

1 Introduction

1.1 Objectif

Le traitement du signal permet non seulement de manipuler du son, de la parole et de l'audio, mais aussi :

- ✓ d'analyser des signaux biomédicaux,
- ✓ d'aborder des problématiques dans le secteur de l'aéronautique liées aux signaux sonar et radar, à la localisation et à la navigation GPS, à la poursuite de missiles,
- ✓ d'étudier des systèmes de communications mobiles,
- ✓ d'analyser des structures mécaniques,
- ✓ de mettre en œuvre des systèmes combinant différents capteurs, etc.

Comme la parole est devenue l'une des applications les plus classiques sur Internet, qu'elle ne requiert pas d'expertise *a priori*, mais reste complexe à maîtriser, l'un des objectifs de ce projet en traitement du signal est de vous familiariser avec ces signaux de parole et de mettre en pratique les enseignements fondamentaux de traitement du signal vus en première année et ce premier semestre, notamment le cours de processus aléatoires, de signal continu, de traitement numérique du signal et de filtres numériques et estimation.

Ce projet vous permet de renforcer vos compétences, en vue d'une découverte de cette thématique ou d'une orientation envisagée pour des options de 3^{ème} année telles qu'I2SC.

Le projet sera effectué en binôme.

Le projet sera mené sous *Matlab*.

1.2 Evaluation

L'évaluation du travail repose sur un rapport, un test sur machine et une note de travail continu.

Le rapport d'une dizaine de pages maximum doit être dactylographié (sous *Word* ou *Latex*). **Le rapport en pdf sera déposé sous Moodle.** Les dates de remise seront renseignées sur *Moodle*.

Sous *Word*, les équations doivent être générées avec l'éditeur d'équations et numérotées. Les commentaires doivent être pertinents et tout résultat justifié. Les programmes *Matlab* peuvent être mis en annexe. Des références bibliographiques peuvent être introduites et seront regroupées dans une section en fin de rapport.

Le rapport devra également comporter un bilan de l'organisation et du déroulement du projet, qui comprendra en particulier la liste des tâches entreprises, avec une évaluation du temps passé par chacun lors de chaque séance et entre les séances.

Une attention particulière devra être apportée à l'orthographe et à la rédaction de l'introduction et de la conclusion. L'introduction doit amener le sujet que vous abordez et poser le problème. Elle doit enfin présenter le plan de votre rapport. La conclusion doit dresser un bilan technique du travail mené.

A titre d'exemple, voici quelques retours sur la forme transmis à la promotion l'an passé :

-Attention aux fautes de français ou aux expressions familières qui peuvent être utilisées à l'oral mais doivent être évitées dans un rapport ; dans certains cas, on peut atteindre une trentaine de fautes pour un rapport de quelques pages. Ce n'est pas acceptable. Voici des exemples (à vous de trouver les fautes !) :

- "Ces résultats indique que le rehaussement",
- "Nous avons ainsi étbalit les périodogrammes".
- "la transformée de fourrier".

-Evitez l'emploi de "en fait" et "on a affaire avec" qui restent des marques d'expression orale. Evitez des phrases du genre "maintenant on va essayer de..."

-Il n'y a pas de majuscule après une virgule. Attention à l'usage des espaces avant et/ou après les ponctuations ; des règles existent mais varient selon la langue utilisée.

-A la place de "..." dans une liste ou une itération, écrivez ", etc."

-Evitez l'emploi d'expressions comme « notre signal » ou « nos échantillons ». On dit plutôt « le signal » ou « les échantillons ».

-Les notations doivent être homogènes : si X est la notation utilisée dans une équation, la même mise en forme doivent être utilisée dans le texte si vous souhaitez la mentionner.

-Evitez le copier/coller d'illustrations ou figures qui ont été proposées par d'autres ! Soyez créatifs et personnalisez vos rapports. Evitez cependant des schémas faits sur une feuille puis pris en photo. Vous disposez de logiciels dédiés.

-Les figures doivent comporter un titre et doivent être numérotées. Lorsque vous faites référence à une figure, au lieu de dire "selon la figure ci-dessous", soyez plus précis et indiquez "selon la figure n°..."

-Evitez le mélange de notations propres au code matlab et de notations mathématiques.

Tout plagiat sera sanctionné, Cf. règlement pédagogique de l'école.

Le test sur machine consistera en un ensemble de questions (type QCM) traitant de l'ensemble des notions en traitement du signal (des signaux à temps continu aux signaux à temps discrets, des signaux déterministes aux signaux aléatoires).

1.3 Contacts

Ce projet est encadré par Lucas Steffanutto.

2 *Première phase du projet : extraction d'une signature caractérisant la dépendance long terme.*

2.1 *Introduction*

2.2 *Préambule*

- ✓ Générer un bruit blanc de moyenne nulle et de variance σ^2 ; représenter sa fonction d'autocorrélation théorique et celle estimée à partir de différents estimateurs dont on rappellera les définitions et les conditions d'utilisation. Représenter le spectre de puissance d'une réalisation de ce bruit blanc ; le comparer à la densité spectrale de puissance du bruit. Commenter.
- ✓ Qu'appelle-t-on Périodogramme de Daniell, Périodogramme de Bartlett et Périodogramme de Welch ? Quels sont leurs intérêts ? Programmer ces périodogrammes sous *Matlab* ; les comparer au spectre de puissance et à la densité spectrale de puissance dans le cas d'un bruit blanc Gaussien centré.
- ✓ Qu'appelle-t-on Corrélogramme ? Le programmer sous *Matlab*. Le comparer au spectre de puissance et à la densité spectrale de puissance dans le cas d'un bruit blanc Gaussien centré.
- ✓ Mettre en œuvre un traitement visant à bruiteur un signal de Weierstrass et un signal de parole (fournis sous moodle) avec un bruit additif pour un rapport signal à bruit (RSB) donné.

Développer une fonction *Matlab*. On utilisera la fonction *randn*. On définit le RSB en fonction des échantillons du signal et du bruit additif. Les cas suivants sont à tester : 5 dB, 10 dB et 15 dB. On rappelle que la définition du RSB a été donné en séance de travaux dirigés.

- ✓ Fournir, sur une même figure, la représentation temporelle et le spectrogramme du signal à l'aide des fonctions *spectrogram* et *subplot*, puis sur une seconde figure la représentation temporelle et le spectrogramme du signal bruité. **Faire en sorte que les représentations soient l'une au-dessus de l'autre et que les axes temporels coïncident pour la représentation temporelle et le spectrogramme.** Le temps définit l'axe des abscisses. Il est donc exprimé en secondes. On considèrera dans ce projet une représentation avec des fréquences normalisées.
- ✓ Dans le rapport, commenter l'allure des spectrogrammes.

2.3 *Extraction d'une signature se fondant sur le DFA (Detrended Fluctuation analysis)*

L'exposant de Hurst, classiquement noté H , permet d'évaluer la dépendance à long terme. De ce fait, un processus est dit :

- à mémoire longue si $0.5 < H < 1$,
- à mémoire courte si $H = 0.5$,
- à mémoire négative, ou anti-persistant, si $0 < H < 0.5$.

De plus, H vaut respectivement -0.5 , 0 et 0.5 pour un bruit blanc, un bruit rose et un bruit brownien. L'estimation de cet exposant s'avère pertinente pour la classification de signaux notamment physiologiques.

Dans le cadre de ce projet, nous nous intéresserons à une méthode visant à estimer H . Différentes méthodes ont été développées dans la littérature : certaines opèrent dans le domaine fréquentiel et d'autres dans le domaine temporel. Historiquement, la première

méthode proposée en 1951 est l'approche nommée *rescaled range* (R/S) [1]. Afin d'estimer le coefficient de Hurst d'un processus mono-fractal, la technique nommée « fluctuation analysis » a été proposée en 1992 [2]. Puis, pour traiter tout type de processus, notamment ceux évoluant autour d'une tendance, Peng *et al.* ont proposé en 1994 la méthode nommée DFA pour *detrended fluctuation analysis* [3]. Cette méthode opère de la manière suivante :

Etape 1 : le signal de taille M est tout d'abord centré puis intégré de manière à obtenir ce que l'on appelle le **profil**.

Etape n°2 : le profil est découpé en parties de taille N . La tendance locale de chacune des parties est estimée ; on peut opter pour une droite de régression. La tendance globale est ensuite déduite comme la concaténation des différentes tendances locales. Si le nombre d'échantillons du profil n'est pas un multiple de N , on ne gardera que les LN premiers échantillons, avec L la partie entière de M/N .

Etape n°3 : La tendance globale, linéaire par morceaux, est alors retranchée au profil pour d'obtenir un nouveau processus appelé résidu. La puissance du résidu est notée $F_2(N)$. Calculée pour différentes valeurs de N , elle est alors nommée **carré de la fonction de fluctuation**.

Etape n°4 : Il a été montré que le logarithme de $F_2(N)$ est une fonction affine de $\log(N)$ dont la pente vaut $\alpha=H+1$. Partant d'un ensemble de valeurs pour N et $F_2(N)$, il s'agit alors d'estimer α au sens des moindres carrés.

- ✓ Formaliser mathématiquement l'ensemble des étapes de la méthode DFA. En d'autres mots, donner :
 - les équations visant à centrer puis intégrer le signal pour obtenir le profil.
 - les équations permettant de déduire une droite de régression. Pour cela, on partira d'une formalisation matricielle. A noter que cette démarche est utilisée à la fois à l'étape 2 et l'étape 4 pour différents jeux de données.
 - l'expression de $F_2(N)$.
- ✓ Ecrire le code *matlab* associé et tester ce dernier sur les signaux fournis. Rassembler les résultats obtenus dans un tableau.
- ✓ Etendre l'approche au cas d'une tendance locale qui serait un polynôme de degré 2.
- ✓ Qu'observez-vous quand vous bruitez le signal à l'aide d'un bruit blanc Gaussien centré, avec un SNR de 5 dB ?

Le rapport devra comporter la **représentation du signal initial, du signal centré puis intégré et celle d'une représentation du profil découpé en segments de taille N . Pour une partie de taille N , on représentera les N valeurs du signal d'origine ainsi que les tendances déduites modélisées par une droite de régression (polynôme de degré 1) et par un polynôme de degré 2.** Une figure devra fournir le profil une fois la tendance globale soustraite. Enfin, une figure devra représenter l'étape 4.

- ✓ Tester le code *matlab* sur une réalisation d'un bruit blanc Gaussien et une réalisation d'un bruit rose¹ évoluant le long d'une tendance que l'on définira (par exemple une tendance exponentielle ou logarithmique). Qu'observez-vous pour les valeurs de H ?
 - ✓ On souhaite la pertinence du DFA à des fins de classification de signaux physiologiques. À partir de la base de données d'enregistrements vocaux RAVDESS [4], tester le code *matlab* :
 - Pour un même acteur et pour une même phrase, observer la valeur de H en fonction de l'émotion utilisée par l'acteur pour prononcer cette phrase.
 - Pour une même phrase prononcée avec la même émotion, observer la valeur de H en fonction de l'acteur qui joue et en fonction de si l'acteur est une femme ou un homme.
 - ✓ Selon vous, sur quels autres types de signaux physiologiques serait-il possible de tenter cette approche ? Répéter le même genre de test sur un autre type de signal physiologique (pour essayer de classer des signaux avec une pathologie ou non par exemple).
- [1] H. E. Hurst. Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116:770–799, 1951.
- [2] C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley. Long-range correlations in nucleotide sequences. Nature, 356:168–170, 1992.
- [3] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. Mosaic organization of DNA nucleotides. Physical Review E, 49, (2):1685–1689, 1994.
- [4] « The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) » by Livingstone & Russo is licensed under CC BY-NA-SC 4.0
<https://www.kaggle.com/datasets/uwrfkaggler/ravdess-emotional-speech-audio>

3 *Seconde phase du projet : utilisation d'un grand modèle de langage par le biais d'un chatbot*

3.1 *Introduction*

Dans le cadre de ce projet, nous vous proposons de vous familiariser avec un grand modèle de langage (aussi appelé « LLM » pour Large Language Model en anglais) par le biais d'un chatbot. En particulier, nous vous suggérons d'utiliser Microsoft Copilot, anciennement Bing Chat (voir Annexe 1). D'autres technologies similaires existent comme chatGPT (Voir Annexe 2). Ce dernier requiert de renseigner des données personnelles.

3.2 *Tâches à réaliser avec un chatbot utilisant un LLM*

À l'aide d'un chatbot fondé sur un LLM, créer une interface graphique en Python pour ce projet. Elle doit avoir les options suivantes :

- ✓ Possibilité de charger et d'afficher le signal initial.
- ✓ Présence d'un bouton pour bruite le signal initial avec un certain RSB. Le RSB doit pouvoir être renseigné sur l'interface.
- ✓ Présence d'un bouton pour recharger le signal initial.

¹ Un bruit rose peut être obtenu en filtrant un bruit blanc à l'aide d'un filtre passe-bas dont la réponse en fréquence est correctement choisie.

- ✓ Présence d'un bouton pour afficher le périodogramme du signal actuel (bruité ou non).
- ✓ Présence d'un bouton pour afficher le profil du signal actuel (bruité ou non).
- ✓ Présence d'un bouton pour afficher le découpage du profil en parties et les tendances du profil, en fonction de la taille N de chaque partie (à renseigner dans l'interface).
- ✓ Présence d'un bouton pour afficher le résidu du profil.
- ✓ Présence d'un bouton pour afficher la courbe $F_2(N)$, sa tendance, et le coefficient calculé (α ou H).

À noter que :

- ✓ L'interface graphique ne doit comporter qu'une seule fenêtre dans laquelle il y a la zone d'affichage du signal et les différents boutons à part.
- ✓ Lorsque l'on clique sur l'un des boutons, le résultat doit remplacer le signal en cours d'affichage.
- ✓ Expliquez votre stratégie pour obtenir la solution et combien de prompts avez-vous rédigé ? (10 lignes au maximum)
- ✓ Quel est le prompt final qui vous a permis d'obtenir la solution ?
- ✓ En combien de temps avez-vous réalisé cette partie ?

Annexe 1 : Au sujet de Microsoft Copilot

Microsoft Copilot est un chatbot de Microsoft, fondé sur le grand modèle de langage GPT-4 (expliqué en Annexe 2). Il peut être utilisé avec la plupart des navigateurs en se rendant à l'adresse suivante :

<https://www.bing.com/chat?q=Microsoft+Copilot&FORM=hpcodx>

En cliquant sur ce lien, vous devriez arriver à l'interface présentée à la figure 2.

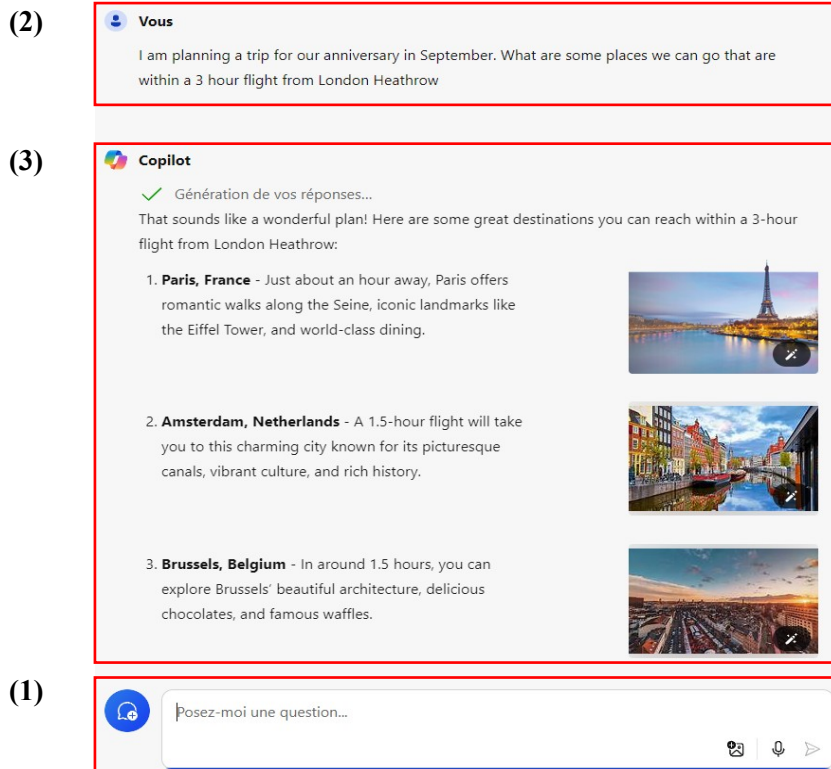


Figure 2 : Interface utilisateur de Microsoft Copilot

En se référant à la figure 2, on identifie :

1. Une barre de saisie de prompt,
2. La question posée à Microsoft Copilot,
3. La réponse de Microsoft Copilot.

Si vous ne l'avez jamais utilisé, posez une question dans la barre de saisie de prompt (1), et observez sa réponse.

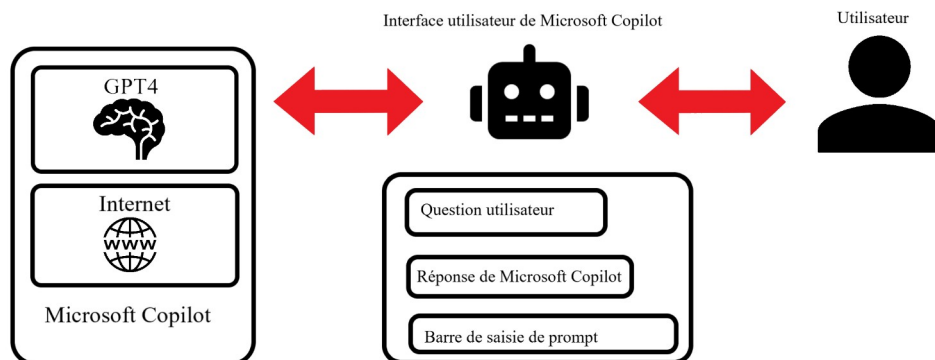


Figure 3 : Résumé de l'architecture de Microsoft Copilot et de ses interactions

Annexe 2 : Au sujet de Chat GPT

A.2. 1 Introduction

Outre Microsoft Copilot, il existe *ChatGPT*. Il s'agit d'une intelligence artificielle générative qui se présente sous la forme d'un agent conversationnel (ou chatbot). Il a été développé par l'entreprise Open AI et a été rendu publique le 30 Novembre 2022.

ChatGPT est fondé sur la technologie de **Transformeur Génératif Pré-entraîné** (GPT pour Generative Pre-trained Transformer). C'est une **famille de grands modèles de langage** (LLM pour Large Language Model), permettant à un **réseau de neurones** d'effectuer des tâches de compréhension et de génération de langage. Les LLMs sont entraînés pour apprendre la syntaxe et la sémantique du langage humain et produire une suite probable à une entrée donnée.

ChatGPT a été entraîné avec des techniques d'apprentissage supervisé et d'apprentissage par renforcement. Pour plus de détails, vous pouvez vous référer au lien suivant :

<https://openai.com/blog/chatgpt>

La base de données sur laquelle il a été entraîné provient d'internet et regroupe sites web, livres, articles, Wikipedia, forums etc. Il a en partie mémorisé ces données. Il possède donc des données sur divers sujets provenant de différentes sources, vérifiées ou non.

On appelle « **prompt** » l'**entrée ou la requête donnée par un utilisateur à un LLM** (GPT-4 pour le cas de *ChatGPT*). L'interface de *ChatGPT* permet de saisir un prompt et de recueillir la réponse du modèle, créer divers contenus tels qu'un mail, un article, un post sur les réseaux sociaux ou bien même du code informatique. ChatGPT peut aussi faire des recherches sur Internet.

Étant seulement au commencement du développement de l'intelligence artificielle générative, il est important de prendre en main cette technologie et de savoir l'utiliser intelligemment et à bon escient.

Attention : *ChatGPT* peut générer du texte incohérent ou imprécis, et peut hériter de biais présents dans les données d'entraînement. De plus, *ChatGPT* se souvient des messages précédents qui lui sont donnés par l'utilisateur au cours d'une même conversation.

Remarque : *ChatGPT* nécessite la création d'un compte et le fait de renseigner des données personnelles (mail et numéro de portable). **Pour des raisons relatives au règlement général sur la protection des données**, Microsoft Copilot qui peut être utilisé sans création de compte est l'option proposée. Ce chatbot de Microsoft est aussi fondé sur le modèle GPT-4 d'OpenAI et est inclus avec le moteur de recherches Microsoft Bing. Microsoft Copilot peut faire des recherches sur internet et cite ses sources.

Remarque 2 : Depuis Mai 2024 le nouveau modèle GPT-4o est disponible. Il est plus performant et est multimodal, c'est-à-dire qu'il peut à la fois traiter des images, des fichiers (pdf, excel, etc), et générer du son et des images. L'option payante permet également de l'utiliser comme assistant vocal. Enfin, d'autres chatbots fondés sur des LLM existent. On compte parmi les plus performants : Claude AI (Anthropic), Gemini (Google Deep Mind), ainsi que d'autres solutions open-source qui peuvent être lancées en local sur un ordinateur comme Llama 3 (Meta), ou encore Mistral (Mistral AI, France).

A.2.2 Prise en main

Les étapes suivantes sont à effectuer :

- ✓ Se rendre à l'adresse <https://chat.openai.com/auth/login>
- ✓ Cliquer sur « sign up » et créer un compte si vous n'en avez pas.
- ✓ Se connecter à son compte.

En vous connectant, vous devriez arriver à l'interface présentée à la figure 3.

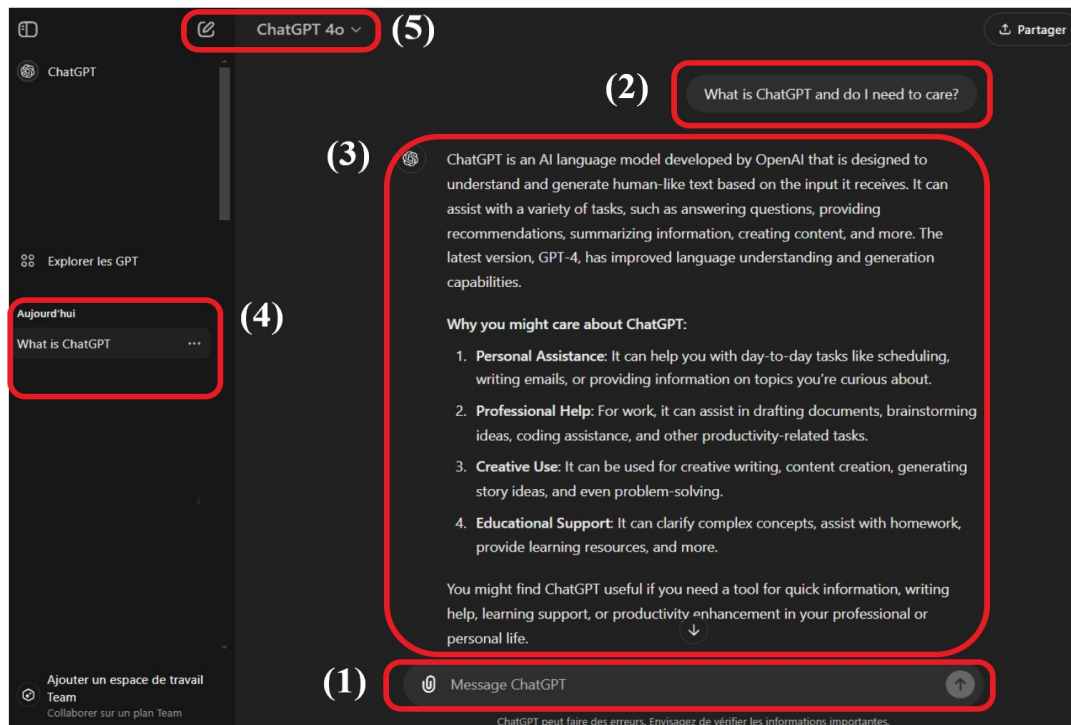


Figure 3 : Interface utilisateur de ChatGPT

En se référant à la figure 3, on identifie :

- (1) La barre de saisie de prompt
- (2) La question posée à ChatGPT.
- (3) La réponse de ChatGPT.
- (4) L'historique des conversations entretenues avec ChatGPT.
- (5) Les autres options de conversation et de choix de modèle.

À ce stade, l'utilisateur peut poser une question dans la barre de saisie de prompt (1).

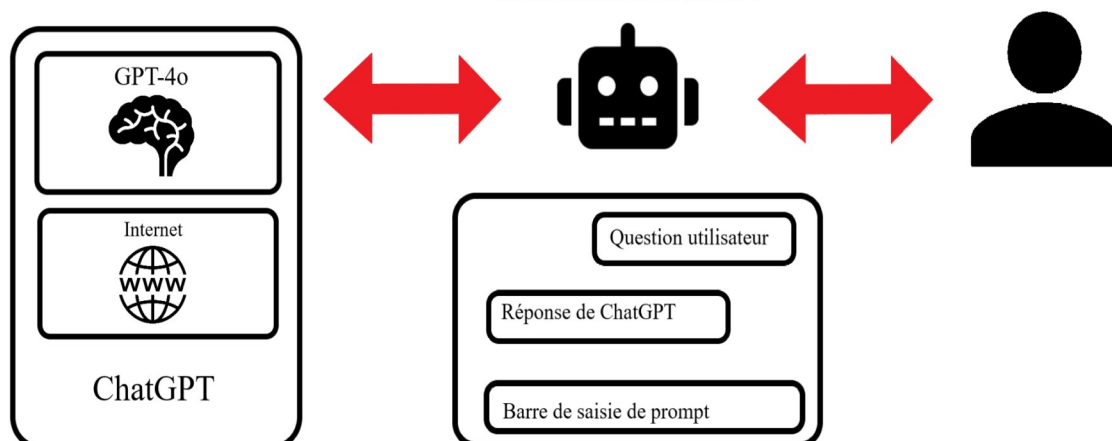


Figure 4 : Résumé de l'architecture de ChatGPT et de ses interactions