

**Faculté des sciences économiques,
sociales, politiques et de communication**

**Stratégies de recherche d'information
numérique à l'ère de ChatGPT et de
vérification de résultats**

Auteur : Hind Abouabdellah
Promoteur : Thibault Philippette
Année académique 2023-2024

Master 120 en Sciences et Technologies de l'Information et de la
communication, à finalité spécialisée en gestion de l'information
numérique

Remerciements

Tout d'abord, je tiens à remercier énormément mon promoteur, Monsieur Philipette Thibault, pour ses conseils et son soutien tout au long de la rédaction de ce mémoire.

Je souhaite également remercier tous les étudiants qui ont accepté de participer à cette étude. Leur coopération et leur engagement ont été indispensables pour la collecte des données et l'analyse des résultats.

Un immense merci à mon mari, pour ses encouragements constants, son soutien moral et sa compréhension tout au long de ces années d'études. Ton amour et ta patience m'ont donné la force de persévérer même dans les moments les plus difficiles.

À mes filles Abla et Yakout, merci pour votre compréhension et votre patience lorsque j'étais occupée par mes études. Votre joie et votre amour m'ont été une source inestimable de motivation.

Je remercie également ma mère pour ses encouragements incessants et son soutien indéfectible. Ton amour et ta foi en moi m'ont donné le courage de poursuivre mes rêves.

Je tiens aussi à remercier mon frère et ma sœur pour leurs conseils et leurs encouragements, surtout dans les moments de faiblesse et de fatigue.

Enfin, ce mémoire est dédié à la mémoire de mon père et de mon frère qui nous ont quitté au cours de ce master. Leur mort a été une épreuve difficile, mais leur souvenir m'a donné la force et la détermination nécessaires pour réussir. Ce travail est un hommage à leur mémoire et à l'inspiration qu'ils m'ont donnée.

Merci à toutes et à tous pour votre soutien et vos encouragements tout au long de cette aventure académique.

Table des matières

1. Introduction	5
2. État de l'art	7
2.1. La recherche d'information RI	7
2.1.1. Définition	7
2.1.2. Histoire et évolution.....	8
2.2. Stratégies de recherche d'information et vérification des résultats	10
2.2.1. Stratégies de recherches.....	10
2.2.2. Vérification des résultats.....	11
2.3. La recherche d'information Conversationnelle	12
2.3.1. Introduction et principes.....	12
2.3.2. Agents conversationnels.....	13
2.3.3. Fouille de texte et fouille de données	15
2.3.4. Fonctionnement.....	17
2.4. ChatGPT	20
2.4.1. Définition	20
2.4.2. Évolution.....	21
2.4.2. ChatGPT et la Recherche d'Information universitaire	23
3. Problématique.....	26
3.1. Question de recherche	26
3.2. Hypothèses	26
4. Méthodologies.....	27
4.1. Questionnaire	27
4.2. Observation	28
4.3. Entretiens	29
4.4. Retours sur la méthodes et justifications	29
5. Résultats.....	30
5.1. Récoltes des données	30
5.1.1. Questionnaire.....	30
5.1.2. Observation.....	31
5.1.3. Entretien.....	32
5.2. Analyse des données.....	33
5.2.1. Questionnaire.....	33
5.2.2. Observation et entretien.....	39
5.3. Discussion des résultats	44
5.3.1. Comparaison utilisation Maroc -Belgique.....	44
5.3.2. Conclusion aux Hypothèses	44
5.3.3. Réponse à la question de recherche	46
6. Limites et perspectives	47
6.1. Limites.....	47
6.2. Perspectives	48
Conclusion	50
Bibliographie.....	52

1. Introduction

Le développement des nouvelles technologies de l'information et de la communication ainsi que l'essor du web, nous ont confrontés à une très grande masse d'informations hétérogènes. Les masses d'informations accessibles n'ont cessé d'augmenter, et les volumes de documents qui les stockent s'accroissent très rapidement. (TH.M.INF.FR.2019.30.pdf, s. d.). Trouver dans cette masse d'information ce que l'on recherche précisément, devient de plus en plus ardu. Le problème n'est plus la disponibilité de l'information mais la capacité de sélection de cette dernière tout en répondant aux besoins précis du chercheur.

Dans notre ère, plusieurs modèles de langages avancés et des outils conversationnels, comme ChatGPT ont été développés offrant un ensemble tout à fait nouveau permettant de générer de réponses rapides, directes et autonomes.

Selon Arthur C. Clark « *Toute technologie suffisamment avancée est indiscernable de la magie.* » (Merveille, 2022). Ceci met la technologie à la même balance que la magie grâce à sa puissance à transformer le monde. Ce qui semblait avant être un fantasme sans espoir est aujourd'hui, grâce à un travail scientifiques et technologiques exceptionnels, une réalité.

L'avènement des technologies numériques et l'explosion des informations disponibles en ligne ont profondément transformé notre manière d'accéder au savoir et de conduire nos recherches. Autrefois dominés par les moteurs de recherche traditionnels, nous voyons aujourd'hui émerger de nouveaux outils grâce au développement de l'intelligence artificielle, notamment les assistants conversationnels qui utilisent le traitement du langage naturel, tels que ChatGPT.

Ce mémoire explore les stratégies de recherche d'informations numériques à l'ère de ChatGPT et de vérification de résultat. L'objectif est d'examiner comment les étudiants de 19 à 26 ans peuvent optimiser leurs méthodes de recherche en utilisant pleinement les fonctionnalités offertes par les systèmes de chat basés sur l'IA, et de comparer ces stratégies de recherche entre des étudiants en Belgique et au Maroc afin d'étendre la recherche sur deux cultures différentes.

Ces jeunes, appartenant à ce qu'on appelle la génération numérique, ont grandi entourés de technologies numériques et sont familiers avec les smartphones, les applications et les plateformes en ligne. Comprendre comment ils utilisent efficacement ChatGPT dans leur recherche d'informations quotidienne est fondamental pour voir l'impact des agents

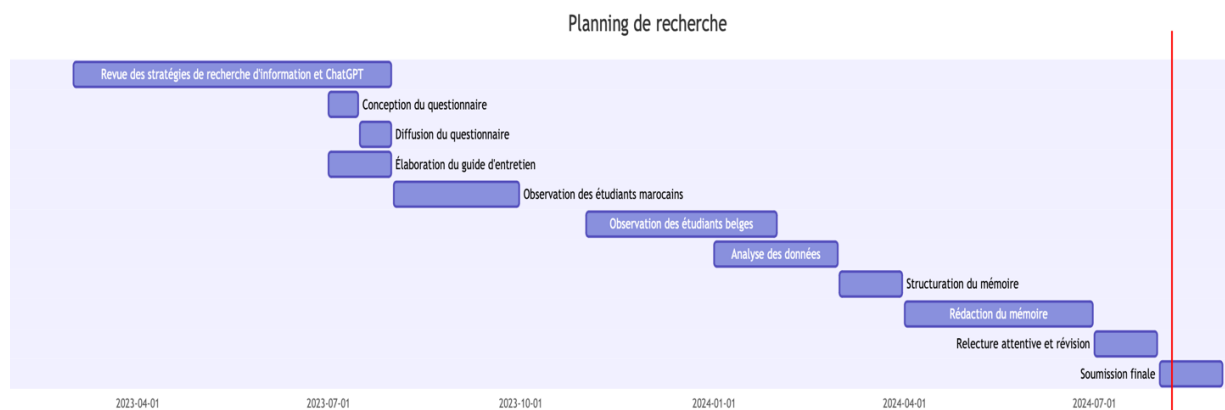
conversationnel et plus particulièrement ChatGPT sur le futur et la qualité de l'enseignant et de la recherche.

Certes, les agents conversationnels tels que ChatGPT ont été réalisés pour aider et faciliter certaines tâches mais il est aussi important de reconnaître et de traiter les défis liés à leur utilisation. Ces défis peuvent inclure la fiabilité et la crédibilité des réponses, ainsi que des problèmes beaucoup plus complexes liés à l'analyse et l'esprit critique des chercheurs.

La présente recherche prend pour objectif l'établissement d'un État des lieux de la recherche d'information, comprendre son évolution et comment on est passés de la recherche d'information classique à la recherche d'information conversationnelle.

En usant du cadre conceptuel offert par les Chatbots et en l'appliquant à un questionnaire et à une observation suivie des entretiens semi-directifs, ce mémoire vise non seulement à comprendre comment les jeunes effectuent leur recherche d'information en utilisant les agents conversationnels, mais aussi à voir leurs stratégies de vérification de résultats obtenus via ces agents conversationnels

Pour mener à bien cette recherche, la figure ci-dessus présente les différentes étapes que nous avons suivies pour réaliser cette recherche :



2. État de l'art

La recherche d'information est un domaine qui existe depuis l'existence de l'information. En effet, dès l'apparition de livres de plus en plus riches, le besoin de trouver des informations précises a été étudié. Il s'agit donc de réussir à trouver de l'information, qui peut être du texte, des sons, ou même des images, au sein d'un ensemble de textes, ou même d'un corpus. Avec l'apparition de l'informatique, l'automatisation de ce domaine est devenue inévitable et essentielle. Cette évolution a été largement facilitée par l'avènement d'Internet et la démocratisation des technologies numériques, modifiant profondément les méthodes de recherche et d'accès à l'information. Dans ce contexte, les outils conversationnels, tels que ChatGPT, émergent comme des innovations prometteuses, capables de révolutionner encore davantage les stratégies de recherche d'information. Cet état de l'art vise à explorer les origines de la recherche d'information, à comprendre l'impact de la technologie sur ces pratiques et à examiner comment l'introduction d'outils conversationnels comme Chatgpt pourrait remodeler les méthodes de recherche universitaires, sans oublier la fiabilité des données recueillies via cet outil.

2.1. La recherche d'information RI

2.1.1. Définition

La recherche d'information (RI) selon Dinet & Rouet « *est l'activité d'un individu qui vise à localiser et traiter une ou plusieurs informations au sein d'un environnement documentaire complexe, dans le but de répondre à une question ou de résoudre un problème.* »(Dinet & Rouet, 2002)

Et selon Chiaramella et Mulhem, « *la RI a pour thème central l'étude de modèles et systèmes d'interaction entre des utilisateurs humains et des corpus de documents numériques, en vue de la satisfaction de leurs besoins d'information* ». (Chiaramella & Mulhem, 2007a)

La recherche d'information est présente dans tous les domaines, allant de l'académie et la recherche scientifique au quotidien des individus cherchant des informations sur Internet. Elle commence par l'identification de l'information à recherché, la formulation de requêtes¹, la recherche, l'évaluation des résultats obtenus, et enfin, l'utilisation de l'information recueillie.

¹ Requête : Une requête constitue l'expression du besoin en information de l'utilisateur. Elle représente l'interface entre le système de Recherche d'Information et l'utilisateur.

Ceci, couvre tous les processus, techniques et technologiques qui facilitent la localisation, l'extraction et l'organisation de l'information depuis le web et diverses bases de données, pour satisfaire les demandes des utilisateurs. S'appuyant sur des algorithmes avancés, de l'intelligence artificielle et des techniques de traitement du langage naturel(TALN), la recherche d'information permet de trier, d'indexer et de retrouver les données les plus pertinentes dans un vaste monde d'informations.

2.1.2. Histoire et évolution

La recherche d'information est un domaine historiquement lié aux sciences de l'information et à la bibliothéconomie qui visent à représenter des documents dans le but d'en récupérer des informations, au moyen de la construction d'index. Avec le développement de l'Informatique, l'Information est devenue numérique. Donc la RI est devenu une préoccupation primordiale de l'Informatique. Cette dernière a permis le développement d'outils pour traiter l'information et optimiser la RI Par l'ouverture de l'information numérique au grand public, le développement de système de RI est devenu nécessaire.

L'évolution du domaine de la recherche d'informations depuis ses origines jusqu'à l'ère de l'intelligence artificielle a suivi un chemin fascinant, qui reflète les progrès technologiques et les changements dans notre manière d'accéder et de gérer les connaissances.

- **1940-1950** : En 1945, Vannevar Bush² a initié à l'idée de **la recherche automatisée** dans les grands dépôts de documents. Trois ans plus tard, et en 1948 le terme « Recherche d'information ou (information retrieval) en anglais» a été inventé par Calvin Mooers (Mooers, 1948) pendant qu'il préparait son mémoire de maitrise dans MIT, reconnaissant l'importance de développer des systèmes capables de retrouver des informations spécifiques au sein d'ensembles de données vastes et complexes(Chiaramella & Mulhem, 2007b). cette période a marqué la grande expansion de la recherche d'information surtout avec l'arrivée des ordinateurs de seconde génération des premiers supports magnétiques³. En 1955, James Perry a réalisé une machine (WRU⁴ Searching Selector) permettant de réaliser jusqu'à dix requêtes

² V. Bush est par ailleurs également considéré comme le père de la notion d'hypertexte... l'ancêtre direct de l'hypermédia et du Web

³ Les premières bandes magnétiques sont apparues en 1951 comme périphériques des ordinateurs UNIVAC I

⁴ WRU : Western Reserve University – Cleveland (US). Il est intéressant de noter que ce projet fut financé par l'industrie (the American Society for Metals).

booléennes en un seul passage sur la bande magnétique (Chiaramella & Mulhem, 2007c).

- **1960** Cette décennie a marqué l'avènement des premiers SRI⁵ automatisés grâce à l'essor des ordinateurs (Jacquinot-Delaunay, 2001) et aussi la fondation la recherche d'information comme domaine scientifique et ce, grâce aux travaux et aux contributions importantes de Gérard Salton, professeur à Cornell University. Il a été parmi les premiers à proposer une approche « modèle » du domaine, qu'il a amplement popularisée grâce à son modèle vectoriel (vector space model) et le système SMART qui est parmi les premiers systèmes de recherche, qui a introduit de nombreuses techniques encore utilisées jusqu'à ce jour (Haddad, s. d.)
- **Les Années 1970-1980** : Des modèles probabilistes ont été introduits, visant à améliorer la pertinence des résultats en utilisant des probabilités pour estimer la pertinence d'un document donné (Rodhain et al., 2010). Aussi, le développement des techniques d'indexation automatique a permis de gérer et de rechercher de grandes quantités de textes.

La naissance de l'informatique documentaire a aussi marqué cette période et elle a connu un énorme développement poussant les spécialistes des sciences de l'information à creuser dans ce domaine pour proposer ainsi des systèmes plus développer. (Nie, s. d.).

- **Les Années 1990** : La croissance exponentielle d'Internet, des moteurs de recherche comme AltaVista, Yahoo! et Google ont révolutionné la manière dont les informations étaient recherchées et récupérées. Dans cette décennie, Google a introduit l'algorithme PageRank, qui classe les pages web en fonction de leur importance, déterminée par le nombre et la qualité des liens entrants.
- **Les Années 2000** : Les moteurs de recherche ont commencé à utiliser des données utilisateur pour personnaliser les résultats de recherche. Une révolution qui a marqué cette période c'est le Big Data. L'augmentation massive des données disponibles a conduit au développement de nouvelles techniques pour traiter et rechercher efficacement de grandes quantités d'informations.
- **Les Années 2010** : L'intégration de l'apprentissage automatique et des modèles de traitement du langage naturel (NLP) a permis de créer des systèmes de recherche beaucoup plus sophistiqués et précis.

⁵ Systèmes de recherche d'information

Recherche Sémantique: Les moteurs de recherche ont commencé à comprendre le contexte et l'intention derrière les requêtes, améliorant ainsi la pertinence des résultats.

- **Les Années 2020 et au-delà :** L'émergence de modèles de langage comme BERT, ChatGPT, et leurs successeurs a transformé la manière dont les systèmes de recherche comprennent et génèrent des réponses. Aussi, l'intégration de différents types de données (texte, image, vidéo) dans les systèmes de recherche ont révolutionnés la recherche d'information.

L'histoire de la RI est donc celle d'une évolution constante, depuis les catalogues manuels des bibliothèques jusqu'à l'intelligence conversationnelle, reflétant les changements dans la manière dont nous accédons et gérons les informations dans un monde de plus en plus numérisé.

2.2. Stratégies de recherche d'information et vérification des résultats

2.2.1. Stratégies de recherches

L'accès à l'information est un domaine très vaste car pour rechercher une information, de nombreuses disciplines peuvent être abordé, comme la documentation, les sciences de l'information, l'informatique documentaire, mais aussi la psychologie cognitive, l'économie de l'information, la communication des organisations, etc. (Zimmer et al., 2007) et aujourd'hui les outils conversationnels.

Pour savoir les méthodes de recherche d'information, il faut d'abord savoir que les sources d'information sont classées en trois catégories : primaire, secondaire et tertiaire.

- Les sources primaires : sont les documents originaux comme, des lettres, des documents officiels, des données brutes, etc. Ces sources sont utilisés dans le cadre de la recherche en histoire, et d'autres domaines académiques.
- les sources secondaires : ce sont les documents qui analysent, synthétisent ou commentent des matières provenant des sources primaires. On peut citer comme exemple des livres, des revues scientifiques, etc.
- Les sources tertiaires : ce sont des ouvrages de référence où sont synthétisés et rassemblés les informations des sources primaires et secondaires.

Par ailleurs, il existe trois méthodes essentielles de recherche d'information :

- Une première méthode de recherche est la recherche bibliographique qui permet de trouver des publications pertinentes sur un sujet spécifique. Cela implique d'explorer des catalogues de bibliothèques, des bases de données et des index pour identifier des livres, des articles de revues, des thèses et d'autres documents. (Karim & El Hassouni, 2023)
- Une deuxième méthode est la méthode la plus connue et la plus utilisée c'est la recherche d'information en ligne qui consiste à utiliser Internet pour trouver des informations adéquates. Cela peut inclure l'utilisation de moteurs de recherche comme google pour effectuer des recherches générales, mais également l'exploration de bases de données en ligne centrées sur des domaines et des sujets particuliers.(Akoka et al., s. d.)
- Une autre méthode est la recherche dans les bases de données. Cette méthode est la plus fiable pour trouver des informations correctes et ciblées. Cette méthode offre un accès à une grande quantité de sources d'information académique, scientifique, et technique, ce qui en fait un outil précieux pour la recherche. Pour une recherche plus satisfaisante, il est essentiel de connaître les bases de données spécifiques à chaque domaine d'étude, afin d'optimiser les recherches.

2.2.2. Vérification des résultats

La crédibilité des sources d'information et des données obtenues est un sujet qui a préoccupé les chercheurs depuis l'avènement d'Internet et qui a connu un énorme développement à travers le temps.(Guével & Pommier, 2012)

En effet, l'évaluation et le jugement de la crédibilité d'un document ou d'un résultat d'une recherche est un problème très complexe mais indispensable pour garantir la qualité et la crédibilité des résultats « *On peut trouver un document intéressant mais trop ancien, tel autre complet mais trop compliqué, celui-là bien écrit et bien documenté, ou encore trop court ou trop simplifié mais disponible dans un format imprimable, un autre intéressant parce qu'il contient une interview, etc.* »(Boubée & Tricot, 2010).

Ne pas vérifier un résultat peut mener à des erreurs et donc compromettre la qualité du travail ou de la recherche réalisée. Pour bien mener cette recherche, des critères d'évaluation de l'information sont essentiels pour garantir la qualité et la fiabilité des sources. Parmi ces critères, on retrouve la pertinence, la précision, la fiabilité, la crédibilité, la date de publication,

l'objectivité, la cohérence et la validité des informations(Renard & Roussiau, 2007). Ces critères permettent d'effectuer un examen approfondi de la documentation, qu'elle soit issue de sources primaires, secondaires ou tertiaires.

Mais est ce que ces critères peuvent être vérifiés dans le cas de la recherche d'information conversationnelles ou les résultats de recherche sont présentés sans sources, sans date, et sans auteur ?

2.3. La recherche d'information Conversationnelle

2.3.1. Introduction et principes

Une nouvelle méthode de la recherche d'information, c'est la recherche d'information conversationnelle (RIC) à l'aide des agents conversationnels ou (Chabots en anglais).

La Recherche d'Information Conversationnelle est une discipline émergente qui se concentre sur l'amélioration de l'interaction entre l'utilisateur et les systèmes de recherche d'information par le biais de dialogues naturels. Connaître le contexte et les intentions d'un utilisateur ne peut pas être exprimé uniquement par le biais de mots-clés dans une requête. Au lieu de cela, la RIC intègre cet aspect en posant plusieurs questions pour comprendre ce dont le demandeur a besoin de savoir.

Ces systèmes de RIC reposent sur des technologies et des algorithmes avancées de traitement automatique du langage naturel (TALN), d'apprentissage automatique et de l'intelligence artificielle pour analyser et interpréter la complexité du langage humain (Aissa et al., 2019) . Pour cela, ils doivent gérer plusieurs requêtes afin de garantir des résultats identique au langage humains. Parmi ces requêtes on cite: l'ambiguïté linguistique, la reconnaissance des questions implicites, et même l'identification des émotions ou les intentions non exprimées directement dans les requêtes. Aussi, la capacité à mener des dialogues cohérents et pertinents avec l'utilisateur est une priorité, ce qui peut le pousser à poser des questions de clarification .

L'un des défis majeurs dans les RIC est de parvenir à une compréhension profonde non seulement du langage mais aussi de perfectionner dans certains domaines comme la médecine, la programmation etc. Cela exige une grande quantité de données d'entraînement et une expertise dans le domaine concerné pour affiner les modèles d'apprentissage automatique et le rendre capable de fournir des réponses ciblées et précises.

En outre, la personnalisation est un enjeux majeur dans l'amélioration de l'efficacité de la RIC. En s'adaptant aux préférences individuelles des utilisateurs tout en enregistrons leurs interactions passées, ce qui peut produire des réponses plus pertinentes et personnalisées. Cette

capacité d'adaptation est essentielle pour créer une expérience utilisateur enrichissante, qui rend la recherche d'information non seulement plus productive mais également plus engageante.

les principes fondamentaux de la RIC sont:

-La compréhension du langage humain : les technologies de la RIC permettent de comprendre les requêtes formulées en langage naturel, et exprimer ainsi les besoins d'information comme dans le cas d'une conversation humaine.

-L'interaction dynamique : La RIC privilégie une interaction bidirectionnelle et dynamique, où le système peut poser des questions pour clarifier les besoins de l'utilisateur.

- L'apprentissage continu : Les systèmes de RIC utilisent l'apprentissage automatique pour s'améliorer au fil du temps, apprenant des interactions passées pour fournir des réponses précises et pertinentes(Risser-Maroux & Kerboua-Benlarbi, 2018).

2.3.2. Agents conversationnels

L'émergence de la recherche d'information conversationnelle a ouvert la voie à des interactions plus naturelles, où les utilisateurs peuvent formuler leurs requêtes comme s'ils conversaient avec un humain. Cette évolution vers des interactions plus naturelles et intuitives trouve son apogée dans le développement des chatbots.

Avant d'aborder le sujet de ChatGPT, il est essentiel de comprendre le contexte historique et l'évolution des agents conversationnels, car cela fournit une toile de fond pour apprécier pleinement les innovations et les avancées que ChatGPT représente dans ce domaine.

Bien que **Turing** lui-même n'ait pas développé de chatbots, son test a établi un cadre conceptuel pour l'évaluation des systèmes d'IA conversationnelle et a inspiré la création des premiers agents conversationnels. Les chercheurs en IA et en TALN ont vu dans le test de Turing un défi stimulant, cherchant à développer des systèmes capables de passer ce test en engageant des humains dans des dialogues crédibles et naturels.

En 1966, Joseph Weizenbaum, professeur germano-américain au MIT, a créé ELIZA(Wilson & Marasoiu, 2022), considéré comme le véritable premier agent conversationnel. ELIZA a été conçue pour imiter le comportement d'un psychothérapeute en utilisant une interface de dialogue, incitant ainsi les utilisateurs à partager leurs pensées par écrit, auxquelles ELIZA répondait. Certains individus ont même développé un lien émotionnel avec cette machine, suggérant ainsi la réussite potentielle d'un test de Turing. Cependant, ELIZA adopte une approche simpliste en se basant uniquement sur la détection d'un mot-clé (*Une brève histoire des chatbots – Partie 2/2 – Les chatbots célèbres*, s. d.-a) dans

la phrase de l'utilisateur pour générer une question en retour. Cette méthode démontre clairement ses limites en termes de complexité et de compréhension des interactions humaines.

ELIZA a été une source d'inspiration pour de nombreux scientifiques, même des décennies plus tard, comme l'informaticien américain Richard Wallace, qui a créé **ALICE** en 1995 (AbuShawar et al., 2015) en adoptant le concept de Joseph Weizenbaum pour améliorer le modèle. ALICE est ainsi en mesure d'identifier des mots-clés dans les échanges avec les utilisateurs, tout en se basant sur des règles plus complexes. Ainsi, le chatbot peut interagir avec un individu en engageant une conversation, au-delà de simplement répondre aux questions contrairement à ELIZA. Grâce à son succès, Alice a été récompensée trois fois par un prix Loebner⁶.

En l'an 2001, le chatbot Eugene Goostman doté d'intelligence artificielle, a été développé par trois informaticiens russes un programme celui-ci a été conçu pour imiter le comportement d'un jeune garçon ukrainien de 13 ans. Plus tard, en 2014, des chercheurs de l'université de Reading ont décidé de soumettre ce chatbot au célèbre test de Turing, (Warwick & Shah, 2016). L'expérience a impliqué 30 examinateurs qui ont été invités à engager une conversation avec le chatbot pendant 5 minutes. Les chercheurs ont établi que le test serait considéré comme réussi si plus de 30 % des examinateurs étaient trompés par le chatbot. Finalement, le résultat a dépassé les attentes en atteignant un taux de réussite de 33 %. Eugene Goostman est le premier programme informatique à avoir réussi le test de Turing, un test de l'intelligence artificielle qui évalue la capacité d'une machine à imiter la conversation humaine de manière indiscernable. (*Une brève histoire des chatbots – Partie 2/2 – Les chatbots célèbres*, s. d.-b).

Une évolution récente dans le domaine des chatbots est l'émergence des assistants vocaux. Ces programmes révolutionnaires sont conçus pour interpréter des instructions données à l'oral et fournir des réponses à travers une voix synthétique. Cette technologie a connu une popularité croissante, Apple étant le précurseur en 2011 avec Siri, intégré notamment dans

⁶ Le **prix Loebner** est une compétition annuelle qui couronne les dialogueurs satisfaisant le mieux les critères du test de Turing (avec lesquels il est le plus difficile de déterminer s'il s'agit d'un robot ou d'un humain).

ses iPhone, suivi par Amazon et son assistant Alexa, Google avec Google Assistant, ainsi que Microsoft et son assistant Cortana.

Au fil des décennies, le développement des chatbots a été guidé par l'objectif du test de Turing. Cela a conduit à l'intégration de technologies avancées, comme l'apprentissage automatique, le traitement du langage naturel, et les modèles de langage basés sur des réseaux de neurones profonds. Grâce à l'intégration de modèles de langage de plus en plus sophistiqués, tels que ceux utilisés dans **ChatGPT**, les chatbots actuels sont capables de mener des dialogues longs et complexes, offrant ainsi une expérience utilisateur enrichie et intuitive.

Les agents conversationnels peuvent adopter une différentes architectures de programmation et d'approches de conception en fonction des objectifs qui leur sont imposées. Certains de ces agents ont des réponses programmées à l'avance pour chaque type de requête introduites par l'utilisateur. En revanche, d'autres agents, plus avancés et plus développés, utilisent des modèles de langage préentraînés(LLM). Ces modèles peuvent apprendre de manière automatique et non supervisée à partir d'un grand ensembles de données textuelles ce qui leur permettra de donner des réponses personnalisées et similaires aux langage humain.

2.3.3. Fouille de texte et fouille de données

Pour clarifier les technologies utilisées dans les chatbots, il est évident d'expliquer deux techniques qui paraissent différentes, mais elles visent un but commun qui est de convertir les données en résultats pratiques se sont la fouille de texte et la fouille de données. En effet, bien que la recherche d'information facilite l'identification de l'information, la fouille de texte et de données révèlent des motifs, des tendances et des liens au sein de ces informations. Selon David Hand « *Data Mining consists in the discovery of interesting, unexpected, or valuable structures in large data sets*» (*Memoire Online - Une contribution du datamining à fÂ la segmentation du marché et au ciblage des offres à fÂ l'aide de la statistique de Bayes - samuel Kalombo, s. d.*) . La fouille de données, également connue sous le nom de data mining, désigne la démarche analytique qui vise à explorer de manière automatique de vastes ensembles de données dans le but de mettre en lumière des tendances et des schémas non évidents à travers une analyse approfondie. Cette pratique est fréquemment associée au Deep Learning et au machine learning, qui sont des techniques complémentaires permettant d'approfondir la compréhension des données et d'optimiser les résultats obtenus.

Data Mining utilise des algorithmes mathématiques sophistiqués pour segmenter les données et évaluer la probabilité d'événements futurs. Data Mining est également connue sous le nom de découverte de connaissances dans les données « *Knowledge Discovery in Data (KDD)* ». (Maimon & Rokach, 2010)

Les propriétés clés de Data Mining sont :

- Découverte automatique des motifs
- Préviation des résultats probables
- Création d'informations exploitables
- Mettre l'accent sur les grands ensembles de données et les grandes bases de données(*Qu'est-ce que le Data Mining, s. d.*).

La fouille de données permet de découvrir des informations cachées dans de grandes quantités de données et en transformant ces données en connaissances exploitables. La fouille de données, en synergie avec la recherche d'information (RI), évolue comme un domaine clé dans le traitement et l'analyse de l'immense volume de données généré dans divers contextes, notamment scientifique. L'intégration de la fouille de données dans la RI vise à exploiter efficacement les données disponibles, y compris les publications et les données expérimentales, pour en extraire des connaissances actionnables et pertinentes. La fouille de texte, ou text mining, est une branche de la fouille de données qui fait l'extraction d'informations utiles à partir de grandes quantités de données textuelles. Elle combine les techniques du traitement automatique du langage naturel (TALN), de la statistique, et de l'apprentissage automatique pour transformer le texte non structuré en données structurées ou en insights significatifs.

« *La fouille de textes (aussi nommée forage de textes, ou encore text mining) est un nouveau domaine de recherche interdisciplinaire dont l'objectif général est le développement de techniques, d'algorithmes, d'applications informatiques et de méthodologies applicatives afin d'extraire et de structurer automatiquement, ou semi-automatiquement, de nouvelles informations à partir de grands corpus de documents textuels non structurés ou semi-structures* »(*Forest, 2009*).

Ainsi, la fouille de texte transforme l'information en une ressource prête à l'emploi chose qui est particulièrement pertinente à l'ère du Big Data, où la quantité d'informations textuelles

disponibles dépasse largement la capacité humaine de les analyser. En creusant les technologies utilisées dans les chatbots dans le chapitre suivant nous explorerons comment ces agents conversationnels intelligents repoussent les limites de la recherche d'information, en utilisant la fouille de texte et la fouille de données pour répondre de manière dynamique et personnalisée aux besoins des utilisateurs.

2.3.4. Fonctionnement

Grace à l'intégration de l'intelligence artificielle et de ses sous technologies, les chatbots sont devenus plus sophistiqués et leurs résultats sont de plus en plus adaptés aux besoins des utilisateurs ce qui a changé complètement nos méthodes de recherche et d'accès à l'information. Dans ce chapitre de notre mémoire, nous allons voir de plus près les différentes technologies que utilisent les chatbots pour nous donner des résultats aussi fascinants et ainsi transformer notre stratégie de recherche d'information en transformant notre manière d'interagir avec le monde numérique.

2.3.3.1. TALN

« On regroupe sous le vocable de traitement automatique du langage naturel (TALN) l'ensemble des recherches et développement visant à modéliser et reproduire, à l'aide de machines, la capacité humaine à produire et à comprendre des énoncés linguistiques dans des buts de communication. »(Yvon, s. d.)

« Le Traitement Automatique des Langues (TAL) ou Traitement Automatique du Langage Naturel (TALN) vise à appliquer des méthodes informatiques pour exploiter des ressources exprimées en langue naturelle. C'est une discipline à l'interface entre la linguistique, l'informatique et l'intelligence artificielle. » (Le TALN appliqué aux chatbots, 2018)

Ainsi, le TALN utilise des algorithmes et des modèles d'apprentissage automatique pour analyser, comprendre et générer du texte dans une langue naturelle. TALN se consacre à la compréhension du langage humain et la génération d'un texte identique à celui des humains, et pour le faire le TALN suit cinq étapes fondamentales(Yvon, s. d.) :

- **la Tokenisation** : Pour toute analyse automatique de texte, il est primordial de procéder à sa segmentation en unités linguistiques(Mots, Phrases, Paragraphes).
- **L'Étiquetage (morpho-)syntaxique** : Après la tokenisation, l'étiquetage syntaxique

permet de marquer chaque token avec des informations grammaticales, ce processus étant connu sous le nom d'étiquetage syntaxique. Cette étape vise à assigner une catégorie grammaticale (par exemple nom, verbe, adjectif) à chaque token. Lorsque des indications sur le genre et le nombre sont également fournies, on parle alors d'étiquetage morphosyntaxique. Ce marquage initial permet de résoudre une première ambiguïté entre un nom et un verbe conjugué qui sont homonymes.

- **La lemmatisation** : cette étape, permet d'associer à chaque mot sa forme neutre.
- **l'analyse en dépendances** : Elle permet d'identifier les relations syntaxiques et grammaticales qui lient les différents mots d'une phrase
- **la désambiguïstation lexicale** : Permet de choisir quelle interprétation de phrase choisir parmi plusieurs . « (...)l'objectif est de prédire automatiquement le sens des mots en contexte, à partir d'un inventaire de sens prédéfini. »(Segonne, 2021)

2.3.3.2. Intelligence Artificielle (IA)

Les Chatbots comme ChatGPT, se basent sur des technologies avancées pour offrir des résultats semblable au langage humain. Parmi ces technologies, la compréhension du langage naturel (NLU) l'apprentissage profond, l'apprentissage automatique et les réseaux de neurones pour déterminer ce que l'utilisateur essaie d'accomplir. Cela améliore leur capacité à prédire les besoins des utilisateurs avec précision et à répondre correctement au fil du temps.

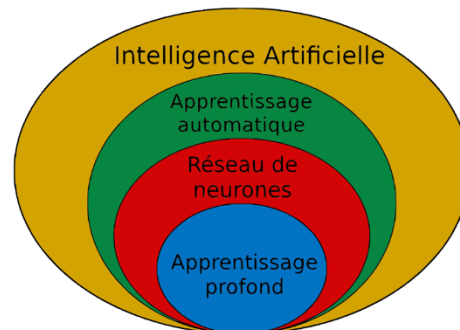


Figure 1: Décomposition du domaine de l'IA (Matteis, s. d. -a)

- La compréhension du langage naturel (NLU)

Le Natural Language Understanding (NLU), fait partie de l'intelligence artificielle et du traitement automatique du langage naturel (TALN), il permet aux chatbots de comprendre et d'interpréter le langage humain. Sa particularité réside dans sa capacité à décoder non seulement le sens littéral des phrases, mais aussi les émotions qui y sont véhiculées. (McShane, 2017).

Selon Raphael Kassel, « (...)le NLU est capable de décrypter les connotations émotionnelles, les intentions sous-jacentes et les objectifs exprimés à l'écrit ou à l'oral. En plus, le natural language understanding est programmé pour comprendre le sens, malgré les erreurs humaines courantes (comme les fautes de prononciation, les fautes d'orthographe ou les transpositions de lettres et de mots) » (Kassel, 2024).

- **La Génération de Langage Naturel (GLN)**

est une branche TALN permettant la transformation des données structurées en texte par l'humain. En évaluant les motifs dans les données, les systèmes de GLN peuvent générer du contenu significatif (*Au cœur du développement de la génération de contenu IA*, 2023) qui s'aligne avec des directives et des exigences de style spécifiques.

- **Apprentissage Automatique**

« (...) l'apprentissage automatique est un champ de l'IA porté sur l'analyse statistiques de données d'apprentissage. Historiquement, cette branche est définie comme le développement de machines capables d'apprendre sans avoir été explicitement programmées à apprendre une tâche. »(Matteis, s. d.-b)

- **Les Réseaux de Neurones**

Les réseaux de neurones d'intelligence artificielle sont inspirés du cerveau humain, ils analysent les entrées des utilisateurs (texte ou voix) pour en déduire l'intention et le contexte. Ces réseaux de neurones permettent l'entraînement d'un chatbot de manière automatique en l'alimentant de données. En effet, pour former un réseau de neurones artificiels, la meilleure façon est de laisser le système deviner, de recevoir un retour d'informations et de deviner, en changeant continuellement les probabilités, quelle approche permet d'obtenir la bonne réponse. En tant que systèmes inspirés du cerveau et conçus pour reproduire la façon dont les humains apprennent, les réseaux de neurones modifient leur code pour trouver le lien entre les entrées et les sorties dans des situations où cette relation est complexe ou peu claire.

Un autre avantage des réseaux de neurones est qu'ils apprennent des expériences passées ce qui leur pousse à ajuster constamment leurs paramètres pour améliorer la précision du modèle et donc plus ils sont utilisés, plus ils s'améliorent.

- L'Apprentissage Profond (Deep Learning)

Le deep learning, ou apprentissage profond, est un sous-domaine de l'intelligence artificielle qui repose sur l'utilisation de réseaux de neurones complexes composés de plusieurs couches pour analyser et traiter des données. Dans le cas des chatbots, le deep learning permet de mieux comprendre les subtilités du langage et de générer des réponses plus naturelles et plus adaptées. Grâce à leurs entraînements sur de vastes quantités de données, ces chatbots apprennent et s'améliorent constamment ce qui leur permet de produire des réponses de plus en plus pertinentes au fil du temps (Kelleher, 2019).

En intégrant l'intelligence artificielle et le deep learning dans le développement des chatbots, on obtient une amélioration continue des performances, une personnalisation des interactions et une meilleure compréhension du contexte conversationnel. ChatGPT, illustre parfaitement cette avancée en combinant des modèles de langage sophistiqués avec des techniques d'apprentissage profond pour créer des échanges qui imitent de manière convaincante la communication humaine.

2.4. ChatGPT

Dans ce chapitre de notre mémoire, on va explorer ChatGPT, une innovation née suite au croisement entre l'IA, le TALN et la recherche d'information numérique conversationnelle. Ce chapitre sert de pont entre les discussions techniques et théoriques sur les chatbots et l'application concrète de ces technologies dans des contextes réels.

2.4.1. Définition

ChatGPT est un agent conversationnel développé par OpenAI. « Chat » de ChatGPT fait référence à la discussion en ligne, tandis que « GPT » signifie « Generative Pre-trained Transformer ».

« Le Generative Pre-trainer Transformer (GPT) est un modèle de deep learning de génération de texte entraîné sur les données disponibles sur Internet. Il peut répondre des questions, résumer un texte, traduire ou même de générer un code informatique. » (Robert, 2024)

ChatGPT est un modèle de langage basé sur l'architecture Transformer lui permettant de comprendre et générer du texte de manière à imiter le langage humain, permettant diverses interactions conversationnelles avec les utilisateurs. L'une des forces de ChatGPT c'est sa capacité à faire de multiples (Bersini, 2023) tâches allant de la génération du texte à la programmation, la traduction, le résumé etc. ChatGPT se distingue des anciens agents

conversationnels car il associe la technologie de traitement du langage naturel (TALN) à des algorithmes d'apprentissage automatique pour générer des réponses semblables à celles des humains, il utilise un texte prédictif pour générer des conversations dynamiques basées sur le contexte.

2.4.2. Évolution

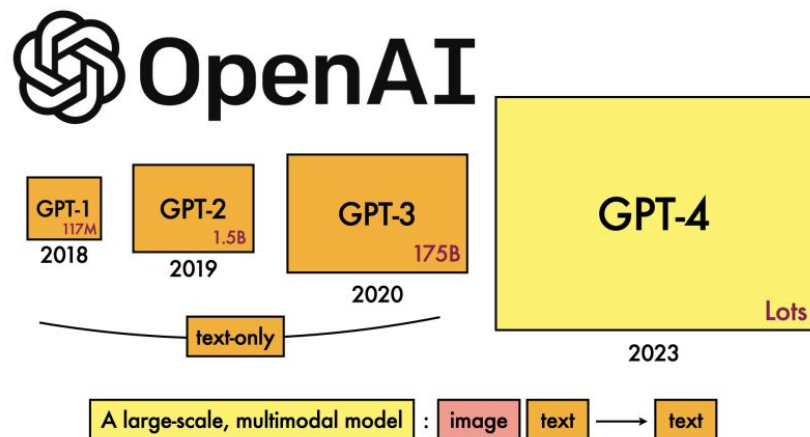


Figure 2: Évolutions GPT(Zafar, 2020)

Avant d'arriver à ChatGPT, GPT est passé par plusieurs évolutions allant de GPT-1 à GPT-4, chaque évolution, a marqué des améliorations pertinentes notamment en génération de texte et à la qualité des réponses fournies et aussi aux types d'Input et d'output (Texte, image, vidéo...).

- GPT-1: est premier modèle d'OpenAI , sortie en 2018. Ce modèle de **117 millions** de paramètres peut générer un texte cohérent marquant ainsi une avancée par rapport aux modèles de langage précédents.

- un an plus tard, OpenAI a lancé GPT-2 avec un entraînement sur un corpus de données plus large, GPT-2 ce qui a amélioré la génération de texte, rendant les réponses plus fluides et contextuellement appropriées et ce, avec 1,5 milliards de paramètres. Mais, ce modèles a soulevé des préoccupations concernant la désinformation.

- GPT-3: lancé en 2020,GPT-3 est un modèle beaucoup plus puissant que GPT-1 et GPT-2, il a marqué un tournant dans le domaine de l'IA conversationnelle avec ses 175 milliards de paramètres. Il a élargi les horizons de ce qui est possible en termes de compréhension du langage et de génération de contenu. GPT-3 est devenu un outil polyvalent pour de nombreuses applications.

- GPT-4⁷: GPT-4 marque déjà une avancée significative dans le domaine de l'IA **multimodale**. Les utilisateurs peuvent interagir avec du texte, mais aussi des images et des vidéos. OpenAI a écrit à son propos : Selon OpenAI : « *Par exemple, si un utilisateur envoie **une photo de l'intérieur de son frigo**, le Virtual Volunteer ne va pas seulement être capable d'identifier ce qui est dedans. En fait, il va aussi extrapoler et analyser ce qui peut être préparé avec ces ingrédients* »(L, 2024).

La nouvelle fonctionnalité du service c'est « *Le Virtual Volunteer* » se basant sur GPT-4 et qui est un outil de reconnaissance d'objets qui peut répondre à des questions sur une image(Brin et al., 2023) d'une manière inédite. Ces réponses de l'IA contiennent des explications très nuancées dans un format convivial et conversationnel.

« *GPT-4, peut résoudre des problèmes difficiles avec une plus grande précision, grâce à ses connaissances générales plus larges et à ses capacités de raisonnement avancées.* »(GPT-4, s. d.). cette version promet d'apporter des améliorations substantielles en termes de précision, de compréhension du contexte, et de réduction des biais, poussant encore plus loin les capacités d'interaction naturelle avec les utilisateurs.

Le schéma ci-dessous, présente une comparaison faite par OpenAI de différentes versions du modèle linguistique GPT, pour voir les performances de ses modèles dans différents domaines.

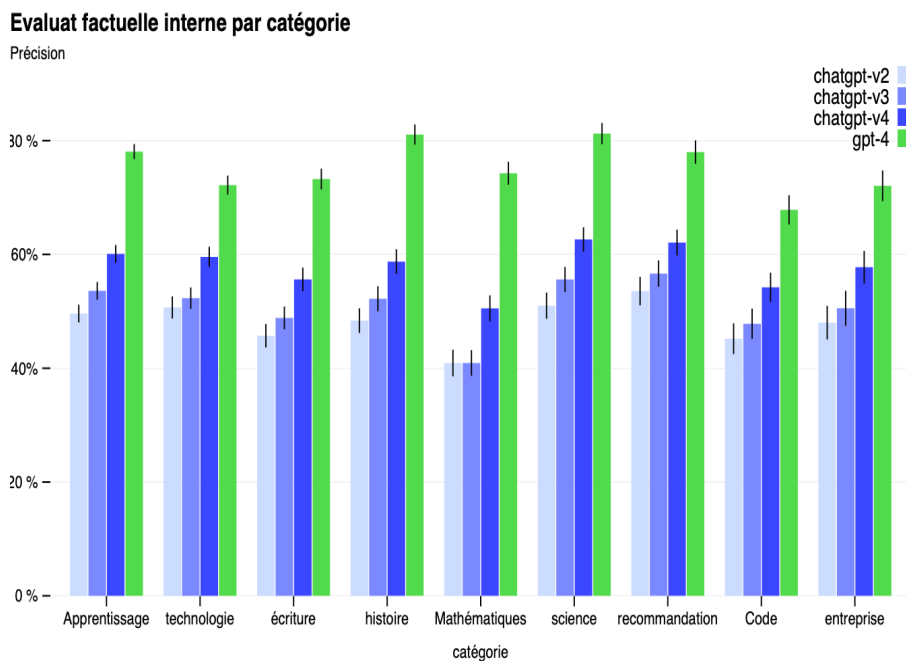


Figure1: Étude faite par OpenAI pour évaluer la performance de GPT-4(GPT-4, s. d.)

⁷ Vu que le domaine de l'IA ne cesse de se développer donc j'ai arrêté ma recherche en février 2024, donc toutes version ou amélioration après cette date ne sera pas mentionnée dans notre mémoire.

À travers ces évolutions, ChatGPT a non seulement repoussé les limites de ce que les modèles de langage peuvent accomplir, mais il a également redéfini les interactions entre les humains et les machines. La capacité de ChatGPT à apprendre de vastes quantités de données textuelles et à générer des réponses qui semblent de plus en plus identiques de celles d'un humain marque une avancée significative dans notre quête d'une intelligence artificielle véritablement conversationnelle. Son évolution continue promet de transformer encore davantage notre manière d'interagir avec la technologie, ouvrant de nouvelles voies pour l'application de l'IA dans notre vie quotidienne

2.4.2. ChatGPT et la Recherche d'Information universitaire

ChatGPT autant que nouvelle stratégie de recherche de recherche d'information met en lumière l'utilisation de l'intelligence artificielle dans la recherche d'information, mais il convient de s'interroger sur la véracité des systèmes de recherche basés sur des modèles d'IA par rapport aux méthodes de recherche traditionnelles comme les moteurs de recherche.

Pour rechercher une information sur ChatGPT, il n'est plus besoin de mettre des requêtes spécifiques et des mots clés, l'utilisateur peut discuter avec une machine comme s'il le fait avec un humain qui comprend ses émotions ce qui modifie fondamentalement les stratégies de recherches d'information.

Face aux nombreux débats autour des avantages et des dangers sur l'utilisation de ChatGPT dans l'enseignement supérieur, il faut savoir que l'intégration de ChatGPT dans la recherche d'information numérique ouvre la porte à des applications révolutionnaires au-delà de la simple acquisition de connaissances.

L'application de ChatGPT dans l'enseignement supérieur promet de transformer le paradigme pédagogique traditionnel, en offrant aux enseignants et aux étudiants des outils innovants pour l'apprentissage et l'enseignement. Tout comme ChatGPT a rendu la recherche d'information plus intuitive et accessible, son utilisation dans l'éducation vise à rendre l'apprentissage plus personnalisé, interactif, et engageant.(Quinio & Bidan, 2023).

Malgré les avantages de ChatGPT, mais celui-ci représente des menaces pour la recherche universitaire. En effet, les données sur lesquelles ChatGPT a été entraîné peuvent contenir des informations biaisées, non mise à jour ou tout simplement incorrectes, ceci est dû au fait que son entraînement s'est arrêté en 2021 (OpenAI, 2022) donc il ne peut ne pas être informé des événements ou informations récentes (Hassani & Silva, 2023). Aussi, autre inconvénient, c'est **le manque d'expertise** dans certains domaines spécialisés ce qui peut générer des réponses moins détaillées. Le manque de citations est aussi un problème car on ne sait pas d'où vient

cette information et donc on n'est sûr de son exactitude ce qui peut biaiser la recherche. Un dernier inconvénient, c'est l'addiction à ChatGPT, trop dépendre de cet outil, peut détruire l'analyse et l'esprit critique du chercheur ou l'étudiant.

Certains pays ont pris la décision de limiter voire d'interdire l'utilisation de ChatGPT dans le domaine de l'enseignement pour diverses raisons. En effet, pour interdire réellement, il est nécessaire dans un premier temps de pouvoir détecter efficacement son utilisation. Cependant, la détection de ChatGPT requiert des outils de détection tout aussi performants que l'outil lui-même. De ce fait, de nombreux éditeurs spécialisés dans le développement de détecteurs d'intelligence artificielle soulignent que ces derniers ne sont pas forcément adaptés pour être utilisés directement dans le but de sanctionner les étudiants.

Mais la question qui se pose, est-ce qu'il y a un cadre légal de cette interdiction ? D'après Régis Meissonnier, professeur à l'université de Montpellier, « *Le vide juridique qui existe pour l'instant autour des usages pouvant être faits de textes générés par des IA confère une zone de flou quant à la recevabilité par un tribunal administratif des sanctions pouvant être prises envers les étudiants fautifs* » (Meissonnier, 2023)

Selon une étude faite à l'UCLouvain en 2023 (Deville & Jacqmot, s. d.), la question « Avez-vous déjà utilisé ChatGPT dans le cadre de vos études ? » a été posée aux étudiants et 88% ont répondu qu'ils l'utilisent pour la rédaction d'un travail avec un grand pourcentage pour ceux qui l'utilisent régulièrement pour l'ensemble des travaux, ce pourcentage est susceptible d'augmenter.

Une autre étude a été faite en Mai (2023) à l'université de Namur (Lobet et al., s. d.), sur les étudiants de première année universitaire, cette étude montre que certes ChatGPT est utilisé mais pas de la même fréquence que d'autres IA comme Google Translate, DeepL ou encore Google Search Image qui sont placés en haut du podium.

En 2024, plusieurs articles ont été faits sur l'impact de ChatGPT sur l'enseignement, certains le voient comme ennemi et d'autres comme un outil parmi d'autres et qu'il faut juste mettre en place des stratégies pour familiariser les étudiants sur son utilisation. « *Concernant les établissements d'enseignement supérieur, trois injonctions peuvent être appliquées. Premièrement, il apparaît essentiel de dispenser à tous les niveaux (de la Licence 1 au Master 2) des cours abordant l'IA. Il s'agira d'une part de démystifier le concept d'IA et, d'autre part, d'être apte à travailler avec une IA.* » (Godé et al., 2023).

Dans l'article « ChatGPT dans la rédaction scientifique » de Jean MOUSSAVOU, celui-ci cite que bien que ChatGPT ne puisse remplacer la pensée critique du chercheur, ni sa capacité à interpréter les résultats d'une recherche, à discuter de ses conclusions sur la base des meilleures

preuves scientifiques disponibles, ChatGPT est néanmoins décrit comme un outil prometteur à la fois dans la rédaction scientifique et dans le processus de recherche lui-même (Moussavou, 2023) ce qui veut dire que pour une bonne utilisation de ChatGPT, il faut l'utiliser dans des tâches qui ne nécessitent pas une analyse et une intelligence humaine (Burger et al., 2023). Par exemple, dans le processus de recherche d'information, ChatGPT peut aider à collecter des données de masse mais il faut toujours vérifier l'exactitude de ces données dans d'autres sources ce qui permettra au chercheur de gagner du temps dans certaines étapes de sa recherche.

Cependant, l'intégration de ChatGPT dans la recherche universitaire offre des avantages, notamment en termes de gain de temps et d'accessibilité à l'information. Cependant, ces avantages doivent être contrebalancés par une prise de conscience des limitations et des risques associés. Les chercheurs d'information doivent utiliser ChatGPT de manière critique et responsable, (Thorp, 2023) en s'assurant que l'IA est un outil complémentaire qui renforce, mais ne remplace pas la rigueur scientifique et l'intégrité académique.

Exploisons au mieux sa valeur ajoutée et introduisons-le intelligemment dans nos enseignements *« Ils doivent continuer à réfléchir, à structurer leur raisonnement et à comprendre comment les choses fonctionnent. Nous devons également réfléchir à l'évaluation des étudiants. Avant nous évaluions une production. Désormais nous devons évaluer la maîtrise qu'ils ont de leur production. C'est un changement fondamental »* affirme Daniel Courivaud enseignant-chercheur à ESIEE Paris – Université Gustave Eiffel. (*ChatGPT dans l'enseignement supérieur*, s. d.).

ChatGPT est présent en force dans notre vie donc il faut bien savoir l'utiliser tout en adoptant des pratiques, des précautions et une réflexion bien réfléchie comme par exemple, veuillez à vérifier les résultats obtenus, faire des journées de sensibilisation dans les universités pour sensibiliser les étudiants au bon usage des outils de l'IA, encourager l'analyse et la pensée critique dans les classes pour ne pas les détruire.

En fin de compte, le succès de l'intégration de ChatGPT dans la recherche dépendra de la capacité des chercheurs à l'utiliser de manière éclairée et responsable. Et donc la bonne équation porte surtout sur l'usage de ChatGPT et non sur GhatGPT en lui-même.

3. Problématique

3.1. Question de recherche

A l'ère de l'intelligence artificielle, plusieurs modèles de langage avancé et des outils conversationnels tel que ChatGPT, ont été développés ce qui présente pour la recherche d'informations numériques des défis uniques en raison de la capacité de ce modèle de langage avancé à générer des réponses à des questions de manière rapide, directe et autonome. En raison de sa vitesse de traitement et de compréhension de grandes quantités d'informations, ChatGPT peut être très utile pour des utilisateurs à la recherche de réponses rapides et précises. Néanmoins, cette technologie présente également des dangers potentiels pour la qualité des résultats de recherche, tels que le manque de fiabilité des sources ou la pertinence des informations trouvées. Par exemple, s'il est mal formé sur des ensembles de données légitimes et de mauvaise qualité, il pourrait produire des réponses incorrectes ou biaisées, ce qui résulterait en des résultats de recherche trompeurs. Par conséquent, pour naviguer dans ces difficultés, les utilisateurs doivent suivre des méthodes de recherche d'information numérique éprouvées. Cela peut inclure l'utilisation de sources réputées et fiables, la vérification des informations à partir de plusieurs sources et la compréhension des limites de l'IA en matière de traitement de l'information.

Ainsi, comment l'utilisation ChatGPT peut influencer recherche d'information en général et la recherche universitaire en particulier? et est-ce que les données récoltées par les chatbots sont toujours fiables ?

3.2. Hypothèses

- a. H1. L'utilisation de CHATGPT permet aux chercheurs de gagner du temps dans leurs processus de recherche d'information.
- b. H2. L'introduction des outils conversationnels pourrait modifier la façon dont les chercheurs interagissent avec les sources d'information, en leur permettant de poser des questions plus précises et d'obtenir des réponses plus ciblées.
- c. H3. L'introduction d'outils conversationnels pourrait changer la façon dont les chercheurs évaluent la qualité des sources d'information.
- d. H4. L'introduction de ChatGPT pourrait entraîner une modification des compétences requises pour la recherche.

4. Méthodologies

L'objectif de la présente recherche est d'examiner les stratégies de recherche d'information à l'ère de ChatGPT en se basant sur les expériences des étudiants universitaires en Belgique et au Maroc. Pour ce faire, une méthodologie mixte a été utilisée pour collecter des données quantitatives et qualitatives visant à obtenir une vision profonde des pratiques de recherche d'information dans ces deux contextes culturels et géographiques. Tout d'abord, un questionnaire en ligne a été partagé dans divers réseaux sociaux pour inviter les étudiants universitaires de Belgique et du Maroc à participer à la recherche. Les résultats collectés via le questionnaire ont permis de recueillir des données quantitatives sur la fréquence d'utilisation de ChatGPT, la pertinence des résultats obtenus et la fiabilité des résultats. D'après les résultats du questionnaire, six participants ont été recrutés pour une observation de leur manière d'effectuer une recherche suivi d'un entretien semi directif pour discuter de leur comportement durant la séance d'observation en mettant l'accent sur l'impact de ChatGPT sur leurs pratiques et stratégies. Cela, nous a permis de collecter des données qualitatives sur le comportement et les habitudes des participants lors de leur processus de recherche. Cette approche méthodologique vise à offrir une étude complète sur les pratiques de recherche d'information à l'ère de ChatGPT, en intégrant des données quantitatives et qualitatives pour une compréhension approfondie .

4.1. Questionnaire

Nous avons créé notre questionnaire sur Google Forms et nous l'avons diffusé à travers nos différents réseaux sociaux. Pour plus de participation, nous avons généré des QR codes qui ont également été pour accéder facilement et rapidement au questionnaire ce qui a été plus pratique surtout pour les participants qui ne sont pas trop actifs dans les réseaux sociaux. Nous avons commencé par demander aux participants de confirmer s'ils étaient actuellement inscrits dans un programme universitaire, leur âge, leur niveau d'études ainsi que leur domaine. Ces informations démographiques sont essentielles pour situer le cadre général des données et comprendre les spécificités des deux groupes d'étudiant tout en abordons plusieurs aspects clés des habitudes de recherche d'information des participants. Ces derniers, ont été invités à indiquer la fréquence de leurs recherches sur internet pour leurs travaux universitaires et les sources d'information qu'ils privilégient dans ce contexte ce qui a permis d'analyser les tendances globales concernant l'utilisation des ressources en ligne et d'identifier les plateformes de recherche les plus fréquemment utilisées.

Une autre partie importante du questionnaire était consacrée à l'évaluation de l'utilisation de ChatGPT dans le processus de recherche d'information. Des questions ont été posées pour évaluer la fréquence d'utilisation de ChatGPT, la crédibilité de ChatGPT en tant que source d'information et la comparaison de sa pertinence avec d'autres sources. Les participants ont également été invités à réfléchir à tout impact potentiel de ChatGPT sur leurs pratiques de recherche, notamment en ce qui concerne les changements dans leurs habitudes ou comportements de recherche après l'utilisation de ChatGPT. Notre questionnaire a été conçu pour mettre l'accent sur l'utilisation de ChatGPT, son impact potentiel et la crédibilité de ses résultats. Les réponses recueillies ont servi de base pour une analyse approfondie des stratégies de recherche d'information à l'ère de ChatGPT, intégrant à la fois des données quantitatives et qualitatives pour une compréhension globale et complète.

4.2. Observation

Après la collecte et l'analyse des résultats obtenues via le questionnaire, nous avons fait des séances d'observation non participante pour approfondir notre compréhension des pratiques de recherche d'information des étudiants, en mettant particulièrement l'accent sur leur utilisation de ChatGPT et ses implications dans leur processus de recherche. Six étudiants ont été sélectionnés, (trois belges et trois marocains), pour participer à ces séances d'observation, permettant ainsi de couvrir une diversité de perspectives culturelles et géographiques.

Durant les séances d'observation, nous avons veillé à ce que ces séances se déroulent dans un environnement familier et confortable pour les participants, favorisant ainsi un sentiment de détente et de naturel. Conformément à une approche d'observation non participante, j'ai adopté une posture d'observation discrète, sans interférer avec les actions ou les décisions des participants.

L'objectif principal de ces séances d'observation était de documenter et d'analyser les interactions des étudiants avec ChatGPT et d'autres outils de recherche en ligne. Nous avons enregistré les actions des participants, par quels outils ils commencent leurs requêtes, leurs prompts, comment ils collectent l'information, et ce, tout au long de leurs sessions de recherche. Cela nous a permis de saisir les stratégies utilisées par les étudiants, leurs préférences en matière de recherche d'information, ainsi que les défis rencontrés lors de l'utilisation de ChatGPT et d'autres outils similaires. Ces informations approfondies enrichiront notre analyse et nos conclusions, en offrant une

perspective concrète et contextuelle sur l'impact de ChatGPT sur les pratiques de recherche d'information des étudiants universitaires en Belgique et au Maroc.

4.3. Entretiens

Parmi les différentes méthodes de collecte de données relevant de l'approche qualitative, nous avons opté pour la technique de l'entretien. Nous considérons cette technique comme étant la plus appropriée pour appréhender pleinement l'expérience des étudiants et obtenir des informations riches et nuancées sur leur stratégies de recherche d'information. Plus précisément, nous avons choisi de mener des entretiens semi-directifs. Situé entre l'entretien directif et l'entretien libre, ce type d'entretien nous permet de mieux saisir la singularité de l'expérience vécue par des individus ou des groupes d'individus dans leurs relations avec les autres, avec les institutions ou plus largement avec certains phénomènes sociaux qui est dans notre cas ChatGPT. Donc, après chaque session d'observation, un entretien semi-directifs complémentaire était fait. Cet entretien avait pour but d'offrir des pistes de compréhension et d'interprétation des résultats collectés suite au questionnaire et l'observation, en plus d'expliquer certaines pratiques des étudiants notés pendant l'observation. Ces entretiens ont fourni l'occasion d'explorer en profondeur les méthodes de recherche d'information des participants, à la fois avant et après l'introduction de ChatGPT dans leur processus de recherche. À travers l'utilisation souple d'une grille de questions, nous offrons à chaque personne interrogée la liberté d'exprimer un point de vue précis sur une question donnée, tout en nous autorisant à la relancer afin d'approfondir certaines réponses et faire émerger de nouvelles hypothèses de travail. Nos entretiens sont ainsi conçus pour tester nos hypothèses sans pour autant adopter une approche excessivement directive.

4.4. Retours sur la méthodes et justifications

Dans notre méthodologie, nous avons adopté une approche variée, combinant trois méthodes, le questionnaire, l'observation et les entretiens. Chose qui nous a permis de capturer à la fois des données objectives (quantitatives) et subjectifs (qualitatifs). Cette combinaison est particulièrement pertinente pour une étude qui cherche à comprendre non seulement les comportements observables mais aussi les perceptions et motivations sous-jacentes des étudiants face à l'utilisation de ChatGPT et donc avoir une vision globale de notre sujet de recherche.

Cependant, nous devons admettre quelques limites à cette méthodologie. Premièrement, le questionnaire était accessible uniquement en ligne, et donc seuls ceux qui ont accès à internet pouvaient y accéder ce qui a réduit le nombre de répondants et donc limiter notre échantillon. Deuxièmement, la durée de mon séjour au Maroc était courte ce qui nous a poussé à faire les trois observation et entretiens dans un intervalle de temps très court chose qui ne nous pas laisser le temps d'analyse et de recul entre les entretiens. Troisième limite, le nombre de participants est assez limité (six), ce qui peut réduire la généralisation des résultats et limiter la diversité des pratiques et des perspectives observées. Quatrième limite, c'est l'effet Hawthorne(Fischer-Lokou & Pinte, 2017) même si nous avons fait une observation non participante et nous avons adopté une posture discrète, les participants peuvent modifier leur comportement simplement parce qu'ils savent qu'ils sont observés, ce qui pourrait affecter l'authenticité des données recueillies. Autre limite concernant le questionnaire, bien que complet, mais il pourrait bénéficier d'une plus grande profondeur dans ses questions pour explorer plus en détail les motivations et les contextes d'utilisation de ChatGPT, durant les observations et les entretiens nous nous sommes rendu compte que d'autres questions ouvertes pourraient être ajoutées au questionnaire.

5. Résultats

5.1. Récoltes des données

5.1.1. Questionnaire

Nous avons diffusé le questionnaire du 15 Juillet 2023 au 30 novembre 2023 et nous avons recueilli 98 réponses.

1. Conception du Questionnaire

Suite à une analyse approfondie de la littérature existante sur le sujet des stratégies de recherche d'information numérique à l'ère de ChatGPT, nous avons créés notre questionnaire qui se compose de quinze questions fermées et deux questions ouvertes, abordant divers aspects de l'utilisation de ChatGPT dans les recherches universitaires, ainsi que les attentes des participants concernant les améliorations futures de cette technologie.

2. Participants

Les participants ont été sélectionnés de manière à assurer une représentation diversifiée des étudiants universitaires, en privilégiant les groupes marocains et belges. Le choix de réaliser un questionnaire se justifiait pour deux raisons principales. D'abord, il allait permettre un premier tour d'horizon du sujet, en recueillant des représentations et pratiques d'utilisation de ChatGPT

auprès du plus grand nombre. Il permettait ensuite de se constituer un carnet de contacts, en autorisant les répondants qui le souhaitaient à laisser des coordonnées afin de les contacter pour une future séance d'observation suivie d'un entretien semi-directif.

Il est également important de souligner que le partage du questionnaire était dans un cadre qui laissait le choix aux répondants de répondre ou non au questionnaire. Un court message avant le questionnaire expliquait clairement le sujet et l'objectif de la recherche. Cette information implique donc que les personnes ayant répondu au questionnaire, évaluaient leur connaissance préalable des agents conversationnels comme suffisamment bonne que pour accepter de répondre à un questionnaire qui se présentait comme universitaire sur le sujet.

3. Analyse des Données

Une analyse descriptive a été réalisée pour examiner la répartition des réponses aux questions fermées. Les réponses aux deux questions ouvertes ont été soumises à une analyse de contenu pour identifier les thèmes émergents qui était l'impact de ChatGPT sur la recherche.

5.1.2. Observation

En fonction des réponses du questionnaire, six profils ont été sélectionnés sur 98 répondants, significatifs selon des critères établis préalablement. Trois critères ont permis de définir les profils-types sur base de la problématique et des hypothèses.

D'abord les répondants du questionnaire devraient être de la tranche d'âge 19-26 ans, donc la valeur de l'âge était une variable à considérer. De plus, ils devraient être étudiants universitaires.

Enfin, le critère était qu'ils ont remarqué l'impact de ChatGPT sur leur recherche. Un critère implicite était également que la personne a laissé une adresse-mail, marquant son accord pour être recontactée pour un éventuel entretien. En croisant les personnes qui correspondaient à chacune de ces variables, 51 personnes pouvaient être contactées. Sur base de la volonté de réaliser un maximum de 6 entretiens (6 sur 51 soit 10,34%), la répartition suivante s'est imposée :

Nomenclature	Localité	Age	Étudiant universitaire ?	Discipline	Impact ChatGPT ?
R1	Maroc	19-26	Oui	Sciences et technologie	Non
R2	Maroc	19-26	Oui	Sciences de la gestion	Oui
R3	Maroc	19-26	Oui	Ingénierie touristique	Oui
R4	Belgique	19-26	Oui	Stic	Oui
R5	Belgique	19-26	Oui	Journalisme	Non
R6	Belgique	19-26	Oui	Ingénieur Civil en Informatique	Oui

Parmi les trois personnes marocaines recontactées, deux seulement ont accepté que nous nous rencontrions afin de mener une séance d'observation de leurs méthodes de recherche. J'ai donc resélectionné une autre personne qui répond à nos critères. Les séances d'observation avec les répondants marocains ont été menées entre Aout et Septembre 2023 lors de la réalisation de mon stage au Maroc, tandis que les séances d'observations avec les répondants Belges sélectionnés ont été réalisés entre fin Novembre2023 et Janvier 2024 ce qui a coïncidé (positivement) avec la période de blocus et des examens.

5.1.3. Entretien

Après les séances d'observation, et comme déjà expliqué aux participants, des entretiens semi-directifs ont été menés.

Le guide d'entretien développé en amont se divisait en quatre grandes catégories :

- Une présentation et rappel du sujet de la recherche et de la confidentialité des données
- Une partie « réflexions sur l'observation de la Recherche » : ou nous avons parlé de certains comportements durant la séance d'observation.
- Expériences avec ChatGPT : ou nous avons discuté de l'intégration de ChatGPT dans le processus de recherche, son utilisation par rapport aux méthodes traditionnelle, ses limites etc...

- Perspectives et suggestions
- Conclusion

Cinq participants ont accepté de faire les entretiens en face à face et directement après les séances d'observation, et la sixième personne R6, et par faute de temps, a préféré le lendemain. Chacun de ces entretiens a duré plus au moins une vingtaine de minutes et a été enregistré avec l'accord préalable des participants. Les enregistrements ont été transcrits et soumis à une analyse qui a permis de clarifier les méthodes de recherche des participants.

Cette analyse a permis de compléter les données quantitatives du questionnaire par des informations qualitatives riches, diversifiées et claires, offrant une vue d'ensemble plus complète des stratégies de recherche des étudiants.

5.2. Analyse des données

5.2.1. Questionnaire

Le partage du questionnaire a non seulement servi de base pour permettre l'émergence de profils-types en prévision des futures séances d'observation suivi par des entretiens, mais a également permis l'obtention de données quantifiables sur l'utilisation de ChatGPT dans la recherche d'information, la fiabilité des données récoltées et la pertinence de celle-ci. C'est principalement sur le questionnaire que se base l'analyse des données récoltées par le biais des entretiens et la mise en relation de celles-ci. Ainsi, 98 personnes ont répondu à notre questionnaire. D'après l'analyse du questionnaire, on peut révéler quelques différences et similitudes entre les répondants marocains et belges notamment dans la pertinence des résultats obtenus par ChatGPT, la fréquence d'utilisation de ChatGPT dans leurs recherches, la fiabilité des résultats obtenus etc.

Comme vous pouvez le voir dans ce graphique, 64,28% des répondants viennent de la Belgique contre 35,71% qui viennent du Maroc

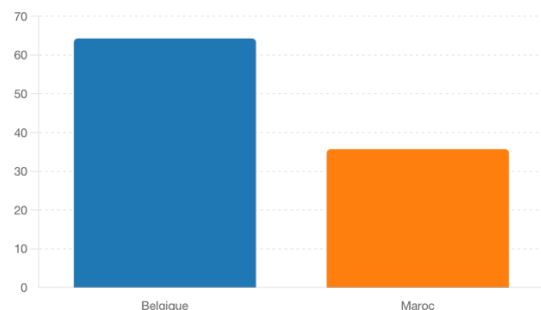


Figure 3: Répartition des répondants

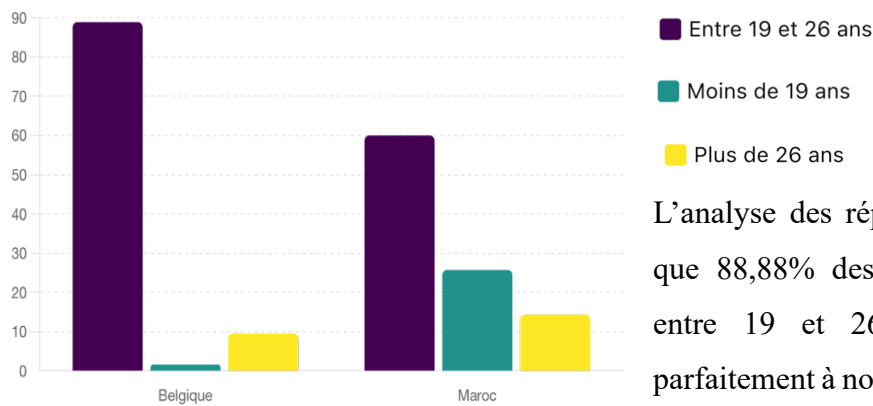


Figure 4:Tranche d'âge des répondants

L'analyse des réponses récoltées montre que 88,88% des répondants belges ont entre 19 et 26 ans ce qui répond parfaitement à nos critères contre 60% des marocains appartenant à cette même tranche d'âge.

Dans les 98 répondants, 79.59% sont inscrits en deuxième cycle universitaire, tandis qu'un plus petit pourcentage est en premier et troisième cycle 20.40%.

Dans cette partie (Pratiques de ChatGPT) nous avons séparé l'analyse des données pour comparer et voir l'impact de ChatGPT dans deux cultures différentes : marocaine et belge. Pour une meilleure analyse, nous avons filtré les réponses des répondants marocains et celles des répondants belges.

Combien de fois par semaine effectuez-vous des recherches sur internet pour votre travail universitaire ?

	Marocains	Belges
- Moins d'une fois par semaine	3,13%	1,59%
- 1-2 fois par semaine	31.25%	4,76%
- 3-6 fois par semaine	59.38%	84,13%
- Tous les jours	6.25%	9,52%

Les résultats montrent que les étudiants belges effectuent principalement des recherches sur Internet "3-6 fois par semaine" (84.13%), tandis que les étudiants marocains ont une répartition plus diversifiée, avec une majorité également dans la catégorie "3-6 fois par semaine" (59.38%) et un pourcentage significatif dans "1-2 fois par semaine" (31.25%).

Dans les deux pays, les étudiants effectuent leur recherches sur internet pour leur travail universitaire majoritairement entre 3 et 6 fois par semaine.

Quelles sources d'information utilisez-vous le plus souvent pour vos recherches universitaires ?

Pour cette question, les répondants ont choisi plusieurs combinaisons axées principalement sur les bases de données en ligne, les bibliothèques universitaire et ChatGPT avec une toute petite minorité pour les réseaux sociaux

- En Belgique, ChatGPT est clairement la source d'information la plus utilisée, avec 40 des 63 répondants (63.49%) indiquant cette source soit seule soit en combinaison avec d'autres sources.
- Au Maroc, les étudiants montrent une diversité plus grande dans les sources d'information utilisées. Les deux sources les plus fréquentes sont "Sites web spécialisés" et "ChatGPT".

Ceci explique que ChatGPT est plus utilisé dans la recherche universitaire en Belgique qu'au Maroc.

Dans quelle mesure utilisez-vous ChatGPT pour vos recherches universitaires ?

	Marocains	Belges
- Toujours	48,57%	15,87%
- Très souvent	17,14%	69,84%
- Rarement	31,42%	13,29%
- Jamais	2,85%	1,1%

85,71 des étudiants belges utilisent ChatGPT toujours ou très souvent contre 65,7% pour les Marocains avec un tous petit pourcentage pour les étudiants qui ne l'ont jamais utilisé dans les deux pays.

Comment évaluez-vous la crédibilité des sources trouvées grâce à ChatGPT ?

	Marocains	Belges
- Très crédible	25,71%	33,33%
- Plutôt crédible	48,57%	63,49%
- Plutôt non crédible	22,85%	3,17%

- Pas du tout crédible	2,85%	0%
------------------------	-------	----

- En Belgique, la majorité des répondants (63.49%) trouvent les sources trouvées grâce à ChatGPT "Plutôt crédibles".
- Au Maroc, bien que "Plutôt crédible" soit également la réponse la plus fréquente, elle est choisie par un pourcentage moindre de répondants (48.57%).

Ces résultats montrent une confiance relativement élevée dans la crédibilité des sources trouvées grâce à ChatGPT parmi les étudiants des deux pays, bien que cette confiance soit plus marquée en Belgique.

Comment trouvez-vous la pertinence des résultats obtenus grâce à ChatGPT par rapport à d'autres sources ?

	Marocains	Belges
- Très pertinents	38,09%	30%
- Plutôt pertinents	51,42%	53,96%
- Peu pertinents	22,85%	16,04%
- Pas du tout pertinents	5,71%	0%

- En Belgique, la majorité des répondants (53.97%) trouvent les résultats de ChatGPT "Plutôt pertinents".
- Au Maroc, la réponse la plus fréquente est également "Plutôt pertinents", mais représente une proportion légèrement inférieure de 51.43%.

Ces résultats montrent que les étudiants des deux pays trouvent généralement les résultats de ChatGPT pertinents par rapport à d'autres sources. Cependant, la Belgique montre une majorité un peu plus marquée en faveur de la pertinence.

Comment évaluez-vous la qualité du résultat obtenus par ChatGPT par rapport à celui généré manuellement ?

	Marocains	Belges
- La qualité est supérieure	59,37%	87,3%
- La qualité est équivalente	21,87%	11,11%
- La qualité est inférieure	18,75%	1,58%

--	--	--

(87.30%) des étudiants belges considèrent que la qualité des résultats obtenus par ChatGPT est supérieure par rapport aux résultats générés manuellement contres (59.38%) pour les marocains.

Ces résultats suggèrent une perception globalement positive de la qualité des résultats de ChatGPT, avec une reconnaissance plus forte en Belgique.

Ceci peut être expliqué par le fait que les belges utilisent ChatGPT plus que les marocains leurs donne plus d'expérience dans la formulations des requêtes et des prompts ce qui donne un résultats plus pertinent en résultat.

Avez-vous remarqué un impact sur votre façon d'effectuer des recherches depuis l'arrivée de ChatGPT ?

	Marocains	Belges
- Oui	68,57%	50.79%
- Non	31,42%	49,20%

La majorité des étudiants des deux pays reconnaissent un impact de ChatGPT sur leur façon d'effectuer des recherches, avec un pourcentage plus fort au Maroc.

Les réponses qualitatives à cette question ont révélé que les étudiants des deux groupes utilisent ChatGPT pour :

- Gagner du temps
- Obtenir des idées de recherche nouvelles et originales.
- Utiliser ChatGPT comme un outil complémentaire aux sources traditionnelles.

Les étudiants marocains ont mentionné plus fréquemment l'utilisation de ChatGPT pour la reformulation des phrases et pour des suggestions des idées.

Pensez-vous que l'utilisation de ChatGPT pour la recherche d'informations pourrait contribuer à réduire les biais cognitifs ou les préjugés dans l'analyse des données ?

	Marocains	Belges
-Oui, fortement	15,62%	53,96%
- Oui, un peu	37,74%	31,74%
- Non, pas vraiment	40,62%	14,28%
- Non, pas de tout	6,25%	0%

Les étudiants belges jugent que ChatGPT peut contribuer à réduire les biais cognitifs et les préjugés dans l'analyse des données, avec 85,7% répondant oui (fortement ou un peu), contre 53,36 % des étudiants marocains.

Quelles sont vos attentes quant à l'amélioration future de ChatGPT ou des technologies similaires pour la recherche d'informations en ligne ?

	Marocains	Belges
- Amélioration de la fiabilité des résultats	25%	48,57%
- Réduction du temps de réponse	13,46%	37,14%
- Meilleure adaptation à différents types de requêtes	25%	5,71%
- Augmentation de la diversité des sources utilisées	21,15%	7,14%
-Amélioration de l'interactivité et de la personnalisation	15,38%	1,43%

Cependant, on peut noter des priorités différentes des étudiants belges et marocains pour les améliorations de ChatGPT dans le futur. Les étudiants belges se concentrent plus sur la fiabilité et la rapidité, tandis que les étudiants marocains mettent l'accent sur l'adaptation aux requêtes variées et l'interactivité.

En conclusion, les étudiants des deux pays utilisent ChatGPT fréquemment et trouvent ses résultats généralement crédibles et pertinents. Cependant, les étudiants belges montrent une adoption plus intensive et une confiance plus marquée dans la qualité des résultats de ChatGPT.

En revanche, les étudiants marocains expriment une plus grande diversité dans leurs opinions et attentes, avec une reconnaissance plus forte de l'impact de ChatGPT sur leurs méthodes de recherche.

Les attentes pour l'amélioration future de ChatGPT révèlent des priorités différentes : les étudiants belges veulent surtout des résultats plus fiables et rapides, tandis que les étudiants marocains cherchent une meilleure adaptation aux requêtes variées.

5.2.2. Observation et entretien

Pour une recherche crédible, nous avons veillé à faire les observations dans l'environnement de travail habituel des étudiants et d'être discret afin de ne pas impacter leur recherche.

Ces observations ont permis de distinguer plusieurs tendances et différences dans leurs pratiques de recherche d'information.

R1 : Étudiant en master sciences et technologies dans une haute école Marocaine, il effectuait son stage et a été observé dans le bureau des stagiaires. L'observation a duré 1h35 suivi d'un entretien de 25 minutes. Sa principale tâche était de rechercher les modèles de langage open source à usage commercial et de les comparer. Il a commencé par utiliser ChatGPT pour demander quels sont les modèles de langage open source. La réponse initiale ne l'a pas pleinement satisfait, il a donc reformulé sa requête pour obtenir des informations plus précises. Il a ouvert plusieurs pages de ChatGPT version gratuite chacune avec une recherche en lien avec le sujet principal.

Ensuite, il a copié les résultats dans un document Word et a commencé à chercher chaque modèle mentionné sur Google, accédant à différentes pages web pour obtenir plus d'information et il ajouté des détails dans son document. ChatGPT est resté un outil central dans sa recherche, il y revenait souvent pour clarifier des concepts techniques ou obtenir des explications rapides.

- *" R1 : (...) ChatGPT est très utile. Mais je dois souvent vérifier et clarifier les détails sur d'autres sites pour être sûr de la qualité des informations car ça m'est déjà arrivé de faire copier des informations incorrectes et je n'ai malheureusement pas vérifier (...)."*

Au cours de l'entretien, R1 nous a expliqué qu'il a déjà eu une mauvaise expérience avec ChatGPT, il ne vérifiait pas l'information avant de l'utiliser, ce qui l'a induit à l'erreur. Cette expérience l'a poussé à utiliser ChatGPT avec plus de prudence et de vérification et moins de confiance.

R2 : Étudiante en master en sciences de la gestion, cette observation s'est déroulée dans un café (c'était son choix) et a duré 1h50 suivi d'un entretien de 18 minutes. Elle devait effectuer des recherches pour une présentation dans le cadre d'un cours de « droit de l'entreprise ». Le but de sa recherche était de comparer le statut interne d'une entreprise publique marocaine avec le code du travail marocain pour déterminer lequel est plus avantageux pour les employés de cette entreprise et de faire un résumé de cette recherche. Elle a commencé par coller dans ChatGPT les articles du statut de l'entreprise, puis ceux du code de travail et a demandé à chatgpt de comparer.

Un message de ChatGPT l'informant d'elle a dépassé le nombre de mots en entrée et elle a donc segmenter sa recherche. Sans lire les résultats de ChatGPT, elle a écrit « *Maintenant que tu as tous les articles, fais une comparaison* » Après, elle a tous copié dans le document word et a fait la mise en page et la recherche est faite.

- " R2 : (...) *J'utilise ChatGPT pour gagner du temps. Au lieu de faire ce travail dans 5 jours et voir plus, grâce à ChatGPT je peux le faire en une journée, il compare effacement les données que je lui donne.*"

R2 utilise ChatGPT pour des comparaisons directes et rapides, ce qui montre une confiance élevée dans l'outil malgré l'absence de vérification des informations. Cette approche rapide permet de gagner du temps mais comporte des risques en termes de précision des données.

R3 : Étudiant en master en ingénierie touristique, il effectuait son stage dans une grande agence de tourisme à Marrakech. L'observation a eu lieu dans un café à Casablanca et a duré 50 minutes suivi d'un entretien de 25 minutes. Il devait chercher des idées pour aider l'agence à fidéliser ses clients. Il a commencé par demander à ChatGPT "*Comment fidéliser les clients*". Après avoir lu les réponses, il a copié et collé certaines phrases dans un document Word. Puis, il a affiné sa requête « *Comment fidéliser les clients d'une agence de voyage marocaine spécialisée dans les destinations Turquie Dubaï et Omra* ». Il a demandé plus de propositions et des détails sur certains points, copiant les idées pertinentes. Ensuite, il a utilisé l'outil QuillBot pour reformuler tout le fichier.

Il est passé du général au plus ciblé, à chaque fois il affinait et reformulait ses requêtes.

- " R3 : *ChatGPT me donne des idées que je n'aurais pas forcément trouvées seul* ».

R3 démontre une utilisation stratégique de ChatGPT pour générer des idées spécifiques, qu'il affine et adapte continuellement. L'utilisation de QuillBot pour la reformulation montre une attention à la qualité et à la clarté des informations présentées et la personnalisation de sa recherche.

R4 étudiante en master STIC à l'UCLouvain, l'observation à durée 2 heures 20minutes, l'étudiante préparait à un examen pour un cours qui selon elle, est trop long et très théorique. Elle m'a informé qu'elle comptait sur ChatGPT pour l'aider à résumer son cours et comprendre certaines notions complexes.

Lors de l'observation, R4 a débuté par ouvrir ChatGPT et lui expliquer qu'elle préparait un cours de X dans le cadre de son master, mentionnant qu'elle avait un examen à venir et qu'elle avait besoin de résumés pour mieux comprendre le contenu du cours.

R4 a commencé son travail en copiant une grande partie du premier chapitre de son cours et en le collant dans ChatGPT, demandant un résumé détaillé.

- *" R4 : (...) J'ai besoin que ChatGPT me donne un résumé clair pour que je puisse comprendre les notions clés sans passer des heures à lire tout le cours et encore plus à le comprendre et faire la synthèse,"*
- *"Je veux être certaine que toutes les informations clés du cours sont bien présentes et bien résumées et je lui demande parfois même de me donner genre questions réponses"*

ChatGPT a généré un résumé qu'elle a ensuite copié dans un document Word, répétant ce processus pour l'intégralité du premier chapitre.

Une fois le résumé du premier chapitre terminé, R4 a pris le temps de lire attentivement les résumés pour vérifier leur exactitude et s'assurer qu'aucune information importante ne manquait.

Elle a ensuite ajusté légèrement le résumé et l'a recopié dans ChatGPT pour demander des explications supplémentaires sur certains points qu'elle trouvait particulièrement complexes.

Finalement, R4 a intégré les explications fournies par ChatGPT avec le résumé initial, combinant les deux pour compléter le premier chapitre. Elle a jugé que cette méthode lui permettait d'avoir un document de révision complet et cohérent, adapté à ses besoins de préparation à l'examen.

Lors de l'entretien, R4 a souligné l'importance de l'utilisation de ChatGPT dans son processus de révision.

- *" R4 : (...) ChatGPT se distingue par sa rapidité et sa capacité à résumer des textes longs et complexes. Cela m'aide énormément à gagner plus de temps par rapport à la lecture et à la prise de notes manuelles, (...) parfois, les résumés de ChatGPT peuvent ne pas être très convaincants. Dans ces cas, je prends quand même le temps de vérifier les informations et les comparer avec le cours et avec mes notes"*

R4 a également évoqué l'impact de ChatGPT sur ses stratégies de recherche d'information.

- *" R4 : Depuis que j'ai commencé à utiliser ChatGPT, j'ai remarqué que je commence souvent mes recherches par l'ouvrir pour obtenir une vue d'ensemble rapide et des résumés. Cela m'aide à structurer ma révision de manière plus efficace et à gagner du temps ainsi que des idées pour l'organisation de mon travail en blocus. "*

Elle a souligné l'importance de la vérification des informations en revenant toujours au cours pour s'assurer de la qualité et de la précision des résumés obtenus.

L'observation et l'entretien mené avec R4 ont mis en lumière un autre objectif d'utilisation de ChatGPT qui est le résumé des cours et l'explication des notions. Donc au lieu de rechercher ces notions dans des sources plus fiables en liens avec la thématique du cours, elle a préféré se référer à ChatGPT.

R5 : Étudiante en Master 1 en journalisme à l'UCLouvain, l'observation à durée 1 heure 20 minutes et s'est déroulé dans le Kot de l'étudiante suivi d'un entretien de 30 minutes. Cette étudiante, devait rédiger un éditorial pour un cours dans le cadre d'un travail de groupe.

Lors de l'observation, R5 a adopté une approche en deux étapes pour rédiger son éditorial. Elle a débuté par chercher des exemples d'éditoriaux sur Google, consultant plusieurs pages web pour s'inspirer du format et du style de ce type d'écrit.

Une fois qu'elle avait une idée claire de la structure d'un éditorial, elle a lancé ChatGPT et lui a demandé de rédiger un éditorial, spécifiant le sujet et le nombre de mots requis conformément aux consignes du cours.

- *" R5 : (...) Je voulais voir comment sont structurés les éditoriaux et m'inspirer de différents styles, pour les comparer avec le style de ChatGPT"*
- *" R5 : (...) J'ai utilisé ChatGPT pour obtenir une première version de l'éditorial afin de gagner du temps et d'avoir une base sur laquelle travailler"*

R5 a ensuite copié le texte généré par ChatGPT dans un document Word et l'a reformulé avec ses propres mots. Pour vérifier si son texte est détectable comme IA, elle a utilisé ZeroGPT à chaque étape de la reformulation pour s'assurer que son texte n'était plus détectable comme généré par une IA.

R5 a pu finalement personnaliser l'éditorial tout en respectant les consignes de son cours, trouvant ainsi un équilibre entre inspiration et originalité.

- *"R5 : (...) Je voulais m'assurer que mon travail soit original et non détectable comme du texte généré par une IA"*

R6 : Étudiant en Master 1 Ingénieur Civil en Informatique à l'ULB, l'observation a duré 40 minutes et s'est déroulée dans un café à Bruxelles et l'entretien s'est déroulé le lendemain dans le même café. R6 devait faire un devoir en programmation Perl qui consistait à écrire un programme répondant à des critères spécifiques.

Lors de l'observation, R6 a commencé par lancer ChatGPT (version payante) et a directement copié les consignes du devoir en demandant : *"fais-moi un programme Perl qui permet de [il a collé les consignes du devoir], fais en sorte que ce programme soit simple et clair."*

ChatGPT a généré le code Perl demandé, que R6 a ensuite lu attentivement. Il a demandé à ChatGPT d'expliquer les lignes du code, ce qui lui a permis de mieux comprendre la structure et le fonctionnement du programme.

- *"R6 : (...) Depuis que j'utilise ChatGPT, j'ai remarqué que je gagne beaucoup de temps et que je peux obtenir des réponses plus précises et personnalisées rapidement, (...)"*

Après avoir pris connaissance des explications, R6 a demandé à ChatGPT de rédiger une version encore plus simple et plus claire du programme. Une fois ce code reçu, il l'a compilé pour vérifier que le résultat correspondait aux consignes du devoir.

Pour finaliser, R6 a personnalisé le programme en ajoutant des commentaires explicatifs dans le fichier texte, puis l'a enregistré. Cette méthode lui a permis de s'assurer que le code était non seulement fonctionnel, mais également bien documenté et compréhensible.

5.3. Discussion des résultats

L'objectif de cette étude a été de vérifier les différentes stratégies de recherche d'information numérique à l'ère de ChatGPT et de vérification de résultats chez les étudiants universitaires de 19 à 26 ans et ce dans deux pays différents, le Maroc et Belgique pour comparer les stratégies de recherche dans deux cultures différentes.

5.3.1. Comparaison utilisation Maroc -Belgique

Cette comparaison a montré des différences quant à la fréquence d'utilisation de ChatGPT dans les deux populations. Cette différence peut être expliquée par le fait que les répondants ne possèdent pas tous les mêmes compétences linguistiques (le français) dans notre étude. Au Maroc, le français est la deuxième langue nationale, ce qui conduit les étudiants à percevoir que ChatGPT produit des résultats en français bien structurés avec un vocabulaire avancé. De plus, les étudiants belges ont commencé à utiliser ChatGPT avant les marocains, ce qui leur donne plus d'expérience dans ce domaine. Autre point, les normes régissant l'intelligence artificielle en Belgique sont plus rigoureuses, ce qui pousse les étudiants belges à formuler des requêtes de recherche et des réponses de manière à obtenir des résultats aussi humains que possible et vérifier avec des outils de détection de l'IA chose qui n'est pas utilisé chez les marocains qui ne s'intéressent pas si le résultat est détectable ou pas par les détecteurs d'IA.

Pour la vérification de la crédibilité des résultats, durant cette étude, les étudiants utilisent le moteur de recherche google pour vérifier les résultats chose qui ne peut pas toujours fournir des réponses correctes.

Cette étude, révèle que dans les deux pays, ChatGPT est perçu comme un outil complémentaire précieux dans les recherches académiques. Les suggestions d'amélioration fournies par les étudiants mettent en évidence des domaines clés où ChatGPT pourrait évoluer pour mieux répondre aux besoins académiques et professionnels. Les différences observées entre les étudiants marocains et belges, notamment en termes de défis technologiques et de préférences pour les environnements de travail, offrent des perspectives intéressantes sur l'adaptation des technologies d'IA à divers contextes culturels et académiques

5.3.2. Conclusion aux Hypothèses

Les différents points abordés lors du questionnaire, des entretiens et des observations, permettent de confirmer ou non nos hypothèses formulées au début de notre recherche.

Première hypothèse (H1) : Dans l'hypothèse où l'utilisation des outils de conversation telle que ChatGPT devrait permettre aux chercheurs de gagner du temps, la plupart des étudiants ont répondu qu'en effet, ChatGPT les a aidés à gagner du temps car « Il permet de faire des recherches plus pertinentes et en moins de temps ». Une déclaration qui a validé donc cette hypothèse.

La deuxième hypothèse H2 indiquait que l'introduction des outils de conversation pourrait également modifier la forme des interactions des chercheurs avec les sources d'information. Ils pourraient demander des questions plus spécifiques et obtenir des réponses plus pertinentes. L'observation a montré que les étudiants demandent des questions précises et reçoivent des réponses ciblées avec ChatGPT. Selon les entretiens, les participants disent qu'ils apprécient comment « ChatGPT » fournit des réponses claires à des questions complexes. Ainsi, l'hypothèse H2 est également vérifiée, ce qui montre comment ces outils facilitent une interaction plus ciblée avec les sources.

La troisième hypothèse, qui porte sur le fait que l'introduction d'outils conversationnels influencerait la manière dont les chercheurs évaluent de la qualité des sources d'information, a également été confirmée. Nous avons observé que les étudiants utilisent ChatGPT pour obtenir rapidement un grand nombre d'informations, mais restent quand même vigilants quant à la crédibilité des données. Ce comportement se manifeste de manière assez prononcée en Belgique, où des mesures supplémentaires sont prises pour vérifier les informations les résultats de ChatGPT. Par conséquent, l'hypothèse est vérifiée dans cette mesure, ce qui signifie que l'utilisation de ChatGPT dans la recherche d'information universitaire conduit à une aptitude à l'autoévaluation des résultats

Enfin, la quatrième hypothèse (H4) qui suggérait que l'intégration d'outils conversationnels pourrait modifier les compétences requises pour la recherche. Les résultats montrent que les étudiants ont adapté leurs compétences pour tirer pleinement parti de ChatGPT. Par exemple, les étudiants marocains affinent souvent leurs questions de manière plus précise pour obtenir des réponses plus adaptées. Cette hypothèse est confirmée, démontrant que la formulation de requêtes, l'évaluation critique et la vérification de l'originalité sont devenues des compétences essentielles à l'ère de ChatGPT.

H1	Validée	ChatGPT aide les étudiants à gagner du temps en fournissant rapidement des informations pertinentes.
H2	Validée	Interaction modifiée avec les sources : ChatGPT permet aux étudiants de poser des questions plus précises et d'obtenir des réponses ciblées.
H3	Validée	Évaluation de la qualité des sources : L'utilisation de ChatGPT incite les étudiants à développer des compétences critiques pour évaluer la crédibilité des informations
H4	Validée	Modification des compétences requises : L'intégration de ChatGPT dans le processus de recherche modifie les compétences nécessaires, en mettant l'accent sur la formulation de requêtes, la vérification de l'originalité et l'évaluation critique des sources.

Ces résultats mettent en lumière les avantages et les défis associés à l'utilisation de ChatGPT dans les environnements académiques au Maroc et en Belgique, et offrent des perspectives sur les améliorations potentielles pour mieux répondre aux besoins des utilisateurs dans divers contextes culturels et académiques.⁸

5.3.3. Réponse à la question de recherche

Après avoir vérifié nos hypothèses, nous devons maintenant répondre à notre question de recherche qui est : « Comment l'introduction d'outils conversationnels comme Chatgpt influence la stratégie de recherche ? et est-ce que les données récoltées par les chatbots sont toujours fiables ? ».

Il ressort de l'étude que les étudiants utilisent ChatGPT car il permet de gagner du temps, génère des résultats pertinents et aussi il est très utiles dans certaines tâches comme le résumé ou encore la traduction ce qui leur a permis de remarquer des améliorations et un impact positif depuis son utilisation ce qui peut répondre clairement à notre question de recherche. ChatGPT a changé la stratégie de recherche d'information, les chercheurs ont tendance à utiliser ChatGPT plus que les autres stratégies traditionnelles. Certes le degré d'utilisation de chatgpt diffère d'une discipline à une autre, d'un pays à un autre et aussi d'un besoin à un autre mais le point commun c'est que ChatGPT est devenu une stratégie de recherche toujours utilisée soit comme source principale soit comme outils complémentaire avec d'autres sources mais il est

⁸ Voir le point 6.2. Perspectives

toujours utilisé dans la recherche universitaire. L'introduction de cet outil a permis aux utilisateurs de passer moins de temps à rechercher ce qu'ils veulent tout en leur offrant des pistes d'amélioration de leur recherche qu'aucune autre stratégie ne le permettait avant. La question de vérification de résultats reste relative. Les répondants savent que les résultats de ChatGPT ne sont pas toujours fiables mais s'ils doutent, ils vérifient sur d'autres sources chose qui peu faussé leur recherche. C'est pourquoi un grand pourcentage de répondant on répondant à la question d'amélioration de chatgpt par afficher les sources pour plus de crédibilité.

6. Limites et perspectives

6.1. Limites

Néanmoins, bien que les résultats de cette étude soient clairs et concluants, plusieurs limites doivent être reconnues et différentes considérations devraient être prises en compte pour obtenir cette recherche. D'abord, l'étude ne s'est focalisée que sur un faible échantillon d'étudiants universitaires. Bien que cela ait permis une analyse approfondie des pratiques individuelles, mais il reste peu diversifié et ne permet pas toujours une analyse concrète. Un échantillon plus large et plus diversifié pourrait fournir une perspective plus représentative des différentes méthodes de recherche et de l'utilisation de ChatGPT. D'une part, une étude regroupant des disciplines supplémentaires pourrait fournir un vue plus complète pour suivre les usages de ChatGPT dans divers domaines. Par ailleurs, une autre limite du travail présenté est le fait que les résultats de l'étude sont en grande partie évalués sur la base de l'auto déclaration des participants à l'entretien et au questionnaire. Il est possible que certains biais de réponse aient influencé les résultats, tels que la désirabilité sociale ou la tendance des participants à répondre de manière cohérente avec ce qu'ils pensent être les attentes de l'enquêteur. Autre limite, c'est que cette étude a été menée sur une période limitée. Vue que les pratiques d'utilisation de ChatGPT évolueront, en particulier en raison de l'introduction de nouvelles fonctionnalités et mises à jour, une étude longitudinale permettrait de fournir des insights plus robustes sur l'évolution des pratiques de recherche. En adressant ces limitations, les recherches futures pourraient donner des insights plus complexes et éclairants sur l'utilisation de ChatGPT et d'autres outils conversationnels dans divers contextes universitaires.

6.2. Perspectives

Dans cette étude, nous avons veillé à sélectionner des participants provenant de domaines académiques différents afin d'explorer la manière dont ChatGPT est utilisé dans des contextes de recherche variés. R1 l'a utilisé pour une recherche traditionnelle lors de son stage, R2 pour comparer des textes, R3 pour générer des idées dans le domaine touristique marocain, R4 pour résumer des cours, R5 pour rédiger un éditorial en journalisme, et R6 pour la programmation. Cette diversité de profils a permis de couvrir un large éventail de pratiques et d'utilisation de ChatGPT.

Mais est-ce que ChatGPT performe dans des domaines plus que d'autres ?

Une question importante à explorer dans les recherches futures est de savoir si ChatGPT offre des résultats plus fiables dans certains domaines plus que d'autres. Par exemple, dans les disciplines où l'originalité et l'analyse critique sont primordiales, comme en journalisme ou en littérature, l'utilisation de ChatGPT pourrait poser des défis éthiques et académiques. En revanche, dans des domaines comme la programmation, où ChatGPT peut servir d'assistant technique, son utilisation pourrait être plus justifiable et bénéfique.

Il est également opportun d'étudier si les professeurs ont remarqué un changement d'humeur chez les étudiants après l'avènement de ChatGPT. Surtout, il sera intéressant de comparer les résultats de l'apprentissage entre les étudiants belges et étrangers et de comprendre s'il existe une différence significative en raison de l'utilisation de ChatGPT. Cela aidera à mieux comprendre l'importance réelle de l'outil pour l'apprentissage et la recherche universitaire. L'étude des professeurs sera menée pour discuter de l'opinion des professeurs sur les compétences et les résultats des étudiants après l'avènement de ChatGPT.

De plus, une autre question pertinente pour les recherches futures concerne les bibliothèques. Est-ce qu'il y a eu des changements dans les habitudes d'emprunt des étudiants depuis l'arrivée de ChatGPT ? Si les étudiants se tournent davantage vers des outils numériques comme ChatGPT pour leurs recherches, cela pourrait réduire leur dépendance aux ressources physiques des bibliothèques. Cette perspective mérite d'être explorée pour comprendre l'impact global de ChatGPT sur les pratiques de recherche universitaire.

Autre perspective, nous pouvons voir que bien que les étudiants utilisent régulièrement ChatGPT, ils sont encore extrêmement sceptique quant à sa crédibilité. Cependant, il serait intéressant pour les développeurs d'intégrer des fonctionnalités d'affichage des sources directement dans l'outil. Cela aiderait à renforcer la crédibilité, à gagner plus de temps en réduisant la charge de vérification des résultats par les chercheurs ce qui pourrait rendre

ChatGPT un outil encore plus fiable et utile pour les recherches universitaires. Cette amélioration garantirait que les informations fournies sont non seulement rapides et accessibles, mais aussi précises, vérifiables et crédibles.

Du côté de l'enseignement, il est essentiel d'intégrer ChatGPT de manière réfléchie et encadrée dans les programmes académiques. Les institutions éducatives pourraient sensibiliser les étudiants à utiliser ChatGPT de manière éthique et efficace. En encourageant une utilisation critique et éclairée de l'outil, les enseignants peuvent aider les étudiants à développer des compétences essentielles en matière de vérification des sources pour assurer la crédibilité de l'outil. Ils pourraient également imposer un taux d'utilisation de l'IA variable en fonction du sujet et du domaine de recherche. Cela permettrait de limiter l'utilisation de l'IA dans certaines recherches et de revenir à une approche où l'analyse et la critique humaines sont prioritaires.

Les résultats de cette étude mettent en lumière les multiples facettes de l'utilisation de ChatGPT dans le contexte universitaire. Pour maximiser les avantages et minimiser les risques associés à cet outil, il est très important de continuer à explorer ces questions dans des recherches futures.

Conclusion

Notre inspiration pour ce travail est basée sur deux réflexions complémentaires : les stratégies de recherche d'informations numériques à l'ère de ChatGPT et la vérification des résultats. Durant cette recherche, nous avons identifié comment l'intégration de ChatGPT dans les pratiques académiques transforme profondément la manière dont les étudiants universitaires, en Belgique et au Maroc, abordent la recherche d'information.

Nos observations, entretiens et analyses ont révélé que ChatGPT représente un atout majeur pour les étudiants en offrant un gain de temps considérable. Il permet d'accéder rapidement à des informations pertinentes, simplifiant ainsi des processus de recherche souvent laborieux. En facilitant une interaction plus ciblée et précise avec les sources d'information, ChatGPT modifie également la dynamique de la recherche, la rendant plus intuitive et accessible. Cependant, cette transformation s'accompagne de défis nouveaux et significatifs, en particulier concernant la vérification de la crédibilité des informations fournies et l'adaptation des compétences nécessaires pour utiliser cet outil de manière optimale.

Si les étudiants apprécient la rapidité et la commodité que ChatGPT apporte, notre étude a aussi montré qu'ils doivent développer des compétences critiques, non seulement pour évaluer la pertinence et la fiabilité des informations qu'ils obtiennent, mais aussi pour vérifier l'originalité des contenus générés. Il est essentiel que les utilisateurs ne se contentent pas des premières réponses fournies, mais qu'ils apprennent à questionner et à approfondir les données pour s'assurer de leur exactitude. Ce besoin de rigueur est d'autant plus important dans un contexte où la dépendance à des outils comme ChatGPT pourrait conduire à une certaine complaisance ou à une érosion des compétences en recherche traditionnelle.

L'intégration de fonctionnalités de vérification des sources directement au sein de ChatGPT serait une avancée significative pour renforcer la crédibilité des informations fournies et pour soutenir les utilisateurs dans leur quête de rigueur académique. Les développeurs d'OpenAI et d'autres outils similaires devraient envisager de telles améliorations pour maximiser l'efficacité de ces outils tout en minimisant les risques associés à une utilisation non critique.

Cette recherche a donc confirmé nos hypothèses initiales tout en soulignant l'importance cruciale d'une utilisation éclairée et éthique de ChatGPT dans le milieu académique. Les

résultats obtenus offrent une perspective enrichissante pour l'avenir, suggérant la nécessité de développer des stratégies pédagogiques adaptées. Ces stratégies devraient viser à préparer les étudiants non seulement à exploiter pleinement les avantages offerts par l'intelligence artificielle, mais aussi à maintenir un esprit critique et une approche méthodique face aux informations qu'ils consomment. Alors que l'intelligence artificielle continue de transformer l'éducation et la recherche, il est impératif que ces évolutions technologiques soient accompagnées par un développement parallèle des compétences humaines en évaluation critique, en vérification des sources, et en réflexion éthique.

En conclusion, ce travail met en lumière non seulement les opportunités, mais aussi les responsabilités accrues qui incombent aux étudiants, aux éducateurs et aux concepteurs de technologies dans cette nouvelle ère numérique. L'adaptation continue à ces nouvelles dynamiques sera essentielle pour garantir que les outils comme ChatGPT soient utilisés non seulement efficacement, mais aussi de manière responsable, afin de soutenir une recherche académique de qualité.

Bibliographie

AbuShawar, B., Atwell, E., AbuShawar, B., & Atwell, E. (2015). ALICE Chatbot : Trials and Outputs. *Computación y Sistemas*, 19(4), 625-632. <https://doi.org/10.13053/CyS-19-4-2326>

Aissa, W., Soulier, L., & Denoyer, L. (2019, mars 27). *Modèle de compréhension du besoin en information pour la RI conversationnelle*. CORIA 2019 - 16ème Conférence en Recherche d'Information et Applications. <https://hal.sorbonne-universite.fr/hal-02150653>

Akoka, J., Comyn-Wattiau, I., & Mouza, C. D. (s. d.). *Évaluer la crédibilité des sources historiques*.

Au cœur du développement de la génération de contenu IA. (2023, juin 17). <https://www.daillac.com/blogue/liberer-la-creativite-une-plongee-au-coeur-du-developpement-de-la-generation-de-contenu-ia/>

Bersini, H., 1961-... (2023). *ChatGPT : il était une fois une IA régressive*. Éditions de l'Université de Bruxelles. <https://ils.bib.uclouvain.be/global/documents/3724260>

Boubée, N., & Tricot, A. (2010). Chapitre II. La recherche d'information à travers les disciplines. In *Qu'est-ce que rechercher de l'information ?* (p. 86-192). Presses de l'enssib. <https://doi.org/10.4000/books.pressesenssib.807>

Brin, D., Sorin, V., Vaid, A., Soroush, A., Glicksberg, B. S., Charney, A. W., Nadkarni, G., & Klang, E. (2023). Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. *Scientific Reports*, 13(1), 16492. <https://doi.org/10.1038/s41598-023-43436-9>

Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, 26(7), 233-241. <https://doi.org/10.1108/EJIM-02-2023-0156>

ChatGPT dans l'enseignement supérieur : « La vraie question n'est pas son interdiction mais son usage ». (s. d.). Consulté 15 août 2024, à l'adresse <https://www.univ-gustave-eiffel.fr/luniversite/pages-speciales/vue-detaillee/chatgpt-dans-lenseignement-superieur-la-vraie-question-nest-pas-son-interdiction-mais-son-usage>

Chiarabella, Y., & Mulhem, P. (2007c). La recherche d'information. De la documentation automatique à la recherche d'information en contexte. *Document numérique*, 10(1), 11-38.

Deville, Y., & Jacqmot, C. (s. d.). *ChatGPT : Menace ou opportunité pour l'enseignement ?*

Dinet, J., & Rouet, J.-F. (2002). La recherche d'information : Processus cognitifs, facteurs de difficultés et dimension de l'expertise. *Interaction homme - machine et recherche d'information*, 133. <https://shs.hal.science/halshs-00009685>

Fischer-Lokou, J., & Pinte, G. (2017). L'effet Hawthorne. *Le Cercle psy*, Novembre/décembre(6), 22.

- Forest, D. (2009). Vers une nouvelle génération d'outils d'analyse et de recherche d'information. *Documentation et bibliothèques*, 55(2), 77-89. <https://doi.org/10.7202/1029091ar>
- Godé, C., Bidan, M., & Lebraty, J.-F. (2023). CHATGPT, ETUDIANTS ET ENSEIGNANTS-CHERCHEURS : SONT-ILS VRAIMENT FELINS POUR L'AUTRE ? *Management & Datascience*, 7(2). <https://doi.org/10.36863/mds.a.23595>
- GPT-4. (s. d.). Consulté 9 avril 2024, à l'adresse <https://openai.com/gpt-4>
- Guével, M.-R., & Pommier, J. (2012). Recherche par les méthodes mixtes en santé publique : Enjeux et illustration. *Santé Publique*, 24(1), 23-38. <https://doi.org/10.3917/spub.121.0023>
- Haddad, H. (s. d.). *Utilisation des Syntagmes Nominaux dans un Système de Recherche d'Information*.
- Hassani, H., & Silva, E. S. (2023). The Role of ChatGPT in Data Science : How AI-Assisted Conversational Interfaces Are Revolutionizing the Field. *Big Data and Cognitive Computing*, 7(2), Article 2. <https://doi.org/10.3390/bdcc7020062>
- Jacquinet-Delaunay, G. (2001). Les sciences de l'éducation et de la communication en dialogue : À propos des médias et des technologies éducatives. *L'Année sociologique*, Vol.51(2), 391-410. <https://doi.org/10.3917/anso.012.0392>
- Karim, O., & El Hassouni, S. (2023). Les conflits interpersonnels et l'interculturalité dans l'entreprise : Une revue de littérature systématique. *Revue Congolaise de Gestion, Numéro 36(2)*, 92-141. <https://doi.org/10.3917/rcg.036.0092>
- Kassel, R. (2024, février 16). NLU (Natural Language Understanding) : Qu'est-ce que c'est ? *Formation Data Science | DataScientest.com*. <https://datascientest.com/nlu-natural-language-understanding-tout-savoir>
- Kelleher, J. D. (2019). *Deep Learning*. MIT Press.
- L, B. (2024, mai 15). *GPT-4 : La nouvelle IA surpuissante de ChatGPT déjà disponible ! Tout savoir*. LEBIGDATA.FR. <https://www.lebigdata.fr/gpt-4-tout-savoir>
- Le TALN appliqué aux chatbots*. (2018, avril 23). ActuIA. <https://www.actuia.com/contribution/sarah-marchand/le-taln-applique-aux-chatbots/>
- Lobet, M., Honet, A., & Wathelet, V. (s. d.). *ChatGPT, l'usage par les étudiants de première année à l'université ChatGPT*.
- Maimon, O., & Rokach, L. (2010). Introduction to Knowledge Discovery and Data Mining. In O. Maimon & L. Rokach (Éds.), *Data Mining and Knowledge Discovery Handbook* (p. 1-15). Springer US. https://doi.org/10.1007/978-0-387-09823-4_1
- Matteis, L. D. (s. d.-b). *Introduction à l'apprentissage*.
- McShane, M. (2017). Natural Language Understanding (NLU, not NLP) in Cognitive

Systems. *AI Magazine*, 38(4), Article 4. <https://doi.org/10.1609/aimag.v38i4.2745>

Meissonier, R. (2023). La Pensée Complexe contre l'Intelligence Artificielle Dégénérative. *Management & Datascience*, 7(3). <https://doi.org/10.36863/mds.a.24107>

Memoire Online—Une contribution du datamining à fÀ la segmentation du marché et au ciblage des offres à fÀ l'aide de la statistique de Bayes—Samuel Kalombo. (s. d.). Memoire Online. Consulté 17 août 2024, à l'adresse

https://www.memoireonline.com/10/13/7549/m_Une-contribution-du-datamining-la-segmentation-du-march-et-au-ciblage-des-offres--l-aide2.html

Merveille, N. (2022). Une réflexion anthropologique sur la cité numérisée. *Vertigo - la revue électronique en sciences de l'environnement, Hors-série 36*, Article Hors-série 36. <https://doi.org/10.4000/vertigo.37323>

Mooers, C. N. (1948). *Application of random codes to the gathering of statistical information* [Thesis, Massachusetts Institute of Technology]. <https://dspace.mit.edu/handle/1721.1/12664>

Moussavou, J. (2023). ChatGPT dans la rédaction scientifique : Des perspectives prometteuses, des préoccupations légitimes. *Management & Datascience*. <https://doi.org/10.36863/mds.a.23913>

Nie, J.-Y. (s. d.). *Le domaine de recherche d'information – Un survol d'une longue histoire. Qu'est-ce que le Data Mining.* (s. d.). Oracle France. Consulté 22 mars 2024, à l'adresse <https://www.oracle.com/fr/database/data-mining-definition.html>

Quinio, B., & Bidan, M. (2023). ChatGPT : Un robot conversationnel peut-il enseigner ? *Management & Datascience*, 7(1). <https://doi.org/10.36863/mds.a.22060>

Renard, É., & Roussiau, N. (2007). Transformation des représentations sociales et persuasion (modèle ELM) : Les effets de la crédibilité de la source. *Bulletin de psychologie, Numéro 489(3)*, 211-224. <https://doi.org/10.3917/bupsy.489.0211>

Risser-Maroux, O., & Kerboua-Benlarbi, S. (2018). *Agents conversationnels pour la recherche d'information. (Traduction de requêtes en langage naturel vers mots clefs).*

Robert, J. (2024, juin 14). Generated Pre-trained Transformer (GPT) : Qu'est-ce que c'est ? Comment ça marche ? *Formation Data Science | DataScientest.com*. <https://datascientest.com/generative-pretrained-transformer>

Rodhain, F., Fallery, B., Girard, A., & Desq, S. (2010). Une histoire de la recherche en systèmes d'information à travers 30 ans de publications. *Entreprises et histoire*, 60(3), 78-97. <https://doi.org/10.3917/eh.060.0078>

Segonne, V. (2021). *Désambiguïsation lexicale automatique des verbes du français* [Phdthesis, Université Paris Cité]. <https://theses.hal.science/tel-04050259>

TH.M.INF.FR.2019.30.pdf. (s. d.). Consulté 13 août 2024, à l'adresse <http://dspace.univ-tiaret.dz/bitstream/123456789/5328/1/TH.M.INF.FR.2019.30.pdf>

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313-313. <https://doi.org/10.1126/science.adg7879>

Une brève histoire des chatbots – Partie 2/2 – Les chatbots célèbres. (s. d.-a). Consulté 18 août 2024, à l'adresse <https://blog.clevy.io/fr/une-breve-histoire-des-chatbots-partie-2-les-chatbots-celebres/>

Une brève histoire des chatbots – Partie 2/2 – Les chatbots célèbres. (s. d.-b). Consulté 13 août 2024, à l'adresse <https://blog.clevy.io/fr/une-breve-histoire-des-chatbots-partie-2-les-chatbots-celebres/>

Warwick, K., & Shah, H. (2016). Can machines think? A report on Turing test experiments at the Royal Society. *Journal of Experimental & Theoretical Artificial Intelligence*. <https://www.tandfonline.com/doi/abs/10.1080/0952813X.2015.1055826>

Wilson, L., & Marasoiu, M. (2022). The Development and Use of Chatbots in Public Health : Scoping Review. *JMIR Human Factors*, 9(4), e35882. <https://doi.org/10.2196/35882>

Yvon, F. (s. d.). *Une petite introduction au Traitement Automatique des Langues Naturelles.*
Zafar, A. (2020, juin 22). 11 Best Natural Language Processing Online Courses. *MLTut*. <https://www.mltut.com/best-natural-language-processing-courses-online-to-become-expert/>

L'usage des plateformes d'intelligence artificielle (IA), notamment ChatGPT, a bouleversé le monde académique depuis décembre 2022. Cette recherche s'intéresse à l'utilisation de ChatGPT par des étudiants universitaires au Maroc et en Belgique. Elle explore comment cet outil conversationnel influence leurs stratégies de recherche d'information et leurs pratiques académiques. En analysant les pratiques d'utilisation dans des contextes culturels variés, cette étude vise à comprendre l'impact de ChatGPT sur la recherche universitaire et à identifier les différences dans l'adoption et l'utilisation de l'outil entre les deux pays et comparer leurs stratégies de vérification de résultats. De plus, elle propose des recommandations pour une intégration éthique et efficace de cette technologie dans les programmes académiques, en mettant l'accent sur l'importance de la vérification des sources et l'utilisation critique de l'IA pour garantir la qualité et l'intégrité de l'apprentissage académique.

Mots-clefs : Intelligence artificielle, ChatGPT, Recherche universitaire, Recherche d'information, Chatbots, Crédibilité des résultats

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des sciences économique, sociales, politiques et de communication

Place Montesquieu, 4 bte L2.05.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/espo