



Rapport

Projet tutoré de 4ème année

CPTS - IA Accès aux soins

Tutrice : Imen MEGDICHE BOUSARSAR

Client : CPTS ECEGEC

Maître d'ouvrage : Léa FOURNIER

FIE4 - ISIS École d'Ingénieurs, partenaire INSA
BEAUMELLE Valentin - BODIN Simon - DORE Mouhamed - GARCIA PEREZ Julien -
LEQUEUX Lisa - ORCIVAL Julie

TABLE DES MATIÈRES

Rapport.....	0
TABLE DES MATIÈRES.....	1
Liste des tables.....	2
Table des figures.....	2
Remerciements.....	3
Mots clés.....	4
Résumé.....	4
Abstract.....	4
Présentation Générale de l'entreprise.....	5
Abréviations.....	6
Glossaire.....	7
I. Présentation du sujet.....	1
A. Introduction.....	1
B. Contexte.....	1
C. Etat de l'art.....	1
II. Analyse des besoins.....	5
A. Bête à corne.....	5
B. SWOT.....	6
C. WBS.....	6
D. Analyse des risques.....	7
III. Cadrage du projet.....	8
A. Benchmark.....	8
B. Propositions de solution.....	9
IV. Structuration de la solution proposée.....	10
V. Choix des technologies.....	11
VI. Présentation des résultats.....	12
A. Benchmark des modèles IA testés.....	12
B. Présentation de l'application.....	14
VII. Gestion de projet.....	15
A. Organisation & analyse.....	15
B. Interaction avec les parties prenantes.....	17
Conclusion.....	18
Perspectives.....	18
Bibliographie.....	19
Annexes.....	20

Liste des tables

Tableau 1 : Analyse SWOT.....	6
Tableau 2 : Analyse des risques & probabilité d'occurrence.....	7
Tableau 3 : Benchmark des modèles de transcription audio.....	8
Tableau 4 : Benchmark des modèles de transcription audio.....	11

Table des figures

Figure 1 : Cartographie des CPTS.....	5
Figure 2 : Territoire de la CPTS ECEGEC.....	5
Figure 3 : Bête à corne.....	5
Figure 4 : WBS.....	6
Figure 5 : Matrice de probabilité et d'impact des risques.....	7
Figure 6 : Schéma général de la solution.....	10
Figure 7 : Fonctionnement d'une IA générative STT.....	10
Figure 8 : Comparaison des performances.....	14
Figure 9 : Interface utilisateur de l'application.....	15
Figure 10 : Diagramme de GANTT.....	16
Figure 11 : KANBAN.....	16
Figure 12 : RACI.....	17

Remerciements

Nous tenons à adresser nos remerciements à tous les acteurs ayant contribué à la première partie de ce projet tutoré de quatrième année.

Nous remercions notre école d'ingénieurs ISIS, partenaire INSA, de nous offrir l'opportunité de travailler sur des projets enrichissants, et notamment, Madame **MEGDICHE BOUSARSAR Imen**, notre tutrice, qui a su se rendre disponible pour nous aiguiller et nous conseiller dans le lancement de ce projet.

Nous tenons également à remercier Madame **FOURNIER Léa** et la **CPTS-ECEGEC** de nous avoir confié ce projet et de nous faire confiance dans la suite de sa réalisation.

A ce titre, nous aimerions remercier les différents docteurs qui nous ont offert leurs disponibilités pour nous permettre d'avancer sur ce projet, en particulier le **Docteur MOHR**, ainsi que les **Docteurs MEKHALDI** et **LEVEILLER**.

Ce projet nous permet de mettre en application les connaissances vues au cours de nos enseignements tout en développant des compétences personnelles et interpersonnelles.

Remerciements spéciaux aux relecteurs et correcteurs de ce rapport, permettant l'aboutissement de celui-ci.

Mots clés

**RGPD - Intelligence artificielle - Reconnaissance vocale - Compte rendu
- Consultation de médecine générale**

Résumé

Le projet de la Communauté Professionnelle Territoriale de Santé Entre Cèze et Gardon en Cevennes (CPTS-ECEGEC) en partenariat avec ISIS ingénieurs vise à développer une solution innovante basée sur la **reconnaissance vocale** (Speech-to-Text) adaptée aux besoins spécifiques des professionnels de santé. En nous inspirant de modèles d'**intelligence artificielle** open-source comme Whisper et DeepSpeech, nous concevons un outil performant, personnalisable et conforme aux réglementations telles que le **RGPD**.

Cette solution sera optimisée pour des scénarios complexes et inclura des fonctionnalités avancées, comme la transcription précise de termes médicaux lors d'une **consultation de médecine générale**, permettant ainsi la génération automatique de **comptes rendus**. Nous mettons également l'accent sur une interface utilisateur intuitive et un environnement technique robuste, garantissant un déploiement facile dans les structures de santé.

Abstract

The project of the Communauté Professionnelle Territoriale de Santé Entre Cèze et Gardon en Cevennes (CPTS-ECEGEC) in partnership with ISIS ingénieur aims to develop an innovative solution based on speech-to-text recognition, tailored to the specific needs of healthcare professionals. Drawing inspiration from open-source artificial intelligence models such as Whisper and DeepSpeech, we are designing a high-performance, customizable tool that complies with regulations such as the RGPD.

This solution will be optimized for complex scenarios and will include advanced functionalities, such as the precise transcription of medical terms during a general medical consultation, enabling the automatic generation of reports. We are also focusing on an intuitive user interface and a robust technical environment, guaranteeing easy deployment in healthcare structures.

Présentation Générale de l'entreprise

“Le principe d'une CPTS est de mettre en place un projet de santé établi par l'ensemble de ses professionnels leur permettant de mieux structurer leurs relations entre eux, de mieux se coordonner et d'assurer une plus grande fluidité des parcours de santé de la population d'un même territoire.” (issu du site de la CPTS ECEGEC <https://www.cpts-ecegec.fr/>)

Présentation de l'entreprise

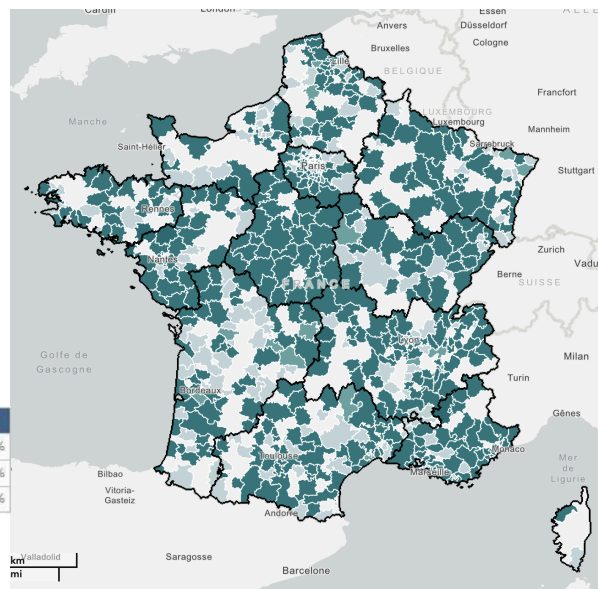
La Communauté Professionnelle Territoriale de Santé Entre Cèze Et Gardon En Cévennes (CPTS ECEGEC) est une organisation dédiée à la coordination des professionnels de santé sur un territoire spécifique. La carte ci-contre représente l'intégralité des CPTS en France. Ces CPTS ont différents statuts : CPTS en fonctionnement, CPTS prochainement en fonctionnement, CPTS en cours de création. Ces organisations sont présentes au niveau national.

	Statut	Nb CPTS	Couverture*
	CPTS en fonctionnement (Contrat ACI signé)	546	66,70%
	CPTS prochainement en fonctionnement (Projet de santé validé)	29	2,52%
	CPTS en cours de création (Lettre d'intention validée)	179	15,78%

Figure 1 : Cartographie des CPTS.

Source :

<https://www.arcois.com/apps/dashboards/6095444b6c8844b1b2ad9ecbb571a631>



Ses missions

La CPTS ECEGEC a diverses missions sur son territoire :

- **Améliorer l'accès aux soins** : Elle s'engage à améliorer la disponibilité et l'accessibilité des services de santé pour les habitants.
- **Coordonner les professionnels de santé** : Elle vise à instaurer des collaborations efficaces entre les différents acteurs de santé pour améliorer la prise en charge des patients dans un souci de continuité et de qualité de soins.
- **Initier des actions territoriales de santé publique** : Elle conçoit et met en œuvre des initiatives de prévention, de dépistage et de promotion de la santé spécifique à la population locale.

Le territoire

La CPTS ECEGEC couvre un territoire de 47 communes regroupant 42 000 habitants et doit faire face à d'importants enjeux de santé identifiés.

En effet, on constate un vieillissement de la population. 40% des habitants ont plus de 60 ans. Ensuite, 12% des patients de 16 ans et plus n'ont pas de médecins traitant. Et parmi les 18 médecins libéraux exerçant sur le territoire, 55% ont plus de 60 ans. Ceci montre une future pénurie de médecins due aux départs en retraite limitant l'accès aux soins.

De plus, il y a une pénurie de certains spécialistes comme les kinésithérapeutes, les sages-femmes et les dentistes. Pour ajouter à cela, ECEGEC est un territoire isolé.

Source : <https://www.cpts-ecegec.fr/>

Figure 2 : Territoire de la CPTS ECEGEC.



Abréviations

API : Application Programming Interface (Interface de programmation)

ASR : Automatic Speech Recognition

CPTS : Communauté Professionnelle Territoriale de Santé

DER : Diarization Error Rate

DPI : Dossier Patient Informatisé

ECEGEC : Entre Cèze Et Gardon En Cévennes

E2E : End-to-End

GAFAM : Google Apple Facebook Amazon Microsoft

GPU : Graphics Processing Unit

HIPAA : Health Insurance Portability and Accountability Act

HDS : Hébergeur de données de santé

IA : Intelligence Artificielle

ISO : International Organization for Standardization

I : Impact

JER : Jaccard Error Rate

KANBAN : Keep Adding New Bottlenecks And Needing improvement

MT : Machine Translation

Po : Probabilité d'occurrence

R : Poids total

RACI : Responsible Accountable Consulted Informed

RGPD : Règlement Général sur la Protection des Données

RTF : Real-Time Factor

SaaS : Software as a Service

SWOT : Strengths Weaknesses Opportunities Threats

STT : Speech-to-Text (Reconnaissance vocale)

WBS : Work Breakdown Structure

WER : Word Error Rate

Glossaire

IA générative : Intelligence artificielle capable de créer du contenu (texte, image, audio) à partir de données existantes.

CPTS ECEGEC : Communauté professionnelle territoriale de santé Entre Cèze et Gardon en Cevennes, structure regroupant les professionnels de santé d'un territoire.

Open-source : Une licence qui permet l'accès, l'utilisation, la modification et la distribution libre du code source d'une solution informatique.

I. Présentation du sujet

A. Introduction

L'intelligence artificielle s'impose aujourd'hui comme l'une des technologies les plus transformatrices. Initialement limitée à des domaines hautement techniques, elle est aujourd'hui présente dans presque tous les aspects de la vie professionnelle et personnelle. Il existe une multitude d'intelligences artificielles, toutes caractérisées par leur rapidité d'exécution, leur précision et leur capacité d'apprentissage et d'adaptation continues. Dans ce contexte, les avancées technologiques en IA se sont accélérées au fil des années, aboutissant à des solutions toujours plus variées.

L'un des secteurs ayant le plus bénéficié du développement de l'intelligence artificielle est le secteur médical. Des avancées significatives ont transformé les pratiques médicales dans plusieurs spécialités comme la radiothérapie et la radiologie, où l'IA est utilisée pour analyser des images médicales plus rapidement et plus précisément. Un autre exemple pourrait être la médecine prédictive, où l'IA est capable de prévoir l'évolution de maladies chroniques en analysant des données cliniques massives.

En somme, l'intelligence artificielle est devenue quasiment omniprésente dans le monde professionnel et particulièrement dans le secteur médical.

B. Contexte

Les Communautés Professionnelles Territoriales de Santé (CPTS) visent à répondre aux besoins de santé des habitants d'un territoire. La CPTS Entre Cèze Et Gardon En Cévennes (CPTS-ECEGEC) s'engage à optimiser le temps médical disponible, notamment en allégeant la charge administrative des médecins généralistes.

Le projet vise à explorer le potentiel d'un outil d'intelligence artificielle (IA) dans le but d'optimiser le temps médical d'une consultation de médecine générale par la retranscription de cette dernière.

L'objectif principal est d'apporter une solution innovante capable d'optimiser la transcription automatique et structurée des échanges entre le patient et le médecin. Cet outil n'interviendrait pas dans le diagnostic ou la prise de décision médicale mais offrirait un gain de temps administratif aux soignants. Ils n'auraient plus à rédiger la consultation, mais simplement à la relire et à la valider. Ceci leur permettra de renforcer le contact humain avec leurs patients tout en réduisant leur fatigue. Le respect de la confidentialité des données de santé est nécessaire pour ce projet.

C. Etat de l'art

Accès au soin & temps médical

La santé des populations relève principalement de l'accès aux soins. L'accès aux soins peut être un facteur d'inégalité sociale (**Pierre Lombrail, 2000**), notamment dans les territoires enclavés et reculés. Apparaît alors la notion d'accessibilité. Le concept d'accessibilité représente une mise à disposition de ressources potentielles (**Pascal Bonnet, 2002**), identifiées ici comme étant l'accès aux soins. Il a été démontré que

l'accessibilité évolue en fonction de l'organisation spatiale et de la typologie des territoires **(Pascal Bonnet, 2002)** Alors, pour un territoire ayant une faible accessibilité, comment favoriser l'accès aux soins?

Il y a plusieurs axes d'action pour favoriser cet accès aux soins : augmenter le nombre de médecins sur le territoire; téléconsultations; ambulances pour conduire les patients de leur domicile aux cabinets de soins; ect. Dans les faits, ces solutions sont non réalisables, non optimales ou coûtent chères. Pour améliorer cet accès aux soins, il faut essayer d'optimiser ce qui existe déjà, notamment en augmentant la capacité de prise en charge des médecins. D'après l'étude scientifique de Marcel Marty, les médecins perdent un temps considérable dans les tâches administratives **(Nicolas Balacheff, 1994)**. En diminuant cette perte de temps, il y aurait un "bénéfice direct pour les patientes et patients" qui serait mieux pris en charge et une "diminution de la charge de travail des médecins" **(Nicolas Balacheff, 1994)**. La dictée et la rédaction des comptes rendus est un axe majeur de perte de temps des médecins. Les nouvelles technologies, comme l'intelligence artificielle, pourraient améliorer ce point.

Intelligence artificielle & santé

L'intelligence artificielle a pour objectif la conception et la réalisation de dispositif informatique capable de simuler des comportements considérés comme intelligents pour un humain **(Nicolas Balacheff, 1994)**. En effet, comme le souligne Balacheff dans **Didactique et intelligence artificielle, Recherches en Didactique des Mathématiques (1994)**, l'intelligence artificielle repose sur sa "*capacité à représenter explicitement des processus de raisonnement ou d'apprentissage*" pour les rendre accessibles. Ce qui est une idée fondamentale dans des applications comme la reconnaissance vocale ou la génération de texte, où l'IA doit être capable de comprendre et de structurer les informations de manière pertinente. Cette définition met donc en évidence la fonctionnalité première de l'intelligence artificielle, qui est de transformer des connaissances humaines en modèles capables de résoudre des problèmes complexes.

D'un autre côté, il est important de noter que "*la conception des systèmes d'IA est étroitement liées aux hypothèses et limites des modèles de connaissances qu'ils intègrent*" **(Nicolas Balacheff, 1994)**. De ce fait, bien que l'intelligence artificielle permette de réaliser des calculs et processus complexes, elle repose encore sur les données et connaissances à partir desquelles on l'alimente. Elle connaît donc des limites, car elle n'a pas accès à ce qui sort du modèle. En santé par exemple, si un cas clinique rare n'est pas présent dans la base de données utilisée pour entraîner le modèle, l'IA pourrait échouer ou fournir une réponse inadéquate.

Comme évoqué précédemment, l'intelligence artificielle joue un rôle de plus en plus important dans le domaine de la santé, offrant des solutions toujours plus innovantes pour améliorer l'accès aux soins, optimiser le temps médical etc... Une IA en santé aide les professionnels de santé à gagner en efficacité et à réduire leur charge cognitive. En effet, l'IA permet avant tout une gestion améliorée des données médicales, tant en termes de précision qu'en rapidité. Comme le souligne l'article **Intelligence artificielle : Une introduction pour les cliniciens (G. Briganti, 2023)**, l'IA est capable d' "*analyser de grandes quantités de données sur les patients et identifier les tendances et les modèles qui peuvent être difficiles à détecter pour les médecins humains*". Elle pourrait donc être une

solution porteuse tant pour *“aider les médecins à gérer leur charge de travail”* que pour contribuer à la recherche scientifique [2].

Intelligence artificielle générative

Différents types d'intelligence artificielle existent pour répondre à divers besoins, tels que *la classification, la détection, la génération de contenu à partir de bases de données, etc.”* (A. Balcerac et al., 2023). L'intelligence artificielle générative permet de créer du contenu, tel que du texte, une image, etc. Pour générer du contenu, ce modèle d'IA se base sur un apprentissage automatique. L'apprentissage automatique est *“basé sur un nombre extrêmement élevé d'exemples, ce qui est très différent de l'approche humaine, qui s'appuie sur l'expérience, le contexte et un système de valeurs.”* (B. Nordlinger et al., 2024).

L'auto-apprentissage se décompose en trois étapes : l'apprentissage à partir des données, la modélisation des probabilités et la génération du contenu. L'apprentissage à partir des données repose sur la compréhension des structures, des patterns et des relations issus de grandes quantités d'informations. La modélisation des probabilités permet de déterminer l'élément le plus susceptible d'apparaître ensuite, qui sera généré par l'étape de génération du contenu. L'intelligence artificielle générative nécessite tout de même une vérification et une validation humaine car le contenu peut être erroné (B. Nordlinger et al., 2024). Dans le domaine de la santé, l'IA ne dispose ni des compétences ni des autorisations médicales. L'Académie nationale de médecine s'est penché sur ces questionnements, elle *“décrit les enjeux d'éthique associés et recommande des points d'actions à mettre en œuvre sans délai.”* (B. Nordlinger et al., 2024).

IA en droit

L'intelligence artificielle a des implications juridiques considérables en matière de protection de données personnelles (encore plus en santé), et de droits fondamentaux. Dans l'article **Les apports du Conseil de l'Europe à une réglementation globale de l'intelligence artificielle (2021)**, il est souligné que *“les GAFAM sont le moteur d'un tout nouvel écosystème, démarrant par la collecte de la donnée jusqu'à diverses offres de services”*. Ce qui montre le rôle central des grandes entreprises technologiques dans la gestion des données, y compris celles liées à la santé, et donc leur rôle dans les défis juridiques liés à la réglementation de l'IA. Cela pose la question de la protection des données de santé et le besoin d'une réglementation stricte pour éviter les risques d'exploitation abusive de ces données.

Le Conseil de l'Europe intervient dans ce contexte, notamment par la convention 108, qui vise à garantir la protection des données personnelles et impose des principes de transparence et de responsabilité. Elle interdit le traitement de ces données sans respect des droits des individus, incluant le droit à l'information, à la rectification et à la transparence.

D'un autre côté, l'IA pose également des questions éthiques sur son utilisation, en particulier en ce qui concerne la transparence des algorithmes. Il est donc important de faire état de la méfiance croissante du public face à ces nouvelles technologies : *“L'inquiétude de la place prise par ces tout nouveaux systèmes automatisés monte donc dans l'opinion publique”* (Yannick Meneceur et Lee Hibbard, 2021). Il est donc important de noter ces défis auxquels les législateurs doivent faire face pour établir un cadre réglementaire qui protège les individus tout en permettant l'innovation technologique.

Speech-to-Text et modèles

La transcription automatique de la parole au texte repose sur des technologies de reconnaissance vocale avancées qui intègrent plusieurs modèles statistiques et linguistiques. Comme le soulignent **A. Hirai et A. Kovalyova** dans l'article **Speech-to-text applications' accuracy in English language learners' speech transcription**, cette reconnaissance vocale repose sur plusieurs étapes clés, notamment *“décomposer le texte parlé en phonèmes et trouver les combinaisons les plus probables de ces phonèmes dans une langue donnée en utilisant un modèle statistique”*.

Le Speech-to-Text vise ainsi à convertir des signaux audio en texte, en s'appuyant sur des modèles sophistiqués d'apprentissage machine. Ces technologies peuvent être divisées en deux approches principales : les systèmes en cascade, et les modèles de bout en bout (E2E : End-to-end). Comme l'explique l'article **Recent Advances in Direct Speech-to-Text Translation (Chen Xu, Rong Ye, 2023)**, *“Les premières solutions pour la traduction vocale consistent à diviser la tâche en sous-tâches plus petites et plus gérables, telles que la reconnaissance vocale automatique (ASR en anglais) et la traduction automatique (MT en anglais)”*. Cette approche, dite approche en cascade, sépare la reconnaissance vocale de la traduction textuelle, ce qui permet une gestion plus précise des étapes. Par contre, cette approche entraîne souvent une latence conséquente et une accumulation d'erreurs entre les modules. D'un autre côté, le modèle de bout en bout offre une solution qui élimine ces étapes intermédiaires. Il a donc l'avantage de réduire la latence et d'améliorer le traitement contextuel. Il pose cependant tout de même des défis en termes de *“cartographie intermodale et interlinguistique (cross-modal & cross-lingual mapping)”* (**C. Xu, R. Ye**), un processus complexe qui demande de combiner de manière précise les données vocales et textuelles.

Comme le mettent en évidence **A. Hirai et A. Kovalyova** dans leur étude, ces systèmes STT (Speech-To-Text), qu'ils soient des systèmes en cascade, des modèles 2E2 ou autre, sont donc composés de plusieurs éléments. Premièrement ils utilisent un modèle dit acoustique, qui contient les exemples de phonèmes et leur fréquence. Deuxièmement, on a le modèle de langage qui répertorie les séquences probables de mots. Ensuite, nous avons un modèle statistique qui analyse les probabilités des enchaînements d'unités linguistiques. Et enfin, des serveurs en cloud se chargent de stocker les données et effectuer les calculs nécessaires.

Dans le domaine de la santé, l'utilisation de ces systèmes d'intelligence artificielle pour la transcription vocale a un impact significatif sur l'efficacité et la rapidité clinique. En automatisant la transcription, ces systèmes remplacent les méthodes manuelles par des solutions rapides et efficaces. En effet, l'intelligence artificielle prend de plus en plus de place dans ces domaines où celle-ci apporte un avantage incontestable, nous pouvons donc affirmer sans trop de risques, que *“Les nouvelles technologies et méthodes de reconnaissance vocale deviendront une partie centrale de la vie future, car elles permettent de gagner beaucoup de temps dans les communications, en remplaçant l'écriture traditionnelle par la voix”* (**N. Shakhovska, O. Basystiuk, K. Shakhovska**). De plus, nous avons déjà fait l'observation de résultats encourageants et satisfaisants de certains Speech-to-Text déjà présents sur le marché. Entre autres, selon l'article **52 A New Way to Address an Old Issue: The Use of Speech-to-Text Dictation to Improve Emergency Medicine Resident Clinical Efficiency**, depuis l'utilisation du logiciel Dragon STT en médecine d'urgence, *“le temps total passé à documenter a diminué de 268,6 à 205,7*

minutes par quart de travail”, et *“le temps par patient et passé de 17,9 à 12,3 minutes”*. L'utilisation d'un logiciel de synthèse vocale peut donc réduire considérablement le temps passé par un soignant à remplir un dossier, tant sur le plan statistique que clinique. Cette capacité à diminuer le temps consacré à la saisie ou à la transcription est particulièrement précieuse dans notre contexte médical, où chaque minute économisée peut être utilisée pour d'autres tâches vitales et urgentes.

Il est cependant également important de noter que les modèles STT rencontrent des limites, notamment par leur incapacité à gérer efficacement les variations de prononciation. En effet, l'une des difficultés majeures rencontrées par les systèmes Speech-to-Text réside dans leur gestion des accents et variations des interlocuteurs. Comme le soulignent **A. Hirai et A. Kovalyova**, *“quand chaque mot n'est pas prononcé de manière intelligible, la transcription STT échoue souvent à produire correctement la suite du discours”*. En d'autres termes, *“la performance des systèmes ASR diminue de manière significative lorsqu'ils sont confrontés à des discours qui divergent de la variété standard”* **Harrington (2023)**. Cette diminution est particulièrement marquée pour les accents moins représentés dans les ensembles de données d'entraînement, ce qui impacte les résultats des systèmes. Par exemple, *“des erreurs de substitution se produisent souvent, où un mot est transcrit par un autre ayant une prononciation proche, mais au détriment du sens global”* **Harrington (2023)**. Dans un cadre médical, il est donc important que ces systèmes soient ajustés pour comprendre le vocabulaire spécifique au domaine, les accents variés des interlocuteurs etc.

II. Analyse des besoins

L'analyse des besoins est une étape indispensable pour assurer la cohérence, l'efficacité et la pertinence du projet. Elle sert à clarifier les attentes en définissant précisément ce que la solution doit accomplir : les fonctionnalités principales pour limiter les risques d'échec, les objectifs visés et les utilisateurs concernés.

A. Bête à corne

La bête à corne est un schéma permettant de répondre à la question “Pourquoi ce projet existe-t-il?”. Elle met en évidence les besoins spécifiques qui doivent être satisfaits. D'après le schéma ci-contre, les utilisateurs finaux sont directement identifiables : les médecins. Cela signifie que notre solution doit être dirigée et adaptée aux médecins. Cette bête à corne indique que la solution doit agir sur le temps médical. Les parties bleues de ce graphique présentent les utilisateurs finaux et le levier d'action du projet. Ensuite, la partie verte décrit le support de la solution : une application d'intelligence artificielle générative speech-to-text (STT).

La partie rouge du schéma explique le problème à résoudre. Cette intelligence artificielle générative a pour objectif de rédiger des comptes-rendus de consultations médicales.

Grâce à cette bête à corne, les besoins spécifiques du projet sont plus clairs : une solution d'IA générative STT, rédigeant les comptes-rendus de consultation, pour optimiser le temps médical des médecins.

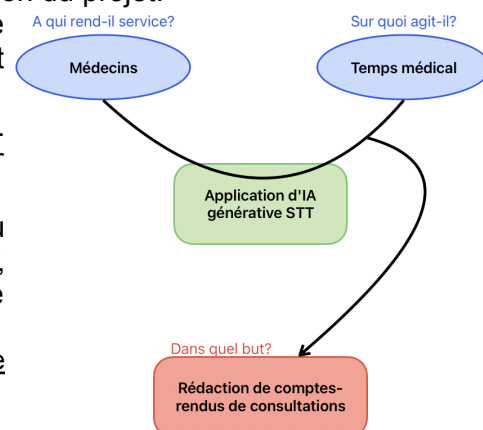


Figure 3 : Bête à corne

B. SWOT

Cette méthode d'analyse des besoins permet de définir les forces, les faiblesses, les opportunités et les menaces de notre projet. Elle guide l'évaluation du contexte, des obstacles et des atouts avant de concrétiser le projet.

Notre SWOT est décrit dans le tableau ci-après. Il est séparé en caractéristiques internes et externes en fonction des soutiens et blocages. Grâce à cette analyse, le projet est parfaitement bien défini ce qui permet de structurer la suite.

	✓	✗
INTERNE	Strengths : <ul style="list-style-type: none"> - Gain de temps médical - Réduction de la fatigue des médecins - Solution innovante - Standardisation des CR 	Weaknesses : <ul style="list-style-type: none"> - Qualité de transcription (erreurs d'interprétations, compréhension) - Dépendance à la technologie et à internet - Adaptation des médecins
EXTERNE	Opportunities : <ul style="list-style-type: none"> - Demande croissante d'outils numériques en santé - Evolution des réglementations - Possibilité de partenariat (intégration aux éditeurs de logiciel médical) - Extensions vers d'autres marchés (hôpital, privé, ...) 	Threats : <ul style="list-style-type: none"> - Risques liés à la confidentialité des données (RGPD) - Acceptation de la part des médecins et patients - Évolution technologique rapide - Maintenance / Matériels

Tableau 1 : Analyse SWOT

C. WBS

L'analyse des besoins permet maintenant de traduire le projet en tâches concrètes via le WBS (Work Breakdown Structure) est une autre méthode d'analyse des besoins permettant de traduire le projet en tâches concrètes, comme un organigramme des tâches.

D'après le WBS suivant, le projet s'articule en 3 grandes parties.

La première permet la rédaction du rapport d'avancement grâce à : l'analyse des besoins, l'état de l'art, le benchmark, la définition de la solution et la gestion de projet.

Ensuite, la seconde partie concerne le développement de la solution : l'analyse du code open source choisi, les modifications de ce code pour l'adapter à notre solution et les tests.

La fin du projet sera marquée par le PV de recette et la soutenance. Cette dernière suppose le rendu de plusieurs livrables comme le diaporama et l'oral, le poster ainsi que le rapport final de projet.

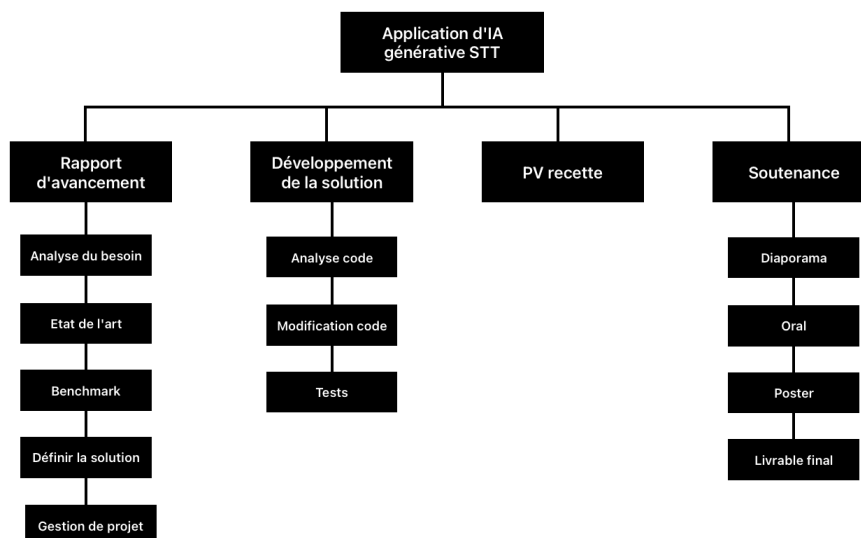


Figure 4 : WBS

Grâce au WBS, l'ensemble du projet est planifié grâce à des tâches principales et secondaires constituant les tâches fonctionnelles et non-fonctionnelles.

D. Analyse des risques

Avant de conclure cette analyse des besoins, une analyse des risques est nécessaire. Elle permet d'identifier les exigences critiques pour ensuite évaluer les risques à l'aide d'une matrice de probabilité. Cela permet d'établir par la suite une gestion des risques.

L'analyse des risques suivante représente un tableau croisé entre celui de l'identification des risques et celui de la probabilité d'occurrence. La probabilité d'occurrence est déterminée à partir de la matrice de probabilité et d'impact.

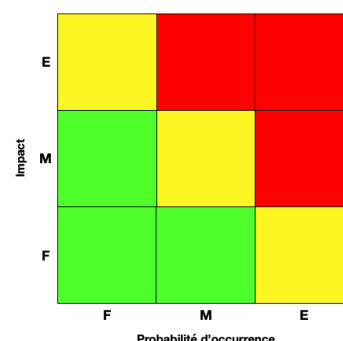


Figure 5 : Matrice de probabilité et d'impact des risques

Numéro	Risque	Probabilité d'occurrence (Po)	Impact (I)	Poids total (R)
Risques scientifiques et techniques				
101	Défaillance de l'application	M	E	ME
102	Mauvaise intégration du code Open Source	M	E	ME
103	Non-respect de tous les objectifs et attentes	F	M	FM
Risques liés aux micros				
201	Dysfonctionnement de l'intégration des données en entrée (sons enregistrés par le micro)	F	E	FE
202	Pas de micro disponible donc pas de réalisation des tests	F	E	FE
Risques humains et organisationnels				
301	Non-respect des délais	F	M	FM
302	Manque de CR pour l'apprentissage de la structuration	F	M	FM
303	Manque de cohésion dans l'équipe	F	M	FM
304	Manque d'entente au sein du groupe	F	M	FM
305	Compréhension difficile du code Open Source	M	E	ME

Tableau 2 : Analyse des risques & probabilité d'occurrence

Le tableau démontre que 3 risques sont à forte probabilité d'occurrence. Il faut donc y porter une attention particulière. Pour éviter la défaillance de l'application, la mauvaise intégration du code Open Source et la compréhension difficile du code, il faut étudier le code afin de comprendre les algorithmes. Cela nécessitera beaucoup de temps mais limitera les risques.

III. Cadrage du projet

A. Benchmark

Un microphone pour capter la conversation est obligatoire pour tous les modèles. Pour toutes les solutions open-source (gratuit), des installations matérielles sont nécessaires. Il faut aussi s'assurer de respecter les RGPD (Règlement Général sur la Protection des Données), les données enregistrées sur les modèles open source ne peuvent pas se retrouver dans la nature. Elles doivent être stockées sur un serveur HDS.

Nom	Payant	Modifications possibles	IA générale / Santé	Stockage des sorties	Conformité RGPD	Facilité d'installation	Environnement de travail nécessaire
Whisper (OpenAI)	Non (open-source)	Oui	IA générale	Fichiers audio, texte, formats JSON	Oui (si hébergement local)	Difficile	Serveur HDS, GPU
DeepSpeech (Mozilla)	Non (open-source)	Oui	IA générale	Fichiers audio, texte, formats JSON	Oui (si hébergement local)	Difficile	Serveur HDS, GPU
Dragon Medical One	Oui (payant, propriétaire)	Non (logiciel fermé)	IA générale	Texte brut ou fichiers structurés	Non (données traitées par Nuance)	Facile (installation standard)	Aucun prérequis matériel spécifiques
Nabla (copilot)	Oui (payant, abonnement)	Non (SaaS, fermé)	IA sante	Données textuelles, EHR compatibles	Oui (spécifiquement conçu pour la santé)	Facile (SaaS, interface web)	Aucun prérequis matériel (service cloud)
Notre modèle	Non	Oui	IA sante si données accessible	Texte brut	Oui	Difficile	Serveur HDS, GPU

Tableau 3 : Benchmark des modèles de transcription audio

Whisper :

Whisper est un modèle de reconnaissance vocale open-source, performant et multilingue développé par Open AI. Il a été entraîné sur une grande variété de données. Il se distingue par sa capacité à reconnaître des accents et à fonctionner dans des environnements variés. L'installation et l'adaptation de Whisper représentent un petit défi technique à relever que notre équipe est prête à relever.

DeepSpeech :

DeepSpeech est un modèle de reconnaissance vocale open-source, offrant une personnalisation similaire à Whisper. Il a été principalement entraîné sur des voix standardisées en anglais, bien qu'une version française soit disponible (avec un tutoriel dédié).

Ce modèle a été conçu à partir de données issues de conversations téléphoniques générales, ce qui limite son adaptation aux besoins spécifiques du secteur de la santé.

Dragon Medical One :

Dragon Medical One est une plateforme de reconnaissance vocale basée sur le cloud, conçue pour les professionnels de santé. Les fonctionnalités de Dragon sont multiples : une reconnaissance vocale précise en temps réel, une application mobile, ainsi qu'une intégration au Dossier Patient Informatisé (DPI). Dragon s'appuie sur des modèles adaptés aux termes médicaux.

Dragon Medical One a l'avantage d'être conforme aux normes RGPD et normes ISO.

Dragon est un logiciel propriétaire, il est donc moins flexible en termes de possibilité.

Nabla (Copilot) :

Nabla Copilot est une solution d'intelligence artificielle développée pour accompagner les professionnels de santé dans leurs activités administratives et cliniques. Elle offre la possibilité de générer, en temps réel pendant les consultations, des comptes rendus médicaux qui peuvent être intégrés directement dans le logiciel médical du praticien.

Comme Dragon, Nabla Copilot est conforme au RGPD ainsi qu'au HIPAA (Health Insurance Portability and Accountability Act).

Notre modèle :

Notre modèle pourrait avoir la caractéristique suivante : transformation d'une conversation en texte.

Le temps de conception, environ trois mois, ne rend pas la tâche possible pour produire un modèle de qualité.

Pour la poursuite de notre benchmark, une étude de la qualité des modèles avec des métriques serait intéressante. Métriques pour comparer les modèles :

- Word Error Rate (WER) : Pour la qualité de transcription.
- DER (Diarization Error Rate) : Pour mesurer les erreurs dans l'identification des locuteurs.
- JER (Jaccard Error Rate) : Pour évaluer la segmentation des locuteurs.
- RTF (Real-Time Factor) : Temps de traitement pour une transcription.
- Robustesse au bruit : Comparer les performances sur des corpus bruités.

Installation DeepSpeech français : voir Annexe 1.

B. Propositions de solution

Nous allons opter pour le fine tuning d'un modèle open source déjà existant comme Whisper ou Deepspeech. Le fine tuning est une technique consistant à spécialiser un modèle d'IA pré-entraîné à l'accomplissement d'une tâche spécifique. Dans ce processus, nous utiliserons comme exemple Whisper, modèle open-source de reconnaissance vocale (Speech-to-Text). Le développement de Whisper repose principalement sur le langage Python, en utilisant des bibliothèques de deep learning telles que PyTorch.

Pour travailler sur ce modèle, nous avons opté pour les environnements de développement VS Code et PyCharm, qui proposent des interfaces appropriées pour gérer les dépendances et explorer le code.

Le prétraitement des fichiers audio (conversion, normalisation, etc.) est réalisé par TorchAudio, tandis que PyTorch assure l'entraînement et l'exécution du modèle.

Afin d'assurer une performance optimale, il est essentiel d'utiliser une GPU puissante, en particulier pour adapter le modèle à des données spécifiques.

Les modifications possibles sont les suivantes :

- Personnalisation des données : Entraînement sur des données spécifiques sur le domaine de la santé en français.
- Fine-tuning : Ajustement des poids du modèle pour des tâches spécialisées.

- Améliorations de l'infrastructure : Intégration avec des APIs ou déploiement dans un service en ligne.
- Résultats produits : Génération d'un compte rendu suite aux conversations entre le praticien de santé et le patient.

IV. Structuration de la solution proposée

Ce schéma illustre le processus de génération automatisée de comptes-rendus médicaux à partir d'une consultation. Les sons captés par un micro, incluant les échanges entre médecins et patients, sont transmis à une IA générative de type Speech-to-Text (STT). Cette IA transcrit les enregistrements audio en texte brut, tout en structurant les informations pour les organiser dans un format cohérent. Enfin, un compte-rendu rédigé et structuré est produit automatiquement, prêt à être utilisé par le médecin pour le dossier médical.

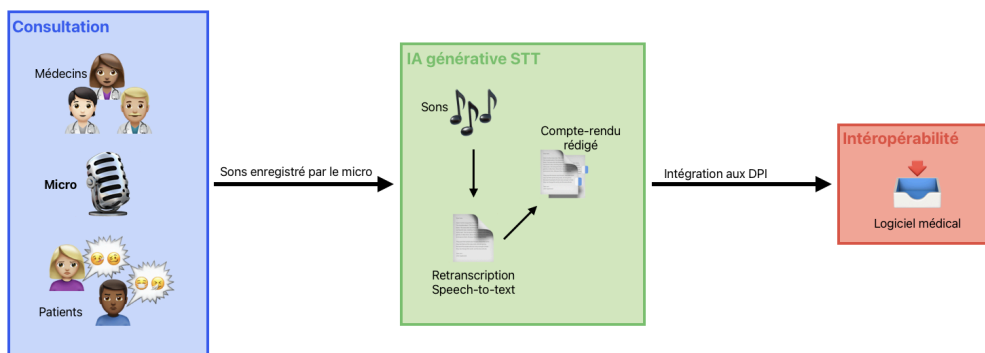


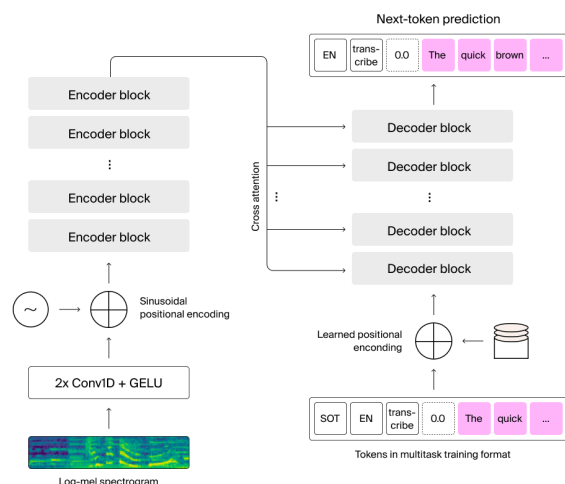
Figure 6 : Schéma général de la solution

Zoomons sur la partie IA générative STT avec comme exemple le modèle d'Open AI. Whisper AI repose sur une architecture de type transformateur, conçue pour le traitement de tâches audio et transcription. Elle se compose de deux principaux modules : un encodeur et un décodeur. L'encodeur convertit les spectrogrammes des signaux audio en une représentation riche et compacte en extrayant des caractéristiques temporelles et spectrales. Ensuite, le décodeur utilise cette représentation pour produire des séquences de texte en exploitant des mécanismes d'attention multi-têtes, permettant de capturer des dépendances globales dans les données. L'ensemble du modèle est pré-entraîné sur une vaste quantité de données multilingues et multitâches, ce qui lui permet de gérer des transcriptions, traductions et autres tâches de reconnaissance vocale dans divers contextes linguistiques.

Figure 7 : Fonctionnement d'une IA générative STT

Source :

<https://www.gladia.io/blog/what-is-openai-whisper>



V. Choix des technologies

Application / Logiciel

Le client a émis le souhait d'avoir une application web. Nous avons conforté son choix car les technologies web sont pertinentes pour un tel projet.

Backend

Framework	Avantages	Inconvénients
Flask (PYTHON)	<ul style="list-style-type: none"> - Langage simple - Léger et flexible - Facile à apprendre et à personnaliser - Langage de référence en termes d'IA 	<ul style="list-style-type: none"> - Nécessite plus de configuration manuelle - Moins de modules intégrés que Django
Spring (JAVA)	<ul style="list-style-type: none"> - Langage connu par toute l'équipe - Autoconfiguration - Déploiement simple 	<ul style="list-style-type: none"> - Langage lourd - Déploiement long
Laravel (PHP)	<ul style="list-style-type: none"> - Création d'application multilingue - Aide de génération URL - Pertinent pour de petit projet 	<ul style="list-style-type: none"> - Langage connu par un seul membre de l'équipe - Manque de continuité entre les versions

Après une étude des avantages et inconvénients des différentes technologies web, nous avons retenu Flask comme framework pour notre projet. Ce choix s'est imposé car Flask est un framework léger, flexible, et parfaitement adapté à la création de petites applications, comme celle que nous développons pour la transcription vocale en texte. De plus, sa simplicité et sa courbe d'apprentissage réduite nous permettent de nous concentrer sur le cœur du projet, à savoir l'intégration des modules d'intelligence artificielle.

Nous avons écarté Laravel, qui bien qu'efficace pour les petites applications, présente un manque de continuité entre les versions, ce qui peut poser problème pour une application évolutive.

Frontend

Framework	Avantages	Inconvénients
Pynecone (PYTHON)	<ul style="list-style-type: none"> - Framework innovant - Performances optimisées 	<ul style="list-style-type: none"> - Framework peu utilisé - Framework non connu de l'équipe - Documentation limité
Angular.js (JS)	<ul style="list-style-type: none"> - Développement rapide - Facilité de maintenance 	<ul style="list-style-type: none"> - Performance pour les petites applications - Framework complexe
Vue.js (JS)	<ul style="list-style-type: none"> - Réactivité des composants - Framework connu par toute l'équipe - Langage rapide et léger 	<ul style="list-style-type: none"> - Faible documentation - Taille de fichier importante

Après une étude des avantages et inconvénients des différentes technologies web, nous avons tout de suite éliminé Pynecone. En effet, ce framework étant très récent et très peu documenté, la formation pour ce framework serait très longue. Angular.js a été très vite écarté. En effet, notre projet consiste à créer une petite application, or, le Framework n'est pertinent que pour les grosses applications. Notre choix s'est porté sur Vue.js.

VI. Présentation des résultats

A. Benchmark des modèles IA testés

Le test s'est fait avec trois modèles open source : Wav2vec (large xlr 53 french), Whisper (large v3) et Seamless (m4t v2 large). Afin de comparer la performance de ces modèles, nous avons retenu deux métriques que nous avons trouvé pertinentes pour ce projet. Il s'agit de Word Error Rate (WER) et RTF (Real-Time Factor).

Word Error Rate (WER)

Word Error Rate (WER) est une métrique utilisée pour évaluer la performance des systèmes de reconnaissance vocale ou de transcription automatique de texte. Elle mesure la quantité d'erreurs dans la transcription produite par rapport à une transcription de référence, c'est-à-dire celle qui est considérée comme correcte ou idéale. WER de 0 % signifie une transcription parfaite (pas d'erreurs). Plus le WER est élevé, plus le système de reconnaissance vocale fait d'erreurs par rapport à la référence.

RTF (Real-Time Factor)

Real-Time Factor (RTF) est une métrique utilisée pour évaluer la vitesse d'exécution d'un système de reconnaissance vocale ou de transcription automatique en temps réel. Il permet de mesurer combien de temps le système prend pour traiter un signal audio par rapport à sa durée réelle. RTF donne ainsi une idée de l'efficacité du système, en particulier lorsqu'il s'agit de transcrire de l'audio en temps réel.

RTF = 1 : Cela signifie que le système prend exactement le même temps pour traiter l'audio que la durée réelle de l'audio. Par exemple, si un fichier audio dure 5 secondes, le système met aussi 5 secondes pour le traiter.

RTF < 1 : Le système est plus rapide que la durée réelle de l'audio, donc il traite l'audio en moins de temps. Par exemple, pour une durée audio de 5 secondes, un système avec un RTF de 0.5 mettrait 2.5 secondes à traiter l'audio.

RTF > 1 : Le système prend plus de temps pour traiter l'audio que la durée réelle de l'audio. Par exemple, si un fichier audio dure 5 secondes, un RTF de 2 signifierait que le système met 10 secondes pour le traiter.

Le résultat des tests se trouve dans le tableau qui suit.

Modèle IA	Moyenne WER	Moyenne RTF
Wav2vec	0,9961609 soit 99% d'erreurs	0,898357189
Whisper	0,1145660747 soit 11% d'erreurs	> 8
Seamless	0,08923759154 soit 8% d'erreurs	1,514579532

Tableau 4 : Benchmark des modèles de transcription audio

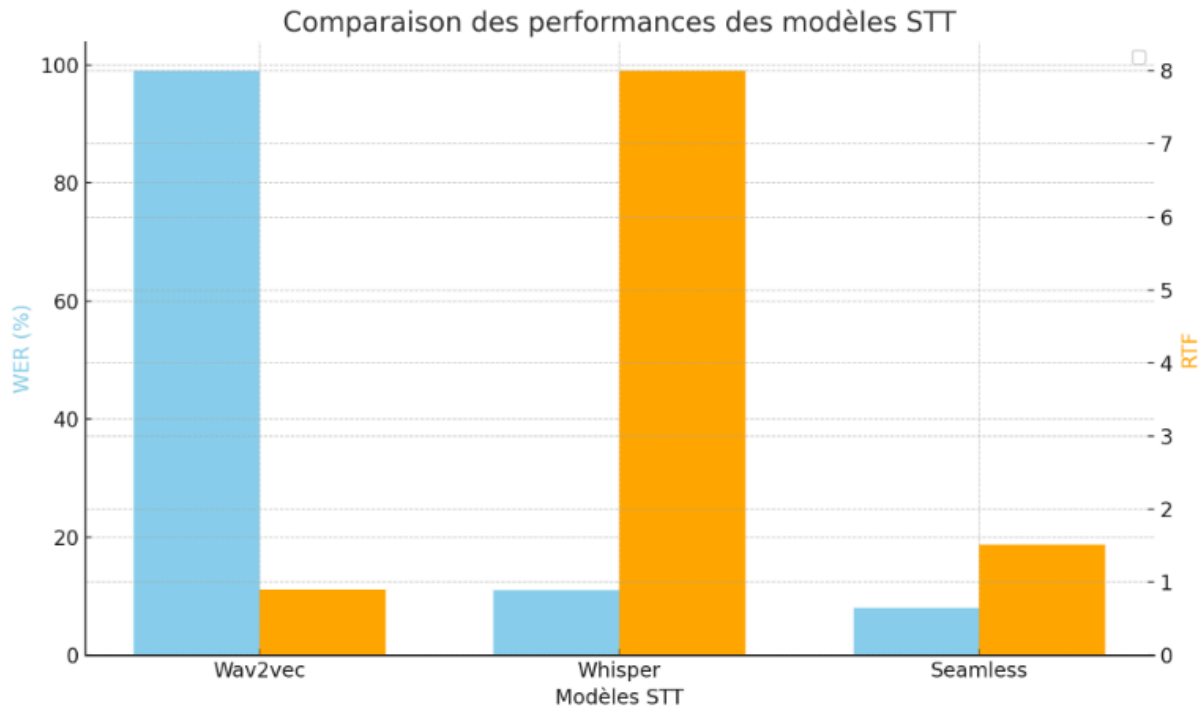


Figure 8 : Comparaison des performances

On voit clairement que :

- Seamless est le plus précis (faible WER) et assez rapide (RTF raisonnable).
- Whisper est assez précis mais très lent.
- Wav2vec est très rapide, mais quasi inutilisable à cause d'un WER très élevé.

Ce texte nous a permis de retenir le modèle Seamless car il a une moyenne d'erreurs inférieure à Whisper et une vitesse de transcription largement supérieure à celle de Whisper.

B. Présentation de l'application

L'application a trois fonctionnalités qui sont : Enregistrer, Transcrire et Télécharger.

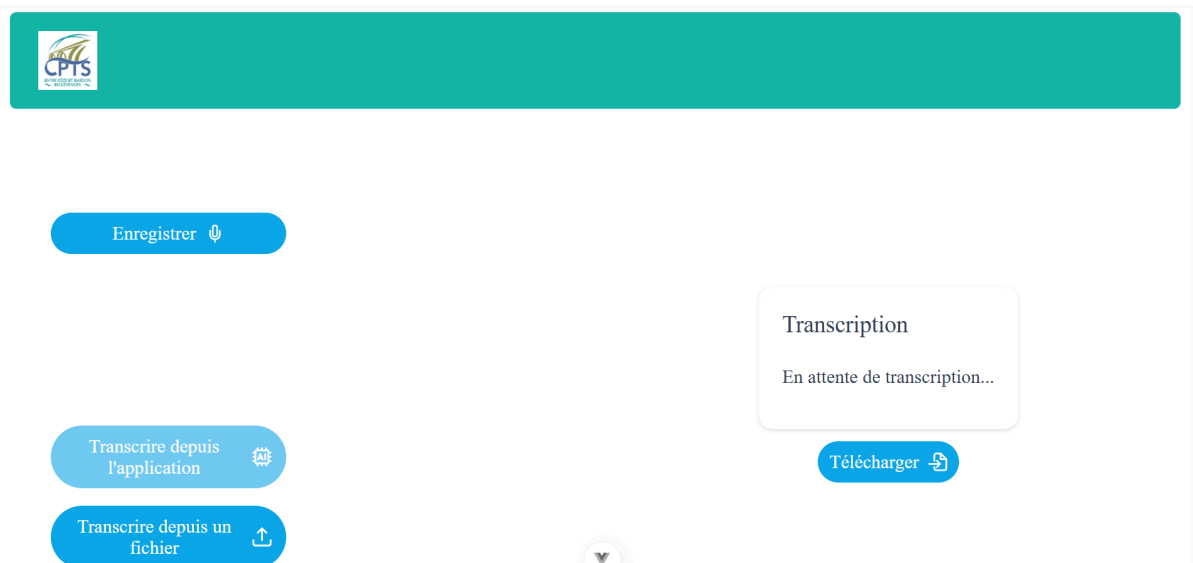


Figure 9 : Interface utilisateur de l'application

La fonctionnalité **Enregistrer** permet au médecin de lancer l'enregistrement lors de la consultation. Après que l'enregistrement soit fait, nous avons la fonctionnalité **Transcrire**. A ce niveau, il y a deux boutons Transcrire. Transcrire depuis l'application qui permet au médecin de lancer la transcription de la conversation qu'il vient d'avoir avec le patient. Le bouton Transcrire depuis un fichier permet de transcrire une conversation enregistrée à partir d'un dictaphone. Et pour finir, la fonctionnalité Télécharger permet de télécharger la conversation transcrite en texte.

VII. Gestion de projet

A. Organisation & analyse

La gestion de projet est essentielle au bon déroulement du projet. Elle consiste à planifier, organiser et suivre les différentes étapes du projet pour atteindre les objectifs définis, en respectant les contraintes de délais, de coûts et de ressources.

L'organisation de notre projet s'est déroulée sur Notion et Google Drive.

Notion est un logiciel permettant de regrouper l'ensemble des ressources en un endroit partagé (**Annexe 2**). Cet espace a été divisé en plusieurs pages pour faciliter notre organisation : contacts, ressources, compte rendu de réunions, to do.

Google Drive a permis de regrouper, partager et rédiger nos livrables (**Annexe 3**). Grâce à sa synchronisation, chaque membre de l'équipe peut accéder aux documents, facilitant ainsi le travail à distance et le suivi des avancées.

Cette organisation a contribué au bon suivi et à une bonne gestion de notre projet.

La première étape a été de déterminer clairement les objectifs et les acteurs du projet pour définir et identifier chaque tâche. Un organigramme de tâches permet de représenter visuellement les tâches à accomplir. Pour chaque tâche, une ressource et une phase de projet sont associées.

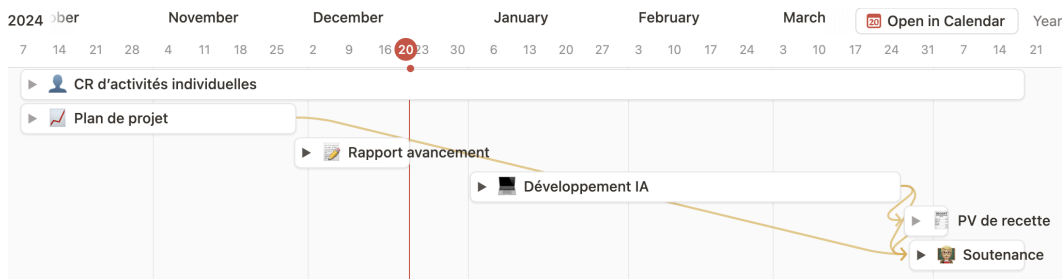


Figure 10 : Diagramme de GANTT

D'après la figure précédente, 3 phases ont été identifiées. La première représente les tâches réalisées au 1er semestre, marquées notamment par le plan de projet et ce rapport d'avancement. La seconde phase regroupe le développement de notre solution et les tests associés. La dernière phase concerne l'ensemble des rendus à réaliser - création d'un poster et rapport de projet - et la soutenance.

Le diagramme de GANTT nous a servi de repère tout au long du premier semestre. Représentant nos tâches et leurs dépendances sur une échelle de temps. Il met en évidence les tâches critiques avec les échéances à respecter.

Pour compléter ce diagramme de GANTT, un tableau KANBAN a été réalisé. C'est une méthode visuelle et flexible représentée en 3 colonnes : non commencé, en cours, fait. Cette méthode permet de suivre le flux des différentes tâches sans imposer de durée d'exécution de tâches. Le KANBAN nous a permis d'avoir une vision globale des tâches à réaliser.

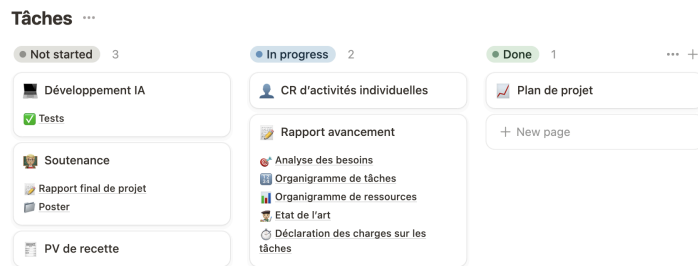


Figure 11 : KANBAN

Le diagramme de GANTT et le KANBAN ont été de vrais repères bénéfiques, notamment pour réaliser et rendre les livrables dans les temps. Sans cette structure, nous aurions eu des difficultés à respecter les dates butoires.

A la suite de cela, un organigramme des ressources a été établi. Nommé RACI, il permet de répartir les ressources nécessaires aux différentes tâches. Certains d'entre nous ont alors été davantage assignés à des tâches techniques, de recherche ou de gestion de projet, en fonction de nos compétences. Chaque activité doit être associée au minimum à : un "responsable" exécutant la tâche, un "approbateur" qui valide les décisions finales, un "consultant" qui donne ses avis et un "informateur" simplement tenu au courant.

Kanban Toutes les tâches **RACI** Gantt Responsables List Phases +

Aa Nom de la tâche	Responsable	Accountable	Consulted	Informed
Plan de projet	julie o Lisa Lequeux	S Simon	C Coordination CPTS-ECEGEC I imen.megdiche@univ-jfc.fr V Valentin Beaumelle	Juliengp M Mouhamed
PV de recette	V Valentin Beaumelle M Mouhamed		C Coordination CPTS-ECEGEC I imen.megdiche@univ-jfc.fr	I imen.megdiche@univ-jfc.fr
Tests Développement IA	M Mouhamed V Valentin Beaumelle		C Coordination CPTS-ECEGEC I imen.megdiche@univ-jfc.fr	
Soutenance	M Mouhamed V Valentin Beaumelle			C Coordination CPTS-ECEGEC I imen.megdiche@univ-jfc.fr
CR d'activités individuelles	julie o Lisa Lequeux M Mouhamed V Valentin Beaumelle Juliengp S Simon		I imen.megdiche@univ-jfc.fr	
Développement IA	V Valentin Beaumelle M Mouhamed		I imen.megdiche@univ-jfc.fr	C Coordination CPTS-ECEGEC
Rapport final de projet Soutenance	M Mouhamed V Valentin Beaumelle	I imen.megdiche@univ-jfc.fr	C Coordination CPTS-ECEGEC	
Etat de l'art Rapport avancement	Lisa Lequeux Juliengp	S Simon	Juliengp julie o M Mouhamed	C Coordination CPTS-ECEGEC I imen.megdiche@univ-jfc.fr
Analyse des besoins Rapport avancement	S Simon M Mouhamed	Lisa Lequeux	julie o Juliengp V Valentin Beaumelle C Coordination CPTS-ECEGEC	I imen.megdiche@univ-jfc.fr
Rapport avancement	julie o Lisa Lequeux	Juliengp S Simon	V Valentin Beaumelle M Mouhamed I imen.megdiche@univ-jfc.fr C Coordination CPTS-ECEGEC	

Figure 12 : RACI

Ce RACI a permis d'équilibrer les ressources en fonction des tâches à réaliser. De la sorte, nous avons tous mis nos compétences à profit pour la réussite de ce projet. Comme chacun des membres de l'équipe avait une tâche assignée, nous avons pu avancer plus rapidement et efficacement, aucune perte de temps sur la répartition des tâches.

Nous pouvons dire que les outils de gestion de projet mis en place ont contribué à la réussite du projet. Pour l'instant, nous n'avons aucun écart par rapport aux plannings. Si nous continuons à entretenir une telle gestion, le projet pourra se dérouler correctement jusqu'à la fin et nous resterons dans les temps.

B. Interaction avec les parties prenantes

Une bonne gestion de projet repose sur une communication claire et efficace entre les parties prenantes et l'équipe projet.

Pour cela, Mme Léa Fournier, commanditaire du projet, a été contactée dès l'attribution de celui-ci. Grâce aux échanges menés, il a été possible de redéfinir les objectifs et les limites du projet. Elle nous a également apporté de nombreuses connaissances qui nous ont été fondamentales dans la construction de ce rapport. Par ailleurs, Mme Fournier nous a transmis de précieux contacts de médecins disposés à participer à ce projet. Notamment, celui du Dr MOHR qui nous a partagé ses recherches sur le sujet. Elles ont participé à la rédaction de notre Benchmark.

En complément, nous avons régulièrement consulté notre tutrice, qui nous a guidés dans le choix du plan de ce rapport d'avancement et nous a aidé à clarifier nos idées lorsque nous faisons face à des doutes.

Conclusion

Ce projet tutoré nous a permis d'explorer les possibilités offertes par l'intelligence artificielle dans un contexte médical, avec pour objectif principal l'optimisation du temps médical via la transcription automatique des consultations. À travers une analyse rigoureuse des besoins, un état de l'art complet, un benchmark technologique pertinent et des choix d'implémentation justifiés, nous avons posé les bases solides d'une solution innovante, adaptable. En développant une application web simple et efficace, intégrant un modèle de transcription vocale comme Seamless, nous avons montré qu'il était possible d'automatiser une partie des tâches administratives des professionnels de santé.

Perspectives

Pour la suite du projet, plusieurs axes d'amélioration peuvent être envisagés :

- Amélioration du modèle IA : l'entraînement du modèle sur un corpus médical francophone vaste et varié permettrait d'augmenter la précision de transcription, notamment sur des termes techniques, accents ou bruits ambiants. Aussi l'ajustement du modèle pour qu'il puisse reconnaître les différents interlocuteurs de la conversation (médecin - patient). Enfin, permettre au modèle de transcrire simultanément (en temps réel) la conversation.
- Déploiement et maintenance : la solution pourrait être proposée sous forme de service web (SaaS) pour faciliter son adoption par d'autres CPTS ou structures de santé confrontées aux mêmes problématiques.

Bibliographie

- Balacheff, Nicolas. 1994. « Didactique et intelligence artificielle ».
- Balcerac, A., B. Tervil, N. Vayatis, et D. Ricard. 2023. « Principes fondamentaux de l'apprentissage automatique pour les neurologues ». *Pratique Neurologique - FMC* 14 (4): 225-36. <https://doi.org/10.1016/j.praneu.2023.10.005>.
- Briganti, G. 2023. « Intelligence artificielle : une introduction pour les cliniciens ». *Revue des Maladies Respiratoires* 40 (4): 308-13. <https://doi.org/10.1016/j.rmr.2023.02.005>.
- Bonnet, Pascal. 2002. « Le concept d'accessibilité et d'accès aux soins ». <https://agritrop.cirad.fr/527450/1/ID527450.pdf>.
- Harrington, Lauren. 2023. « Incorporating Automatic Speech Recognition Methods into the Transcription of Police-Suspect Interviews: Factors Affecting Automatic Performance ». *Frontiers in Communication* 8 (juillet). <https://doi.org/10.3389/fcomm.2023.1165233>.
- Hogrefe, C. P., A. S. Nugent, K. K. Harland, et H. R. House. 2012. « 52 A New Way to Address an Old Issue: The Use of Speech-to-Text Dictation to Improve Emergency Medicine Resident Clinical Efficiency ». *Annals of Emergency Medicine* 60 (5): S179-80. <https://doi.org/10.1016/j.annemergmed.2012.07.075>.
- Iancu, Bogdan. 2019. « Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources ». *Informatica Economica* 23 (1/2019): 17-25. <https://doi.org/10.12948/issn14531305/23.1.2019.02>.
- Lombrail, Pierre. 2000. « 26. Accès aux soins ». In *Les inégalités sociales de santé*, 403-18. La Découverte. <https://doi.org/10.3917/dec.fassi.2000.01.0403>.
- Meneceur, Yannick, et Lee Hibbard. 2021. « Les apports du Conseil de l'Europe à une réglementation globale de l'intelligence artificielle:Revue des instruments juridiques du Conseil de l'Europe relatifs à l'intelligence artificielle et des enjeux particuliers en matière de santé et de biomédecine ». *Droit, Santé et Société* 3 (3): 55-63.
- Nordlinger, Bernard, Claude Kirchner, et Olivier de Fresnoye. 2024. « Rapport 24-03. Systèmes d'IA générative en santé : enjeux et perspectives ». *Bulletin de l'Académie Nationale de Médecine* 208 (5): 536-47. <https://doi.org/10.1016/j.banm.2024.03.005>.
- Saint-Affrique, Diane de. 2022. « Intelligence artificielle et médecine : quelles règles éthiques et juridiques pour une IA responsable ? ». *Médecine & Droit* 2022 (172): 5-7. <https://doi.org/10.1016/j.meddro.2021.09.001>.
- « Speech-to-Text Applications' Accuracy in English Language Learners' Speech Transcription ». s. d. Consulté le 20 décembre 2024. https://www.lltjournal.org/item/10125-73555/?utm_source=chatgpt.com.
- Xu, Chen, Rong Ye, Qianqian Dong, Chengqi Zhao, Tom Ko, Mingxuan Wang, Tong Xiao, et Jingbo Zhu. 2023. « Recent Advances in Direct Speech-to-text Translation ». arXiv. <https://doi.org/10.48550/arXiv.2306.11646>.

Annexes

Liste des Annexes :

Annexe 1 : Lien d'installation DeepSpeech francais.....	ii
Annexe 2 : Lien du Notion.....	ii
Annexe 3 : Lien du Google Drive.....	ii

Annexe 1 : Lien d'installation DeepSpeech francais

<https://linuxfr.org/forums/linux-debian-ubuntu/posts/voix-vers-texte-en-francais-avec-deepspeech-commonvoice-ubuntu-20-04>

Annexe 2 : Lien du Notion

<https://www.notion.so/invite/d9f62ad8a78453ceda8944df9143fb748e7bafa5>.

Annexe 3 : Lien du Google Drive

https://drive.google.com/drive/folders/1WDg6OBW4nm0-aoyb6_a73hPsxNYQme_n?dmr=1&ec=wgc-drive-globalnav-goto