

RAPPORT DE PROJET

Détection d'extraits TV par métadonnées

Année 2022-2023 DIIS

Historique des modifications		
Version	Date	Description
1	23/01/2023	Création du document
2	14/02/2023	Modifications post soutenance

Validation		
Nom	Date	Validation
<u>Emetteur</u> : Guillaume Mekrouda		
<u>Concerné</u> : Matthieu Delalandre		

Encadré par : Matthieu DELALANDRE

Réalisé par : Guillaume Mekrouda

TABLE DES MATIERES

1. Remerciements	3
2. Préambule	4
3. Références	4
4. introduction.....	5
4.1. Contexte du projet	5
4.2. Objectifs du projet.....	5
5. Méthodologie de gestion projet.....	6
5.1. Choix organisationnel.....	6
5.2. Planification prévisionnelle	6
6. Analyse fonctionnelle	8
6.1. Volume de données vidéo à traiter.....	8
6.2. Architecture applicative	8
7. Réalisation	10
7.1. Livrables réalisés.....	10
7.2. Logiciel.....	11
7.3. Etudes réalisées.....	11
7.3.1. Etude du nombre de scènes.....	11
7.3.2. Etude du filtrage SEO.....	13
7.4. Stratégie de Test.....	13
8. Etat du projet à date.....	15
8.1. Fonctionnalités du projet	15
8.2. Fonctionnalités restantes à implémenter	15
9. Bilan.....	17
9.1 Bilan de réalisation	17
9.2. Gant réel	18
9.2. Difficultés rencontrées	18
9.3. Conclusion	19
Table des illustrations.....	20

1. Remerciements

Je souhaiterais remercier Mr Delalandre pour avoir encadré ce projet et de s'être rendu disponible pour m'aider dans les choix à mettre en place tout au long du projet.

Je souhaiterais aussi remercier Mr Chauvin pour m'avoir apporté de l'aide pour la rédaction des différents livrables.

Je souhaiterais aussi remercier l'ensemble de l'équipe pédagogique de Polytech'Tours pour m'avoir permis de travailler sur ce projet de fin d'étude.

2. Préambule

Ce document permet de synthétiser le projet **Détection d'extraits TV par métadonnées**. Il décrira l'ensemble des éléments mis en place, les points positifs et négatifs ainsi que les points critiques. Il présentera aussi les différents choix effectués durant le projet.

- Ce projet est encadré par Mr DELALANDRE qui assurera le rôle du Maître d'Ouvrage (MOA)
- La Maîtrise d'œuvre (MOE) sera assurée par Guillaume MEKROUDA

3. Références

Ce document est lié à différents livrables écrits dans le cadre de la mise en œuvre de ce PFE, les livrables sont les suivants :

Nom	Version
Cahier de spécifications	2
Plan de développement	2
Fiche d'Ouverture de Projet (FOP)	2
Cahier d'analyse	1
Cahier de reprise	1
Manuel utilisateur	1
Manuel développeur	1

Figure 1 - Livrables de référence du rapport

4. introduction

4.1. Contexte du projet

Le laboratoire LIFAT (Laboratoire d'Informatique Fondamentale et Appliquée de Tours) est un laboratoire chargé de concevoir et de développer des modèles. Il est spécialisé dans le domaine de la fouille et visualisation de données, l'apprentissage automatique, le traitement des langues naturelles et des images ou encore l'optimisation combinatoire. Le laboratoire est organisé en 3 groupes de recherche :

- Base de données et Traitement Automatique du Langage Naturel (BDTLN)
- Recherche Opérationnelle, Ordonnancement et Transport (ROOT)
- Reconnaissance de Formes et Analyse d'Images (RFAI)

Le projet Détection d'extraits TV par métadonnées est lié à la station TV développée depuis 2018 à Polytech Tours par le laboratoire LIFAT. La station TV est une plateforme de calcul parallèle pour le traitement automatique des chaînes de télévision de la TNT. L'objectif de la station est le déploiement d'applications de traitement d'images, de vidéos et d'intelligence artificielle du laboratoire LIFAT.



Figure 2 - Logo LIFAT

4.2. Objectifs du projet

Dans le cadre du projet Antenne TV, L'objectif principal du projet est de concevoir un logiciel afin d'extraire des séquences d'intérêt pour les internautes à partir des extraits vidéos provenant de la station TV. L'idée serait, pour l'utilisateur de cette application, de pouvoir extraire des séquences, de quelques secondes idéalement, des programmes TV sur la TNT qui ont généré un important trafic.

Une partie logicielle sera ajoutée au logiciel existant pour pouvoir traiter cette problématique d'extraction de scènes d'intérêt pour les internautes. Ce projet sera livré avec l'ensemble des documentations nécessaires afin de faciliter la reprise du projet. Ces documentations décriront l'ensemble des solutions envisagées, les diagrammes, l'architecture etc... Un manuel d'utilisation sera aussi livré pour que les utilisateurs futurs puissent utiliser l'application le plus facilement possible.

Une machine de capture Dell 5820 a été mise en place au sein de Polytech'Tours équipée d'un disque dur 38 To pour pouvoir capturer en temps réel le flux vidéo issu de l'antenne TV. L'objectif final de ce projet serait de mettre en place l'ensemble des différentes couches logicielles sur cette machine pour réaliser le traitement vidéo en temps réel.

5. Méthodologie de gestion projet

5.1. Choix organisationnel

La méthodologie suivie durant ce projet est une méthodologie type cycle en V. La description de chaque phase est réalisée dans la section 5 du plan de développement.

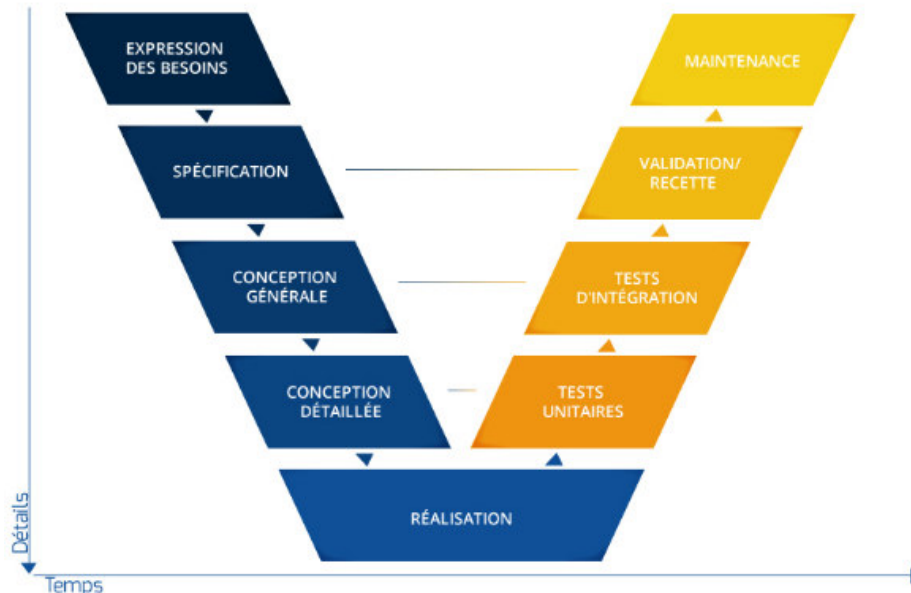


Figure 3 - Description du cycle en V

Une méthode « cycle en V » m'a paru plus judicieux à mettre en place étant tout seul sur le projet. De plus, étant donné que ce projet est un PoC, avoir une méthode formelle est un bon moyen pour ne pas s'éloigner de l'objectif du projet en spécifiant l'ensemble des fonctionnalités à implémenter dès le début du projet.

5.2. Planification prévisionnelle

Le Gantt prévisionnel mis en place est disponible ci-dessous :

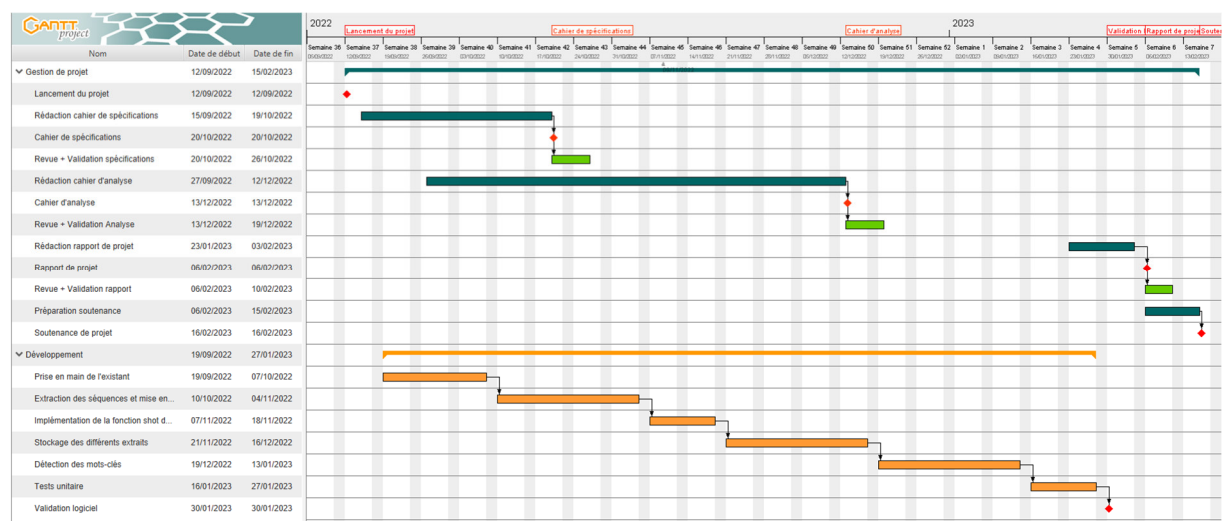


Figure 4 - Planning prévisionnel

La mise en place de Jalons pour la remise des livrables m'a aidé à mieux concevoir ce planning prévisionnel. Les différentes tâches pour la réalisation du logiciel et l'implémentation des

fonctionnalités ont été mises à la suite puisque les fonctionnalités étaient dépendantes entre elles. La rédaction des différents documents a été mise en parallèle des fonctionnalités.

6. Analyse fonctionnelle

6.1. Volume de données vidéo à traiter

Une capture vidéo a été réalisée sur la période du 17/06/2022 jusqu'au 16/07/2022 sur différentes chaînes françaises (France 2, France 3, France 5) afin d'être traité. Le volume de données vidéo traité est résumé dans le tableau suivant :

Intitulé	Description
Channel	France 2, France 3, France 5
Capture	20h par jour
Durée	30 jours (du 17/06/2022 au 16/07/2022) soit une durée de 1 800h
Vidéo	Résolution SD (720x576) à 1,6Mbps à 30 fps
Audio	256 kbps
Mode	Asynchrone audio / video
Taille	Environ 1,4 TiB
Storage	HDD02
Horodaté	no

Figure 5 - Description des données issues de la capture

Comme expliqué dans le tableau, une première problématique à avoir été traitée sur ce projet a été le ré-horodatage de la base de fichiers vidéo. Cette tâche a pu être réalisée grâce à un script qui m'a été donné ainsi que l'utilisation d'un logiciel nommé « Avidemux ».

De plus, étant donné que la capture vidéo et audio a été réalisée en asynchrone, une séparation de la partie vidéo et audio a été nécessaire à mettre en place pour pouvoir stocker les différentes parties.

Il a été fait le choix, dans le cadre de ce projet, que toute la capture effectuée ne soit pas traitée pour des raisons de performances matérielles.

6.2. Architecture applicative

Ce projet est la suite d'un stage réalisé en DI4, l'application a été programmée en Python qui est un outil intéressant pour pouvoir prototyper un programme rapidement, c'est donc le cas pour ce PoC.

La spécification du fonctionnement du programme a été réalisée en collaboration avec la MoA. Cette réunion m'a permis de mieux comprendre les attendus du projet ainsi que les technologies que j'allais être amené à utiliser. Le fonctionnement attendu de l'application est décrit dans le schéma ci-dessous :

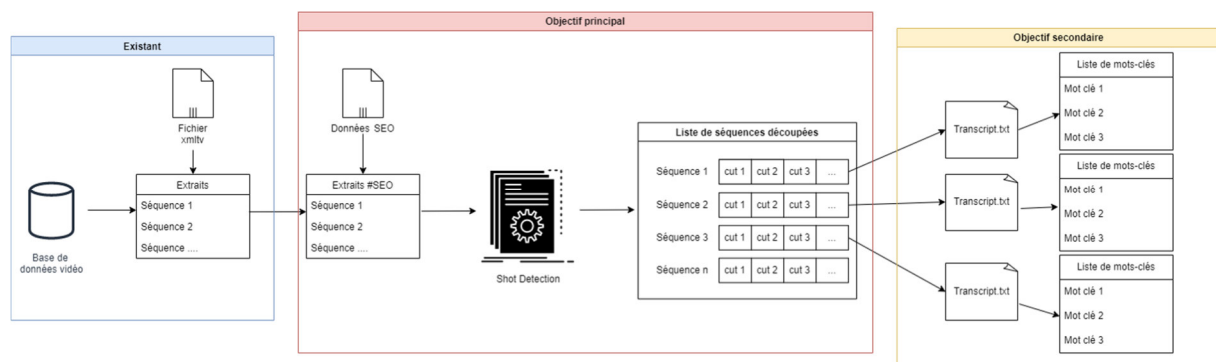


Figure 6 - Modèle de fonctionnement demandé

On distingue 3 parties différentes sur ce projet :

1. L'existant : L'existant contient l'analyse du fichier xmltv au format xml pour pouvoir récupérer l'ensemble des informations disponibles pour extraire une émission.

Le format d'un fichier xmltv est le suivant :

```
<programme start="20220711134000 +0200" stop="20220711143500 +0200" channel="C47.api.telerama.fr">
  <title>Les docs du Mag de la santé</title>
  <category lang="fr">magazine de la santé</category>
  <length units="minutes">55</length>
  <icon src="https://focus.telerama.fr/1111x625/2022/06/15/c271caed66b7081590153826725a0792eb68ca4e.jpg" />
  <rating system="CSA">
    <value>Tout public</value>
  </rating>
</programme>
```

Figure 7 - Extrait d'un fichier xmltv

On retrouve l'ensemble des informations dont nous avons besoin tel que le timestamp de démarrage de l'émission, le timestamp de stop, la chaîne sur laquelle est diffusée le programme ainsi que le nom de l'émission.

2. L'objectif principal : L'objectif principal du projet est d'à partir de la liste d'émissions TV récupérée, de faire matcher cette liste avec la liste d'émissions contenues dans le fichier SEO au format csv. Dans le cas où une des émissions présentes dans la base de données vidéo n'aurait pas de score SEO associé, celle-ci sera automatiquement rejetée de l'algorithme. Une fois ce filtrage effectué, l'idée serait d'extraire les différentes scènes d'une émission décrites par un changement de plan en fonction de l'algorithme utilisé :

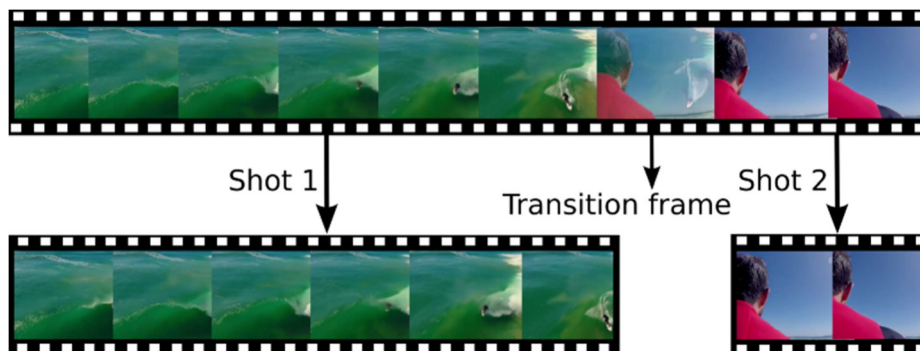


Figure 8 - Détection de scènes

L'utilisateur pourra définir une durée de scène minimum pour avoir accès à plus ou moins de scènes.

3. L'objectif secondaire : Une technique permettant d'identifier la liste des mots-clés et qui m'a été suggéré par la MoA était le TF-IDF (Term Frequency – Inverse Document Frequency). Cette méthode utilisée dans la recherche d'informations est décrite dans le cahier d'analyse section 3.4.2 du cahier d'analyse. J'ai pu donc m'informer sur les différentes bibliothèques qui existaient et identifier celle qui serait la plus adaptée pour ce projet. Le but étant de caractériser une scène en fonction des mots-clés détectés grâce à la transcription texte.

A l'issue de la réunion, j'ai pu commencer à travailler sur le cahier de spécifications. Le principal élément sur lequel je devais définir une technologie à utiliser a été l'identification des scènes.

7. Réalisation

7.1. Livrables réalisés

Le tableau ci-dessous recense l'ensemble des livrables réalisés ainsi que leurs descriptions :

Nom	Description
Cahier de spécifications	Spécifie la demande du MOA et recense l'ensemble des fonctionnalités à implémenter pour la réalisation de ce projet. Ce document pourra être repris par d'éventuels projets futurs pour pouvoir comprendre les choix techniques mis en œuvre. Les contraintes de développement et l'existant.
Plan de développement	Le plan de développement explique l'organisation globale du projet avec l'estimation de la charge de travail pour chaque tâche. Il recense aussi l'ensemble des jalons du projet avec leurs dates, les outils de développement utilisés et enfin le cycle de développement. Le cycle de développement permet de préciser quels sont les différents éléments d'entrée et de sorties pour chacune des phases du projet dans le cycle en V.
Fiche d'Ouverture de Projet (FOP)	La Fiche d'Ouverture de Projet permet de chiffrer le coût du projet au global pour les différentes tâches à réaliser pour un Chef de projet ou un ingénieur débutant. Il permet d'avoir une vision globale des différents coûts.
Cahier d'analyse	Le cahier d'analyse explique la conception globale et détaillée du projet. Dans le cadre de ce projet, il permet de mettre en avant les choix mis en place de fonctionnement du programme et de l'architecture mis en place.
Cahier de reprise	Le cahier de reprise explique ce qui n'a pas pu être réalisé dans le cadre de ce projet et ce qu'il reste à faire avec des propositions de fonctionnement. Il pourra être utilisé pour de futurs projets.
Manuel utilisateur	Le manuel utilisateur explique comment configurer et mettre en place le projet pour le type d'utilisateur définit.
Manuel développeur	Le manuel développeur explique comment configurer et mettre en place le projet ainsi que comment est structuré le code source du projet.
Poster	Le poster est un document permettant d'expliquer en quelques mots-clés et images le projet afin d'être présenté pour des événements organisés par Polytech.
Vidéo	La vidéo permet de montrer le programme en fonctionnement et les différentes entrées/sorties du logiciel.
S.P.E.R	Le S.P.E.R est un document au format Excel permettant le suivi du projet. Il permet d'effectuer une analyse sur les coûts prévisionnels comparés au coût effectif du projet. De même pour connaître l'écart de charge à l'issue du projet.
Cahier de tests	Le cahier de test permet de tester, à partir d'un jeu de données mises en place, le bon fonctionnement du programme et son comportement attendu. Cet outil pourra être réutilisé pour les prochains projets afin de savoir si les tests sont toujours passants, une fois que des fonctionnalités ont été rajoutées.

Figure 9 - Description des livrables du projet

7.2. Logiciel

En termes de fonctionnalités, le schéma suivant représente ce qui a pu être implémenté au cours de ce projet (encadré en vert) :

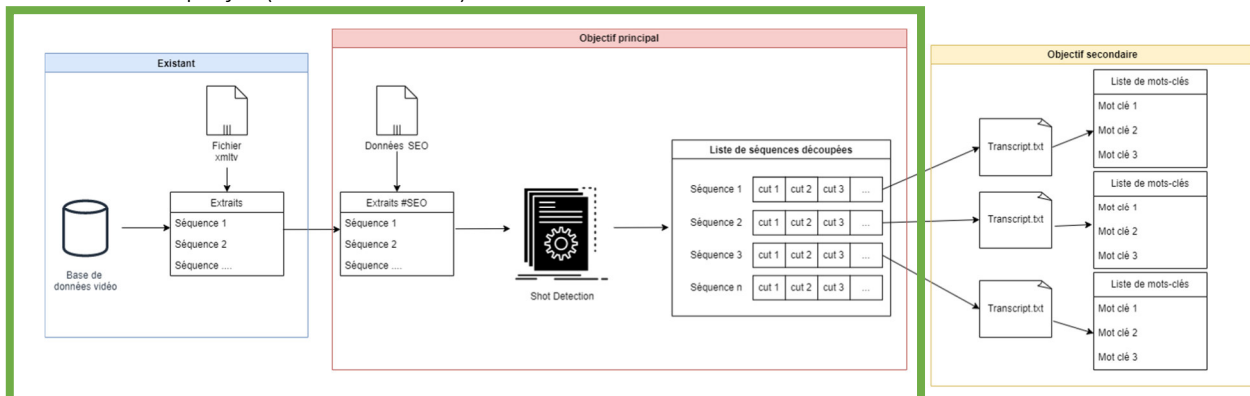


Figure 10 - Fonctionnalités implémentées

Durant ce projet j'ai pu traiter :

1. La partie filtrage à partir du score SEO (cf : cahier de spécifications section 5). Cette partie consistait au matching des données issues d'un document au format xml contenant les scores SEO pour une liste d'émissions TV par rapport à la liste d'émissions mise à disposition à l'issue de la capture vidéo. Ce filtrage est fonctionnel.
2. L'implémentation de la détection de scènes pour chaque émission TV extraites
3. L'organisation des fichiers pour pouvoir stocker les audios et vidéos segmentées par le programme

Une problématique apparue durant la prise en main du code existant a été le manque de trace d'exécution dans le programme pour pouvoir retracer d'éventuels problèmes. C'est pour cela que j'ai mis en place des logs pour pouvoir aider les prochains développeurs à comprendre le fonctionnement du projet.

De plus, j'ai pu mettre en place un fichier requirements.txt pour pouvoir gérer plus facilement les dépendances du projet (les télécharger, les supprimer, les ajouter etc...) ainsi qu'un fichier gitignore si les prochains développeurs souhaitent mettre en place un outil de gestion de versions.

7.3. Etudes réalisées

7.3.1. Etude du nombre de scènes

Pour rappel la librairie de détection des scènes dispose de plusieurs algorithmes :

- **ContentDetector**: detects fast changes/cuts in video content.
- **ThresholdDetector**: detects changes in video brightness/intensity.
- **AdaptiveDetector**: similar to *ContentDetector* but may result in less false negatives during rapid camera movement.

Figure 11 - Détecteurs de scènes disponibles

L'analyse ci-dessous, montre le nombre de scènes détectées avec les différents algorithmes mis à disposition. Il a été fait le choix d'exclure le Threshold Detector car il ne nous donnait pas de résultats cohérents comparés aux autres algorithmes.

Avec le Content Detector :

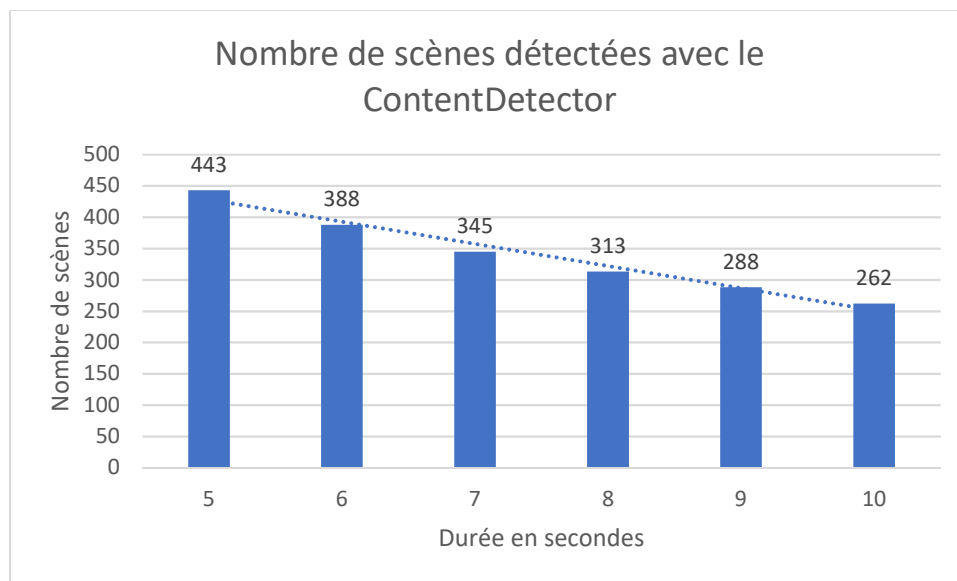


Figure 12 - Graphique d'évolution du nombre de scènes pour le Content Detector

Avec le Adaptive Detector :

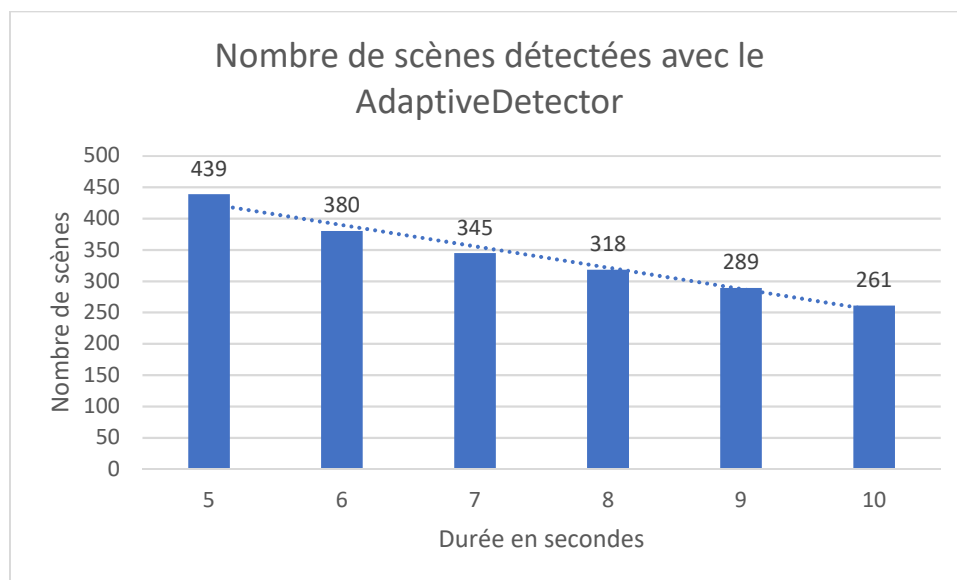


Figure 13 - Graphique d'évolution du nombre de scènes pour le Adaptive Detector

Les 2 algorithmes de détection de scènes sont très similaires, c'est pour cela que le nombre de scènes est sensiblement le même entre les 2 algorithmes. On peut néanmoins voir que l'Adaptive Detector en détecte moins, cela est dû au fait celui-ci peut détecter des « fausses scènes », c'est-à-dire des mouvements rapides de caméra.

Cette analyse a été faite sur une distribution de durée variée de 5 à 10 secondes car d'après la loi concernant la diffusion d'extraits audiovisuel sur internet :

« Pour avoir le droit de citer gratuitement un extrait audiovisuel (reportage, émission, film) dans un autre film (ou dans une exposition, un spectacle, un DVD), il faut que la citation d'extrait **soit brève, qu'elle soit justifiée "par le caractère critique, polémique, pédagogique, scientifique ou d'information de l'oeuvre"** à laquelle est incorporé l'extrait, et il faut **indiquer la source de l'extrait** (CPI art. L. 122-5 et 211-3). »

C'est pour cette raison que le contenu des extraits vidéo doit être traité pour pouvoir respecter le critère « Polémique, pédagogique, scientifique ou d'information à l'œuvre ».

7.3.2. Etude du filtrage SEO

Une étude a été réalisée afin de déterminer les performances du filtrage SEO sur la base de données vidéo de 1600h. Le graphique ci-dessous montre l'évolution de la quantité d'informations à traiter en heure en fonction du score SEO exprimé en k requêtes / mois :

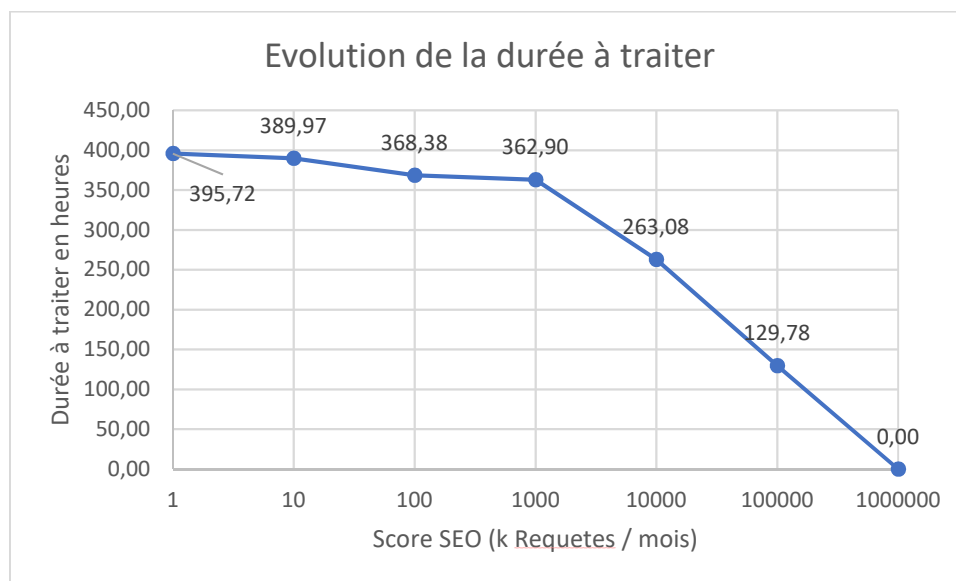


Figure 14 - Graphique d'évolution de la durée totale restante à traiter

On peut se rendre compte à travers ce graphique que sur les 1600h de vidéo, il ne reste que très peu de contenu à traiter en augmentant le seuil SEO. Cette durée de traitement est fortement impactée en fonction du nombre d'émissions répertoriées dans le fichier SEO. De plus, ce fichier contient des émissions toutes chaînes confondues ce qui rend le filtrage d'autant plus important.

7.4. Stratégie de Test

Concernant la stratégie de test, j'ai pu essentiellement tester le filtrage SEO. Ces tests sont répertoriés dans un cahier de test au format Excel, il répertorie le jeu de données utilisé avec les émissions à traiter.

PFE -Détection d'extraits TV										
Table de test - Filtrage score SEO										
Identifiant de test	Doc de réf	ID	Pré-condition	Résultat attendu	Résultat obtenu	Verdict	Sévérité (1:Grave, 2:Très grave, 3:Critique)	Probabilité (1:Peu probable, 2:Probable, 3:Très probable)	Priorité	Commentaire (dans le cas Non-Applicable ou Non-Exécuté)

Figure 15 - Champs du cahier de test

Le tableau répertorie l'identifiant du test, le doc de référence dans lequel est spécifié la fonctionnalité, l'identifiant de la fonctionnalité défini dans le plan de développement, la précondition d'exécution du test, le résultat attendu et obtenu, le verdict (OK, NOK, NA (Non-Applicable)). La sévérité ainsi que la probabilité d'apparition permettent de calculer la priorité du test.



Identifiant de test	Doc de réf	ID	Pré-condition	Résultat attendu
TU01	Plan de développement	4	Le seuil SEO définit est de 0 et la durée d'émission minimum est de 0 min	Le nombre d'émissions conservées est de 15
TU02	Plan de développement	4	Le seuil SEO définit est de 1200000 et la durée d'émission minimum est de 0 min	Le nombre d'émissions conservées est de 0
TU03	Plan de développement	4	Le seuil SEO définit est de 10000 et la durée d'émission minimum est de 0 min	Le nombre d'émissions conservées est de 12
TU04	Plan de développement	4	Le seuil SEO définit est de 0 et la durée d'émission minimum est de 60 min	Le nombre d'émissions conservées est de 7

Figure 16 - Exemple de test pour le filtrage SEO

	A	B	C	D	E	F	G	H
	keyword	isAttributed	Reference	isOptimal	ReferenceTraffic	CalculatedTraffic	CsvIsValid	StepNumber
2	H	Yes	zidane	Yes	275000	1122526	No	2
3	Autour de	Yes	google	Yes	19200000	1029626	No	2
4	Maison à vendre	Yes	zidane	Yes	275000	984441	No	2
5	Malcolm	Yes	doctolib mon compte	Yes	212000	246989	No	2
6	Boris	Yes	zidane	Yes	275000	243404	No	2
7	Urgences	Yes	zidane	Yes	275000	231994	No	2
8	Bones	Yes	doctolib mon compte	Yes	212000	140453	No	2
9	Kiri	Yes	avengers endgame	Yes	83500	122473	No	2
10	Sissi	Yes	tiktok	Yes	1410500	121735	No	2

Figure 17 - Jeu de données utilisé pour les tests

Le jeu de données est répertorié dans une autre page de l'Excel avec des enregistrements utilisés dans le programme.

8. Etat du projet à date

8.1. Fonctionnalités du projet

Le projet à date est fonctionnel mais certaines fonctionnalités n'ont pas pu être implémentées à temps. C'est le cas pour l'implémentation de la détection des mots-clés qui n'a pas pu être réalisé.

L'utilisateur peut néanmoins charger une base de données vidéo et audio avec un score SEO attribué à chaque émission TV. L'application fera donc automatiquement le tri en fonction de ce critère. La détection des changements de scènes se fait aussi automatiquement pour chacune des émissions à tester. Il reste aussi la transcription des scènes à faire grâce à une couche logicielle déjà existante dans ce projet.

Un indicateur disponible est l'évolution de l'écart de charge ci-dessous :

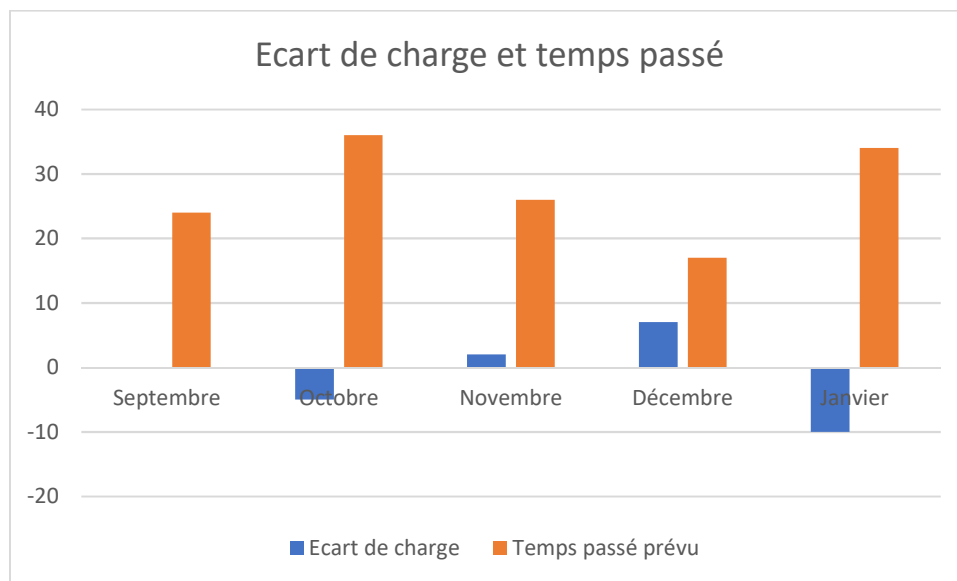


Figure 18 - Ecart de charge au cours du projet

On remarque que l'écart de charge augmente drastiquement entre novembre et décembre, ce qui s'explique par l'implémentation du « shot detection » qui m'a pris plus de temps qu'initialement prévu. A l'inverse, l'écart de charges diminue en janvier à cause d'une des fonctionnalités non implémentées.

8.2. Fonctionnalités restantes à implémenter

Ce qui n'a pas pu être réalisé durant ce projet est représenté dans le schéma suivant, il concerne toute la partie optionnelle du projet concernant l'extraction des mots-clés des scènes :

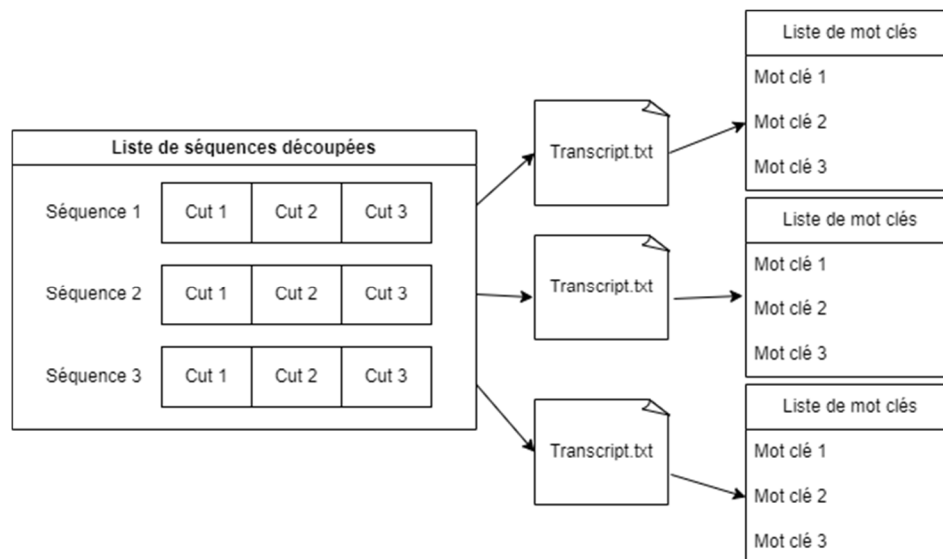


Figure 19 - Fonctionnalités non-implémentées du projet

Comme expliqué dans le cahier de reprise, toute la partie transcription des scènes et extraction des mots-clés n'a pas pu être traitée dans ce projet. Néanmoins des pistes ont été données dans le cahier de reprise pour pouvoir traiter cette partie pour de futurs projets comme les librairies à utiliser permettant l'implémentation du TF-IDF ainsi que d'autres librairies plus performantes et plus modernes pour traiter la transcription de toutes les scènes extraites. Car en effet, le traitement de l'ensemble des scènes est très long (On compte une centaine de scènes sur des émissions de 1h en fonction de la durée des scènes choisies). D'où l'intérêt de paralléliser ce traitement si l'application venait à être « industrialisée ».

Une deuxième problématique à traiter qui sortait du périmètre de ce projet serait la pertinence des scènes extraites. En effet de nombreuses scènes extraites n'ont pas vraiment d'intérêt, comme des scènes sans dialogues ou des dialogues de personnages qui n'ont pas de rapport direct avec l'histoire dans le cas de série TV. Ces scènes pourraient être traitées pour pouvoir les supprimer et en conséquence réduire la masse de données générée par l'algorithme.

9. Bilan

9.1 Bilan de réalisation

Les exigences du MOA concernant le fonctionnement de l'application ont été globalement satisfaites. Un comparatif l'évolution du pourcentage de réalisation dans le temps est disponible dans le graphique ci-dessous représentant le pourcentage de réalisation du projet :

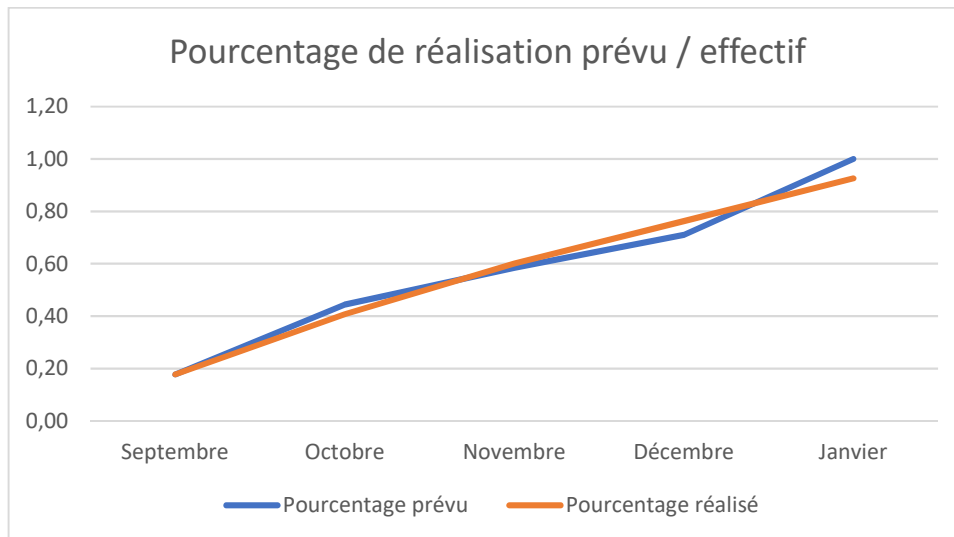


Figure 20 - Evolution du pourcentage de réalisation du projet

On peut voir que le pourcentage de réalisation atteint est d'environ 93%, cela s'explique par la non-implémentation des transcriptions de scènes et de l'absence de la détection des mots-clés. L'évolution de la courbe reste néanmoins constante tout au long du projet. Dans l'ensemble je suis plutôt satisfait de l'évolution globale du projet.

En matière de coût du projet, le graphique ci-dessous représente la différence de coût prévu et effectif à l'issue du projet :

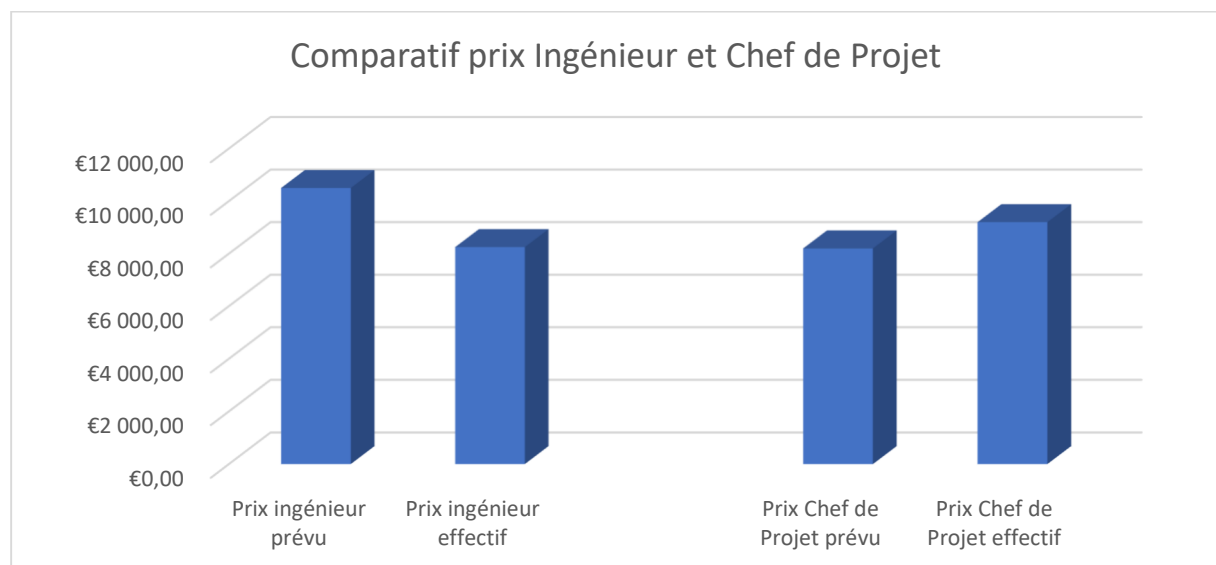


Figure 21 - Comparatif du budget prévu par rapport au budget effectif

On remarque que le prix ingénieur prévu est supérieur au prix effectif car une des fonctionnalités n'a pas été implémentée. Le budget prévu alloué au Chef de Projet est quant à lui inférieur à celui effectif car j'ai passé plus de temps que prévu à rédiger l'ensemble de la documentation.

L'ensemble des calculs est disponible dans le S.P.E.R du projet.

9.2. Gantt réel

Le Gantt réel à l'issue du projet est disponible ci-dessous :

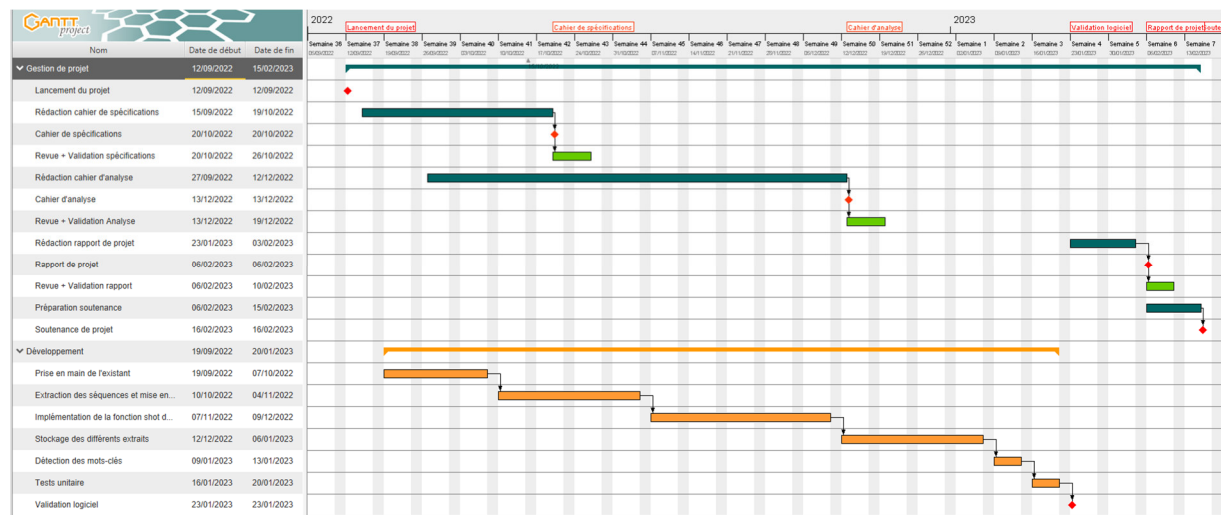


Figure 22 - Gantt réel du PFE

Le Gantt réel est sensiblement le même que le Gantt prévisionnel. Une des seules différences notables par rapport au planning prévisionnel se situe sur la durée de la tâche « Implémentation du shot detection ». Etant donné que toutes les tâches de réalisation sont interdépendantes entre elles, les tâches venant après une tâche retardée seront fortement impactées.

9.2. Difficultés rencontrées

La principale difficulté rencontrée durant ce projet a été la prise en main de l'existant car ce projet est issu de plusieurs projets réalisés par différentes personnes. La coordination de toutes les couches logicielles n'a pas été facile à décoder. De plus, j'ai rencontré quelques difficultés à rédiger l'ensemble des documentations exigées pour ce projet, en effet, j'ai fortement sous-estimé le temps de rédaction des différents documents. Néanmoins j'ai pu apprendre à mieux les formaliser.

9.3. Conclusion

Pour conclure sur ce projet de fin d'étude, cette expérience m'a permis de me rendre compte du travail d'un ingénieur en intégrant plus de gestion de projet au sein de celui-ci. Malgré le fait que le projet n'est pas arrivé complètement à son terme. Gérer un projet seul contrairement aux derniers projets de 3A et 4A m'a permis d'acquérir une meilleure constance dans mon travail.

L'écriture des différentes documentations m'a aussi permis de m'améliorer dans le domaine de la formalisation de documents. Les compétences que j'ai acquises m'ont permis de mener à bien ce projet et à en acquérir de nouvelles dans le développement logiciel en Python. En travaillant seul sur ce projet j'ai été en mesure de gérer tous les aspects du projet, de la planification à la mise en œuvre, en passant par les tests et la documentation. Je suis convaincu que les leçons apprises au cours de ce projet me seront utiles pour mes futurs projets professionnels.

Si je devais changer quelque chose dans la manière d'organiser le projet, j'aurais alloué plus de temps à la rédaction des documentations et à les rédiger plus tôt. Mis à part ce point, je suis plutôt satisfait de la manière dont j'ai géré ce projet de bout en bout concernant la répartition de ma charge de travail dans le temps. D'un point de vue plus technique, j'aurais voulu remettre plus en cause ce qui a été fait dans les projets précédents concernant la performance de certaines librairies. Il s'agit d'un point sur lequel j'aimerais m'améliorer pour mes projets futurs.

Table des illustrations

Figure 1 - Livrables de référence du rapport.....	4
Figure 2 - Logo LIFAT	5
Figure 3 - Description du cycle en V	6
Figure 4 - Planning prévisionnel	6
Figure 5 - Description des données issues de la capture	8
Figure 6 - Modèle de fonctionnement demandé	8
Figure 7 - Extrait d'un fichier xmltv	9
Figure 8 - Détection de scènes	9
Figure 9 - Description des livrables du projet.....	10
Figure 10 - Fonctionnalités implémentées.....	11
Figure 11 - Détecteurs de scènes disponibles	11
Figure 12 - Graphique d'évolution du nombre de scènes pour le Content Detector	12
Figure 13 - Graphique d'évolution du nombre de scènes pour le Adaptive Detector	12
Figure 14 - Graphique d'évolution de la durée totale restante à traiter	13
Figure 15 - Champs du cahier de test.....	13
Figure 16 - Exemple de test pour le filtrage SEO	14
Figure 17 - Jeu de données utilisé pour les tests	14
Figure 18 - Ecart de charge au cours du projet	15
Figure 19 - Fonctionnalités non-implémentées du projet	16
Figure 20 - Evolution du pourcentage de réalisation du projet	17
Figure 21 - Comparatif du budget prévu par rapport au budget effectif.....	17
Figure 22 - Gantt réel du PFE.....	18