|  |  |
| --- | --- |
|  | Stage de 2ème année |
| Rédigé par Gustave Richter | Créer le 03/09/2024 |
| Référent : Nisrine | Tuteur : Yassine |
| Effectué du | 18 juin au 31 aout 2024 |

CY TECH

Stage de 2ème année au sein de BS Digital

# Table des matières

[Table des matières 2](#_Toc176333565)

[Remerciements 3](#_Toc176333566)

[Introduction 4](#_Toc176333567)

[L’entreprise d’accueil 5](#_Toc176333568)

[Mes Mission – le stage en détail 6](#_Toc176333569)

[Bilan 7](#_Toc176333570)

[Conclusion 8](#_Toc176333571)

# Remerciements

Je tiens à remercier Alexandre Stojanovic ainsi que Emmanuel Briquet de m’avoir fait confiance pour ce stage.  
Merci à mon maitre de stage Yassine Sefrioui pour m’avoir mis sur des missions intéressantes et chalangeuses.   
Merci aussi a Rémi Courtade, Antoine Logeais et Charles Tesson qui mon tous les trois accompagné et répondu à la plus pars de mes questions.

Et je remercie aussi toutes l’équipe de BS Digital qui ma accueillit si rapidement et chaleureusement.

Merci pour l’accompagnement et le suivi tout au long du stage.

# Introduction

Pour la recherche du stage je suis passé par plein de site pour envoyer mon CV et j’avais obtenu une réponse favorable, mais à la suite d’un quiproquo je n’ai pas pu faire de stage avec eux, ils cherchaient un alternant et moi juste un stage donc j’ai dû reprendre mes recherches et j’ai obtenu mon stage grâce à Nadège qui m’avait envoyé un mail disant qu’il cherchait un stagiaire/alternant et j’ai eu la chance qu’ils acceptent de me prendre en stage sans faire par la suite une alternance.

Dès les premier jour Charles m’a principalement montré sur quoi j’allais travailler, la mission principale de mon stage consistait à faire une automatisation des tests pour L’IA SAMI, le but étais de noter l’IA avec une certaine base de données de question, afin d’éviter plusieurs problèmes.

N’ayant pas spécialement touché à l’IA l’intérêt du stage a été de d’approfondir la découverte de l’univers de l’IA via une approche que dont je sous-estimer l’importance.

Le stage s’est déroulé en plusieurs projet, tout d’abord le plus gros celui de faire une automatisation des tests, puis lorsque ce dernier fut fini je me suis attaqué a d’autre petit projet.

Je vous présenterai la société dans un premier temps, puis les missions que j’ai pu effectuer et pour finir mon bilan et ma perception.

# L’entreprise d’accueil

La société dans laquelle j’ai fait mon stage se nomme BS Digital



Il s’agit d’une entreprise crée en 2017 par Alexandre Stojanovic et Briquet Emmanuel, elle est basée dans Hélioparc à Pau.

Juridiquement, il s’agit d’une société SAS : Société par Action Simplifiée.

Elle est dirigée par le groupe ISAGRI depuis fin 2018.

On y comptait juste 2 personnes avant qu’ISAGRI les rachètent et depuis aujourd’hui, on y compte environ 30 salariés, un nombre qui ne devrait qu’évoluer d’ici les prochaines années.

L’entreprise est en pleine transformation hiérarchique, elle se sépare en trois grande partie, une équipe exploit, une équipe dev et une équipe IA dans laquelle je me trouvais.

Son activité principale est la vente de leur logiciel AMI Compta avec une assistance d’une IA SAMI. AMI Compta est un logiciel pour les experts-comptables qui les aide à la saisie comptable et afin de faire une dématérialisation des factures.

# Mes Mission – le stage en détail

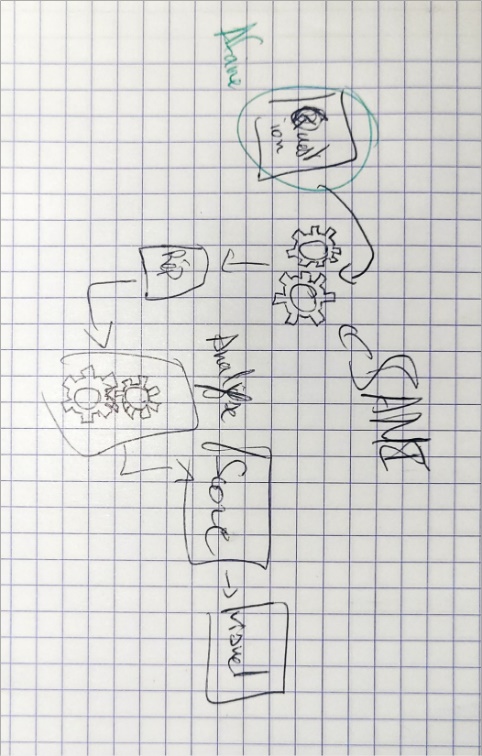
Mon stage s’est déroulé sur 11 semaines, il s’est déroulé tout du long en présentiel, avec l’équipe IA nous étions au début dans des locaux séparés du reste de l’équipe, puisque l’entreprise ne faisait que grandir et qu’ils n’avaient pas prévu une telle expansion ils ont dû trouver des locaux rapidement, mais nous n’étions pas très loin du reste de l’équipe et les retrouvions à chaque pause ainsi que pour manger les midis. Puis vers début/mi-aout BS Digital a récupéré des locaux adjacents a ceux que le reste de l’équipe a et nous avons pu les rejoindre. J’avais à ma disposition sur place un ordinateur.

## 1ere mission : Testing\_chatbot

Testing\_chatbot est le nom du projet que j’ai moi-même donné peut-être pas le plus inspiré a ce moment je n’en doute pas, au début de mon projet on m’avait donnée une liste de question sur lesquelles SAMI se basais pour faire son apprentissage, j’ai commencé par essayer de trouver des manières de noter une IA parce que je n’y connaissais rien sur le système de notation d’IA.

Mais j’avais un problème sur lequel je me suis buté pendant 1-2semaines s’en poser de question au début croyant que c’était voulu, dans la base question je n’avais pas de réponse attendue.

Je suis donc partie du principe qu’il fallait que je fasse une prédiction de ce qu’allait me sortir SAMI, puis je me suis dis pourquoi pas essayer de voir si l’on retrouve un certain nombre de mot clé de la question dans la réponse cela devrait fonctionner… Sans surprise cela n’était pas très fructueux et lorsque j’en ai parlé avec Yassine il m’a tout de suite fais comprendre que je partais beaucoup trop loin et que la réponse attendue aurait dû être là. Il m’a donc passé le bon fichier et j’ai dû recommencer une partie du travail.



Je me suis donc fais plusieurs schémas pour essayer de décortiquer le projet en plusieurs petits problème plus simple à commencer.

### J’ai choisi de coder en Python car c’est le langage utilisé par l’équipe IA, ce qui facilitait l’intégration et la collaboration

### Les questions

Elles étaient au format ceux csv qui me semblais plus complexe à utiliser qu’un format json, j’ai donc fait un petit bout de code ‘transform.py’ afin d’avoir une homogénéisation du format de question. Après l’utilisation de ce code les questions avaient ce format :

Une image contenant texte, capture d’écran, logiciel, affichage

Description générée automatiquement Id : …  
Question : « … »  
link : « … »  
reponse : «

-content : « … »  
 -role : « … »  
 - type : « … »

 »

Après avoir récupérer la base de question sur laquelle nous voulions tester SAMI je lui envoie les questions et obtiens ainsi la réponse de SAMI que j’enregistre dans un autre dossier ‘réponse’ et je passe ensuite à la plus grosse partie du projet l’analyse.

### Une image contenant texte, capture d’écran, logiciel, affichage Description générée automatiquementL’analyse

Pour l’analyser les réponses de SAMI, j’ai comparé les réponses attendues, issues du fichier de référence, avec celles générées par l’IA. Cette analyse s’est décomposée en trois parties principales :

#### Lien

Ici je test si SAMI me renvoi bien le bon lien, je le détecte avec la clé « type » qui a comme valeur « link » et le lien que je veux obtenir se trouve dans la valeur de la clé « link ».  
Ne voulant pas d’erreur sur le lien renvoyé je fais un simple de test pour voir si l’une des valeurs renvoyées dans la réponse est exactement la même.

#### Function

Pour le test des fonctions cela se passe un peu de la même manière puisque l’on attend une fonction précise en retour

#### Text

Le texte est sûrement celui qui ma posé le plus de problème mais lorsque j’ai eu les réponses attendues je me suis dit que de faire un embedding des deux réponses puis comparer les deux vecteurs obtenus avec un similarité de cosinus serai la meilleur et la plus simple idée à implémenter.

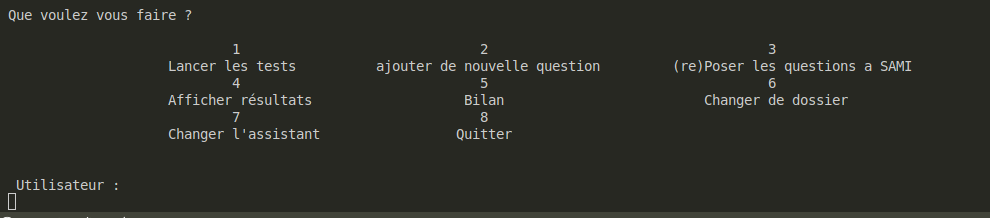
### Assignation du score

Pour cette partie cela fut assez simple dans les trois cas j’enregistrais dans un premier temps le model utilisé d’embedding, j’enregistrais ensuite l’id de la question tester, le temps de réponse de SAMI et la ‘distance’ obtenue entre les deux vecteurs avec le calcul de la similarité des cosinus, pour ce qui est du calcul de la distance des liens ou des fonctions j’assignais tout simplement 0 ou 1 lorsque le résultat était bon.

Et pour finir et gagner du temps lors de l’affichage des résultats je calcul directement le temps de réponse moyen ainsi que la distance moyenne obtenu.

### Affichage Graphique

Pour cette partie au début je croyais en discuter avec un dev-front afin d’implémenter ma solution directement sur AMI Compta,  
Au final après discutions avec Yassine, les développeur IA étant sûrement les seuls a vouloir utiliser cette solution un affichage graphique dans le terminal était amplement suffisant

Avant d’arriver sur le menu l’utilisateur choisi au préalable 2 fichiers, le fichier de questions, de réponse.

Ainsi dès qu’il arrive sur le menu il peux choisir de faire toutes les 7 possibilités qui lui sont présenté.

(1) lancera les test de similarité entre les fichiers qu’il a choisi s’il n’a pas choisi deux fichiers compatible une erreur surviendra.

(2) lui permettra d’ajouter une question selon toutes ses préférence a la base de question qu’il a sélectionner plus tôt.

(3) Permet de relancer les questions a SAMI fortement recommandé après l’ajout d’une question sinon une erreur lors des tests sera affiché

(4) Affichera la totalité des résultats plus utiles pour voir rapidement les résultats obtenus, cette fonctionnalité peux possiblement évolué pour affichée juste les meilleurs ou pire résultats du fichier choisi

(5) Demande de choisir parmi un des fichiers de score enregistrer et affiche un récapitulatif du fichier avec les moyennes calculé plus tôt .

(6) Permet de changer de fichier pour faire une autre série de test sans avoir a relancer totalement le programme.

(7) Permet de changer facilement d’assistant, il s’agit d’une fonction qui ne sers pour le moment pas a grand-chose mais implémenté afin de facilité la comparaison future de différent model de SAMI

(8) Quitter…

**Fin de stage**  
Durant mon stage, ma mission Testing\_chatbot était prévue pour occuper toute la durée du stage. Étant très impliqué dans celle-ci, j'ai réussi à la terminer 10 jours avant la fin. Charles et Yassine ont donc dû me trouver une autre mission. En attendant, Charles m'a proposé de continuer à travailler sur Testing\_chatbot, plus précisément en élargissant la base de questions à tester.

J'ai commencé par suivre les propositions de Charles, notamment en reformulant la base de questions existante avec différents styles de langage. J’ai créé un langage médiéval, un langage soutenu, un registre de langue parlée, et j’ai simplement reformulé certaines questions. J'ai également créé une base de questions que j'ai appelée « piège », contenant des questions n’ayant aucun rapport avec la comptabilité, ainsi que des questions illégales formulées de manière à tenter d'obtenir des réponses.

Comme il me restait trop peu de temps pour commencer une toute nouvelle mission, j'ai passé mes derniers jours à aider Lisa dans sa mission. J'ai principalement effectué des relectures de code, quelques retouches, ainsi que proposé des idées et des solutions.

## **Conclusion & Bilan** Mon stage chez BS Digital a été une expérience très enrichissante. J’ai trouvé ma mission tellement intéressante que j’ai réussi à la terminer plus rapidement que prévu. J’ai apprécié travailler sur une véritable IA, et non sur une simple version montée rapidement pour répondre à quelques questions basiques.

Ce stage m’a également permis d’approfondir mes connaissances en Python. J'ai pris conscience des nombreuses différences avec le C, notamment dans mes réflexes de codage qui ne fonctionnaient pas de la même manière en Python. J’ai également découvert un environnement de programmation orienté IA, un domaine que je n’avais pas encore exploré.

J'ai rencontré quelques difficultés au début du projet, principalement dues à un manque de communication de ma part. Cependant, le stage m'a permis de découvrir le travail en équipe en présentiel, après avoir fait mes précédents stages en distanciel. J’ai rapidement ressenti la différence : pouvoir demander de l’aide directement à un collègue ou partager mes idées pour obtenir un avis externe a été un réel plaisir.

Cette expérience a renforcé mon envie de faire partie d’une équipe avec laquelle je peux échanger en face à face. Même si je ne pense pas me spécialiser dans l’IA, ce stage m’a permis de découvrir l’importance de la sécurisation des IA, un sujet qui m’a beaucoup intéressé.

