

Utilisation d'un LLM pour la synthèse et l'exploration des ressources médicales

Rapport de stage de 1^{ère} année effectué du 2 juin au 1^{er} août 2025

Sous la direction du Dr Romain GRIFFIER

CHU DE BORDEAUX

12 rue Dubernat
33404 Talence, France

05 56 79 56 79

Adélaïde BROUCAS

ING1 GMF4

01/08/2025

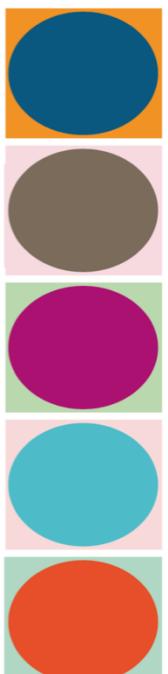


Table des matières

Remerciements	2
Introduction	4
I – Présentation de l’entreprise d’accueil	5
a. Le Centre Hospitalier Universitaire (CHU) de Bordeaux.....	5
b. L’Unité d’Informatique et d’Archivistique Médicales (IAM)	5
c. L’Entrepôt de Données de Santé (EDS) : un outil stratégique	6
d. Lien avec les enjeux contemporains	7
II – Ma mission, mes résultats, mon analyse	7
a. Ma mission.....	7
b. Mes résultats	9
c. Mon analyse.....	12
III – Bilan du stage	13
Conclusion.....	16
Annexes.....	18
Annexe 1 : Lexique.....	18
Annexe 2 : Schéma explicatif du processus d’intégration des données « ETL »	19
Annexe 3 : Schéma explicatif du fonctionnement d’un LLM.....	20
Annexe 4 : Recommandations du CERS pour la rédaction de résumés grand public	20
Annexe 5 : Formule d’évaluation du score de Flesch-Kincaid et tableau de correspondance	21
Annexe 6 : Schémas explicatifs du fonctionnement d’un RAG	22
Annexe 7 : Fiche d’évaluation par le tuteur de l’entreprise	23
Annexe 8 : Attestation de fin de stage	27
Bibliographie	29

Remerciements

Avant tout développement sur cette expérience professionnelle, il me paraît important de commencer ce rapport de stage en exprimant mes remerciements envers celles et ceux qui m'ont beaucoup appris et qui ont permis de faire de ce stage une expérience des plus enrichissante.

En premier lieu, je tiens à remercier le **professeur Rodolphe THIEBAUT**, chef du service d'information médicale du CHU de Bordeaux, pour avoir présenté ma candidature aux membres de l'Unité IAM¹.

J'adresse mes sincères remerciements au **Docteur Moufid HAJJAR**, chef de l'Unité d'Informatique et d'Archivistique Médicales (IAM), pour son accueil chaleureux. Je ne peux rédiger ces remerciements sans adresser une mention particulière au **Docteur Romain GRIFFIER**, mon maître de stage, pour son encadrement attentif, sa gentillesse, sa bonne humeur quotidienne, sa disponibilité et ses conseils méthodologiques. Je remercie également le **Docteur Vianney JOUHET** pour son accompagnement quotidien ainsi que **M. Loïc RONFLETTE**, dont l'expertise en intelligence artificielle et en machine learning m'a beaucoup aidée à progresser et à mieux comprendre les enjeux techniques liés aux modèles de langage.

Je tiens aussi à remercier l'ensemble de mes **collègues de l'Unité IAM**, ainsi que **les membres du COREVIH** (Comité de coordination de la lutte contre les infections sexuellement transmissibles et le virus de l'immunodéficience humaine), pour leur bienveillance et avec qui j'ai eu la chance de partager de nombreux déjeuners et soirées conviviales. Ces moments de détente et de bonne humeur resteront parmi mes plus beaux souvenirs et ont largement contribué à rendre cette expérience aussi agréable que formatrice.

D'autre part, je souhaite exprimer ma gratitude à **l'équipe pédagogique de CY Tech** ainsi que **le pôle des Relations Entreprises de CY Tech** sans qui je n'aurais pas pu réaliser ce stage mais aussi pour son suivi attentif et ses conseils précieux dans la recherche de stage ainsi que la réalisation de ce rapport.

Enfin, j'adresse toute ma reconnaissance à ma famille, et en particulier à **mes parents**, pour leur soutien constant et pour m'avoir transmis dès mon plus jeune âge un intérêt pour le milieu médical. Leur accompagnement m'a donné la motivation et la confiance nécessaires pour mener à bien ce stage. J'adresse également une mention toute particulière à **ma sœur**

¹ Unité d'Informatique et d'Archivistique Médicale du CHU de Bordeaux

Astrid BROUCAS. Ayant elle-même effectué son stage de première année à CY Tech au sein de l'Unité IAM, c'est elle qui m'a encouragée à postuler à mon tour, forte de son expérience très positive. Tout au long de mon stage, elle a su me rassurer, me conseiller et m'accompagner, ce qui a été pour moi une aide précieuse et un véritable soutien moral.

À toutes celles et ceux qui ont contribué, de près ou de loin, à faire de ce stage une expérience professionnelle et personnelle enrichissante, je souhaite vous adresser mes plus sincères remerciements.

Introduction

Au cours de ma première année en cycle ingénieur à CY Tech, le stage en entreprise constitue une étape essentielle de la formation. C'est dans ce cadre que j'ai eu l'opportunité d'intégrer le Centre Hospitalier Universitaire (CHU) de Bordeaux, au sein de l'Unité d'Informatique et d'Archivistique Médicale (IAM) du 2 juin au 1^{er} août 2025 en tant que stagiaire **data engineer** sous la tutelle du Docteur Romain GRIFFIER, Médecin de Santé Publique spécialisé en informatique médicale.

Mon choix s'est naturellement porté sur le domaine de la santé et de l'intelligence artificielle. Ayant grandi dans une famille de médecins, j'ai été sensibilisée très tôt aux enjeux liés aux soins, aux pratiques hospitalières et à l'importance de la recherche médicale. Ce contexte familial a nourri mon intérêt pour les innovations technologiques appliquées au soin, et m'a incitée à chercher un stage où se croisent informatique, données et santé publique. Cet intérêt a été renforcé par le module « Éthique et sciences techniques » suivi à CY Tech, au cours duquel j'ai réalisé un exposé sur les dilemmes éthiques de l'intelligence artificielle en médecine. Mon stage m'a ainsi permis de prolonger cette réflexion académique en observant concrètement la manière dont ces questions trouvent un écho dans les projets hospitaliers.

L'Unité IAM, rattachée au Service d'Information Médicale du CHU, a pour mission principale la gestion de l'Entrepôt de Données de Santé (EDS). Ce dispositif stratégique centralise et valorise les données hospitalières au service de la recherche biomédicale et de l'amélioration des soins. C'est dans ce cadre que mes missions se sont inscrites : expérimenter le déploiement local de modèles de langage (LLM) et évaluer différentes méthodes de vulgarisation scientifique, avant de développer un premier prototype de pipeline Retrieval-Augmented Generation (RAG)².

Ce stage m'a ainsi permis de découvrir le fonctionnement d'un grand établissement hospitalier, d'acquérir de nouvelles compétences techniques et méthodologiques, et de réfléchir aux enjeux éthiques liés à l'usage de l'IA en santé.

Afin de refléter au mieux cette expérience dans le milieu professionnel, il me semble approprié de se pencher en premier lieu sur la place de l'unité IAM au sein du CHU de Bordeaux. Ensuite, nous nous intéresserons aux missions qui m'ont été confiées. Enfin, nous dresserons un bilan de cette expérience personnelle et professionnelle.

² Cf. Annexe 1 : Lexique

I – Présentation de l’entreprise d’accueil

a. Le Centre Hospitalier Universitaire (CHU) de Bordeaux

Le Centre Hospitalier Universitaire (CHU) de Bordeaux est l’un des plus grands établissements hospitaliers de France et constitue un acteur majeur de la santé publique au niveau régional et national. Il exerce une triple mission : assurer les soins aux patients, participer à la formation des futurs professionnels de santé et contribuer activement à la recherche médicale. Avec plusieurs milliers de lits répartis sur trois sites principaux (Pellegrin, Haut-Lévêque, Saint-André), le CHU est reconnu pour son excellence médicale et son rôle de référence dans de nombreuses spécialités. Il constitue également le premier employeur de la région Nouvelle-Aquitaine, avec plus de 14 000 employés.

Au-delà de sa mission sanitaire, l’établissement s’inscrit également dans une démarche responsable à travers une stratégie de développement durable et de sobriété écologique ambitieuse. Consciente de l’impact environnemental des activités hospitalières, la direction a élaboré une feuille de route intégrant la réduction de l’empreinte carbone, l’optimisation énergétique des bâtiments, la gestion durable des déchets, ainsi que la mise en place de projets « unités durables » favorisant l’efficacité des soins tout en limitant l’impact écologique. Le CHU a également lancé des initiatives en faveur d’une sobriété numérique, encourageant par exemple la suppression des mails inutiles et le transfert de dossiers médicaux anciens vers des serveurs moins énergivores. Cette orientation traduit la volonté du CHU de concilier excellence médicale, innovation et responsabilité environnementale.

b. L’Unité d’Informatique et d’Archivistique Médicales (IAM)

Au sein de cet établissement, j’ai intégré l’Unité d’Informatique et d’Archivistique Médicales (IAM). Elle est rattachée au Service d’Information Médicale, au sein du pôle de Santé Publique, et placée sous la direction du Dr Moufid Hajjar. Située sur le site hospitalier de Pellegrin, cette unité se distingue par sa nature hybride, à la fois médicale et technique.

L’équipe est pluridisciplinaire et regroupe des médecins de santé publique spécialisés en informatique médicale, des data scientists, des ingénieurs de données, ainsi que des développeurs et des responsables techniques. Cette diversité de profils permet de répondre

aux enjeux complexes liés à la gestion et à l'exploitation des données de santé. L'Unité IAM accueille également des internes et des doctorants, renforçant ainsi le lien avec la recherche académique et universitaire.

Les missions principales de l'Unité IAM sont les suivantes :

- La gestion de l'Entrepôt de Données de Santé (EDS), cœur stratégique de l'unité,
- La gestion des archives médicales, qu'elles soient papier ou numériques, garantissant la traçabilité et la conservation des données,
- La cellule d'Identito-Vigilance et Données de Santé (CIDS), chargée de veiller à la cohérence et à la fiabilité de l'identification des patients dans les systèmes informatiques hospitaliers.

Cette organisation témoigne d'un rôle clé : assurer la fiabilité, la sécurité et l'exploitation des données de santé, qui constituent aujourd'hui un levier essentiel pour l'amélioration de la qualité des soins et l'innovation médicale.

c. L'Entrepôt de Données de Santé (EDS) : un outil stratégique

L'Entrepôt de Données de Santé (EDS) est au cœur des activités de l'Unité IAM. Il s'agit d'une plateforme technique dont l'objectif est de centraliser, structurer et harmoniser les données issues des différents systèmes d'information hospitaliers, afin de les rendre exploitables pour la recherche, les études épidémiologiques et l'évaluation des pratiques de soins.

Le fonctionnement de l'EDS repose sur un **processus appelé ETL** (Extract – Transform – Load)³ :

- Extraction (*Extract*) : les données sont extraites automatiquement des logiciels-métiers de l'hôpital (dossiers patients informatisés, prescriptions, systèmes de laboratoire, etc...),
- Transformation (*Transform*) : elles sont ensuite nettoyées, standardisées et enrichies pour garantir leur cohérence et leur qualité. Cela inclut l'harmonisation des formats, la correction des erreurs et la fusion des sources multiples,
- Chargement (*Load*) : les données ainsi préparées sont intégrées dans la base finale de l'EDS, qui permet une représentation relationnelle des informations médicales.

³ Cf. Annexe 2

Ce processus garantit que les données, initialement dispersées et hétérogènes, deviennent une ressource fiable et exploitable. L'EDS constitue ainsi un socle unique permettant d'alimenter des projets de recherche clinique, de développer des outils de médecine prédictive et d'explorer de nouvelles applications de l'intelligence artificielle dans le domaine médical.

Au-delà de la technique, l'EDS se distingue par des valeurs fondatrices :

- Servir l'intérêt général dans une logique de service public,
- Favoriser l'open science, en rendant les données accessibles à la communauté scientifique dans un cadre sécurisé,
- Travailler de manière collaborative et transparente, en cultivant l'autonomie et l'innovation au sein de l'équipe.

d. Lien avec les enjeux contemporains

Le rôle de l'Unité IAM et de l'EDS s'inscrit pleinement dans les débats actuels sur l'avenir du numérique en santé. La centralisation et l'exploitation des données médicales soulèvent des questions essentielles de confidentialité, de souveraineté numérique et d'éthique. Le CHU de Bordeaux, en développant un tel outil, cherche à concilier innovation et responsabilité, dans le respect des droits des patients.

Par ailleurs, cette dynamique rejoint la stratégie de développement durable du CHU. L'optimisation des ressources numériques et la valorisation des données permettent d'éviter les redondances, de rationaliser l'organisation des soins et de contribuer à une recherche plus efficace et plus respectueuse de l'environnement.

En définitive, l'Unité IAM du CHU de Bordeaux apparaît comme un acteur clé dans la transformation numérique et éthique de la santé. Par son rôle dans la gestion de l'EDS, elle incarne la volonté d'associer rigueur scientifique, innovation technologique et responsabilité sociale au service des patients et de la recherche biomédicale.

II – Ma mission, mes résultats, mon analyse

a. Ma mission

Mon stage au sein de l'Unité IAM du CHU de Bordeaux avait pour objectif de contribuer à l'exploration des usages de l'intelligence artificielle en santé, à travers deux missions principales. La première consistait à déployer un **modèle de langage de grande taille**⁴(LLM) en environnement local et à tester différentes stratégies de **pré-prompts** pour produire des résumés scientifiques vulgarisés. La seconde visait à développer un premier pipeline de type **Retrieval-Augmented Generation**⁵(RAG), permettant de contextualiser les réponses d'un modèle à partir de documents médicaux internes.

Ces deux missions s'inscrivaient dans une dynamique exploratoire : l'Unité IAM cherchait à évaluer concrètement comment les modèles de langage pouvaient faciliter la vulgarisation scientifique et l'accès aux données de santé, tout en respectant les impératifs éthiques et méthodologiques.

Avant de pouvoir entamer ces missions, une phase essentielle de recherche et de découverte des outils a été nécessaire. En effet, je n'avais jamais manipulé auparavant ni l'infrastructure **vLLM**, ni l'ordonnanceur de tâches **SLURM**, ni les bibliothèques spécifiques comme LangChain, HuggingFace Embeddings ou FAISS. J'ai donc consacré une partie importante de mon stage à l'autoformation : étude de documentations, tests en environnement local, compréhension des fondements du **machine learning** et des **réseaux de neurones**, ainsi que des concepts liés aux LLM. Cette étape m'a permis de mieux appréhender le rôle des embeddings, le fonctionnement du processus ETL pour la préparation des données, et la logique d'indexation vectorielle. Cette étape préalable a été décisive pour me permettre d'aborder mes missions de manière autonome et structurée.

La première mission a débuté par le déploiement d'un modèle Mistral via vLLM, exécuté sur des GPU du cluster interne. Après plusieurs essais, le modèle retenu fut *Magistral-Small-2506*, qui présentait un équilibre entre capacité de raisonnement et suivi des consignes. Une fois ce modèle opérationnel, j'ai conçu différentes familles de pré-prompts :

- Pré-prompts à règles structurées
- Pré-prompts par exemples
- Pré-prompts guidés par étapes
- Pré-prompts utilisant des questions implicites
- Pré-prompts contraints par un score de lisibilité Flesch-Kincaid

⁴ Cf. Annexe 3

⁵ Cf. Annexe 6

Chaque pré-prompt avait pour objectif de transformer un résumé scientifique dense en un texte de 250 mots accessible au grand public. Nous avons pour cela suivi les consignes de rédaction pour les résumés grand public imposés par le CERS⁶ (Centre Éthique & Recherche en Santé). Pour avoir accès aux projets déclarés par le personnel médical sur la plateforme de l'EDS, mon maître de stage m'a fourni l'url d'une **API** interne au CHU de BORDEAUX regroupant toutes les données relatives à ces projets. Avant de commencer à exploiter les résumés scientifiques, j'ai donc dû réaliser en amont un traitement important des données afin de conserver uniquement les informations essentielles. Cela m'a également permis de me familiariser avec les données sur lesquelles j'allais travailler.

La seconde mission consistait à développer un pipeline RAG. Pour cela j'ai mis en place un flux complet comprenant l'ingestion des documents (PDF et CSV), leur segmentation en morceaux cohérents, la génération d'**embeddings** locaux avec HuggingFace, l'indexation vectorielle de ces représentations dans **la base vectorielle FAISS**, puis la récupération contextuelle au moment de l'interrogation du modèle. Le tout était orchestré via LangChain et connecté à Mistral par l'API vLLM. Ce système permettait d'interroger des corpus médicaux spécifiques.

b. Mes résultats

Les résultats de la première mission ont été obtenus grâce à une méthodologie d'évaluation rigoureuse. Dix projets scientifiques ont servi de corpus de test, et chaque résumé généré par les pré-prompts a été évalué par deux à trois collègues selon les critères du Centre Éthique de la Recherche en Santé (CERS) : compréhension facile, fidélité au contenu scientifique, langage accessible, absence d'exagération, et absence d'informations ajoutées. Un classement global a été établi pour chaque pré-prompt, et la cohérence des jugements a été validée par le calcul du coefficient de l'accord inter-évaluateurs, nous avons choisi **le coefficient de concordance de Kendall** qui semble être le plus approprié à la situation. Nous avons obtenu un coefficient de **0,463** ce qui indique une concordance inter-évaluateurs modérée mais fiable.

⁶ Cf. Annexe 4

```

def calculer_kendall_global(tab_eval):
    """
    Calcule un accord inter-évaluateurs global (tous projets confondus)
    """
    pivot = tab_eval.pivot_table(index=["Résumé", "Lettre"], columns="Evaluateur", values="Rang")
    if pivot.shape[1] < 2:
        return None

    accords = []
    for eva1, eva2 in combinations(pivot.columns, 2):
        tau, _ = kendalltau(pivot[eva1], pivot[eva2])
        accords.append(tau)

    moyenne_tau = np.nanmean(accords)
    return moyenne_tau

```

Figure 1 : Fonction python déterminant le coefficient de concordance de Kendall

(Valeur comprise entre 0 et 1 : 0 discordance entre les évaluateurs

1 concordance parfaite entre les évaluateurs)

Les résultats issus de l'évaluation manuelle réalisée par mes collègues mettent en évidence plusieurs tendances fortes. Le pré-prompt basé sur des exemples a obtenu le meilleur classement et la meilleure note moyenne. Il s'est distingué par sa capacité à produire des textes à la fois accessibles et fidèles au contenu scientifique. Les questions implicites ont également donné de bons résultats, bien qu'une légère perte de précision ait été observée. Les règles structurées ont montré de bonnes performances en lisibilité, mais leur rigidité limitait parfois la qualité du fond scientifique. Les étapes guidées ont produit des résultats très variables selon les textes sources. Enfin, le score de Flesch-Kincaid, bien que garantissant une lisibilité élevée, a donné les pires performances, car il produisait des résumés appauvris, parfois infidèles, et contenant des ajouts non souhaités.

Pré-prompt	Classement	Compréhension	Fidélité scientifique	Langage accessible	Absence d'exagération	Absence d'informations ajoutées	Note moyenne globale
Règles structurées	3	4,45	3,45	4,61	4,35	4,19	4,21
Exemples	1	4,48	3,84	4,42	4,30	4,42	4,32
Étapes guidées	4	3,87	3,71	3,87	4,13	4,00	3,92
Questions implicites + exemples	2	4,62	3,77	4,26	4,43	4,34	4,28
Flesch-Kincaid	5	1,84	2,97	2,19	2,39	1,97	2,27

Figure 2 : Tableau récapitulatif des évaluations obtenues pour chaque résumé généré par les 5 pré-prompts (Notes croissantes allant de 0 à 5)

Parmi les cinq stratégies testées, l'approche fondée sur le **score de Flesch-Kincaid**⁷ mérite une attention particulière. Cet indicateur, développé initialement pour l'évaluation des textes scolaires anglophones, vise à mesurer la lisibilité d'un texte écrit. Il se base sur deux paramètres principaux :

- La longueur moyenne des phrases, exprimée en nombre de mots
- La longueur moyenne des mots, exprimée en nombre de syllabes

La formule combine ces deux éléments pour attribuer un score compris entre 0 et 100, où un score élevé correspond à un texte plus facile à lire, un score supérieur à 60 correspond à un texte adapté au grand public, tandis qu'un score bas indique un texte complexe, un score inférieur à 30 sera attribué aux textes destinés aux étudiants universitaire. Cet outil est utilisé pour calibrer automatiquement la difficulté lexicale et syntaxique des textes et s'assurer que leur niveau est adapté au public visé.

⁷ Cf. Annexe 5

Dans le cadre de mon stage, j'ai conçu un pré-prompt demandant explicitement au modèle de générer des résumés atteignant un score supérieur à 50. L'objectif était de produire des textes fluides et accessibles. Si, en pratique, les textes obtenaient effectivement de bons scores de lisibilité, les évaluateurs ont constaté une perte de fidélité scientifique et l'apparition d'ajouts non souhaités. Autrement dit, le texte devenait facile à lire, mais au prix d'une simplification excessive, parfois jusqu'à déformer le contenu scientifique initial.

Ce résultat illustre une limite importante : la lisibilité mécanique ne garantit pas la qualité informationnelle d'un résumé scientifique. La vulgarisation demande un compromis entre accessibilité et exactitude, que les pré-prompts « par exemples » et « par règles structurées » ont mieux su préserver. Ce constat rejoint une réflexion plus large issue de mes cours d'éthique : les indicateurs quantitatifs sont utiles, mais ils ne peuvent se substituer au jugement critique humain, surtout lorsqu'il s'agit de santé publique.

La deuxième mission a démontré la faisabilité d'un RAG médical local. Les premiers tests qui portaient sur la question des contre-indications médicamenteuses en cas d'insuffisance rénale, ont montré que le système pouvait extraire les documents pertinents, les segmenter, et produire une réponse contextualisée citant les sources. Cette approche garantit une traçabilité des réponses, essentielle en contexte médical, et constitue une avancée significative par rapport à une génération de texte libre, souvent difficile à valider scientifiquement.

c. Mon analyse

La réalisation de ces missions a été jalonnée de difficultés techniques et méthodologiques, qui m'ont conduit à développer des stratégies de résolution de problèmes.

L'une des principales difficultés a été la prise en main de vLLM et du cluster GPU via SLURM. Ces outils, que je n'avais jamais utilisés auparavant, nécessitaient une compréhension fine des scripts de soumission de jobs et des contraintes de ressources. J'ai surmonté cette difficulté en m'appuyant sur la documentation interne, en testant des scripts simples, puis en les complexifiant progressivement.

Un autre défi a concerné la compréhension des concepts de machine learning et de leur mise en œuvre pratique. Le passage de la théorie à la pratique, notamment dans la manipulation d'embeddings ou la logique des bases vectorielles, n'était pas immédiat. J'ai

résolu ce problème par un travail de recherche approfondi (documentation scientifique, cours en ligne, lectures spécialisées, etc...) et par des expérimentations incrémentales.

J'ai également rencontré des limites liées aux pré-prompts. Certains pré-prompts donnaient des résultats imprévisibles ou incohérents, ce qui m'a amenée à formaliser une méthode d'évaluation plus stricte et à recourir à des outils statistiques comme le coefficient de Kendall pour objectiver les résultats. De même, la stratégie basée sur le score de Flesch-Kincaid, pourtant séduisante sur le plan théorique, s'est révélée peu adaptée au contexte médical, ce qui m'a appris l'importance de confronter des hypothèses à l'expérimentation concrète.

Une autre difficulté majeure à laquelle j'ai été confrontée est apparue lors de la mise en place du pipeline RAG. En effet, cela a posé des difficultés en termes de gestion des données (formats hétérogènes rendant nécessaire un nettoyage total des données) et d'intégration logicielle (compatibilité entre LangChain, FAISS et vLLM). Ces obstacles ont été levés grâce à une approche incrémentale, en construisant le pipeline étape par étape (ingestion, embeddings, indexation, génération) et en validant chaque brique indépendamment avant de les assembler.

Enfin, tout au long de mon stage, je devais me montrer particulièrement attentive afin d'éviter de pousser sur le Git du CHU de Bordeaux des informations sensibles, comme le code du token nécessaire pour se connecter à HuggingFace. Dans ce cadre-là, mon maître de stage ainsi que mes collègues au sein de l'Unité IAM m'ont appris à me familiariser avec l'interface Git et m'ont sensibilisé à l'utilisation du fichier `.gitignore`. L'attention accordée à cette problématique a permis d'assurer l'intégrité des données manipulés mais également faciliter la réutilisation future de mes codes.

En définitive, ces difficultés se sont révélées formatrices. Elles m'ont permis de développer non seulement des compétences techniques, mais aussi des capacités d'adaptation et de rigueur méthodologique. Elles m'ont également confortée dans une conviction centrale : l'IA en santé ne peut être envisagée sans un regard critique constant, où les enjeux techniques et éthiques sont étroitement imbriqués.

III – Bilan du stage

Ce stage a constitué une expérience formatrice, tant sur le plan professionnel que personnel. Il m'a permis de découvrir le fonctionnement concret d'un grand établissement

hospitalier, de découvrir les facettes du métier de data engineer et surtout d'acquérir de nouvelles compétences techniques et méthodologiques. Ce stage m'a aussi permis de mieux cerner mes capacités, mes aspirations et d'affiner ma vision du futur métier d'ingénieur.

Tout d'abord, ce stage m'a appris à travailler dans un contexte où la recherche scientifique et l'innovation technique s'inscrivent dans une mission de service public. J'ai découvert le rôle central de l'Unité IAM dans la gestion et la valorisation des données de santé, et plus largement, la manière dont un hôpital universitaire combine soins, recherche et enseignement. Le rôle de l'ingénieur appliqué à la santé m'est apparu comme exigeant mais aussi profondément utile, puisqu'il consiste à développer des outils numériques qui doivent être à la fois performants, éthiques et adaptés aux besoins des soignants. Mon intégration au sein d'une équipe pluridisciplinaire, composée de médecins, d'ingénieurs et de chercheurs, m'a permis de constater la richesse du dialogue entre des profils complémentaires, chacun apportant une expertise indispensable à l'avancée des projets. J'ai également observé l'importance de la rigueur organisationnelle dans un environnement où la confidentialité et la qualité des données sont cruciales.

Sur le plan technique, ce stage m'a beaucoup apporté. J'ai manipulé des outils que je n'avais jamais utilisés auparavant, tels que les LLM, l'interface SLURM, ainsi que les librairies LangChain, HuggingFace Embeddings et FAISS. J'ai appris à les combiner pour concevoir des pipelines complets. J'ai expérimenté en pratique le fonctionnement des modèles de langage de grande taille (LLM), leur potentiel mais aussi leurs limites, ainsi que les apports spécifiques des approches de type Retrieval-Augmented Generation (RAG). Le déploiement de modèles, la conception de pré-prompts et surtout l'évaluation des résultats m'ont permis d'acquérir une méthodologie expérimentale rigoureuse : formulation d'hypothèses, mise en place de critères objectifs (grille CERS), recours à des mesures statistiques (coefficient de concordance de Kendall) pour valider la robustesse des observations. Ces acquis techniques sont directement transférables à mon futur métier d'ingénieur : ils m'ont appris que l'intelligence artificielle n'est pas seulement une boîte noire génératrice de texte, mais un outil qui demande une mise en œuvre méthodique, des choix techniques précis et une évaluation critique permanente.

Au-delà de l'aspect technique, j'ai progressé sur le plan humain et relationnel. Ce stage m'a appris à travailler en autonomie dans un cadre exploratoire, où beaucoup de solutions restaient à inventer. J'ai souvent dû me former par moi-même, en consultant des documentations, en testant des scripts ou en reformulant mes approches. Cette capacité d'autoformation a renforcé ma confiance en mes aptitudes à m'adapter et à surmonter des difficultés. J'ai aussi pris conscience de l'importance de la communication : expliquer de

manière claire mes démarches et mes résultats, que ce soit à des ingénieurs familiers des concepts techniques ou à des médecins peu habitués à l'IA, a été un exercice exigeant voire difficile mais très enrichissant. J'ai compris que la vulgarisation ne concerne pas seulement les patients ou le grand public, mais aussi la collaboration entre des experts de disciplines différentes.

Enfin, ce stage m'a permis de mieux me connaître et d'apprendre sur moi-même. J'ai découvert que je pouvais assimiler rapidement des outils complexes et produire des résultats concrets dans un temps limité. J'ai mesuré à quel point je suis motivée par les projets situés à l'intersection de la technologie et de la santé, où les enjeux éthiques sont aussi importants que les aspects techniques. J'ai aussi confirmé mon intérêt pour les problématiques de responsabilité et de transparence dans l'usage de l'IA, ce qui fait directement écho à l'exposé que j'avais réalisé en cours d' « Éthique et sciences techniques ».

Si c'était à refaire, je changerai deux éléments principaux. D'une part, je veillerais à formaliser davantage la documentation de mes expérimentations afin de faciliter leur réutilisation par de futurs stagiaires ou chercheurs de l'unité. D'autre part, si le temps l'avait permis j'aurais aimé confronter plus systématiquement mes résultats à l'avis d'experts médicaux, afin de mieux mesurer leur adéquation aux besoins concrets des soignants et des patients.

En définitive, ce stage a été pour moi une expérience riche et marquante. J'y ai acquis des compétences techniques solides, des méthodes de travail rigoureuses et une meilleure compréhension de l'écosystème hospitalier et de la recherche en santé. Mais j'y ai surtout appris que la technologie ne prend tout son sens que si elle est accompagnée d'une réflexion éthique et d'un souci constant de responsabilité. Ce stage s'inscrit ainsi dans la continuité de mon projet professionnel, en confirmant mon intérêt pour les applications de l'intelligence artificielle dans des environnements sensibles et en m'encourageant à poursuivre dans la voie d'un ingénieur capable de conjuguer savoir-faire technique et engagement éthique.

Conclusion

Ce stage de première année d'ingénieur à CY Tech, réalisé du 2 juin au 1^{er} août au sein de l'Unité IAM du CHU de Bordeaux, a constitué une étape décisive de ma formation. Il m'a offert une immersion concrète dans le monde hospitalier et de la recherche médicale, et m'a permis de confronter mes acquis théoriques aux réalités d'un service en pleine transformation numérique.

Les deux missions principales qui m'ont été confiées reflètent bien les enjeux actuels de l'intelligence artificielle appliquée à la santé. Le déploiement d'un modèle de langage local et l'expérimentation de stratégies de pré-prompts m'ont permis d'aborder la question complexe de la vulgarisation scientifique, entre lisibilité et fidélité. La mise en place d'un pipeline RAG a montré la faisabilité d'outils capables de produire des réponses contextualisées et traçables à partir de documents médicaux internes. Ces réalisations, bien qu'exploratoires, ouvrent la voie à de futurs développements utiles pour la recherche biomédicale et l'aide à la décision médicale.

Sur le plan personnel et professionnel, ce stage m'a beaucoup appris. J'ai renforcé mes compétences techniques tout en développant une méthode de travail rigoureuse basée sur l'expérimentation et l'évaluation. J'ai également progressé en savoir-être : autonomie, adaptabilité, communication claire auprès d'interlocuteurs variés. J'ai découvert l'importance de la collaboration interdisciplinaire et de l'éthique dans la mise en œuvre de technologies appliquées à la santé.

Ce stage a surtout conforté mon projet professionnel. Il a confirmé mon intérêt pour les projets à l'intersection de la technologie et des enjeux humains, en particulier dans le domaine de la santé. Il m'a montré que le rôle de l'ingénieur ne se limite pas à la maîtrise d'outils techniques, mais exige une réflexion constante sur les implications sociales et éthiques des solutions développées. Si certaines limites sont restées, comme l'absence de confrontation plus systématique avec l'avis d'experts médicaux, l'expérience a été une véritable réussite et a dépassé mes attentes en termes d'apprentissage.

À l'avenir, je souhaite approfondir mes compétences en intelligence artificielle et en data science, tout en poursuivant ma réflexion éthique et mon intérêt pour les applications sensibles comme la santé publique. Ce stage m'a permis de mieux définir mon orientation : devenir un ingénieur capable de conjuguer expertise technique, responsabilité sociale et capacité de vulgarisation scientifique.

En définitive, cette expérience m'a permis de progresser techniquement et humainement, et elle a renforcé ma conviction qu'innovation et responsabilité doivent avancer de pair. Elle constitue une base solide pour la suite de mon parcours académique et professionnel.

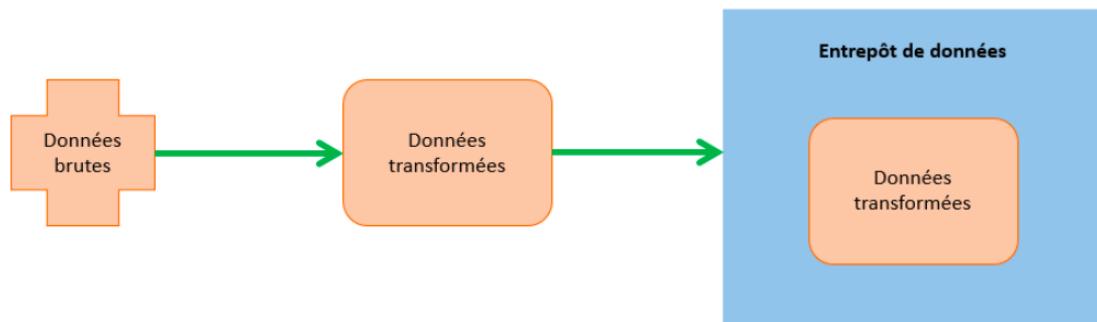
Annexes

Annexe 1 : Lexique

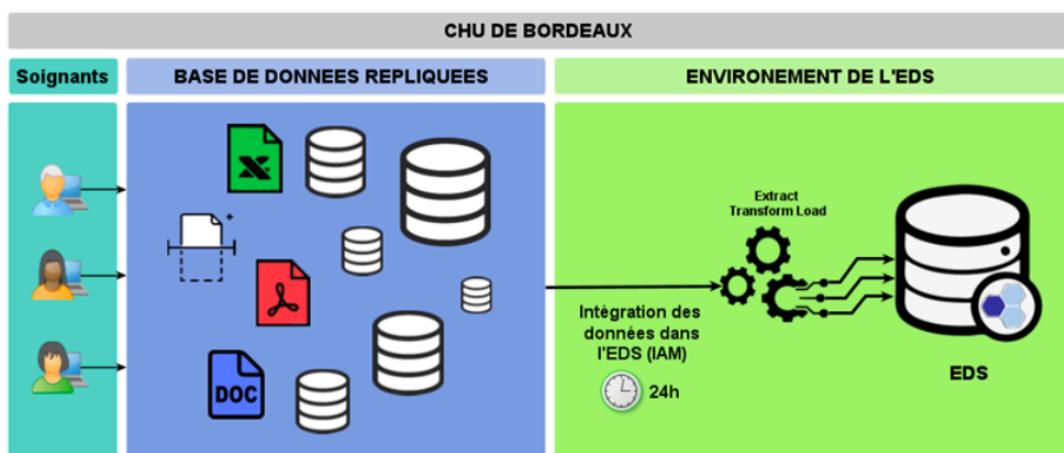
Termes spécifiques	Explications
Processus ETL	Méthode utilisée pour rassembler des données dispersées, les nettoyer et les mettre au bon format avant de les stocker dans une base de données centralisée.
Large Language Model (LLM)	Modèle d'intelligence artificielle entraîné sur de très grandes quantités de textes, capable de comprendre et de générer du langage naturel.
Retrieval-Augmented Generation (RAG)	Méthode qui permet à un modèle d'IA de s'appuyer sur des documents externes pour répondre à une question, en citant les sources utilisées.
Application Programming Interface (API)	Interface qui permet à deux programmes informatiques différents de communiquer entre eux. Par exemple, une API permet à une application d'envoyer une question à un modèle d'IA et de recevoir une réponse.
Pré-prompt	Texte ou consignes préparatoires envoyés à un modèle de langage avant la question principale, pour orienter la manière dont il doit répondre
vLLM	Logiciel spécialisé qui permet de faire fonctionner des modèles de langage (LLM) plus rapidement et efficacement sur des serveurs équipés de cartes graphiques.
SLURM	Programme qui organise et gère l'utilisation partagée des ressources informatiques (serveurs, cartes graphiques) dans un centre de calcul.
Machine learning	Branche de l'intelligence artificielle où l'ordinateur apprend à reconnaître des modèles à partir de données, sans être explicitement programmé pour chaque tâche.
Réseaux de neurones	Type de modèle de machine learning inspiré du fonctionnement du cerveau humain, constitué de "couche" de calculs qui transforment progressivement les données pour en extraire des informations.
Embeddings	Représentation numérique (sous forme de vecteurs) d'un mot, d'une phrase ou d'un document, permettant à l'ordinateur de comparer leur sens ou leur proximité.
Base vectorielle	Base de données spécialisée qui stocke les embeddings et permet de rechercher rapidement des documents proches en sens ou en contenu.
Coefficient de concordance de Kendall	Mesure statistique qui permet de savoir à quel point plusieurs personnes sont d'accord entre elles lorsqu'elles classent ou notent des éléments.
Score de lisibilité de Flesch-Kincaid	Indice qui évalue la facilité de lecture d'un texte, en fonction de la longueur des phrases et des mots. Plus le score est élevé, plus le texte est simple à lire.
Data engineer	Professionnel chargé de concevoir, organiser et maintenir les systèmes qui collectent, transforment et stockent de grandes quantités de données pour qu'elles puissent être exploitées par les chercheurs ou les analystes.

Annexe 2 : Schéma explicatif du processus d'intégration des données « ETL »

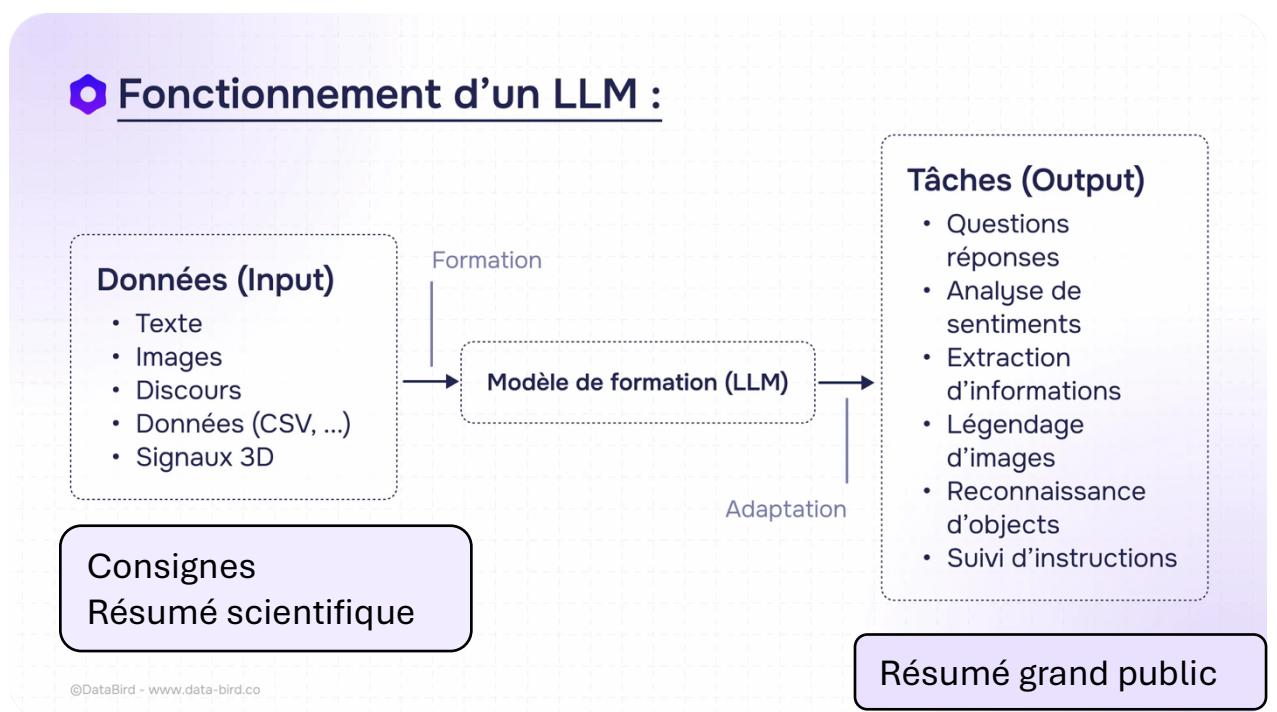
Le processus d'intégration de données ETL (Extract, Transform, Load) est un processus essentiel dans le cadre de la gestion d'un entrepôt de données. Il permet de collecter, transformer et charger des données provenant de sources hétérogènes vers un unique entrepôt de données.



Principe de l'ETL (Extract – Transform – Load)



Annexe 3 : Schéma explicatif du fonctionnement d'un LLM



Annexe 4 : Recommandations du CERS pour la rédaction de résumés grand public

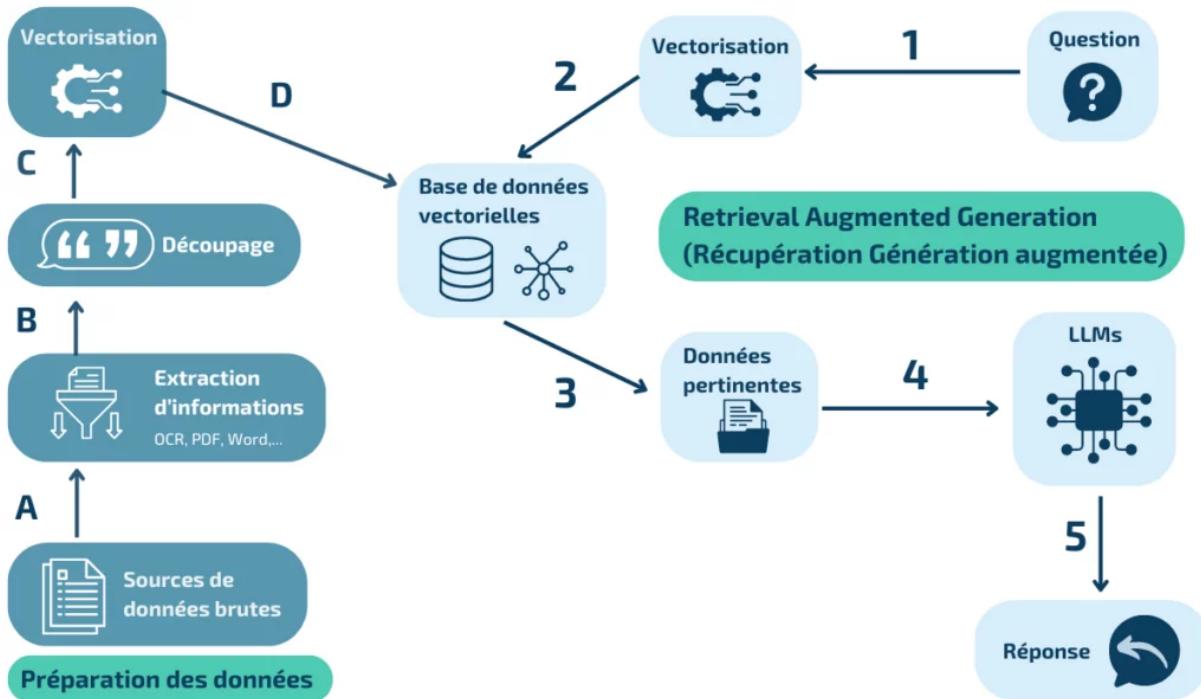
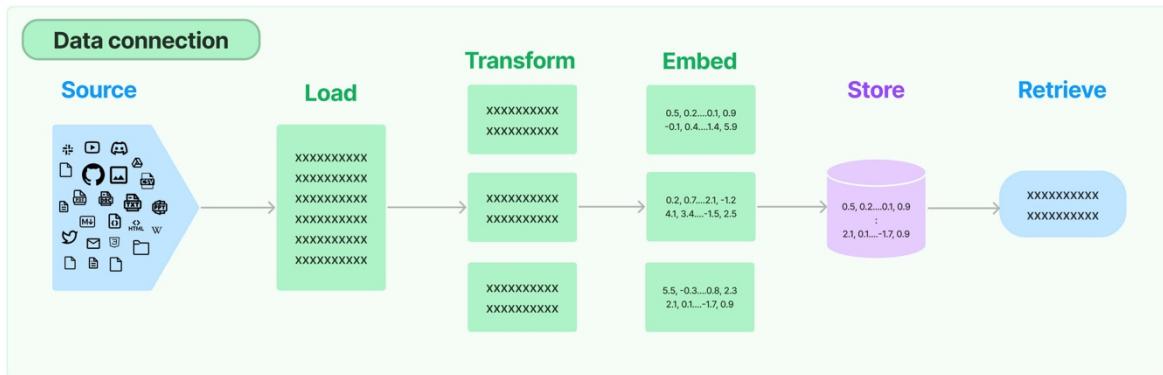
Recommandations de RÉDACTION d'une INFORMATION GRAND PUBLIC d'un projet scientifique	
EN PRATIQUE	
CONTENU	FORME
POURQUOI ? <ul style="list-style-type: none"> Contexte scientifique du projet de recherche. Objectif de l'étude - Question scientifique - Hypothèse(s). 	CONCISION Objectif → 250 mots
COMMENT ? <ul style="list-style-type: none"> Schéma de l'étude (ex : cohorte historique - comparaison de cohortes). Dimension de l'étude : monocentrique - multicentrique - préciser si promotion externe. 	INTELLIGIBILITÉ <ul style="list-style-type: none"> Formulation distincte du texte scientifique. Phrases courtes, explicites. Eviter les termes trop techniques, les abréviations. Privilégier des termes neutres. → Eviter les termes connotés, ambigus ou sujets à interprétations.
QUI ? <ul style="list-style-type: none"> Décrire la population concernée : nombre, critères d'inclusion et, si besoin, d'exclusion. 	FORMULATION ADAPTÉE des RETOMBÉES ATTENDUES Objective et proportionnée à la nature de l'étude.
QUOI ? <ul style="list-style-type: none"> Périodes concernées par le recueil des données (années concernées). Données recueillies, précisions sur les données à risque de discrimination. Objectifs de l'étude (si non précisés précédemment). 	
À QUELLE FIN ? <ul style="list-style-type: none"> Aborder les possibles retombées attendues : scientifiques - techniques - professionnelles. 	
<i>Si recherche comparant des stratégies diagnostiques ou thérapeutiques</i> → Rappeler que les stratégies ou soins ont été faits dans le cadre du soin courant au moment de leur réalisation, dans l'intérêt des personnes, et conformément aux recommandations et bonnes pratiques.	
<i>Si recherche rapportant des stratégies diagnostiques ou thérapeutiques, réalisées hors recommandations</i> → Rappeler que les stratégies ou soins ont été appliqués dans l'intérêt des personnes au moment de leur réalisation.	
RELECTURE POSSIBLE PAR LE CERS Relecture systématique pour les projets mis en place avec l'Entrepôt de Données de Santé.	

Annexe 5 : Formule d'évaluation du score de Flesch-Kincaid et tableau de correspondance

$$206.835 - 1.015 \left(\frac{\text{mots totaux}}{\text{phrases totales}} \right) - 84.6 \left(\frac{\text{syllabes totales}}{\text{mots totaux}} \right)$$

Readability Score	Understanding Position
90-100	Very Easy
80-89	Easy
70-79	Fairly Easy
60-69	Standard
50-59	Fairly Difficult
30-49	Difficult
0-29	Very Confusing

Annexe 6 : Schémas explicatifs du fonctionnement d'un RAG



Annexe 7 : Fiche d'évaluation par le tuteur de l'entreprise

Évaluation de Stage & Alternance

Questionnaire pour le maître de stage et à remettre dans le rapport de stage

Bonjour, vous avez accueilli un(e) étudiant(e) CY TECH pour un stage et nous vous en remercions. Dans le cadre de son évaluation, nous souhaiterions connaître votre avis concernant son passage au sein de votre entreprise. Cette évaluation qualitative sera prise en compte dans la note écrite de stage. Nous vous remercions donc de bien vouloir remplir le questionnaire ci-dessous qui prendra quelques minutes.

Il y a 20 questions dans ce questionnaire. Barème : Très bien (1 pt) ; Bien (0.75 pts) ; Moyen (0.5 pts) ; Passable (0.25 pts) ; Insuffisant (0 pt) ; Sans objet compte pour 1 point.

Nom et prénom de l'étudiant: BROUCAS Adélaïde

Formation : INGÉNIEUR MATHÉMATIQUES FI CERGY 1

Veuillez sélectionner une seule des propositions suivantes :

APPRECIATIONS GENERALES SUR LE/LA STAGIAIRE

1. Niveau des connaissances théoriques

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

2. Niveau des connaissances techniques

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

3. Niveau de réflexion (par rapport à vos attentes)

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

4. Motivation et implication dans le travail

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

5. Assiduité/Ponctualité durant la période de bureau

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

6. Présentation

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

7. Respect du règlement intérieur

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

8. Aptitude au travail en groupe

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

9. Aptitude au travail en autonomie

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

10. Intérêt manifesté à son activité

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

ACTIVITE EN STAGE**11. Compréhension de l'environnement**

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

12. Capacité à maîtriser les techniques utilisées

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

13. Capacité à proposer et à rechercher des solutions (initiative)

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

14. Qualité du travail effectué

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

15. Utilisabilité du travail réalisé

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

16. Qualité des communications écrites

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

17. Qualité des communications orales

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant

Évaluation de l'étudiant(e) : 18,25/20

Commentaire :

Adélaïde a travaillé sur l'exploration de l'utilisation de LLM (large language model)

au sein du CHU de Bordeaux sur 2 cas d'usages:

- Génération de résumés de projets de recherche pour le grand public
- Utilisation d'un RAG pour explorer des ressources biomédicales

Adélaïde a beaucoup progressé au cours de son stage, tant sur la compréhension théorique des LLM que sur leur implémentation. En particulier, Adélaïde a su mener du début à la fin un projet incluant l'évaluation des différentes méthodes développées.

Très bon stage.

Les étudiants de CY Tech partent en stage, en contrat de professionnalisation et en contrats d'apprentissage.

Pour toute offre, toute demande : service-entreprises@cy-tech.fr

Merci d'avoir complété ce questionnaire.

19. Capacité à tirer un bilan critique du travail

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

20. Capacité à communiquer en anglais à l'oral

- Très bien
- Bien
- Moyen
- Passable
- Insuffisant
- Sans objet

Visa maître de stage

Tampon de l'entreprise

Romain GRIFFIER



Annexe 8 : Attestation de fin de stage

ATTESTATION DE STAGE

À remettre au stagiaire à l'issue du stage

ORGANISME D'ACCUEIL

Nom ou Raison sociale : CHU de Bordeaux.....
Adresse : 12 Rue Dubernat.....
CP : 33404..... Ville : Talence.....
Téléphone : 0556795679.....

Certifie que

LE/LA STAGIAIRE

Nom : BROUCAS..... Prénom : Adélaïde..... Sexe : F M Née le : 24/10/2005
Adresse : 33 Avenue de la Belle Étoile.....
CP : 33270..... Ville : BOULIAC.....
Téléphone : 0649715567..... e-mail : ade.broucas@icloud.com.....

ETUDIANT EN (intitulé de la formation ou du cursus de l'enseignement supérieur suivi par le ou la stagiaire) :

CY TECH - CERGY (AVENUE DU PARC) - INGÉNIEUR - INGÉNIEUR MATHÉMATIQUES FI CERGY 1 - DI02A1.....

AU SEIN DE : CY TECH

a effectué un stage prévu dans le cadre de ses études

DUREE DU STAGE

Dates de début et de fin du stage : Du 02/06/2025..... (JJ/MM/AAAA) au 01/08/2025..... (JJ/MM/AAAA)

Représentant une durée totale de 8..... mois / semaines (rayer la mention inutile))

La durée totale du stage est appréciée en tenant compte de la présence effective du stagiaire dans l'organisme, sous réserve des droits à congés et autorisations d'absence prévus à l'article L.124-13 du code de l'éducation (art. L.124-18 du code de l'éducation). Chaque période au moins égale à 22 jours de présence consécutifs ou non et considérée comme équivalente à un mois de stage et chaque période au moins égale à 7 heures de présence consécutives ou non est considérée comme équivalente à un jour de stage.

MONTANT DE LA GRATIFICATION VERSE AU STAGIAIRE

Le/la stagiaire a perçu une gratification de stage pour un montant total brut de 0 €

L'attestation de stage est indispensable pour pouvoir, sous réserve du versement d'une cotisation, faire prendre en compte le stage dans les droits à retraite. La législation sur les retraites (loi n°2014-40 du 20 janvier 2014) ouvre aux étudiants dont le stage a été gratifié la possibilité de faire valider celui-ci dans la limite de deux trimestres, sous réserve du versement d'une cotisation. La demande est à faire par l'étudiant(e) dans les deux années suivant la fin du stage et sur présentation obligatoire de l'attestation de stage mentionnant la durée totale du stage et le montant total de la gratification perçue. Les informations précises sur la cotisation à verser et sur la procédure à suivre sont à demander auprès de la caisse primaire d'assurance sociale (code de la sécurité sociale art. L.351-17 – code de l'éducation art.D.124-9).

FAIT A BORDEAUX..... LE 01/08/2025.....

Nom, fonction et signature du représentant de l'organisme d'accueil

Romain GRIFFIER, praticien hospitalier



Bibliographie

LLM et apprentissage automatique

- Hugging Face. *Large language model course*. <https://huggingface.co/learn/llm-course/chapter0/1>
- Google Developers. *Transformers – Machine Learning Crash Course*. <https://developers.google.com/machine-learning/crash-course/llm/transformers>
- LBKE. *Lexique des LLM*. <https://www.lbke.fr/ressources/lexique-des-llm>
- OpenClassrooms. *Initiez-vous au machine learning*. <https://openclassrooms.com/fr/courses/8063076-initiez-vous-au-machine-learning>
- MathWorks. *Réseaux de neurones*. <https://fr.mathworks.com/discovery/neural-network.html>
- Elements of AI. *Cours d'introduction à l'intelligence artificielle*. <https://course.elementsofai.com/fr/>

Déploiement local et API

- vLLM. (2024). *Serving models with OpenAI-compatible server*. https://docs.vllm.ai/en/v0.4.1/serving/openai_compatible_server.html
- Mistral AI. (2024). *Endpoint instruct – Code generation*. https://docs.mistral.ai/capabilities/code_generation/#instruct-endpoint
- IT-Connect. (2024). *Utiliser l'API Mistral AI en Python*. <https://www.it-connect.fr/guide-pour-utiliser-api-mistral-ai-en-python/>

Optimisation et quantization

- Hugging Face. (2024). *Quantization – Optimum guide*. https://huggingface.co/docs/optimum/concept_guides/quantization

RAG et bases vectorielles

- Mistral AI. (2024). *Guide RAG*. <https://docs.mistral.ai/guides/rag/>
- NextStart.ai. (2024). *Comment fonctionne le RAG*. <https://www.nextstart.ai/2024/01/19/comment-fonctionne-le-rag-en-ia-generative/>
- Rodrigues, A. (2023). *Retrieval-augmented generation with LangChain and FAISS*. Medium. <https://medium.com/@alexrodriguesj/retrieval-augmented-generation-rag-with-langchain-and-faiss-a3997f95b551>
- Dev.to. (2024). *The ultimate guide to RAG*. <https://dev.to/codemaker2015/the-ultimate-guide-to-retrieval-augmented-generation-rag-5e6e>
- Hackernoon. (2024). *Créer une application RAG avec LangChain*. <https://hackernoon.com/lang/fr/tutoriel-complet-sur-la-cr%C3%A9ation-d%27une-application-rag-%C3%A0-l%27aide-de-langchain>

Prompt engineering

- Jedha. (2024). *10 techniques de prompt engineering à tester*. <https://www.jedha.co/formation-ia/10-techniques-de-prompt-engineering-a-tester>

Outils Python et environnements

- Stéphane Robert. (2024). *Environnements virtuels en Python*. <https://blog.stephane-robert.info/docs/developper/programmation/python/environnements-virtuels/>

Éthique et IA en santé

- Université de Lille. (2024). *Mémoire de médecine – 2024ULILM159* (PDF). https://pepite-depot.univ-lille.fr/LIBRE/Th_Medecine/2024/2024ULILM159.pdf
- Académie de Versailles. (2024). *Éthique et intelligence artificielle*. <https://drane-versailles.region-academique-idf.fr/spip.php?article1167>

Présentations internes

- Broucas, Adélaïde. (2025, juillet). *Méthodologie de rédaction et d'évaluation des pré-prompts pour la vulgarisation scientifique* Diaporama
- Unité IAM – CHU de Bordeaux. (2025, avril). *Structuration de l'organisation EDS* (Diaporama)

Liens des images de ce rapport

- Blent AI. (2024). *RAG : tout savoir sur le Retrieval-Augmented Generation*. <https://blent.ai/blog/a/rag-tout-savoir>
- Little Big Things. (2024). *Grand modèle de langage (LLM) : tout comprendre*. <https://littlebigthings.fr/grand-modele-langage-llm>
- Neovision. (2024). *Les RAG : une révolution dans l'intelligence artificielle*. <https://neovision.fr/les-rag-une-revolution-dans-lintelligence-artificielle/>
- D'autres visuels et schémas techniques ont été gracieusement fournis par l'Unité IAM du CHU de Bordeaux, et sont utilisés avec leur autorisation à titre pédagogique dans le cadre de ce rapport de stage.