

PROJET STATION TV



PLAN DE DEVELOPPEMENT

SUIVI DES MODIFICATIONS DU DOCUMENT			
Version	Date	Validation	Commentaire
1.0	18/10/2025	Frédéric Chauvin	Rédaction initiale

REDACTEUR : Dorian BRISSON	VALIDATEUR : Frédéric CHAUVIN
CLIENT : Polytech Tours	ENCADRANT : Mathieu DELALANDRE

OBJET DU DOCUMENT

L'objectif du présent document est de présenter le plan de développement du projet « Transcription audio haute performance sur Station TV », mené dans le cadre du module PRI de 5^e année à Polytech Tours (filière ISIE).

Ce document a pour but de fournir un cadre structuré au déroulement du projet, en décrivant :

- Les objectifs généraux et le contexte scientifique du travail,
- Les ressources matérielles et logicielles mobilisées,
- L'organisation du développement (cycle, sprints, livrables),
- La gestion des risques et le suivi de l'avancement,
- Ainsi que les outils utilisés pour la conception, la mise en œuvre et la documentation.

Le projet s'inscrit dans la continuité du programme Station TV, développé par le Laboratoire LIFAT (équipe RFAI). Il vise à améliorer les performances de transcription audio (Speech-To-Text) sur la station Dell 5820, en exploitant l'outil Whisper (OpenAI) et en optimisant le pipeline de traitement pour le rendre plus rapide, plus stable et plus scalable sur de larges volumes de données audiovisuelles issus de la TNT.

Ce document permettra de clarifier les objectifs, les résultats attendus et les indicateurs de réussite, tout en définissant les étapes nécessaires à la réalisation complète du projet.

Il servira également de référence commune entre l'étudiant, l'encadrant du LIFAT et le client (DSI Polytech), afin d'assurer un pilotage efficace, une traçabilité claire des livrables et une maîtrise des risques.

CONTEXTE

Description du besoin et de l'existant

Le Laboratoire LIFAT (Laboratoire d'Informatique Fondamentale et Appliquée de Tours), à travers son équipe RFAI (Reconnaissance des Formes et Analyse d'Images), développe depuis 2019 un projet de recherche ambitieux intitulé Station TV. Ce projet vise à concevoir une station de calcul parallèle dédiée à la capture, au traitement et à l'analyse automatique de flux télévisuels issus de la TNT. La plateforme regroupe aujourd'hui de multiples services numériques : capture vidéo intelligente, détection de publicités, analyse de contenus, fact-checking, scraping de données TV et reconnaissance sémantique multimodale.

Dans ce contexte, la transcription audio automatique (Speech-To-Text, ou STT) est devenue une brique centrale du système.

Elle consiste à convertir les pistes audios des programmes TV en texte brut, afin de faciliter la recherche sémantique, l'annotation automatique, la génération de résumés ou encore la détection de citations.

Une première version d'un système de transcription a été développée en 2024–2025 par un précédent étudiant (T. Bourdeau), basée sur l'outil Whisper (OpenAI). Cette solution a démontré de bonnes performances qualitatives, mais s'est heurtée à des limites matérielles liées à la capacité mémoire et à la stabilité lors du traitement de modèles de grande taille (*Whisper medium et large*). Ces contraintes empêchent la station de transcrire efficacement des volumes supérieurs à plusieurs centaines d'heures d'audio.

Le besoin principal de ce projet est donc de renforcer la performance et la robustesse du système STT, en agissant à deux niveaux :

- Évolution matérielle : passage de la machine DELL Precision 5820 à une configuration 256 Go RAM ECC, afin de supporter des traitements haute intensité sur CPU multi-thread.
- Optimisation logicielle : amélioration du pipeline Whisper existant (gestion des threads, réduction du “trashing” mémoire, monitoring QoS, batch processing).

À terme, cette évolution doit permettre la transcription à très grande échelle, ouvrant la voie à une exploitation automatisée et continue des données audiovisuelles collectées par la Station TV.

Ressources

Le projet est réalisé par Dorian Brisson, étudiant ingénieur en 5^e année ISIE, en tant que chef de projet et ingénieur d'études.

Le projet est encadré par M. Mathieu Delalandre, maître de conférences au LIFAT, et suivi techniquement par la DSI Polytech Tours pour la partie matérielle.

Les ressources matérielles et logicielles mobilisées sont les suivantes :

Matériel :

- Station DELL Precision 5820 (EPU-2019006288)
 - CPU : Intel Xeon W-2295 (18 cœurs / 36 threads)
 - RAM : 256 Go DDR4 ECC (4 × 64 Go Kingston KTD-PE432/64G)
 - Stockage : 512 Go SSD + 38 To RAID interne + 190 To NAS externe
 - OS : Windows Server 2022 / Ubuntu 22.04
- Station DELL PowerEdge T640 (serveur parallèle de traitement temps réel).
- Antenne et tuners TNT pour la capture de flux audiovisuels.

Logiciel :

- Whisper (OpenAI) – moteur principal de transcription audio.
- Python 3.10 avec bibliothèques : *Torch*, *NumPy*, *Pandas*, *Multiprocessing*.
- FFmpeg – conversion et segmentation des fichiers audio.
- Scripts LIFAT internes : *BasicTestWhisper.py*, *RunBatchWhisper.py*, *ComputeQoS.py*.
- Outils de supervision et monitoring (CPU, RAM, temps de réponse).

Données :

- Base de test *Hello World* (588 heures audio issues de 7 chaînes TNT : France 2, France 3, France 5, Arte, W9, TF1, C8).
- Fichiers audio au format MP3 – 256 kbps / 48 kHz, découpés en lots de 4 à 20 minutes.
- Ensemble de scripts et de résultats de tests précédents (archives du rapport Bourdeau 2025).

L'ensemble du matériel est disponible en salle serveur du LIFAT ou sur la plateforme Station TV, et peut être utilisé à tout moment avec accord du responsable d'infrastructure.

ORGANISATION

Cycle de développement

Le développement du projet Station TV – Transcription audio haute performance repose sur une méthodologie Agile allégée, adaptée à un projet individuel à forte composante technique. Cette approche privilégie la flexibilité et la progression par incrément fonctionnels, en alternant des phases de préparation, de validation documentaire et de développement expérimental.

Chaque sprint correspond à une période de 2 à 3 semaines, permettant de concentrer le travail sur un objectif précis et mesurable.

Le projet a débuté par une phase de préparation et d'installation technique (Sprint 0), suivie d'une phase documentaire (Sprint 1). Les phases de développement logiciel et de tests Whisper débuteront à la suite de la validation des documents officiels (FOP, CDC, SPER), prévue le 28 octobre 2025.

L'ensemble du projet s'étendra jusqu'à janvier 2026, avec un enchaînement logique de sprints espacés de 2 à 3 semaines chacun.

Description des sprints

SPRINT 0 — Préparation, mise en place et validation matérielle (*terminé*)

Période : du 16/09/2025 au 04/10/2025

Objectif : préparer l'environnement matériel et logiciel nécessaire à la suite du projet, et valider la bonne installation des nouvelles ressources sur la station Dell 5820.

Tâches réalisées :

- Installation des 4 barrettes DDR4 ECC (4 × 64 Go Kingston KTD-PE432/64G).
- Mise à jour du BIOS et vérification de la compatibilité Quad-Channel.
- Validation du fonctionnement ECC et des 256 Go reconnus.
- Création d'un dépôt Git
- Relecture des documents existants

Livrables :

- Station Dell 5820 opérationnelle et validée (RAM, BIOS, stabilité).
- Environnement logiciel fonctionnel pour les futurs tests Whisper.

SPRINT 1 — Rédaction et validation documentaire (FOP, CDC, SPER, Plan de développement)

Période : du 07/10/2025 au 28/10/2025

Objectif : produire et faire valider l'ensemble des documents de cadrage nécessaires au lancement officiel du projet.

Tâches prévues :

- Rédaction de la Fiche d'Ouverture de Projet (FOP).
- Rédaction du Cahier des Charges (CDC).
- Rédaction du SPER (Spécification Préliminaire des Exigences et Résultats).
- Finalisation du Plan de Développement (présent document).
- Rédaction SPEC

Livrables attendus :

- Ensemble des documents de cadrage validés au plus tard le 28/10/2025.
- Lancement du développement logiciel à la suite de la validation.

SPRINT 2 — Réappropriation du logiciel Whisper et premiers tests

Période : du 29/10/2025 au 14/11/2025

Objectif : se réapproprier le code existant, l'adapter à la nouvelle configuration matérielle et réaliser les premiers tests unitaires sur un échantillon réduit.

Tâches prévues :

- Étude du code Whisper utilisé dans le rapport 2025.
- Compréhension du pipeline existant et de ses paramètres.
- Test du modèle *medium*.
- Analyse des performances initiales (débit, charge CPU/RAM).
- Documentation des observations et premières pistes d'optimisation.

Livrables attendus :

- Rapport de réappropriation logicielle.
- Première exécution stable du pipeline Whisper sur Dell 5820.
- Graphiques de performance de base (baseline).

SPRINT 3 — Optimisation du pipeline et gestion mémoire avancée

Période : du 17/11/2025 au 05/12/2025

Objectif : améliorer les performances de transcription grâce à la parallélisation, la gestion mémoire optimisée et le multi-process.

Tâches prévues :

- Refonte des scripts Python pour exécution parallèle (module *multiprocessing*).
- Gestion dynamique de la mémoire selon la taille du modèle.
- Tests du modèle *small/medium*.
- Analyse des goulets d'étranglement (CPU, I/O, RAM).
- Enregistrement automatique des métriques QoS.

Livrables attendus :

- Pipeline multi-thread stable et documenté.
- Journal de performances (CPU/RAM, throughput).
- Rapport comparatif avant / après optimisation.

SPRINT 4 — Mise en place des métriques QoS et supervision

Période : du 08/12/2025 au 19/12/2025

Objectif : définir et automatiser les indicateurs de performance du système.

Tâches prévues :

- Création de *ComputeQoS.py* et intégration des mesures (temps, débit, WER).
- Automatisation du suivi CPU/RAM.
- Export CSV et visualisation graphique.

Livrables attendus :

- Tableaux et graphiques QoS
- Rapport d'évaluation intermédiaire.

SPRINT 5 — Tests intensifs sur modèles Whisper large & validation

Période : 05/01 → 23/01 2026

Objectif : valider la stabilité du système et les performances sur les modèles haute qualité.

Tâches principales :

- Exécution complète du modèle *large* sur longues séquences.
- Mesure détaillée CPU/RAM → évaluation des seuils de saturation.
- Analyse du temps total et du débit effectif.
- Ajustement du pipeline (buffers, batch size).
- Rédaction du rapport final + documentation technique.

Livrables :

- Rapport de test haute intensité
- Courbes de stabilité mémoire
- Comparatif des modèles Whisper
- Transcription intégrale
- Rapport de performance global
- Guide d'intégration Station TV

Synthèse globale des sprints

Sprint	Intitulé	Période	Statut
0	Préparation & validation matérielle	16/09 → 04/10	Terminé
1	Documents de cadrage (FOP, CDC, SPER)	07/10 → 28/10	En cours
2	Réappropriation du logiciel Whisper	29/10 → 14/11	À venir
3	Optimisation pipeline et mémoire	17/11 → 05/12	À venir
4	Mise en place des métriques QoS	08/12 → 19/12	À venir
5	Tests Whisper large & validation finale	05/01 → 23/01	À venir

Comité de pilotage du projet

Le suivi du projet Station TV – Transcription audio haute performance est assuré par un comité de pilotage regroupant :

Rôle	Nom	Fonction
Encadrant scientifique	Mathieu DELALANDRE	Maître de conférences – LIFAT (équipe RFAI)
Chef de projet / développeur	Dorian BRISSON	Étudiant 5A ISIE – Polytech Tours

Le comité a pour mission d'assurer :

- La validation progressive des livrables à chaque sprint,
- Le suivi des performances techniques (RAM, CPU, stabilité),
- Le contrôle qualité des documents.

Fréquence des réunions et points de suivi

Les échanges avec l'encadrant se déroulent selon le calendrier suivant :

- **Réunions de suivi** : à la fin de chaque sprint (toutes les 2 à 3 semaines), pour présenter les livrables techniques et les résultats de tests.
- **Points intermédiaires rapides** : par échanges Teams ou mail (état d'avancement, difficultés techniques, résultats partiels).
- **Réunion finale de validation** : prévue à la fin du SPRINT 5 (janvier 2026) pour la présentation et la validation du rapport final de performances.

Les réunions se tiennent de préférence en présentiel dans les locaux de Polytech, ou à distance en cas d'indisponibilité.

Validation des livrables

Chaque sprint donne lieu à la production d'un ou plusieurs livrables, validés par l'encadrant.

Les critères de validation sont les suivants :

- Conformité au cahier des charges,
- Stabilité technique et absence d'erreurs majeures,
- Documentation claire (tests, résultats, conclusions).

La validation se matérialise par un compte rendu de sprint, signé ou validé par mail, précisant :

- Les livrables remis,
- Les écarts éventuels,
- Les ajustements à prévoir pour le sprint suivant.

Suivi administratif et documentation

Le chef de projet (D. Brisson) est responsable de la gestion documentaire et du versionnement des fichiers.

Tous les documents (rapports, scripts, résultats de tests) sont centralisés dans une équipe Teams et sur un Git.

Chaque sprint fera l'objet d'un rapport de synthèse, incluant :

- Les points techniques abordés,
- Validation des résultats,
- Les perspectives d'amélioration.

ANALYSE DES RISQUES

L'analyse des risques a pour objectif d'identifier les événements pouvant compromettre le bon déroulement du projet, d'en estimer la probabilité et l'impact, puis de mettre en place des plans d'atténuation.

Compte tenu du caractère technique et expérimental du projet, plusieurs risques sont liés à la charge mémoire, à la stabilité du pipeline Whisper et à la dépendance à un seul intervenant.

Risques identifiés

N°	Risque identifié	Description synthétique
1	Surcharge mémoire / crash système	Le pipeline Whisper peut saturer la RAM lors du traitement des modèles <i>medium</i> et <i>large</i> , entraînant un plantage.
2	Dépendance à une seule personne	Le projet est mené par un seul étudiant. Toute absence ou surcharge de travail peut stopper temporairement le développement.
3	Instabilité du modèle Whisper	Les grands modèles peuvent produire des erreurs imprévisibles ou des lenteurs dues à des dépendances logicielles mal gérées.
4	Perte ou corruption de données	Risque de perte du code source ou des résultats de tests suite à une mauvaise manipulation ou un crash disque.
5	Panne matérielle (RAM ou station Dell 5820)	Le matériel utilisé est unique et critique. Une panne ou un défaut de compatibilité peut bloquer le projet.
6	Erreur d'interprétation des métriques QoS	Risque de mauvaise analyse des indicateurs de performance (CPU, RAM, WER), faussant les conclusions.
7	Retard sur la validation documentaire	Les documents FOP, CDC et SPER doivent être validés avant le 22 octobre. Un retard décalerait l'ensemble du calendrier.

Corrections / Solutions apportées

Risque 1 — Surcharge mémoire / crash système

Plan d'action :

- Limiter la taille des batches audio lors des tests.
- Prioriser les tests sur *medium* avant *large*.
- Surveiller la consommation RAM en temps réel avec *psutil*.
- Éteindre et redémarrer le service avant chaque exécution longue.

Risque 2 — Dépendance à une seule personne

Plan d'action :

- Documenter chaque étape et stocker les scripts sur Teams/Git.
- Maintenir un journal d'avancement (log de sprint).
- Organiser des points réguliers avec l'encadrant pour transférer les connaissances.

Risque 3 — Instabilité du modèle Whisper

Plan d'action :

- Tester les modèles un par un avant le lancement batch.
- Réinstaller Torch / Whisper en environnement isolé (`virtualenv`).
- En cas d'instabilité, fallback sur le modèle *medium*.

Risque 4 — Perte ou corruption de données

Plan d'action :

- Sauvegarde automatique quotidienne du code et des résultats sur Teams/Git.
- Export régulier des CSV et rapports vers un support externe (NAS LIFAT).

Risque 5 — Panne matérielle (RAM ou station Dell 5820)

Plan d'action :

- Minimiser les manipulations physiques sur la machine.
- Tenir un contact direct avec la DSI en cas d'anomalie matérielle.

Risque 6 — Erreur d'interprétation des métriques QoS

Plan d'action :

- Définir clairement les indicateurs (temps, débit, WER, charge CPU).
- Vérifier la cohérence des résultats sur plusieurs runs.
- Soumettre les analyses à validation par l'encadrant avant interprétation finale.

Risque 7 — Retard sur la validation documentaire

Plan d'action :

- Fixer les dates butoirs de relecture (FOP, CDC, SPER) avant le 28/10/2025.
- Réserver un créneau hebdomadaire pour relecture / validation encadrant.
- Geler le périmètre documentaire à la date prévue pour éviter les retards.

5.3 Indicateurs de risques

L'évaluation de chaque risque est réalisée selon la matrice de criticité 5×5, croisant la probabilité et l'impact :

Impact / Probabilité	Très improbable (1)	Peu probable (2)	Possible (3)	Probable (4)	Très probable (5)
Insignifiant (1)	1	2	3	4	5
Mineur (2)	2	4	6	8	10
Modéré (3)	3	6	9	12	15
Majeur (4)	4	8	12	16	20
Catastrophique (5)	5	10	15	20	25

Classification des risques

N°	Risque	Probabilité	Impact	Criticité	Niveau
1	Surcharge mémoire / crash système	4	4	16	Élevé
2	Dépendance à une seule personne	3	3	9	Moyen
3	Instabilité du modèle Whisper	3	4	12	Moyen
4	Perte ou corruption de données	2	4	8	Modéré
5	Panne matérielle (RAM / Dell 5820)	2	5	10	Moyen
6	Erreur d'interprétation QoS	2	3	6	Modéré
7	Retard de validation documentaire	2	2	4	Faible

Indicateurs et suivi

Pour assurer le pilotage et la maîtrise des risques, les indicateurs suivants seront suivis mensuellement :

Indicateur	Unité	Seuil d'alerte	Mode de suivi
Charge mémoire moyenne	% RAM utilisée	> 90 %	Mesure automatique (script psutil)
Temps d'exécution moyen	min/h audio	+20 % du seuil cible	Rapport QoS
Version des outils logiciels	numéro	Obsolète	Audit mensuel environnement
Sauvegardes automatiques	fréquence	> 7 jours	Teams / Git
Avancement documentaire	% validé	< 80 %	Suivi par encadrant

L'objectif de ce suivi est de réduire la criticité globale du projet à un niveau acceptable avant la phase finale (Sprint 5).

OUTILS

Les outils présentés ci-dessous couvrent l'ensemble du cycle de développement du projet **Station TV – Transcription audio haute performance**. Ils permettent de gérer la conception, le développement, la planification et la documentation du projet dans un environnement cohérent, stable et reproductible.

Les outils retenus ont été choisis en fonction de :

- Leur compatibilité avec la station Dell 5820 (Windows Server / Ubuntu 22.04),
- Leur capacité à traiter des projets de grande échelle en IA / traitement du signal.

Conception

La conception du projet ne concerne pas ici la modélisation d'éléments mécaniques, mais la structure logicielle et l'architecture du pipeline de transcription.

Les outils utilisés pour la conception et l'organisation du projet sont :

- **GitHub Projects** → visualisation du flux de tâches et backlog des sprints.
- **Markdown** → rédaction de la documentation technique interne (fichiers README.md et notes de sprint).

L'ensemble des schémas et diagrammes est versionné avec le code source afin d'assurer la traçabilité entre la conception et le développement.

Développement

Le cœur du projet repose sur la mise en place, l'optimisation et le suivi du pipeline de transcription Whisper (OpenAI).

Environnement de développement

- **Langage principal** : Python 3.10
- **IDE** : Visual Studio Code (extension Python, Jupyter, GitLens)
- **Système d'exploitation** : Windows Server 2022 et Ubuntu 22.04

Bibliothèques et dépendances

- **Whisper (OpenAI)** – moteur principal de transcription audio.
- **Torch (PyTorch)** – calculs tensoriels et chargement des modèles.

- **Pandas / NumPy** – traitement des données et indicateurs QoS.
- **Matplotlib / Seaborn** – visualisation des métriques de performance.
- **psutil** – monitoring CPU/RAM en temps réel.
- **Multiprocessing / Threading** – parallélisation du pipeline de transcription.

Scripts principaux

- BasicTestWhisper.py : test de base des modèles Whisper.
- RunBatchWhisper.py : exécution multi-thread.
- ComputeQoS.py : calcul et export automatique des métriques QoS.

Chaque script est versionné et documenté dans le dépôt Git.

Un journal de modifications (CHANGELOG.md) suit les évolutions des fonctions et paramètres.

Planification et documentation

La planification et la documentation suivent une logique Agile simplifiée avec des points fixes de validation à chaque sprint.

Outils de planification

- **GanttProject** : gestion du calendrier des sprints et du diagramme de Gantt global.
- **GitHub Projects** : suivi des tâches par sprint (Do / Doing / Done).
- **Teams** : planification des réunions et partages de documents avec l'encadrant.

Outils de documentation

- **Microsoft Word / Excel** : rédaction du Cahier des Charges, Plan de Développement, SPER, FOP et tableaux de performance.
- **Markdown / PDF** : documentation technique interne et archivage des versions finales.
- **Teams** : espace partagé pour dépôt des documents officiels et rapports de sprint.

L'ensemble des documents est centralisé, daté et versionné pour garantir la traçabilité complète du projet.

FICHE D'OUVERTURE DE PROJET

La Fiche d'Ouverture de Projet (FOP) associée à ce document a été rédigée et validée dans le cadre du module PRI 5A ISIE – année 2025/2026. Elle formalise les objectifs généraux, le budget prévisionnel et les engagements de réalisation du projet.

Cette fiche constitue le document de référence pour la planification administrative et la validation du cadrage technique du projet Station TV – Transcription audio haute performance.

Client : M. DELALANDRE

Projet : Station TV – Transcription audio haute performance

Membre de l'équipe :

- Chef de projet : M. BRISSON

Prestations intellectuelles

Rôle/Nom	Prix/heures (€)	Quantité (H)	Total (€)
Dorian - Chef de projet	100 €	240	24000
Total : 24 000,00 €			

Autres chiffres d'affaires

Désignation	Coût Unit. (€)	Qté	Total (€)
Extension RAM DELL 5820 (4 × 64 Go DDR4 ECC Kingston)	326,25	4	1 305,00
Matériel de test et stockage (disques, câbles, adaptateurs)	150,00	1	150,00
Sous-total			1 455,00 €
Plate-forme informatique	<u>S.O</u>	5%	1 200,00 €
Total ACA			2 655,00 €
Analyse des risques	<u>S.O</u>	10%	2 400,00 €
Total budget interne			26 655,00 €
Marge prévisionnelle (25%)			6 663,75 €
Prix de vente			35 718,75 €

PLANNING PRÉVISIONNEL

