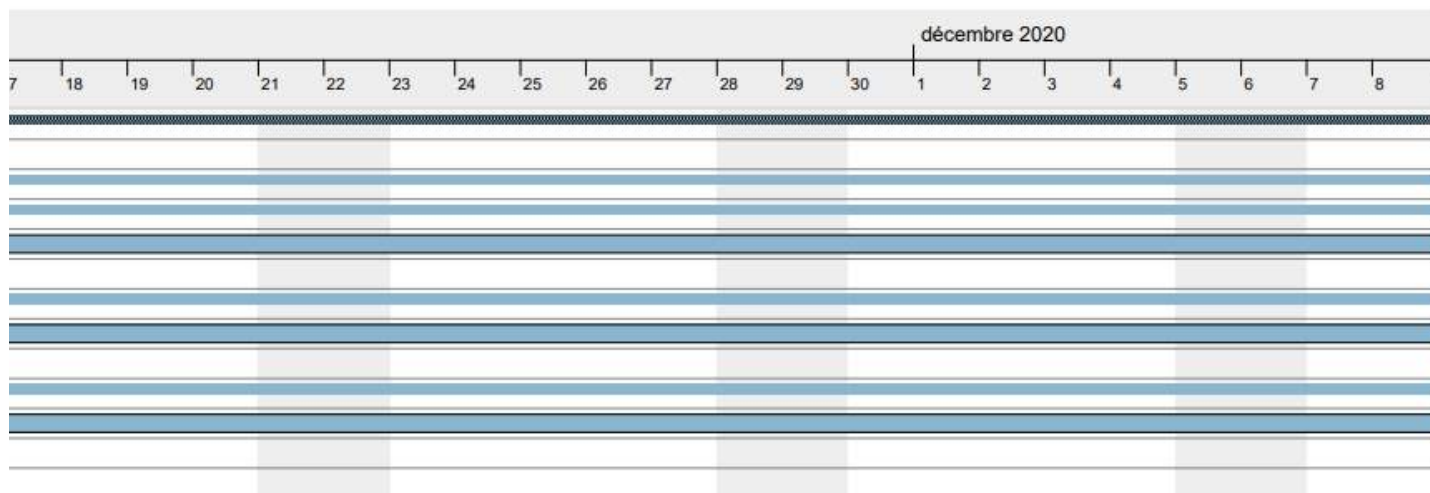


Figure 6 : Diagramme de Gantt du projet partie 2

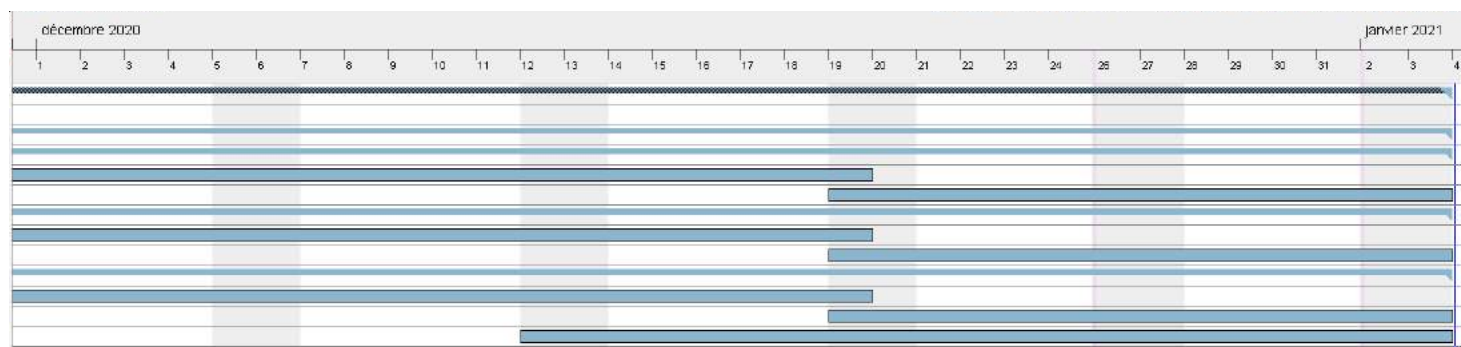
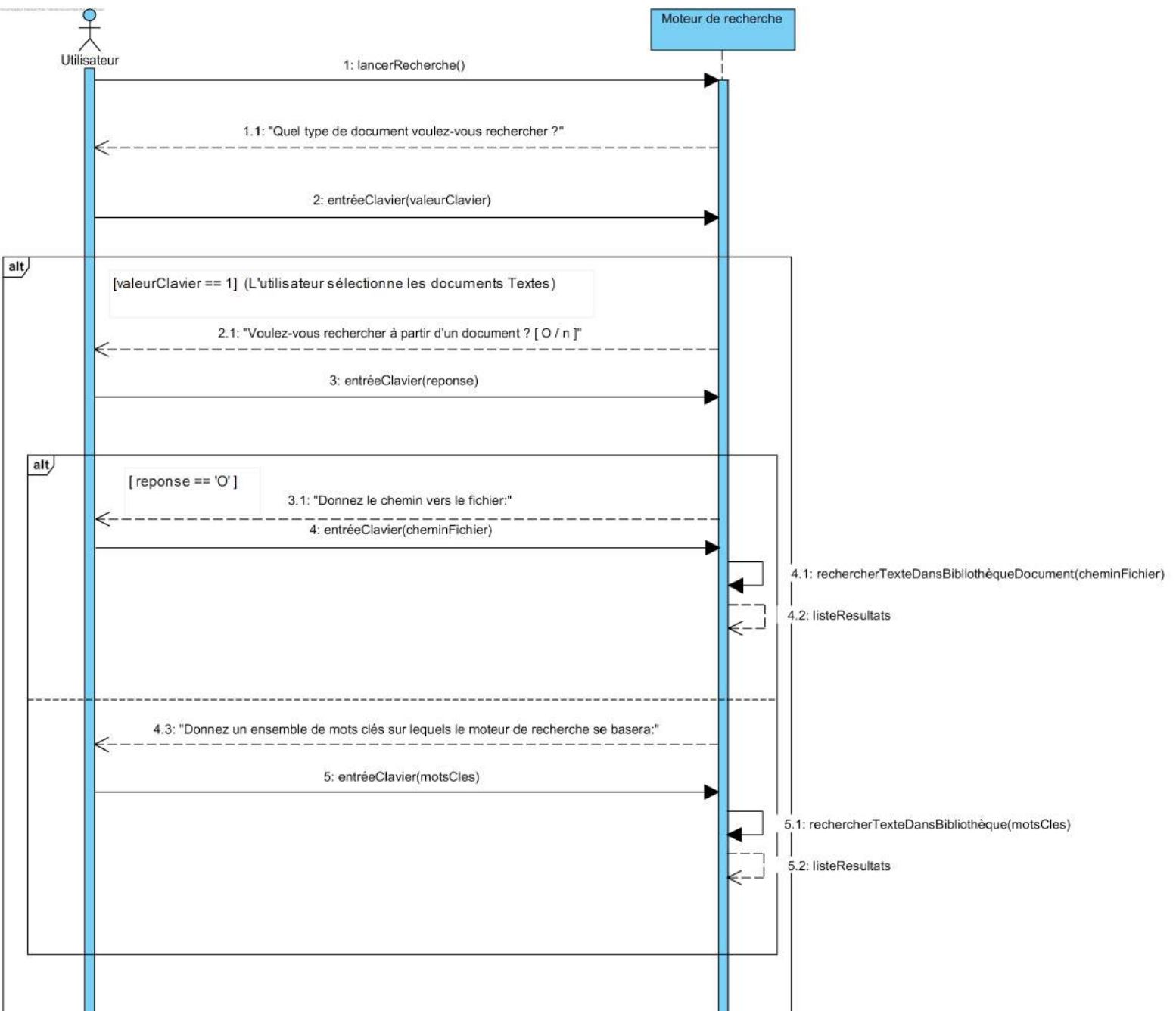
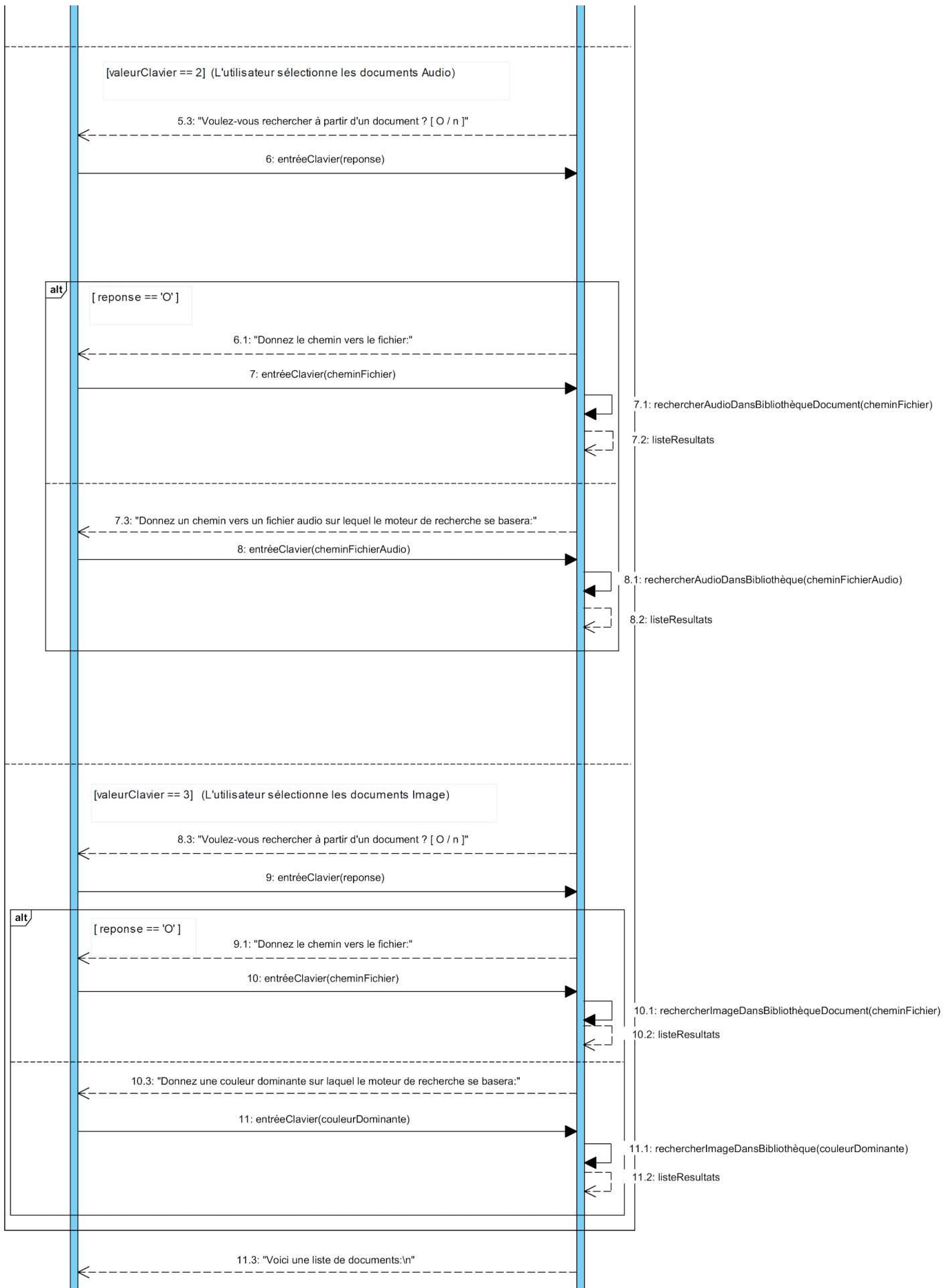


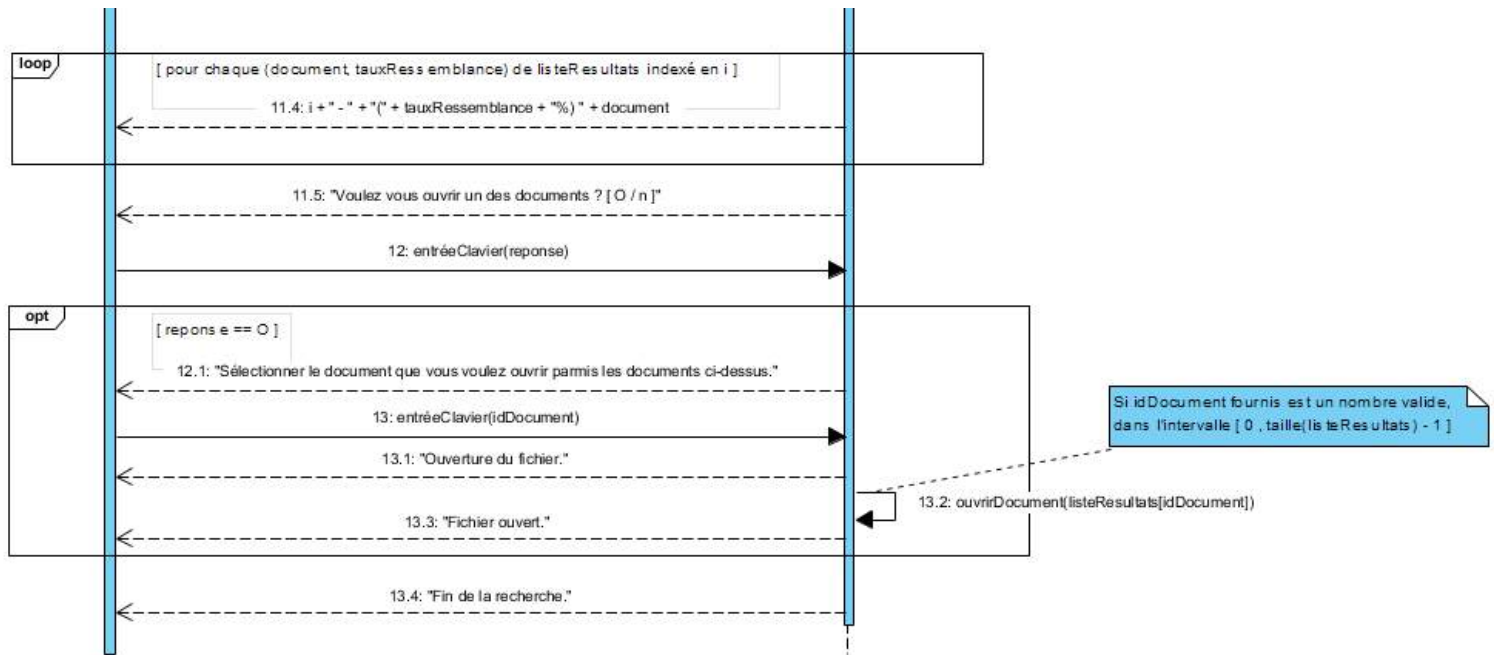
Figure 7 : Diagramme de Gantt du projet partie 3

# Annexes

## Annexe 1. Diagramme de séquence - Moteur de recherche







*Diagramme de séquence mettant en évidence le fonctionnement détaillé de l'interaction entre l'utilisateur et le moteur de recherche.*

## Annexe 2. Diagramme de séquence - Indexation

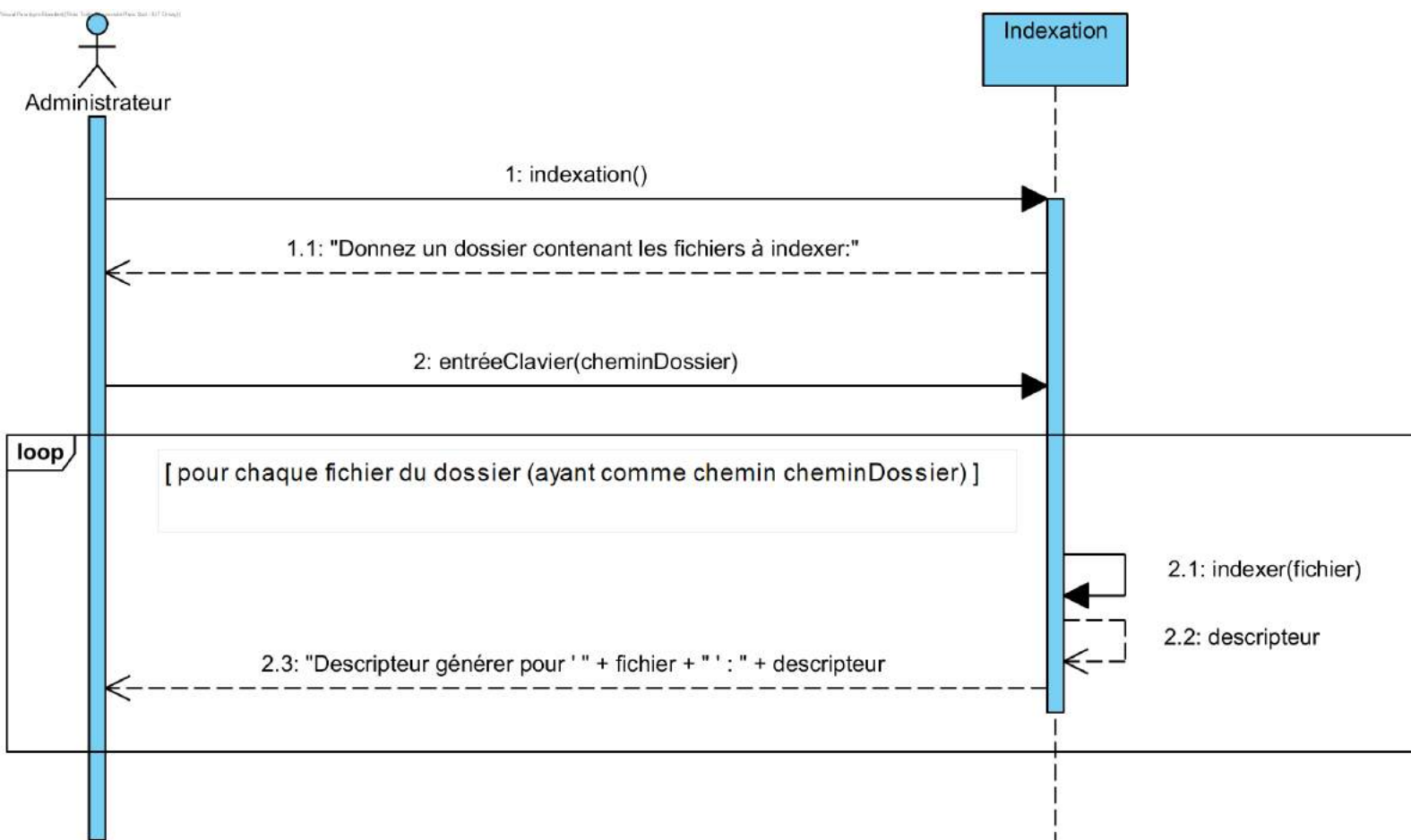
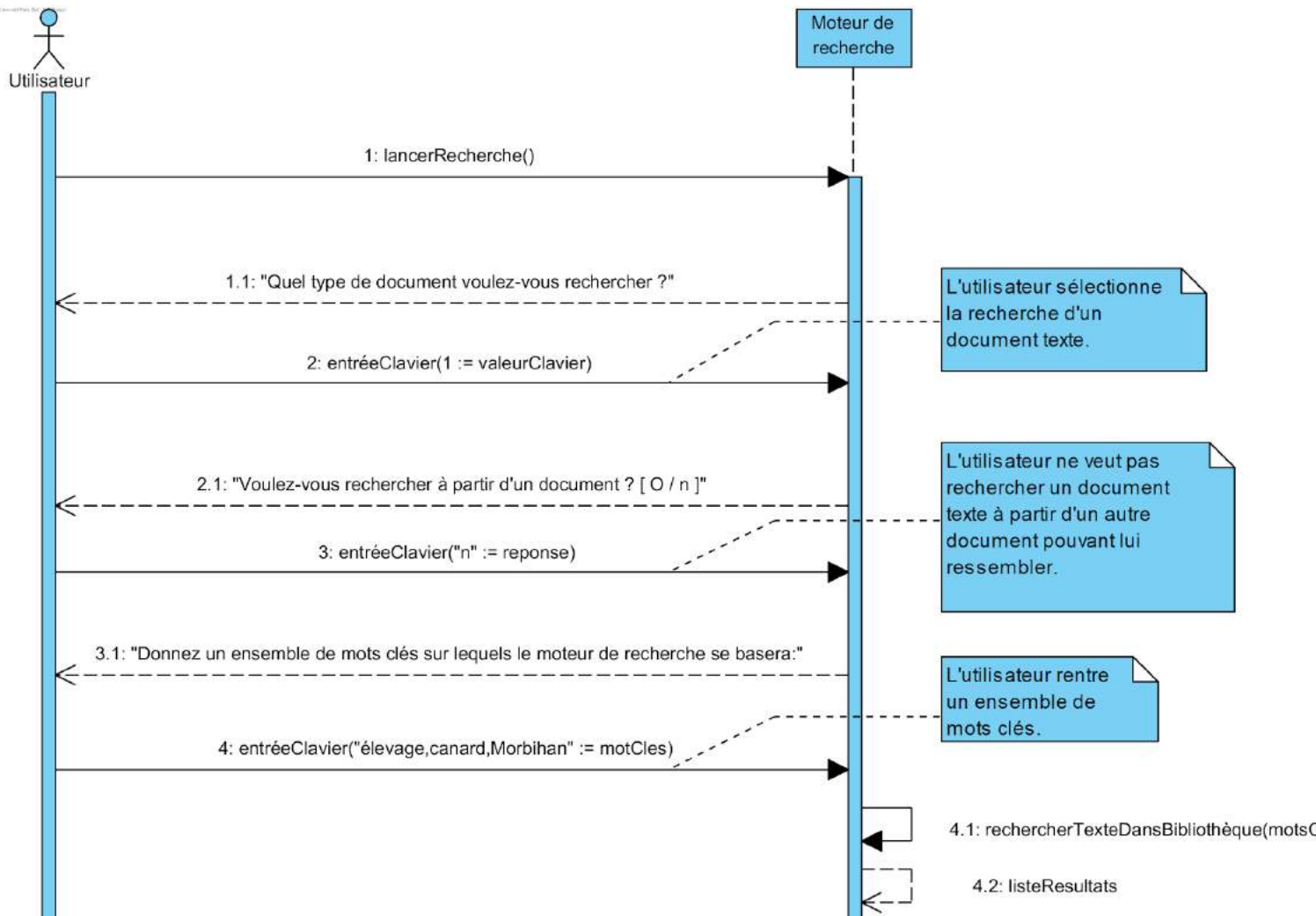
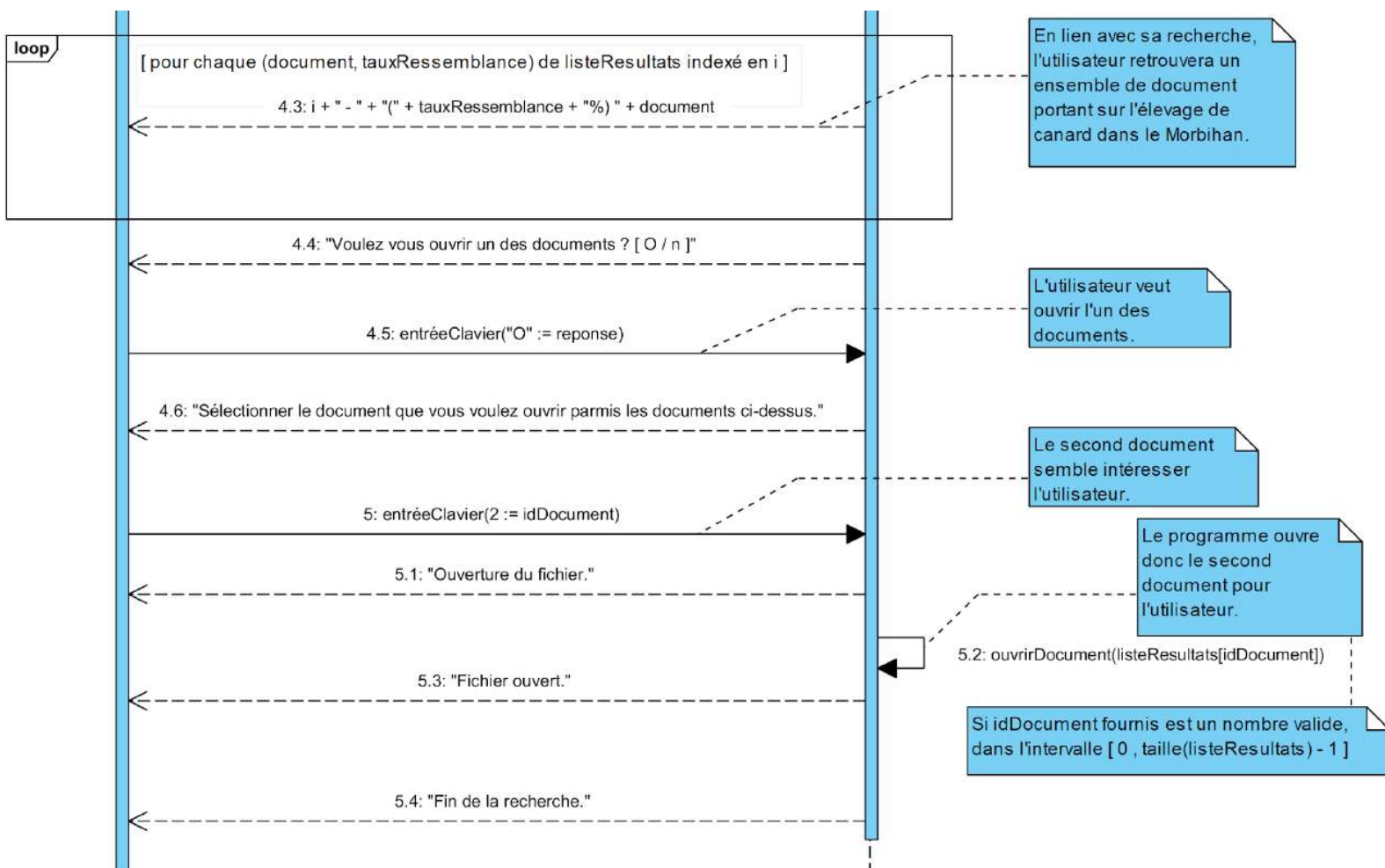


Diagramme de Séquence mettant en évidence le fonctionnement détaillé de l'interaction entre l'administrateur et le système d'indexation.

### Annexe 3. Scénario - Moteur de recherche Texte

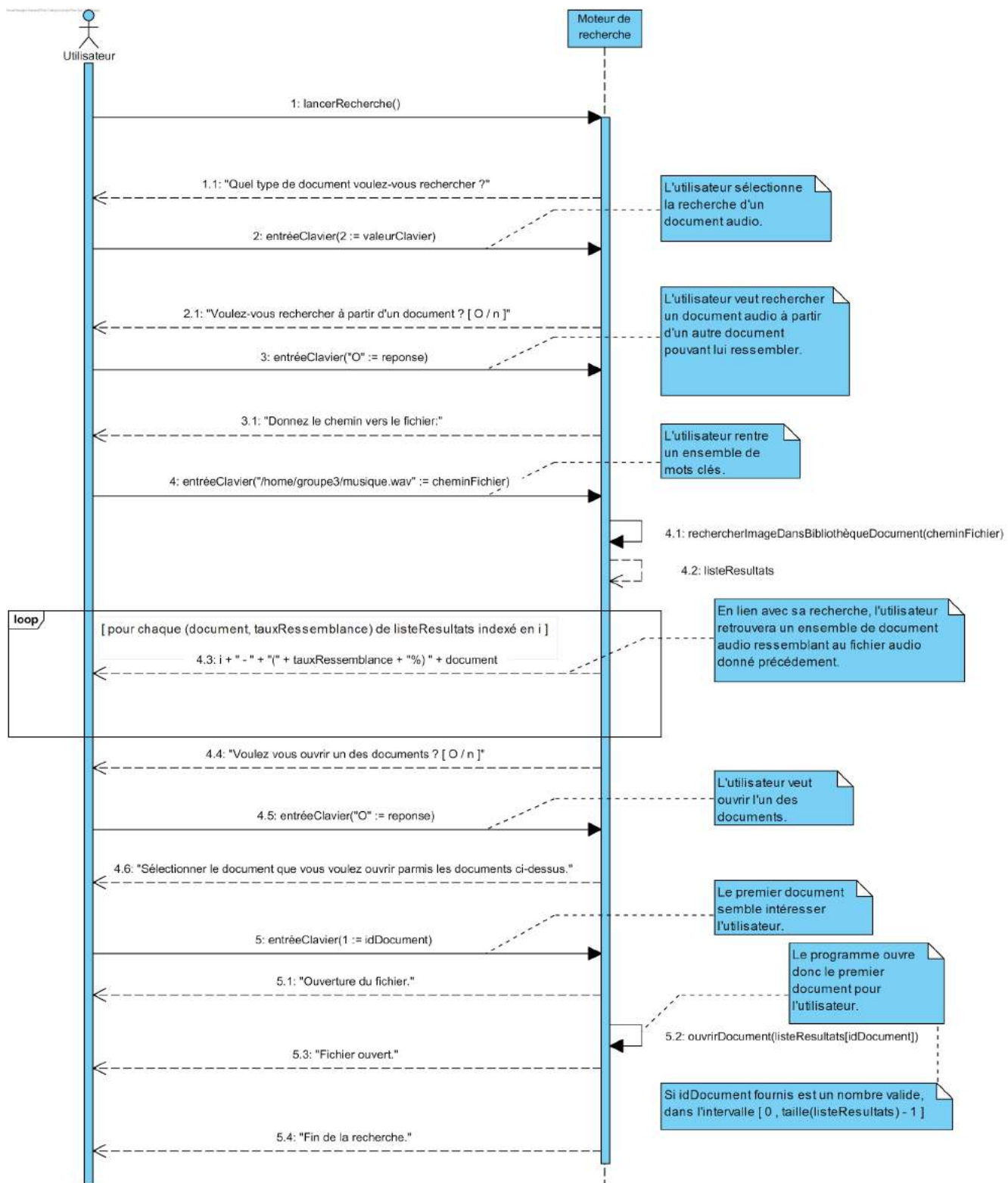




Scénario mettant en scène le fonctionnement détaillé de l'interaction entre l'utilisateur et le moteur de recherche pour la recherche d'un mot clé dans un document texte.

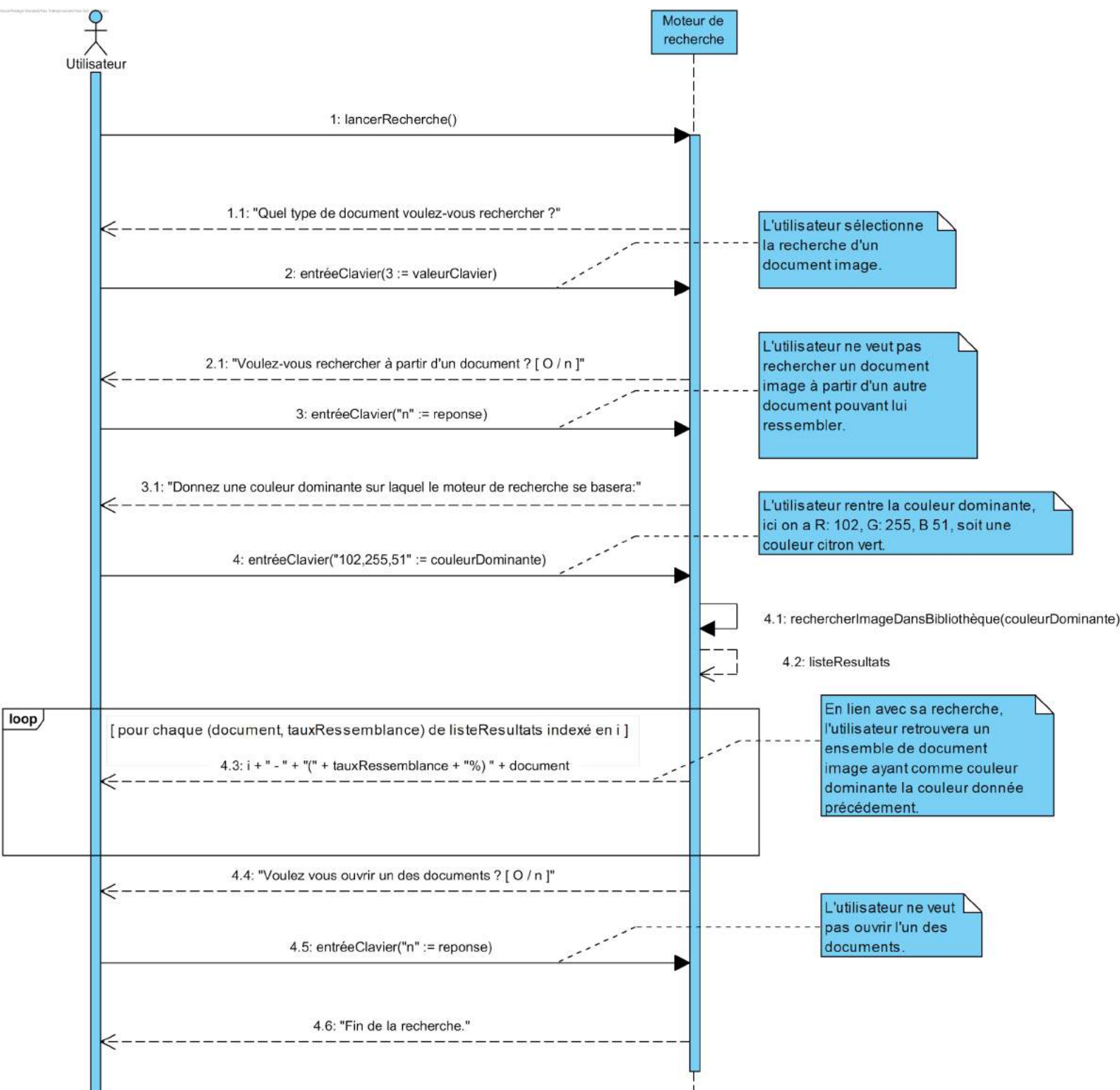


## Annexe 4. Scénario - Moteur de recherche Audio



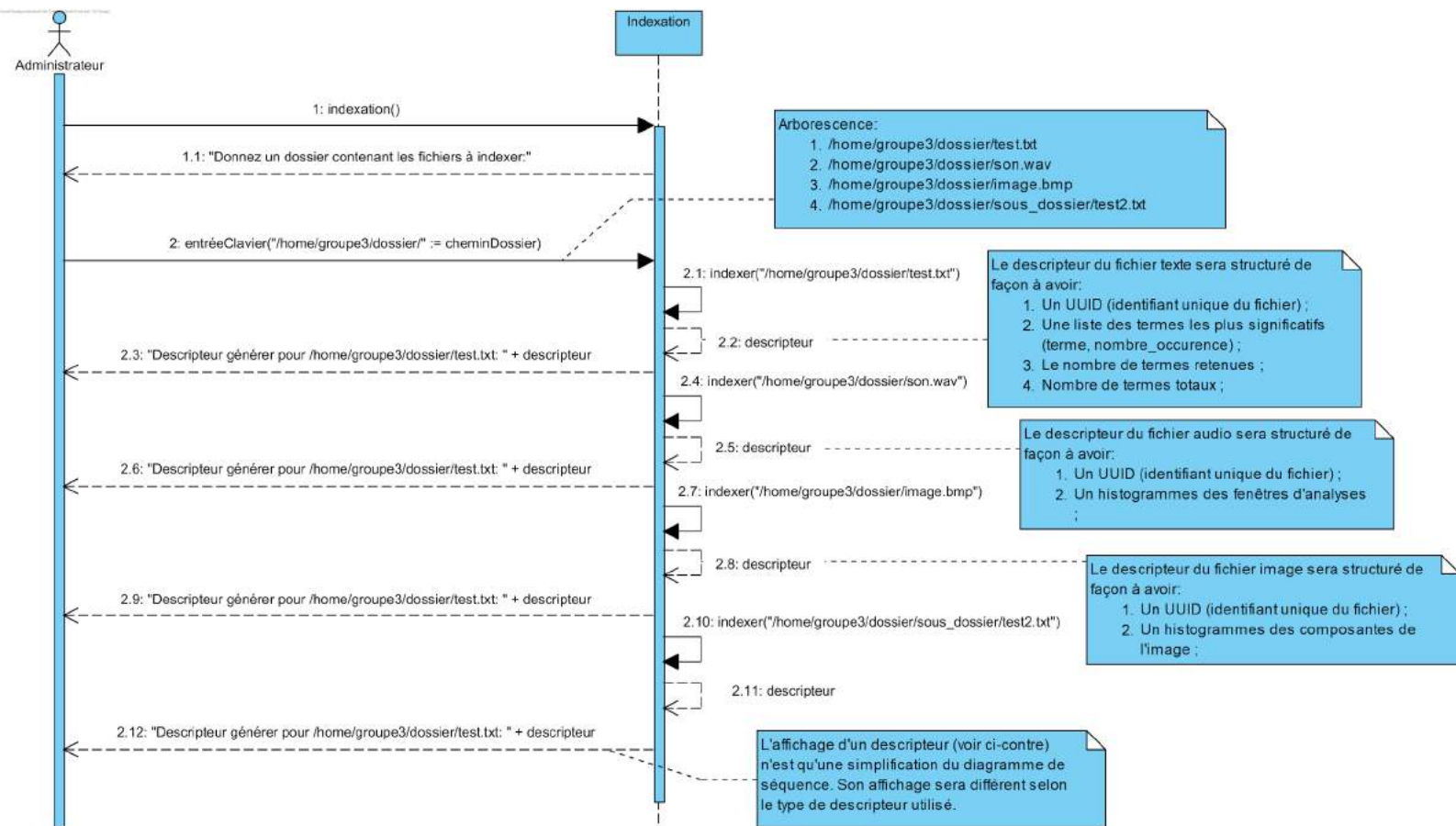
Scénario mettant en scène le fonctionnement détaillé de l'interaction entre l'utilisateur et le moteur de recherche pour un document audio.

## Annexe 5. Scénario - Moteur de recherche Image



Scénario mettant en scène le fonctionnement détaillé de l'interaction entre l'utilisateur et le moteur de recherche pour la recherche d'une couleur dominante dans un document image.

## Annexe 6. Scénario - Indexation



Scénario mettant en scène le fonctionnement détaillé de l'interaction entre l'administrateur et le système d'indexation.

### Accord du maître d'ouvrage

Bon pour accord le 19/11/2020

*[Signature]*