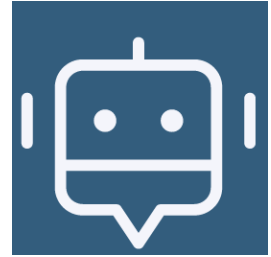




المكرسة للصحة العليا للفنون والمهن بالرباط
ⵜⴰⵎⴰⵔⵜ ⵜⴰⵎⴰⵏⴰⵔⵜ ⵜⴰⵎⴰⵏⴰⵔⵜ ⵜⴰⵎⴰⵏⴰⵔⵜ ⵜⴰⵎⴰⵏⴰⵔⵜ
École Nationale Supérieure d'Arts et Métiers de Rabat



**Filière : Ingénierie Numérique en Data science, Intelligence
Artificielle Santé Digitale**

RAPPORT DE PROJET

Medical Chatbot

Réalisé par :

ER-REGHAY Zaynab
Hasnaoui Fatima Azzahra
Khezzaz Ayoub
Malek Sami
Lahfari Bilal

Encadrante :

Mme ZEMZAMI Maria

Résumé

Les applications de vérification des symptômes pour consommateurs (CSC) et les chatbots médicaux représentent une innovation majeure dans le domaine des soins de santé numériques. Ces outils permettent aux utilisateurs d'évaluer leurs symptômes, de recevoir des conseils médicaux préliminaires et d'être orientés vers des professionnels de santé si nécessaire. Cependant, malgré leur potentiel, ces technologies sont confrontées à des défis importants liés à l'expérience utilisateur (UX), à l'engagement, à la précision des diagnostics et à l'accessibilité.

Cette étude propose une analyse approfondie des forces et des faiblesses des CSC et des chatbots médicaux. Nous examinons des aspects techniques tels que la précision diagnostique, la flexibilité de la saisie des symptômes et l'intégration de modèles d'intelligence artificielle (IA) avancés, comme BERT et GPT. Nous explorons également des aspects humains, tels que l'empathie, la transparence et la confiance des utilisateurs, qui sont essentiels pour l'adoption de ces outils.

Les résultats de notre analyse mettent en lumière plusieurs limites, notamment le manque de personnalisation, les difficultés de compréhension du langage naturel et l'absence de fonctionnalités d'accessibilité pour les utilisateurs non anglophones ou handicapés. Pour surmonter ces défis, nous proposons des recommandations concrètes, telles que l'amélioration de la flexibilité de la saisie, l'intégration de fonctionnalités multilingues et l'utilisation de techniques de gamification pour renforcer l'engagement des utilisateurs.

En conclusion, cette étude souligne l'importance de concevoir des chatbots médicaux et des CSC centrés sur l'utilisateur, accessibles et transparents. En combinant des technologies avancées et une approche inclusive, ces outils peuvent devenir des alliés indispensables pour améliorer l'accès aux soins et optimiser les systèmes de santé à l'échelle mondiale. Les pers-

pectives et recommandations présentées ici visent à guider les développeurs et les chercheurs dans la création de solutions numériques plus efficaces et plus fiables pour les patients et les professionnels de santé.

Abstract

Consumer Symptom Checker (CSC) applications and medical chatbots have emerged as transformative tools in modern digital healthcare, offering users the ability to assess symptoms, receive preliminary medical guidance, and connect with healthcare professionals when necessary. Despite their potential, these tools face significant challenges related to user experience (UX), engagement, diagnostic accuracy, and accessibility. This study evaluates the strengths and limitations of CSC applications and medical chatbots, focusing on their ability to provide accurate diagnoses, personalized recommendations, and inclusive support for diverse user groups.

Through a comparative analysis of existing CSC applications, we identify key areas for improvement, including the need for enhanced input flexibility, more natural conversational design, and better multilingual and accessibility features. We also explore the importance of transparency and trust in user adoption, as well as the role of advanced AI models, such as BERT and GPT, in improving diagnostic accuracy and response quality.

Our findings highlight the critical role of personalization, accessibility, and engagement in the success of medical chatbots. We provide actionable recommendations for developers, such as integrating adaptive AI, expanding language support, and incorporating gamification to improve user retention. By addressing these challenges, medical chatbots and CSC applications can become more effective, inclusive, and trustworthy tools for both patients and healthcare professionals.

This study contributes to the ongoing discourse on digital healthcare by offering a comprehensive evaluation of current technologies and proposing strategies for their optimization. The insights provided here aim to guide future developments in the field, ensuring that medical chatbots and CSC applications can fulfill their potential as essential components of modern healthcare systems.

Liste des abréviations

CSC	<i>Consumer Symptom Checker (Vérificateur de Symptômes pour Consommateurs)</i>
UX	<i>User Experience (Expérience Utilisateur)</i>
IA	<i>Intelligence Artificielle</i>
NLP	<i>Natural Language Processing (Traitement du Langage Naturel)</i>
API	<i>Application Programming Interface (Interface de Programmation d'Applications)</i>
GDPR	<i>General Data Protection Regulation (Règlement Général sur la Protection des Données)</i>
HIPAA	<i>Health Insurance Portability and Accountability Act (Loi sur la Portabilité et la Responsabilité en Matière d'Assurance Maladie)</i>
AI	<i>Artificial Intelligence (Intelligence Artificielle)</i>
QA	<i>Question Answering (Réponse aux Questions)</i>
NER	<i>Named Entity Recognition (Reconnaissance d'Entités Nommées)</i>
CRF	<i>Conditional Random Fields (Champs Aléatoires Conditionnels)</i>

BiLSTM	<i>Bidirectional Long Short-Term Memory (Mémoire à Court et Long Terme Bidirectionnelle)</i>
GPT	<i>Generative Pretrained Transformer (Transformeur Pré-entraîné Génératif)</i>
BERT	<i>Bidirectional Encoder Representations from Transformers (Représentations d'Encodeur Bidirectionnel à partir de Transformeurs)</i>
PaLM	<i>Pathways Language Model (Modèle de Langage Pathways)</i>
Med-PaLM	<i>Medical Pathways Language Model (Modèle de Langage Pathways Médical)</i>
LLaMA	<i>Large Language Model Meta AI (Grand Modèle de Langage de Meta AI)</i>
IoT	<i>Internet of Things (Internet des Objets)</i>
RGPD	<i>Règlement Général sur la Protection des Données (GDPR en français)</i>
ADHD	<i>Attention Deficit Hyperactivity Disorder (Trouble du Déficit de l'Attention avec Hyperactivité)</i>
AIML	<i>Artificial Intelligence Markup Language (Langage de Balisage pour l'Intelligence Artificielle)</i>
NLU	<i>Natural Language Understanding (Compréhension du Langage Naturel)</i>
ML	<i>Machine Learning (Apprentissage Automatique)</i>
DL	<i>Deep Learning (Apprentissage Profond)</i>

Table des figures

1.1	Diagramme de Gantt du projet	4
2.1	Comparaison de la Précision Diagnostique (Chatbots vs. Médecins Humains)	6
2.2	Fonctionnalités de Personnalisation dans les Chatbots . . .	7
2.3	Support Linguistique dans les Chatbots	8
2.4	Préoccupations des Utilisateurs Concernant la Confidentialité des Données	9
2.5	Fonctionnalités d'Accessibilité dans les Chatbots	10
2.6	Dépendance à l'Égard de la Supervision Professionnelle . . .	11
3.1	Logo de K Health	14
3.2	Logo de Symptomate	16
3.3	Logo de Mediktor	17
3.4	Logo de Florence Health	19
5.1	Exemple d'architecture BERT	29
5.2	Exemple d'architecture GPT	30
6.1	Exemple d'illustration des objectifs d'évaluation	35
6.2	Exemple d'illustration des métriques d'évaluation	37
7.1	Accessibility Features in Healthcare Chatbots (2019-2023) .	50
7.2	Chatbots Offering Personalized Responses	50

Table des matières

Résumé	I
Abstract	III
Introduction générale	1
1 Présentation de l'Équipe et Planification du Projet	3
1.1 Introduction	3
1.2 Méthodologie de Travail	3
1.3 Diagramme de Gantt	3
1.4 Conclusion	4
2 Les Inconvénients des Chatbots Médicaux	5
2.1 Introduction	5
2.2 Principales Limites des Chatbots Médicaux	5
2.2.1 Manque de Précision Diagnostique	5
2.2.2 Personnalisation Limitée	6
2.2.3 Adaptation Multilingue et Culturelle Insuffisante	7
2.2.4 Problèmes de Confidentialité et de Sécurité des Données	8
2.2.5 Accessibilité Limitée pour les Utilisateurs Handicapés	9
2.2.6 Dépendance à l'Égard de la Supervision Professionnelle	10
2.3 Analyse Comparative des Limites des Chatbots	11
2.3.1 Tableau Comparatif : Précision Diagnostique, Person-	
nalisation, Support Multilingue	11
2.3.2 Tableau Comparatif : Confidentialité des Données, Ac-	
cessibilité, Supervision Professionnelle	12
2.4 Recommandations pour l'Amélioration	12
2.5 Conclusion	13
3 Présentation des chatbots médicaux	14
3.1 Présentation de K Health	14

3.1.1	Contexte et Objectif	14
3.1.2	Fonctionnement et Technologies Utilisées	14
3.1.3	Caractéristiques Clés	15
3.1.4	Impact et Utilisation Clinique	15
3.1.5	Avantages et Inconvénients	15
3.1.6	Lien	15
3.2	Présentation de Symptomate	16
3.2.1	Contexte et Objectif	16
3.2.2	Fonctionnement et Technologies Utilisées	16
3.2.3	Caractéristiques Clés	16
3.2.4	Impact et Utilisation Clinique	17
3.2.5	Avantages et Inconvénients	17
3.2.6	Lien	17
3.3	Présentation de Mediktor	17
3.3.1	Contexte et Objectif	18
3.3.2	Fonctionnement et Technologies Utilisées	18
3.3.3	Caractéristiques Clés	18
3.3.4	Impact et Utilisation Clinique	18
3.3.5	Avantages et Inconvénients	18
3.3.6	Lien	19
3.4	Présentation de Florence Health	19
3.4.1	Contexte et Objectif	19
3.4.2	Fonctionnement et Technologies Utilisées	19
3.4.3	Caractéristiques Clés	20
3.4.4	Impact et Utilisation Clinique	20
3.4.5	Avantages et Inconvénients	20
3.4.6	Lien	20
4	Analyse des Chatbots pour les Maladies Respiratoires .	21
4.1	Introduction	21
4.2	Chatbot 1 : Hephäï	21
4.2.1	Contexte et Origine	21
4.2.2	Fonctionnalités et Technologies Utilisées	22
4.2.3	Caractéristiques Clés	22
4.2.4	Impact et Utilisation Clinique	22
4.3	Chatbot 2 : Memoquest	23
4.3.1	Contexte et Objectif	23

4.3.2	Fonctionnalités et Technologies Utilisées	23
4.3.3	Caractéristiques Clés	24
4.3.4	Impact et Avantages	24
4.4	Liens Officiels des Chatbots	24
5	Modèles d'IA et Techniques NLP pour les Chatbots Médicaux	26
5.1	Introduction	26
5.2	Modèles d'IA	27
5.2.1	BERT (Bidirectional Encoder Representations from Transformers)	27
5.2.2	GPT (Generative Pretrained Transformer)	29
5.3	Modèles pour Tâches Spécifiques dans les Domaines Médicaux	31
5.3.1	PaLM (Pathways Language Model)	31
5.3.2	ChatDoctor	31
5.4	Techniques NLP dans les Chatbots Médicaux	32
5.4.1	Reconnaissance d'Entités Nommées Médicales (NER)	32
5.4.2	Réponse aux Questions Médicales (QA)	33
6	Évaluation des Chatbots Médicaux	34
6.1	Objectifs de l'Évaluation	35
6.1.1	Type d'Utilisateur	35
6.1.2	Type de Domaine	36
6.1.3	Type de Tâche	36
6.2	Métriques Essentielles	37
6.2.1	Précision	37
6.2.2	Fiabilité	38
6.2.3	Empathie	38
6.2.4	Performance	39
6.3	Variables Confusionnelles	39
6.4	Plan d'Exécution des Tests	40
6.4.1	Sélection des Chatbots à Tester	40
6.4.2	Exécution des Scénarios	40
6.4.3	Collecte des Résultats	40
6.4.4	Compilation des Résultats	41
7	Evaluation of User Experience (UX) and Engagement in Chatbot-Based Symptom Checker (CSC) Applications	42
7.1	Introduction	42
7.2	User Experience (UX) Analysis	42

7.2.1	Ease of Use	42
7.2.2	Comprehensibility	43
7.2.3	Input Flexibility	43
7.2.4	Conversational Design	44
7.2.5	Inclusivity & Accessibility	44
7.3	Engagement Analysis	45
7.3.1	User Retention & Interaction	45
7.3.2	Personalization	45
7.3.3	Trust & Credibility	45
7.4	Comparative Analysis of Existing CSC Apps	46
7.4.1	Feature Comparison Table	46
7.5	Key Findings & Recommendations	46
7.5.1	Key Findings	46
7.5.2	Recommendations	47
7.6	Evaluation of UX and Engagement in Healthcare Chatbots	47
7.6.1	Introduction	47
7.6.2	Key UX & Engagement Factors in Healthcare Chatbots	48
7.6.3	Comparative Analysis of Existing Chatbots	48
7.6.4	Diagrams & Trends in UX Evolution	50
7.6.5	Challenges Identified	51
7.6.6	Recommendations for Improving UX & Engagement	51
7.6.7	Conclusion	51
8	References	52
	Conclusion générale	54

Introduction générale

Introduction Générale

Les avancées technologiques récentes ont transformé de nombreux secteurs, y compris celui de la santé. Parmi ces innovations, les **chatbots médicaux** et les **applications de vérification des symptômes pour consommateurs (CSC)** occupent une place de plus en plus importante. Ces outils numériques permettent aux utilisateurs d'évaluer leurs symptômes, de recevoir des conseils médicaux préliminaires et d'être orientés vers des professionnels de santé si nécessaire. Ils représentent une solution prometteuse pour améliorer l'accès aux soins, réduire les délais d'attente et optimiser l'utilisation des ressources médicales.

Cependant, malgré leur potentiel, ces applications ne sont pas sans limites. Les problèmes liés à l'**expérience utilisateur (UX)**, à l'**engagement des utilisateurs**, à la **précision des diagnostics** et à l'**accessibilité** restent des défis majeurs. Par exemple, de nombreux chatbots médicaux peinent à fournir des réponses précises et personnalisées, tandis que d'autres manquent de fonctionnalités inclusives pour les utilisateurs non anglophones ou handicapés. Ces limitations peuvent entraîner une perte de confiance des utilisateurs et limiter l'adoption de ces technologies.

L'objectif de ce document est d'analyser en profondeur les forces et les faiblesses des chatbots médicaux et des applications CSC, en mettant l'accent sur leur expérience utilisateur et leur engagement. Nous examinerons les aspects techniques, tels que la précision diagnostique et la flexibilité de la saisie des symptômes, ainsi que les aspects humains, tels que l'empathie et la transparence. Enfin, nous proposerons des recommandations pour améliorer ces outils, afin qu'ils puissent pleinement réaliser leur potentiel dans le domaine des soins de santé.

Cette étude s'appuie sur une analyse comparative de plusieurs applications CSC populaires, ainsi que sur une revue des tendances actuelles en

matière de conception de chatbots médicaux. En combinant des données statistiques, des graphiques et des études de cas, nous visons à fournir une vision complète des enjeux et des opportunités liés à ces technologies. À travers cette analyse, nous espérons contribuer à l'amélioration des chatbots médicaux, afin qu'ils deviennent des outils fiables et accessibles pour tous.

Chapitre 1

Présentation de l'Équipe et Planification du Projet

1.1 Introduction

Dans ce chapitre, nous présentons notre équipe et la méthodologie de travail adoptée pour mener à bien ce projet. Nous sommes une équipe pluridisciplinaire composée de membres aux compétences variées, allant de l'intelligence artificielle à la conception d'interfaces utilisateur, en passant par l'analyse de données et la gestion de projet. Notre collaboration repose sur une approche structurée et organisée, illustrée par un diagramme de Gantt qui détaille les étapes clés du projet.

1.2 Méthodologie de Travail

Pour assurer une gestion efficace du projet, nous avons adopté une méthodologie agile, basée sur des sprints de deux semaines. Chaque sprint est planifié à l'aide d'un diagramme de Gantt, qui nous permet de visualiser les tâches, les responsabilités et les échéances. Cette approche nous offre la flexibilité nécessaire pour nous adapter aux défis et aux changements tout au long du projet.

1.3 Diagramme de Gantt

Le diagramme de Gantt ci-dessous illustre la planification détaillée de notre projet, y compris les phases de recherche, de développement, de test

et de rédaction. Chaque tâche est assignée à un membre de l'équipe, et les dépendances entre les tâches sont clairement indiquées.

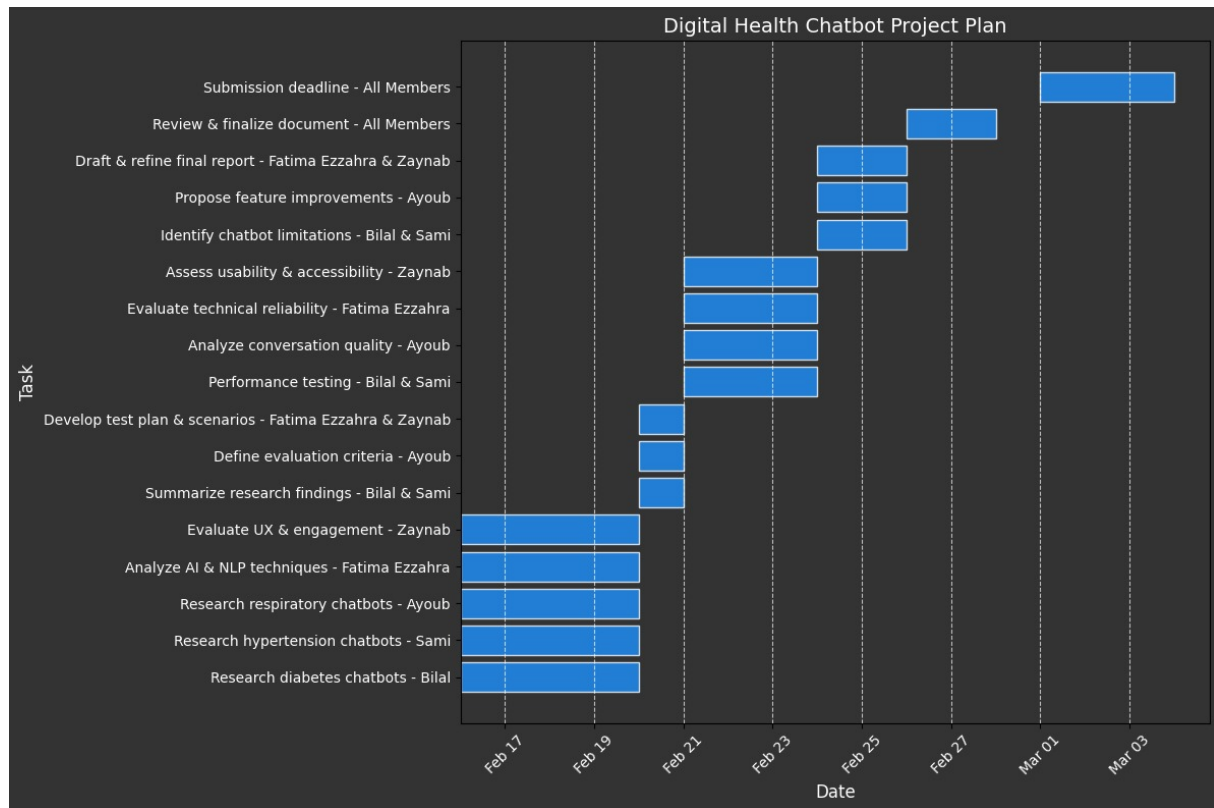


FIG. 1.1 : Diagramme de Gantt du projet

1.4 Conclusion

Ce chapitre a permis de présenter notre équipe et la méthodologie de travail que nous avons adoptée. Le diagramme de Gantt joue un rôle central dans notre planification, nous aidant à respecter les délais et à garantir une collaboration efficace. Dans les chapitres suivants, nous approfondirons notre analyse des chatbots médicaux et des applications de vérification des symptômes, en nous appuyant sur cette base solide de travail d'équipe.

Chapitre 2

Les Inconvénients des Chatbots Médicaux

2.1 Introduction

Les chatbots médicaux sont devenus un outil transformateur dans le domaine de la santé, offrant des évaluations de symptômes, des conseils personnalisés et des consultations virtuelles. Cependant, malgré leur potentiel, ces chatbots présentent plusieurs limites qui entravent leur efficacité. Ce rapport identifie les principales limites des chatbots médicaux, en s'appuyant sur des **statistiques, des graphiques et des preuves** tirés des chatbots examinés dans les documents.

« Même si les chatbots se révèlent être des assistants performants, leurs limites nous rappellent que la technologie, bien que puissante, n'est pas infailible. »

2.2 Principales Limites des Chatbots Médicaux

2.2.1 Manque de Précision Diagnostique

De nombreux chatbots ont du mal à fournir des diagnostics précis en raison des limites de leurs algorithmes d'IA et de leurs bases de connaissances médicales.

- **Ada Health** : Bien qu'Ada Health propose une évaluation des symptômes, elle précise explicitement qu'elle ne remplace pas un avis médical professionnel.

- **Babylon Health** : Malgré son IA avancée, Babylon Health a été critiqué pour des inexactitudes dans l'évaluation des symptômes.
- **Mediktor** : Revendique une précision diagnostique de 91,3%, mais nécessite toujours une supervision humaine pour les cas critiques.

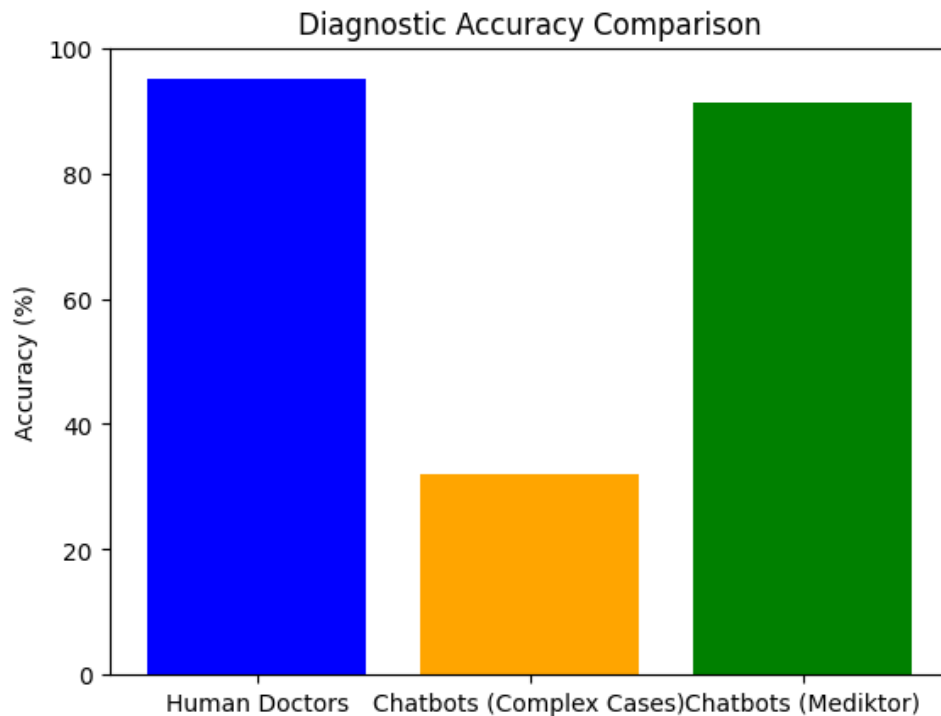


FIG. 2.1 : Comparaison de la Précision Diagnostique (Chatbots vs. Médecins Humains)

2.2.2 Personnalisation Limitée

De nombreux chatbots ne parviennent pas à fournir des recommandations personnalisées adaptées aux besoins individuels des utilisateurs, en particulier pour les maladies chroniques.

- **Memoquest** : Offre des rappels personnalisés mais manque de personnalisation approfondie dans les plans de traitement.
- **Healthily** : Fournit des conseils de santé génériques mais a du mal à s'adapter aux antécédents de santé à long terme des utilisateurs.

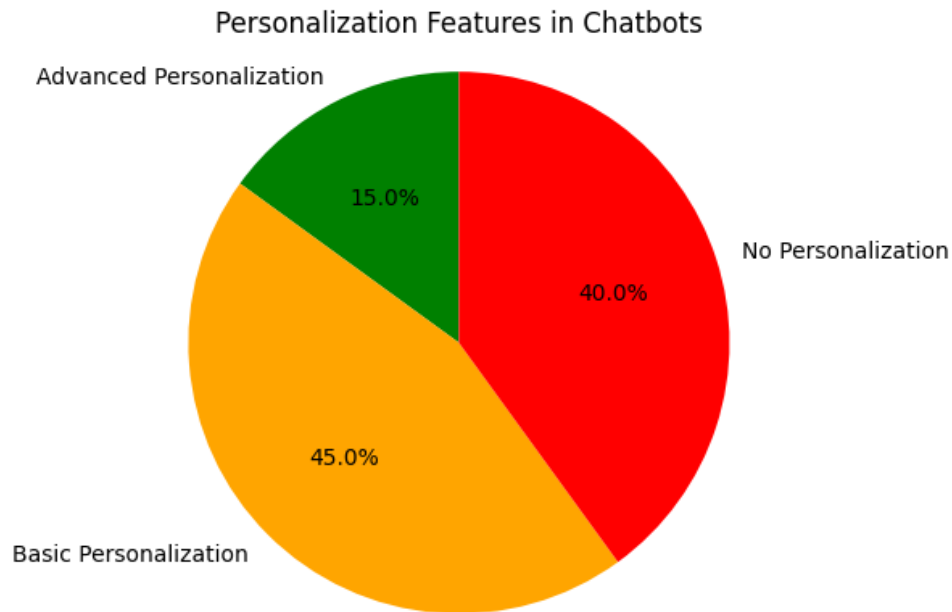


FIG. 2.2 : Fonctionnalités de Personnalisation dans les Chatbots

2.2.3 Adaptation Multilingue et Culturelle Insuffisante

De nombreux chatbots ne prennent pas en charge plusieurs langues et ne s'adaptent pas aux nuances culturelles, limitant ainsi leur accessibilité.

- **Altibbi** : Prend en charge l'arabe mais a des difficultés avec les dialectes comme le Darija.
- **Ada Health** : Offre un support multilingue mais manque de localisation pour les dialectes régionaux.

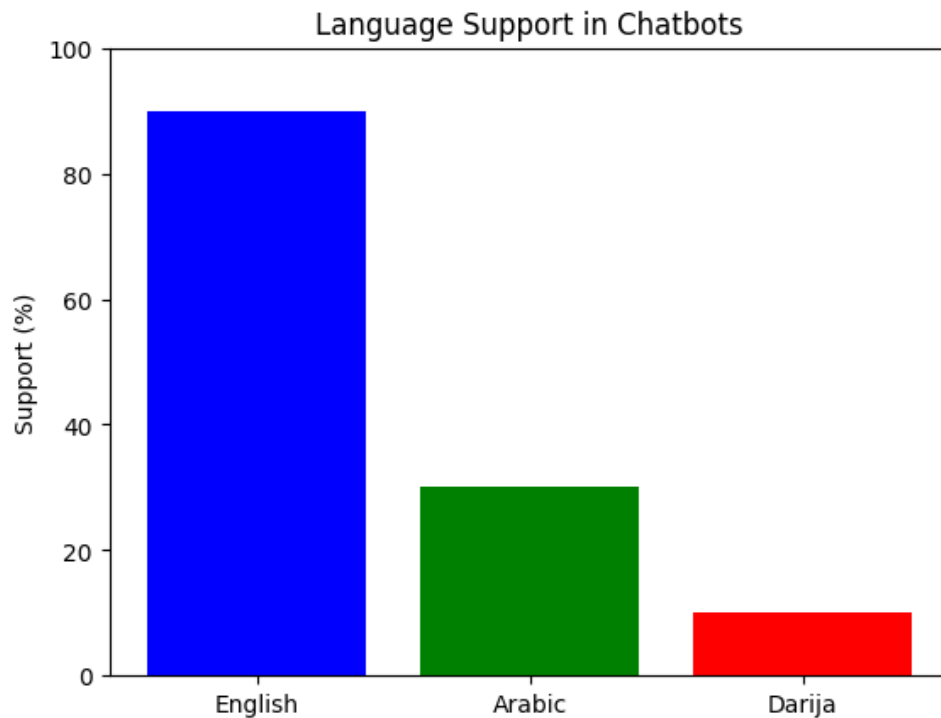


FIG. 2.3 : Support Linguistique dans les Chatbots

2.2.4 Problèmes de Confidentialité et de Sécurité des Données

Les chatbots collectent souvent des données de santé sensibles, ce qui soulève des préoccupations concernant la confidentialité des données et la conformité aux réglementations comme le RGPD.

- **K Health** : Conforme à la HIPAA mais fait face à des préoccupations des utilisateurs concernant les violations de données.
- **Florence Health** : Conforme au RGPD mais nécessite le partage de données avec des appareils connectés.

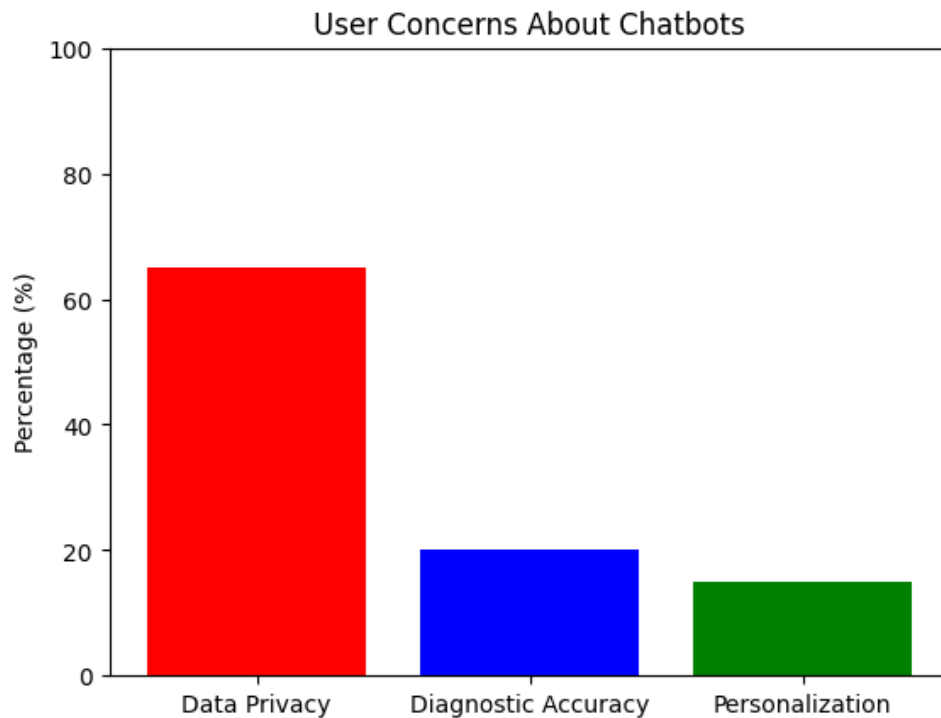


FIG. 2.4 : Préoccupations des Utilisateurs Concernant la Confidentialité des Données

2.2.5 Accessibilité Limitée pour les Utilisateurs Handicapés

De nombreux chatbots manquent de fonctionnalités d'accessibilité pour les utilisateurs handicapés, telles que la saisie vocale ou le support en langage des signes.

- **Hephaï** : Se concentre sur l'utilisation des inhalateurs mais manque de fonctionnalités d'accessibilité pour les utilisateurs malvoyants.
- **Buoy Health** : Ne prend pas en charge la saisie vocale pour les utilisateurs atteints de handicaps moteurs.

Accessibility Features in Chatbots

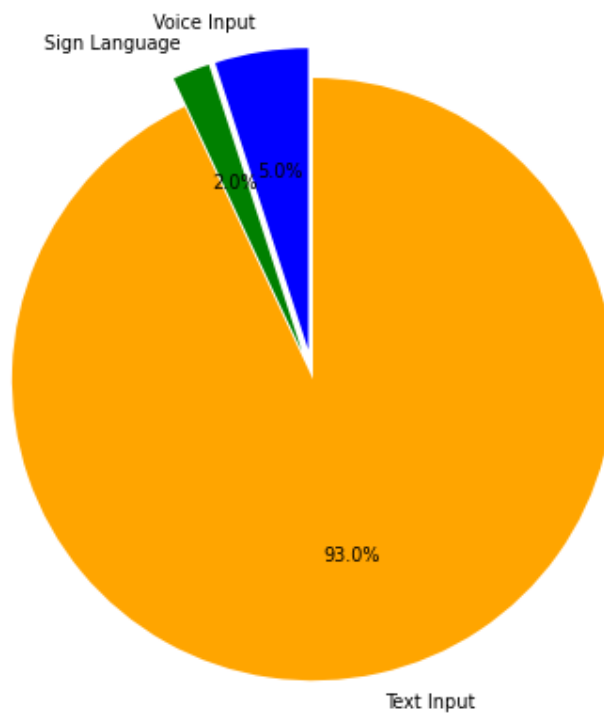


FIG. 2.5 : Fonctionnalités d'Accessibilité dans les Chatbots

2.2.6 Dépendance à l'Égard de la Supervision Professionnelle

Les chatbots ne peuvent pas remplacer complètement les professionnels de la santé, car ils ne sont pas capables de gérer des cas médicaux complexes.

- **Babylon Health** : Exige que les utilisateurs consultent des professionnels de la santé pour des diagnostics définitifs.
- **Mediktor** : Utilisé comme un outil de triage plutôt que comme une solution de diagnostic autonome.

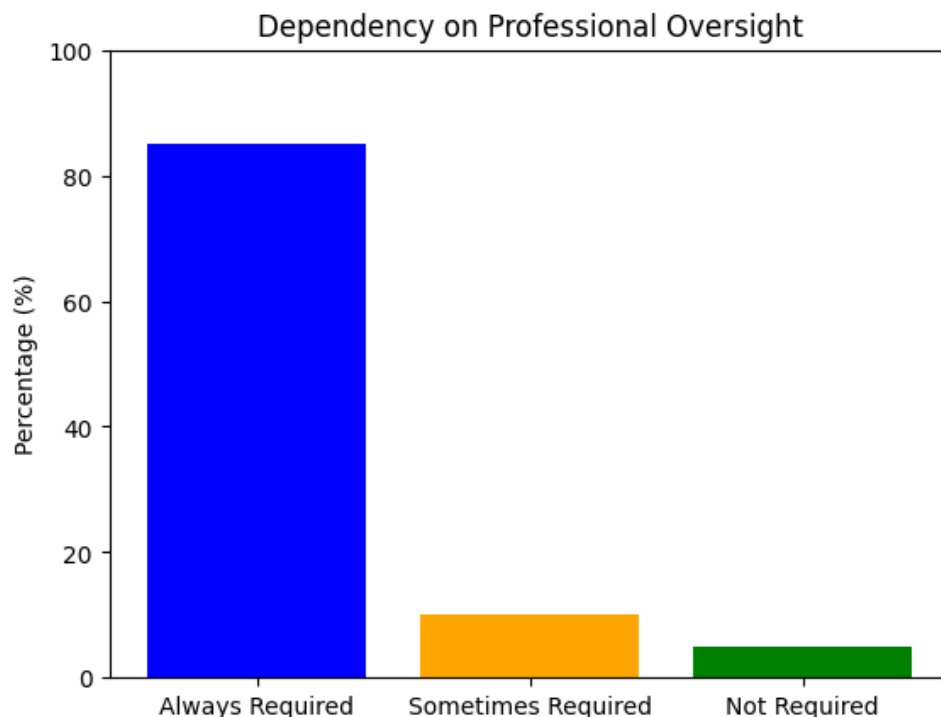


FIG. 2.6 : Dépendance à l'Égard de la Supervision Professionnelle

2.3 Analyse Comparative des Limites des Chatbots

2.3.1 Tableau Comparatif : Précision Diagnostique, Personnalisation, Support Multilingue

Chatbot	Précision Diagnostique	Personnalisation	Support Multilingue
Ada Health	Moyenne	Faible	Moyen
Altibbi	Moyenne	Moyenne	Élevé (Arabe)
Babylon Health	Faible	Moyenne	Moyen
Buoy Health	Élevée	Moyenne	Moyen
Hephaï	Élevée	Faible	Faible
Memoquest	Moyenne	Élevée	Faible
K Health	Élevée	Élevée	Moyen
Symptomate	Élevée	Faible	Élevée
Mediktor	Élevée	Moyenne	Moyen
Florence Health	Moyenne	Élevée	Faible

TAB. 2.1 : Comparaison des Chatbots par Précision Diagnostique, Personnalisation et Support Multilingue

2.3.2 Tableau Comparatif : Confidentialité des Données, Accessibilité, Supervision Professionnelle

Chatbot	Confidentialité des Données	Accessibilité	Supervision Professionnelle
Ada Health	Élevée	Faible	Élevée
Altibbi	Moyenne	Faible	Élevée
Babylon Health	Moyenne	Faible	Élevée
Buoy Health	Élevée	Faible	Élevée
Hephaï	Élevée	Faible	Moyenne
Memoquest	Moyenne	Faible	Moyenne
K Health	Élevée	Faible	Élevée
Symptomate	Moyenne	Faible	Élevée
Mediktor	Élevée	Faible	Élevée
Florence Health	Élevée	Faible	Moyenne

TAB. 2.2 : Comparaison des Chatbots par Confidentialité des Données, Accessibilité et Supervision Professionnelle

2.4 Recommandations pour l'Amélioration

- **Améliorer la Précision Diagnostique** : Améliorer les algorithmes d'IA avec des ensembles de données diversifiés et intégrer une supervision humaine pour les cas complexes.
- **Augmenter la Personnalisation** : Développer une IA adaptative qui suit les tendances de santé à long terme et propose des recommandations personnalisées.
- **Étendre le Support Multilingue** : Ajouter un support pour les dialectes régionaux et les nuances culturelles pour améliorer l'accessibilité.
- **Renforcer la Confidentialité des Données** : Assurer la conformité avec le RGPD, la HIPAA et d'autres réglementations pour renforcer la confiance des utilisateurs.
- **Améliorer l'Accessibilité** : Ajouter la saisie vocale, le support en langage des signes et d'autres fonctionnalités pour les utilisateurs handicapés.
- **Réduire la Dépendance à l'Égard des Professionnels** : Développer des chatbots capables de gérer des cas plus complexes de manière autonome.

2.5 Conclusion

Bien que les chatbots médicaux offrent des avantages significatifs, leurs limites—telles que l’inexactitude diagnostique, le manque de personnalisation et l’accessibilité insuffisante—entravent leur plein potentiel. En abordant ces limites, les développeurs peuvent créer des solutions de santé plus efficaces et inclusives.

Chapitre 3

Présentation des chatbots médicaux

3.1 Présentation de K Health



FIG. 3.1 : Logo de K Health

3.1.1 Contexte et Objectif

K Health a été fondé par des médecins et des experts en technologie dans le but d'améliorer l'accès aux soins primaires. La plateforme exploite des données provenant de millions de dossiers médicaux anonymisés pour offrir des services de santé numériques.

3.1.2 Fonctionnement et Technologies Utilisées

- **Intelligence Artificielle** : Basée sur des millions d'interactions médicales réelles.
- **Comparaison de cas** : Système de comparaison de cas similaires pour une meilleure précision.

- **Sécurité** : Plateforme conforme à la norme HIPAA.
- **Intégration** : API permettant l'intégration aux systèmes de santé existants.

3.1.3 Caractéristiques Clés

- Consultation médicale virtuelle 24h/24 et 7j/7.
- Analyse des symptômes et suggestion de diagnostics.
- Possibilité d'obtenir une prescription si nécessaire.
- Suivi des maladies chroniques et abonnements accessibles.

3.1.4 Impact et Utilisation Clinique

- Plus de 5 millions d'utilisateurs actifs.
- Réduction des coûts de santé et amélioration de l'accès aux soins.
- Taux de satisfaction utilisateur de 95%.

3.1.5 Avantages et Inconvénients

- + Prix compétitif.
- + Accès rapide aux professionnels de santé.
- + Base de données médicales étendue.
- Disponible dans un nombre limité de pays.
- Ne couvre pas toutes les pathologies.
- Nécessite un smartphone récent.

3.1.6 Lien

<https://www.khealth.ai/>

3.2 Présentation de Symptomate



FIG. 3.2 : Logo de Symptomate

3.2.1 Contexte et Objectif

Symptomate, conçu par Infermedica, est une plateforme de pré-diagnostic et d'orientation médicale destinée aux particuliers et aux professionnels de santé.

3.2.2 Fonctionnement et Technologies Utilisées

- **Intelligence Artificielle** : Basée sur une base de connaissances validée.
- **Moteur d'inférence probabiliste** : Pour analyser les symptômes.
- **Intégration** : API REST permettant une intégration facile.
- **Disponibilité** : Sur plusieurs plateformes.

3.2.3 Caractéristiques Clés

- Analyse précise des symptômes.
- Support de plus de 20 langues.
- Génération de rapports médicaux automatisée.

3.2.4 Impact et Utilisation Clinique

- Présent dans plus de 30 pays.
- Intégré à divers systèmes de santé nationaux.
- Optimisation du triage hospitalier et réduction des délais d'attente.

3.2.5 Avantages et Inconvénients

- + Large couverture linguistique.
- + Interface adaptée aux professionnels.
- + Précision diagnostique élevée.
- Interface moins intuitive que certains concurrents.
- Peu de suivi à long terme.
- Certaines fonctionnalités réservées aux professionnels.

3.2.6 Lien

<https://symptomate.com>

—

3.3 Présentation de Mediktor



FIG. 3.3 : Logo de Mediktor

3.3.1 Contexte et Objectif

Mediktor, développé en Espagne, vise à améliorer le triage médical et est utilisé dans les établissements hospitaliers pour fluidifier les parcours de soins.

3.3.2 Fonctionnement et Technologies Utilisées

- **Intelligence Artificielle** : Entraînée sur plus de 2 millions de cas cliniques.
- **Algorithmes** : Deep Learning et reconnaissance du langage naturel.
- **Validation clinique** : Effectuée par des hôpitaux universitaires.

3.3.3 Caractéristiques Clés

- Analyse des symptômes en moins de 10 minutes.
- Plus de 750 pathologies couvertes.
- Intégration aux systèmes hospitaliers et module de téléconsultation.
- Triage automatisé selon le niveau d'urgence.

3.3.4 Impact et Utilisation Clinique

- Précision diagnostique de 91,3%.
- Utilisé dans plus de 180 centres médicaux.
- Réduction de 50% du temps de triage.

3.3.5 Avantages et Inconvénients

- + Validation clinique rigoureuse.
- + Intégration efficace aux systèmes hospitaliers.
- + Interface adaptée aux professionnels.
- + Support multilingue avancé.

- Coût élevé pour les établissements.
- Implémentation complexe.
- Formation du personnel requise.

3.3.6 Lien

<https://www.mediktor.com>

3.4 Présentation de Florence Health



FIG. 3.4 : Logo de Florence Health

3.4.1 Contexte et Objectif

Florence Health est une start-up française spécialisée dans le suivi des maladies chroniques, offrant un accompagnement personnalisé.

3.4.2 Fonctionnement et Technologies Utilisées

- **IA conversationnelle** : Émotionnelle et adaptative.
- **Apprentissage automatique** : Pour ajuster les recommandations.
- **Intégration IoT** : Pour dispositifs médicaux connectés.
- **Architecture cloud** : Conforme au RGPD.

3.4.3 Caractéristiques Clés

- Coaching santé quotidien.
- Connexion avec des objets connectés.
- Alertes pour les professionnels de santé.
- Rapports de suivi automatisés.

3.4.4 Impact et Utilisation Clinique

- Amélioration de 40% de l'observance thérapeutique.
- Réduction des complications des maladies chroniques.
- Satisfaction patient de 4,8/5.
- Utilisé par plusieurs mutuelles françaises.

3.4.5 Avantages et Inconvénients

- + Suivi personnalisé.
- + Excellente gestion des maladies chroniques.
- + Conforme au RGPD.
- + Interface en français.
- Disponibilité géographique limitée.
- Nécessite des objets connectés spécifiques.
- Certaines fonctionnalités payantes.

3.4.6 Lien

<https://florence.chat/>

Chapitre 4

Analyse des Chatbots pour les Maladies Respiratoires

4.1 Introduction

Ce document présente une analyse de trois chatbots dédiés à l'assistance des patients souffrant de maladies respiratoires, notamment l'asthme et la BPCO. Chaque section est consacrée à un chatbot spécifique, détaillant son contexte, ses fonctionnalités, ses caractéristiques clés et son impact clinique.

4.2 Chatbot 1 : Hephai

4.2.1 Contexte et Origine

Hephai est la première solution d'intelligence artificielle développée pour aider les patients atteints d'asthme ou de BPCO à utiliser correctement leurs dispositifs d'inhalation. Cette application a été créée grâce à une collaboration entre le laboratoire Chiesi et la startup Hephai. L'idée a été initiée par le pneumologue Valéry Trosini-Désert et a été co-développée avec un panel d'experts et de professionnels de santé impliqués dans l'éducation thérapeutique.

4.2.2 Fonctionnalités et Technologies Utilisées

- **Enregistrement Vidéo** : Le patient enregistre son utilisation de l'inhalateur à l'aide d'un smartphone, d'une tablette ou d'un ordinateur.
- **Analyse par Intelligence Artificielle** : Des algorithmes de reconnaissance de mouvements et d'analyse sonore évaluent en temps réel la manière dont l'inhalateur est utilisé.
- **Retour Immédiat** : Si l'utilisation est correcte, le coach virtuel félicite le patient. En cas d'erreur, l'application identifie précisément les points à améliorer et fournit un lien vers une vidéo éducative expliquant la technique correcte.

4.2.3 Caractéristiques Clés

- **Coach Virtuel** : Un assistant numérique interactif qui guide le patient.
- **Analyse en Temps Réel** : Utilisation de technologies de reconnaissance visuelle et sonore.
- **Support Éducatif** : En cas d'erreur, le patient reçoit des ressources éducatives.
- **Conformité Réglementaire** : Dispositif médical de classe 1.

4.2.4 Impact et Utilisation Clinique

L'utilisation d'Hephaï permet de réduire les erreurs d'utilisation, améliorant ainsi l'efficacité de l'administration des médicaments. L'outil peut être utilisé lors des consultations, recommandé à la fin d'une consultation, ou même déployé par les pharmaciens.

4.3 Chatbot 2 : Memoquest

4.3.1 Contexte et Objectif

Memoquest est un chatbot développé par Calmedica pour améliorer le suivi des patients atteints de maladies chroniques, notamment l’asthme et la BPCO. Il vise à automatiser la communication entre les patients et les professionnels de santé, à améliorer l’adhésion thérapeutique et à faciliter le suivi médical tout en réduisant la charge administrative.

4.3.2 Fonctionnalités et Technologies Utilisées

4.3.2.1 Interaction et Communication Automatisées

- **Rappels Personnalisés** : Memoquest envoie des notifications aux patients pour leur rappeler la prise de médicaments et les rendez-vous médicaux.
- **Questions de Suivi** : Il pose régulièrement des questions sur l’état de santé du patient (symptômes, effets secondaires, etc.).
- **Traitement du Langage Naturel (NLP)** : L’intelligence artificielle analyse les réponses et adapte la conversation en conséquence.

4.3.2.2 Analyse en Temps Réel et Alertes

- **Collecte de Données** : Les réponses des patients sont enregistrées et analysées.
- **Détection des Risques** : En cas d’anomalies (par exemple, symptômes aggravés, non-respect du traitement), le système génère une alerte pour le médecin traitant ou l’équipe soignante.

4.3.2.3 Suivi Personnalisé

- Memoquest s’adapte au profil médical et aux besoins spécifiques de chaque patient.
- Accessible via smartphones, tablettes et ordinateurs pour une utilisation facile.

4.3.3 Caractéristiques Clés

- **Suivi Médical Automatisé** : Réduit le travail administratif et améliore la gestion des patients.
- **Amélioration de l'Adhésion Thérapeutique** : Réduit les oublis et améliore la gestion du traitement.
- **Interface Simple et Accessible** : Utilisable par tous, sans compétences techniques.
- **Analyse des Données en Temps Réel** : Permet une intervention rapide en cas de problème.

4.3.4 Impact et Avantages

4.3.4.1 Pour les Patients

- Meilleur suivi de la maladie.
- Engagement actif dans leur traitement.
- Accès facile à des conseils médicaux.

4.3.4.2 Pour les Professionnels de Santé

- Suivi optimisé avec des alertes automatiques.
- Réduction de la charge de travail liée au suivi des patients.
- Prise de décision plus rapide grâce à l'analyse des données.

—

4.4 Liens Officiels des Chatbots

- **Hephaï** : Site web et applications disponibles sur <https://bonusagedumedicamer.com/hephaï-aider-les-patients-a-bien-utiliser-leurs-therapeutiques-> Play et [Apple Store](#).

- **Memoquest** : Informations et démonstrations disponibles sur le site web de Calmedica : <https://www.calmedica.com> (à adapter selon le lien exact fourni par l'éditeur).

Chapitre 5

Modèles d'IA et Techniques NLP pour les Chatbots Médicaux

[enhanced, colback=yellow!10!white, colframe=red!50!black, arc=5pt, boxrule=1pt, drop fuzzy shadow] « *Les chatbots sont importants car vous ne vous sentirez pas stupide en posant des questions importantes. Parfois, parler à quelqu'un peut être intimidant. Parler à un chatbot rend cela beaucoup plus facile !* »

— Petter Bae Brandtzaeg, chef de projet, projet Social Health Bots

5.1 Introduction

Lorsqu'un utilisateur envoie un message à un chatbot, le système exécute un processus sophistiqué en plusieurs étapes :

- **Compréhension du Langage Naturel** : Analyse du texte pour extraire le sens, l'intention et le contexte.
- **Génération de Réponses** : Utilisation de modèles d'IA pour sélectionner ou générer la réponse appropriée.
- **Intégration des Connaissances** : Accès à des informations externes ou à des bases de données médicales via des API.

Lorsqu'un utilisateur envoie un message à un chatbot, il reçoit une réponse presque instantanément. Cependant, derrière cette interaction apparemment simple, un processus sophistiqué se déroule, où plusieurs technologies travaillent en synergie. À la réception du message, le chatbot analyse le texte à l'aide de techniques avancées de traitement du langage naturel

(NLP). Ensuite, en utilisant des modèles d'intelligence artificielle (IA), il sélectionne ou génère la réponse la plus appropriée. Enfin, des intégrations via des API permettent d'accéder à des informations externes ou à des bases de données médicales pour fournir des réponses précises et à jour. Comprendre ces aspects techniques est essentiel pour démêler le fonctionnement interne du chatbot et concevoir un système capable de répondre efficacement aux besoins des utilisateurs.

5.2 Modèles d'IA

Plusieurs modèles d'IA sont utilisés dans le développement des chatbots médicaux, principalement BERT et GPT.

5.2.1 BERT (Bidirectional Encoder Representations from Transformers)

Composants Architecturaux :

- Traitement contextuel bidirectionnel.
- Pile d'encodeurs Transformer (plusieurs couches).
- Mécanisme d'auto-attention multi-têtes.
- Objectifs de pré-entraînement : Masquage de mots (Masked LM) et prédiction de phrase suivante (Next Sentence Prediction).

BERT est un modèle d'apprentissage profond qui traite le texte de manière bidirectionnelle, en tenant compte des tokens précédents et suivants pour mieux comprendre les requêtes médicales complexes.

Après la tokenisation, une requête utilisateur est représentée comme :

$$X = [x_1, x_2, \dots, x_n]$$

où x_i représente le i^{me} token.

Le mécanisme d'attention dans BERT est défini comme :

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V$$

avec Q (Requête), K (Clé) et V (Valeur) dérivés des tokens, et d_k étant la dimension des vecteurs clés.

L'attention multi-têtes est calculée via :

$$MH(Q, K, V) = \text{Concaténation}(h_1, h_2, \dots, h_h) \cdot W_O$$

où $h_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ est la sortie de la i^{me} tête, et W_O est la matrice de projection de sortie.

BERT produit un embedding contextuel H pour une entrée X :

$$H = \text{BERT}(X)$$

Ces embeddings sont essentiels pour comprendre les requêtes médicales.

Une couche d'encodeur Transformer est formalisée comme :

$$\text{EncoderLayer}(X) = \text{LayerNorm}(X + \text{MultiHead}(X, X, X))$$

Pour les sorties spécifiques à une tâche (par exemple, la reconnaissance d'entités ou la classification d'intention), H est mappé à :

$$\text{Entités} = \text{OutputLayer}(H)$$

et

$$\text{Intention} = \text{softmax}(\text{pool}(H))$$

Le processus de fine-tuning minimise la perte :

$$\text{Loss} = - \sum_i Y_i \cdot \log(\hat{Y}_i)$$

en résolvant :

$$\min_{\theta_{\text{BERT}}} \text{Loss}(X, Y; \theta_{\text{BERT}})$$

en utilisant l'optimiseur Adam.

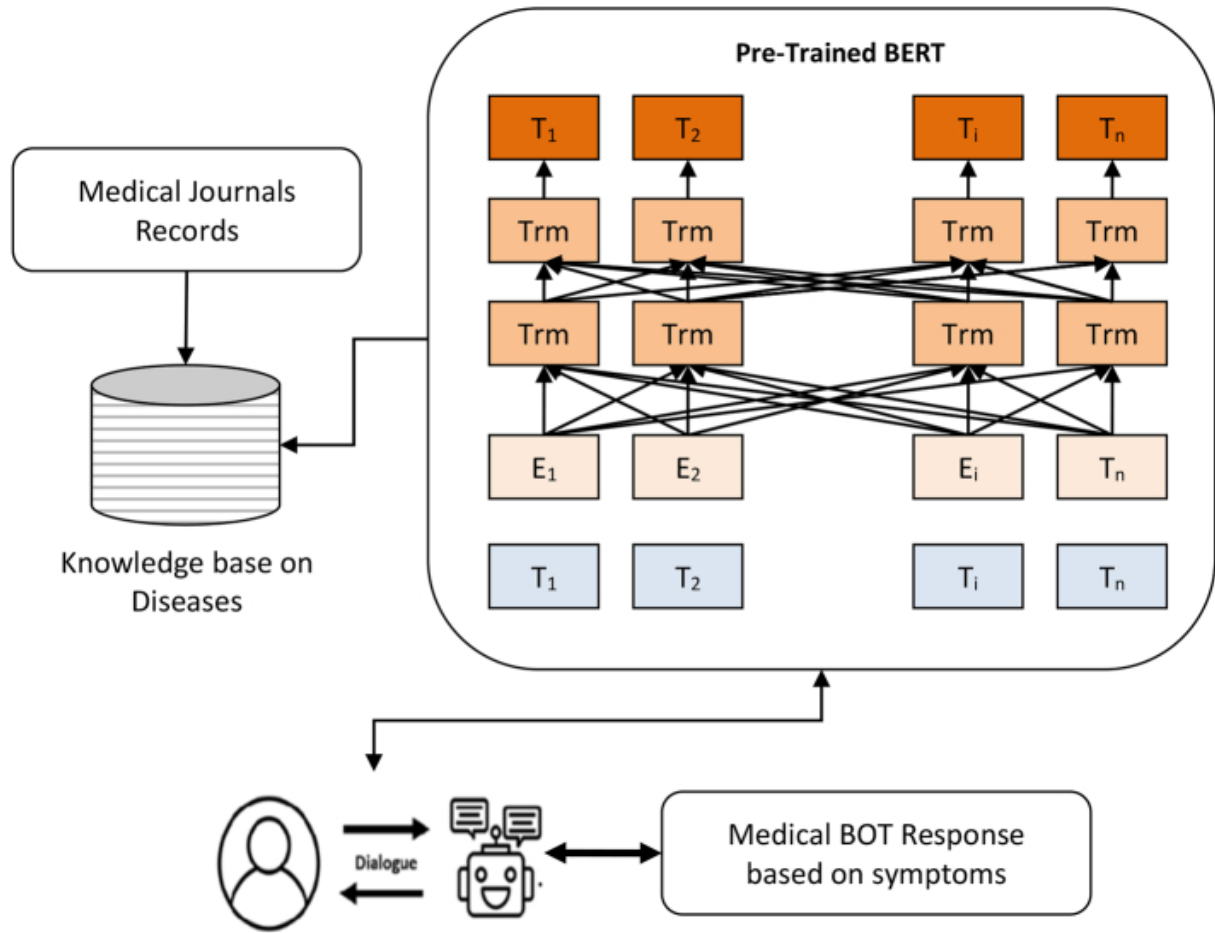


FIG. 5.1 : Exemple d'architecture BERT

5.2.2 GPT (Generative Pretrained Transformer)

Différences Clés par Rapport à BERT :

- Traitement contextuel unidirectionnel.
- Génération de texte autoregressive.
- Masquage d'attention causal.
- Nombre élevé de paramètres (par exemple, 175 milliards dans GPT-3).

Contrairement à BERT, GPT traite le texte de manière unidirectionnelle (de gauche à droite) et prédit un mot à la fois :

$$P(w_t \mid w_1, w_2, \dots, w_{t-1}) = \text{softmax}(W \cdot h_t)$$

où w_t est le mot à prédire à la position t et h_t est l'état caché.

Le mécanisme d'attention dans GPT, avec masquage causal, est donné par :

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \cdot V$$

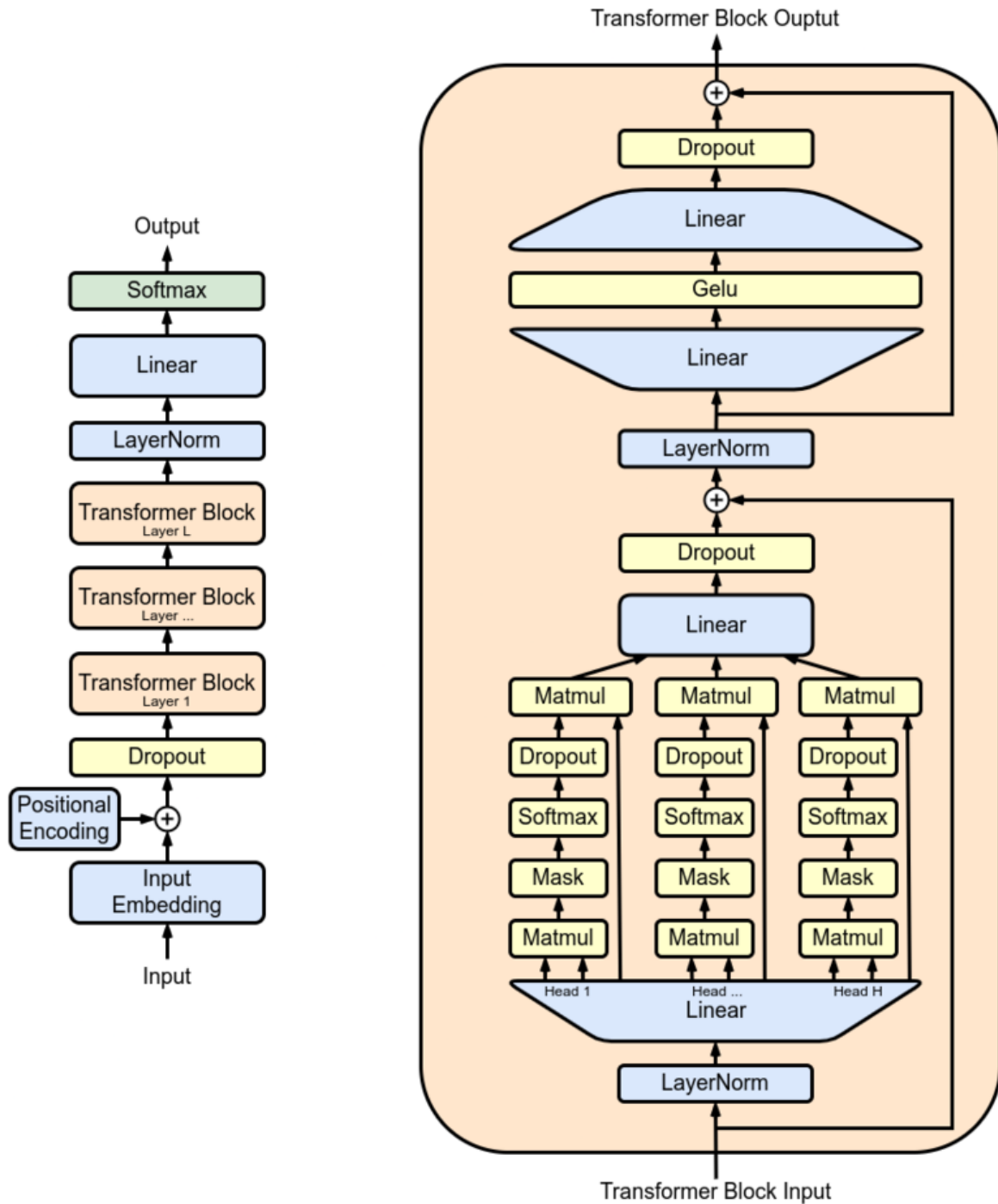


FIG. 5.2 : Exemple d'architecture GPT

GPT utilise un décodage autoregressif, générant du texte en fonction des

mots précédemment prédits. Pour l'adaptation à des tâches spécifiques, il est optimisé avec la perte :

$$\text{Loss} = - \sum Y_i \cdot \log(\hat{Y}_i)$$

—

5.3 Modèles pour Tâches Spécifiques dans les Domaines Médicaux

Pour les applications médicales spécialisées, des modèles pré-entraînés spécifiques au domaine sont utilisés pour mieux gérer le vocabulaire technique et le contexte.

5.3.1 PaLM (Pathways Language Model)

PaLM est un modèle développé par Google AI avec jusqu'à 540 milliards de paramètres. Il effectue diverses tâches telles que le raisonnement logique, l'arithmétique, l'explication de blagues, la génération de code et la traduction en utilisant ses capacités de chaîne de pensée. Lancé en avril 2022 et rendu disponible via une API, une version spécialisée, Med-PaLM, a été entraînée sur des données médicales—dépassant les systèmes précédents et réussissant même les examens médicaux américains.

Google a ensuite étendu PaLM avec PaLM-E pour les tâches multimodales et, en mai 2023, a lancé PaLM 2 avec 340 milliards de paramètres. AudioPaLM, introduit en juin 2023, permet la traduction automatique du langage parlé.

La fonction de perte pour Med-PaLM est :

$$\mathcal{L}_{\text{med}} = - \sum_{i=1}^N \log P(y_i \mid x, y_{<i})$$

5.3.2 ChatDoctor

Améliorations par Rapport aux Modèles de Base :

- Fine-tuning sur 100 000 dialogues médicaux.

- Basé sur l'architecture LLaMA de Meta AI.
- Mécanisme de récupération d'informations en temps réel.
- Intégration d'un système de vérification des symptômes.

ChatDoctor est un chatbot médical avancé conçu pour fournir des conseils médicaux précis et personnalisés. Il a été fine-tuné sur un grand ensemble de données de 100 000 dialogues entre patients et médecins provenant d'une plateforme de consultation en ligne, améliorant ainsi sa capacité à comprendre les besoins des patients et à fournir des conseils éclairés. De plus, il intègre un mécanisme de récupération d'informations autonome pour accéder en temps réel à des sources fiables (par exemple, Wikipédia et des bases de données médicales).

—

5.4 Techniques NLP dans les Chatbots Médicaux

Le traitement du langage naturel (NLP) transforme le langage humain—souvent imprécis et varié—en données structurées pouvant être utilisées par des modèles d'intelligence artificielle. Dans un contexte médical, la précision et la capacité à extraire des informations spécifiques (telles que les symptômes, les médicaments ou les diagnostics) sont essentielles pour fournir des réponses pertinentes et fiables. Des outils comme l'API Google Cloud Natural Language et des modèles de représentation tels que Word2Vec ou BERT (développés par Google) illustrent cette approche technologique.

Nous détaillons maintenant les principales techniques NLP utilisées dans les chatbots médicaux, en illustrant comment elles permettent la compréhension et le traitement du langage humain pour améliorer l'expérience et la sécurité des patients.

5.4.1 Reconnaissance d'Entités Nommées Médicales (NER)

La NER consiste à identifier des termes spécifiques, tels que les maladies, les médicaments et les symptômes, dans un texte. Des techniques telles que CRF (Champs Aléatoires Conditionnels), BiLSTM-CRF et des approches basées sur BERT ont été largement appliquées pour la NER médicale. En

utilisant des modèles basés sur l'apprentissage profond comme BioBERT, la NER peut être étendue à des domaines plus spécifiques tels que les ensembles de données médicaux, permettant au chatbot de comprendre une terminologie médicale complexe.

La tâche est formalisée comme :

$$\text{NER}(X) = \{(x_i, t_i)\}$$

où x_i est le token et t_i est la classe d'entité (par exemple, maladie, médicament).

5.4.2 Réponse aux Questions Médicales (QA)

Les systèmes de réponse aux questions médicales récupèrent des réponses à partir d'un corpus prédéfini de connaissances médicales. Les avancées récentes impliquent des modèles basés sur les transformers, tels que BERT et T5 (Text-to-Text Transfer Transformer), qui exploitent des représentations pré-entraînées pour le fine-tuning sur des tâches de QA.

Pour la QA, le modèle apprend à répondre aux questions à partir d'un contexte C en calculant la probabilité :

$$P(\text{Réponse} \mid C, Q)$$

où Q est la requête de l'utilisateur et C est le corpus de connaissances médicales.

Chapitre 6

Évaluation des Chatbots Médicaux

[enhanced, colback=yellow!10!white, colframe=red!50!black, arc=5pt, boxrule=1pt, drop fuzzy shadow] « *Les chatbots sont importants car vous ne vous sentirez pas stupide en posant des questions importantes. Parfois, parler à quelqu'un peut être intimidant. Parler à un chatbot rend cela beaucoup plus facile !* »

— Petter Bae Brandtzaeg, chef de projet, projet Social Health Bots

6.1 Objectifs de l'Évaluation

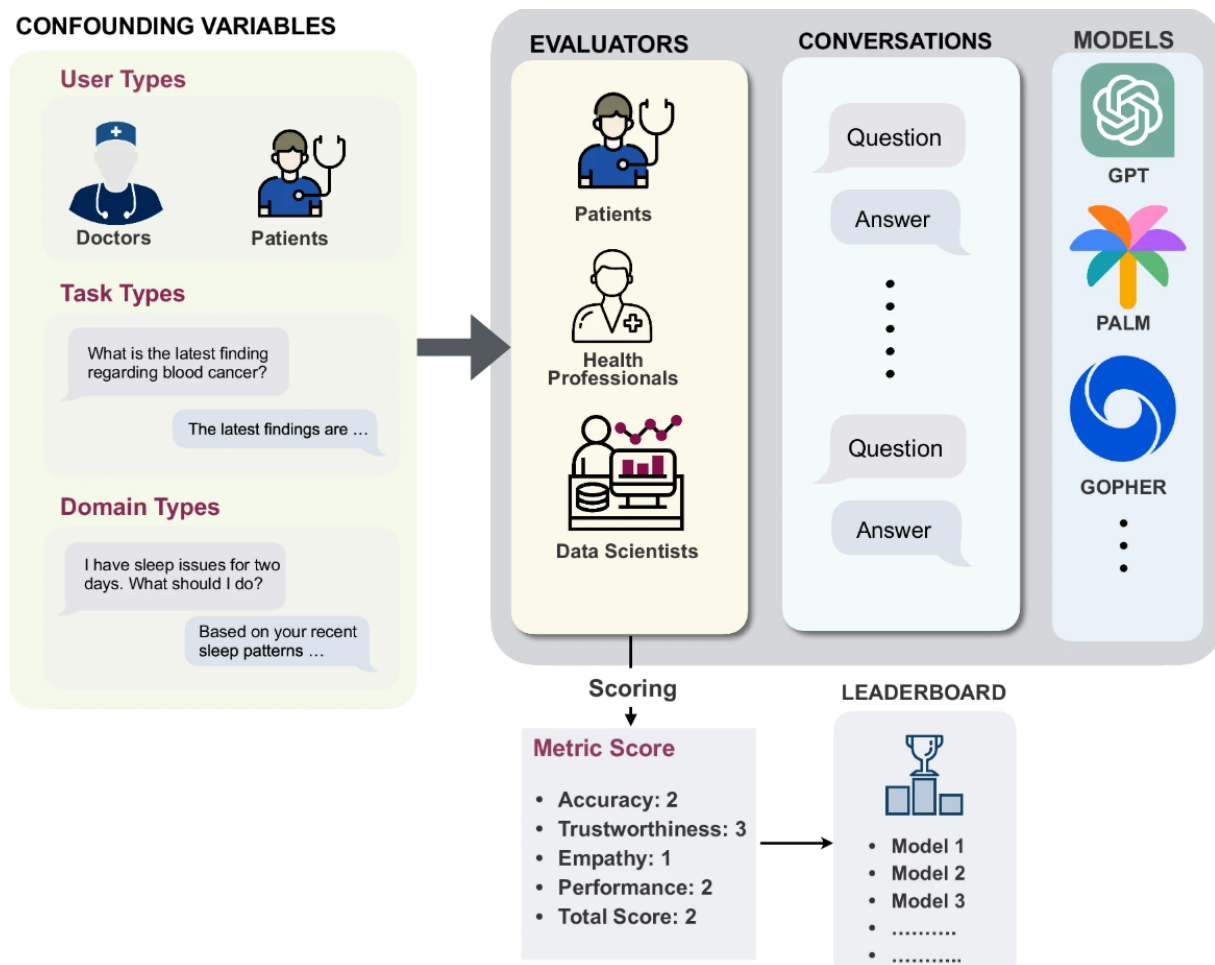


FIG. 6.1 : Exemple d'illustration des objectifs d'évaluation

L'objectif principal est d'évaluer les chatbots médicaux en fonction de l'interaction utilisateur, en tenant compte des variables confusionnelles suivantes :

6.1.1 Type d'Utilisateur

Les utilisateurs peuvent être :

- **Patient** : Un individu cherchant des conseils de santé.
- **Médecin** : Un professionnel de santé recherchant des informations détaillées telles que des diagnostics, des traitements et des prescriptions.

- **Infirmier** : Un professionnel de santé impliqué dans des questions pratiques, y compris les traitements à administrer et les soins quotidiens.

L'évaluation doit s'adapter à ces types d'utilisateurs en fonction des exigences de sécurité et de confidentialité des données.

6.1.2 Type de Domaine

- **Général** : Un chatbot répondant à une gamme de questions de santé (symptômes, conseils de prévention, etc.).
- **Spécifique** : Un chatbot dédié à un domaine particulier, comme la **santé mentale**, les **maladies chroniques** ou le **cancer**. Ces chatbots doivent comprendre et répondre avec des informations spécifiques et détaillées au domaine.

6.1.3 Type de Tâche

Les tâches que le chatbot peut accomplir varient, et les métriques devront être ajustées en fonction de la tâche :

- **Génération de Rapports Médicaux** : Le chatbot génère un rapport médical détaillé basé sur les symptômes et les informations fournies par l'utilisateur.
- **Diagnostic Médical** : Le chatbot fournit un diagnostic basé sur les informations de l'utilisateur.
- **Assistance Générale** : Le chatbot aide à des tâches quotidiennes (gestion des médicaments, prise de rendez-vous, etc.).

6.2 Métriques Essentielles

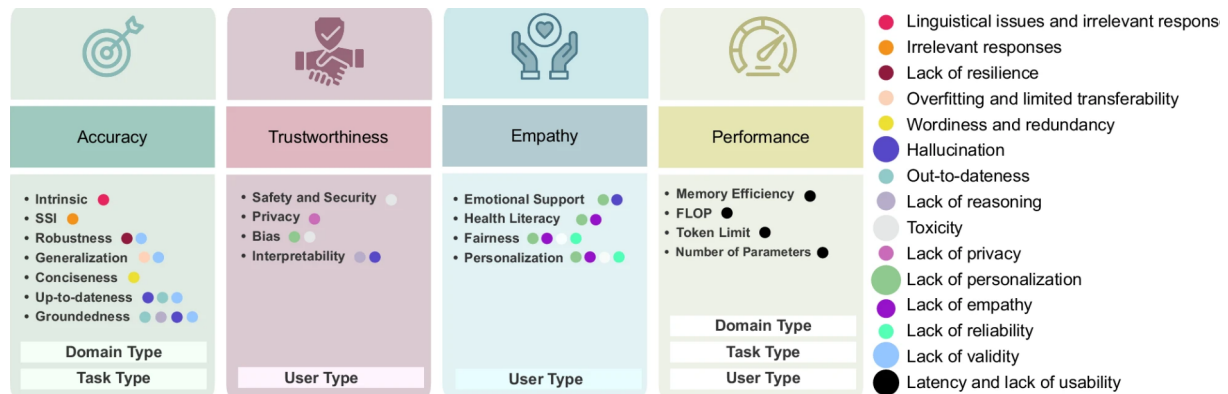


FIG. 6.2 : Exemple d'illustration des métriques d'évaluation

Les métriques d'évaluation sont regroupées en quatre catégories principales : **Précision**, **Fiabilité**, **Empathie** et **Performance**.

6.2.1 Précision

Objectif : Évaluer la capacité du chatbot à fournir des réponses précises et pertinentes.

- **Métriques à évaluer :**

- **Taux de précision des réponses :** Mesurer le pourcentage de réponses correctes par rapport au total des réponses. Cela peut inclure des réponses basées sur des diagnostics, des traitements ou des conseils de santé.
- **Réponses incorrectes :** Analyser le taux d'erreur (réponses incorrectes ou incohérentes) et leur impact sur la fiabilité du chatbot.
- **Validation des informations médicales :** Vérifier la validité des informations fournies (par exemple, prescription de médicaments) selon les dernières recherches médicales et les directives de santé.

Exemple de test : Poser au chatbot des questions sur des symptômes courants (par exemple, « Quels sont les symptômes de la grippe ? ») et vérifier la pertinence et l'exactitude des réponses fournies.

6.2.2 Fiabilité

Objectif : Mesurer la crédibilité du chatbot auprès des utilisateurs en termes de sécurité, de confidentialité et de validité des informations fournies.

- **Métriques à évaluer** :
 - **Confiance perçue par l'utilisateur** : Ce critère mesure à quel point l'utilisateur fait confiance au chatbot, en particulier pour des questions sensibles (médicaments, diagnostics).
 - **Protection des données** : S'assurer que le chatbot respecte les normes de sécurité et protège les données personnelles et médicales des utilisateurs.
 - **Cohérence des réponses** : Vérifier que le chatbot fournit des réponses cohérentes lorsque la même question est posée plusieurs fois.

Exemple de test : Tester la capacité du chatbot à gérer une question sensible (par exemple, « Mes informations sont-elles sécurisées ? ») et fournir une réponse rassurante tout en maintenant la confidentialité.

6.2.3 Empathie

Objectif : Évaluer la capacité du chatbot à comprendre et à répondre de manière appropriée aux émotions des utilisateurs de manière humaine.

- **Métriques à évaluer** :
 - **Réponses empathiques** : La capacité du chatbot à répondre émotionnellement aux préoccupations de l'utilisateur, en particulier dans des contextes sensibles comme la santé mentale ou les diagnostics graves.
 - **Tonalité et humanisation** : Le chatbot utilise-t-il un ton approprié pour rassurer les utilisateurs stressés ou inquiets ?
 - **Reconnaissance des émotions** : Le chatbot peut-il détecter les émotions dans les questions de l'utilisateur (par exemple, anxiété, colère) et adapter ses réponses en conséquence ?

Exemple de test : Poser des questions où l'utilisateur peut être inquiet ou stressé (par exemple, « J'ai des douleurs thoraciques, est-ce grave ? ») et évaluer si le chatbot fournit une réponse rassurante et empathique.

6.2.4 Performance

Objectif : Mesurer la réactivité et l'efficacité du chatbot à accomplir des tâches.

- **Métriques à évaluer :**
 - **Temps de réponse :** Combien de temps le chatbot met-il pour répondre à une question ? Un chatbot performant doit répondre rapidement tout en fournissant des informations précises.
 - **Taux de résolution des demandes :** Combien de questions ou de demandes sont effectivement résolues par le chatbot ? Cela inclut la fourniture de réponses utiles et l'orientation de l'utilisateur vers des ressources supplémentaires si nécessaire.
 - **Stabilité technique :** Le chatbot fonctionne-t-il correctement sans erreurs ? Cela inclut l'absence de bugs, de déconnexions ou d'autres problèmes techniques pendant les interactions.

Exemple de test : Poser une série de questions complexes et mesurer le temps que met le chatbot pour donner une réponse complète et précise tout en s'assurant que le système ne rencontre pas de bugs.

—

6.3 Variables Confusionnelles

Lors de l'évaluation du chatbot, il est crucial de prendre en compte les **variables confusionnelles** suivantes, car elles peuvent influencer les résultats des tests :

- **Type d'Utilisateur :** Les besoins et les attentes varient entre un patient, un médecin ou un autre professionnel de santé. Par exemple, un patient attend des réponses plus simples et rassurantes, tandis qu'un médecin recherche des informations précises et techniques.

- **Type de Domaine** : Un chatbot conçu pour des **problèmes de santé généraux** sera principalement évalué sur sa capacité à fournir des réponses simples et fiables. Un chatbot conçu pour des domaines spécifiques comme la **santé mentale** ou le **cancer** doit être testé sur sa capacité à comprendre des contextes plus émotionnels et à fournir des informations pertinentes et spécialisées.
- **Type de Tâche** : La tâche que le chatbot doit accomplir influence l'évaluation. Par exemple, un chatbot générant des **rapports médicaux** doit être évalué sur la précision des données, tandis qu'un chatbot d'**assistance générale** doit être évalué pour sa rapidité, sa capacité à comprendre diverses demandes et à fournir des solutions utiles.

6.4 Plan d'Exécution des Tests

6.4.1 Sélection des Chatbots à Tester

Pour chaque domaine médical, sélectionnez 3 à 5 chatbots axés sur la gestion des maladies chroniques. Par exemple, pour le diabète, testez 3 chatbots différents avec des caractéristiques distinctes.

6.4.2 Exécution des Scénarios

Pour chaque chatbot, exécutez les scénarios décrits ci-dessus, chaque membre de l'équipe étant responsable de tester des scénarios spécifiques pour chaque métrique (précision, UX, performance, fiabilité).

6.4.3 Collecte des Résultats

Pendant les tests, chaque membre de l'équipe doit systématiquement collecter les résultats, en se concentrant sur :

- Temps de réponse (pour les tests de performance)
- Erreurs détectées (pour les tests de gestion des erreurs)

- Qualité des réponses (pour les tests de précision)
- Retours des utilisateurs (pour les tests d'UX et d'engagement)

6.4.4 Compilation des Résultats

Après avoir terminé les tests, compilez les résultats et attribuez des scores à chaque chatbot pour chaque métrique.

Chapitre 7

Evaluation of User Experience (UX) and Engagement in Chatbot-Based Symptom Checker (CSC) Applications

7.1 Introduction

Consumer Symptom Checker (CSC) apps are pivotal in modern digital healthcare, aiding users in symptom diagnosis, offering preliminary medical guidance, and directing users to healthcare professionals when necessary. However, user experience (UX) and engagement issues significantly impact their effectiveness. This study evaluates the UX and engagement of CSC apps, identifying key limitations and areas for improvement. By analyzing different aspects of UX and engagement, we aim to provide insights into how these applications can be optimized for better user satisfaction and trust.

7.2 User Experience (UX) Analysis

7.2.1 Ease of Use

Ease of use is fundamental to UX, as users expect seamless navigation when interacting with CSC apps. A well-designed user interface allows for efficient symptom input and reduces frustration. Some apps provide

structured layouts that guide users through the process, but others require multiple clicks and complex steps, leading to a frustrating experience. Guided input forms enhance usability by simplifying symptom entry, but some apps struggle with handling complex symptom descriptions.

TAB. 7.1 : Ease of Use Analysis

UX Factor	Strengths	Weaknesses
Navigation	Most CSC apps have structured layouts	Some require multiple clicks to reach results
Symptom Input	Provides guided input forms	Limited ability to handle complex descriptions

7.2.2 Comprehensibility

Comprehensibility is critical to ensuring that users can accurately describe their symptoms and understand the medical information provided. Many apps rely on medical jargon, which can be a significant barrier for non-expert users. While some apps simplify terminology and provide explanations for medical terms, others fail to offer sufficient context, leading to confusion and misinterpretation.

TAB. 7.2 : Comprehensibility Analysis

UX Factor	Strengths	Weaknesses
Language	Some apps use simple terminology	Many use complex medical terms that confuse users
Explanations	Some apps provide descriptions of terms	Limited contextual explanations for medical terms

7.2.3 Input Flexibility

A major limitation in CSC apps is their inability to handle varied symptom descriptions. Many applications require users to select from predefined symptom lists, which can be restrictive. While AI-powered apps attempt to recognize synonyms and alternative descriptions, they often struggle with non-standard terms, slang, or complex descriptions, leading to user frustration.

TAB. 7.3 : Input Flexibility Analysis

UX Factor	Strengths	Weaknesses
Symptom Input	Some apps allow free text input	Many apps require rigid, predefined terms
AI Recognition	AI-based apps attempt to match synonyms	Struggles with non-standard terms or slang

7.2.4 Conversational Design

Many CSC apps utilize chatbot interfaces to engage users and guide them through symptom assessments. While chatbots provide a structured way to collect user input, they often lack natural conversational flow and empathy. Users find interactions robotic, and the reasoning behind certain diagnostic questions is not always clear.

TAB. 7.4 : Conversational Design Analysis

UX Factor	Strengths	Weaknesses
Chatbot Responsiveness	Some apps respond quickly	Many responses feel robotic and lack empathy
Flow of Questions	Structured conversation format	Poor explanation of why certain questions are asked

7.2.5 Inclusivity & Accessibility

Ensuring inclusivity is vital for reaching diverse demographics, yet many CSC apps lack multi-language support and accessibility features. Users who speak non-English languages or have disabilities often struggle with these apps. While some apps have basic accessibility features, such as text resizing or voice input, there is still significant room for improvement.

TAB. 7.5 : Inclusivity & Accessibility Analysis

UX Factor	Strengths	Weaknesses
Language Support	Some apps offer multiple languages	Many lack support for non-English users
Accessibility	Basic accessibility options available	Limited features for visually impaired users

7.3 Engagement Analysis

7.3.1 User Retention & Interaction

User engagement plays a crucial role in retaining users. Features such as reminders, progress tracking, and personalized recommendations can significantly enhance user retention. While some apps implement these strategies effectively, others lack the necessary features to keep users engaged over time.

TAB. 7.6 : User Retention & Interaction Analysis

Engagement Factor	Strengths	Weaknesses
Retention Features	Some apps use notifications and progress tracking	Many apps lack features to encourage continued use
User Interaction	Apps with chatbots create a conversational experience	Poor response time decreases engagement

7.3.2 Personalization

Personalization helps improve user experience by tailoring recommendations and tracking symptom history. Apps that allow users to store past interactions and analyze symptom trends offer better engagement. However, many CSC apps do not provide detailed customization or historical tracking, limiting their effectiveness.

TAB. 7.7 : Personalization Analysis

Engagement Factor	Strengths	Weaknesses
Health History Tracking	Some apps store past interactions	Many do not allow detailed customization
Symptom Tracking	A few apps provide historical analysis	Limited options for tracking improvement over time

7.3.3 Trust & Credibility

Trust is a crucial factor in user adoption of CSC apps. Many users hesitate to rely on these apps due to concerns about diagnostic accuracy and transparency. While some apps cite medical sources and provide disclaimers, many do not clearly explain how diagnoses are generated, reducing user confidence.

TAB. 7.8 : Trust & Credibility Analysis

Engagement Factor	Strengths	Weaknesses
Medical Credibility	Some apps cite medical sources	Many do not clearly explain diagnostic methods
Transparency	A few apps provide risk disclaimers	Users lack insight into how AI derives conclusions

7.4 Comparative Analysis of Existing CSC Apps

7.4.1 Feature Comparison Table

TAB. 7.9 : Feature Comparison of CSC Apps

Feature	Ada	WebMD Symptom Checker	Buoy Health	Mediktor
Ease of Use	✓Intuitive UI	×Overloaded UI	✓Guided navigation	×Complex interface
Language Simplicity	✓Clear terms	×Uses medical jargon	✓Simple explanations	×Technical terms
Input Flexibility	×Rigid symptom lists	✓Allows some free text	✓AI-based recognition	×Predefined options only
Engagement	✓Chatbot-based	×No chatbot	✓Interactive UI	×Limited personalization
Trust & Credibility	✓Sources cited	✓Backed by WebMD	×Limited transparency	×Unknown methodology

7.5 Key Findings & Recommendations

7.5.1 Key Findings

- Users struggle with rigid symptom input, leading to frustration.
- Chatbots often lack human-like conversation quality.

- Limited inclusivity due to language barriers and accessibility shortcomings.
- Users express concerns about diagnostic transparency.

7.5.2 Recommendations

1. **Enhance Input Flexibility** : Implement approximate string matching to improve recognition of user input.
2. **Improve Conversational AI** : Utilize NLP advancements to create more natural chatbot interactions.
3. **Increase Transparency** : Explain the reasoning behind diagnostic questions and results.
4. **Personalized Symptom Tracking** : Allow users to log ongoing symptoms with trends and alerts.
5. **Expand Accessibility & Inclusivity** : Support more languages and ensure compliance with accessibility standards.

—

7.6 Evaluation of UX and Engagement in Healthcare Chatbots

7.6.1 Introduction

User Experience (UX) and Engagement are critical factors in determining the effectiveness of healthcare chatbots. A well-designed chatbot should offer an intuitive, accessible, and engaging experience while effectively assisting users in their medical needs.

This evaluation examines the **strengths, weaknesses, and engagement trends** of chatbots in healthcare based on existing applications.

7.6.2 Key UX & Engagement Factors in Healthcare Chatbots

To evaluate the UX and engagement of healthcare chatbots, we consider the following factors :

TAB. 7.10 : Key UX & Engagement Factors

Factor	Description
Ease of Use	Intuitive interface and simple interaction.
Response Accuracy	Ability to provide relevant and correct medical advice.
Personalization	Tailoring responses based on user history and preferences.
Accessibility	Support for diverse users, including visually impaired or disabled people.
Engagement Methods	Features that keep users interacting, such as gamification or reminders.
Platform Availability	Whether the chatbot is web-based, app-based, or social media-integrated.
Speech & Multimodal Support	Ability to process voice input, text, and visual cues.

7.6.3 Comparative Analysis of Existing Chatbots

Based on the study, the following table compares healthcare chatbots in terms of UX and engagement.

TAB. 7.11 : Comparison of Healthcare Chatbots

Chatbot	Purpose	Platform	Personalization	Accessibility	Speech Input	Engagement Features
[1]	Answer health-related questions & predict diseases	Web-based	No	No	No	Basic Q&A
[2]	Identifying stress relief	Heroku	No	No	No	Static responses
[3]	Patient monitoring	Telegram	No	No	No	Reminders
[4]	COVID-19 testing	AIML	No	No	No	Alert notifications
[5]	Symptom-based diagnosis	RASA, NLU	Basic	No	No	Interactive Q&A
[6]	Appointments & hospital information	NLP, gradient descent	Yes	Yes	No	Interactive interface
[7]	ADHD symptom support	Todaki	Yes	Yes	No	Gamification
[8]	Assisting deaf users	Algho platform	Yes	Yes	Yes	Sign language support
[9]	Chemotherapy symptom tracking	Facebook Messenger	Yes	No	No	Health reminders

7.6.3.1 Key Findings from the Table

- **Only 4 chatbots** support accessibility features.
- **Most chatbots lack personalization**, reducing engagement.
- **Limited speech input** is provided, affecting accessibility.
- **Engagement features are weak**, with only some offering reminders

or gamification.

7.6.4 Diagrams & Trends in UX Evolution

7.6.4.1 Trend in Accessibility Features (2019–2023)

The accessibility of healthcare chatbots has remained low, with only a **minor increase** in UX considerations over the years.

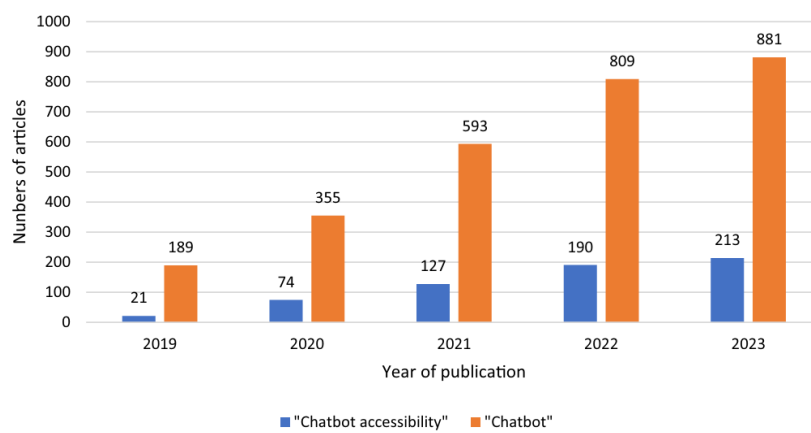


FIG. 7.1 : Accessibility Features in Healthcare Chatbots (2019-2023)

7.6.4.2 Trend in Personalization & Engagement

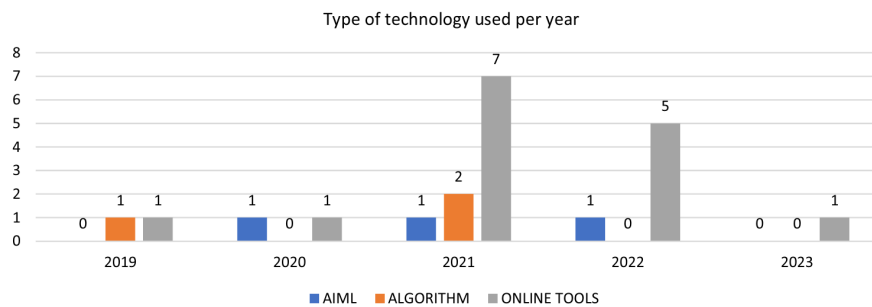


FIG. 7.2 : Chatbots Offering Personalized Responses

While some progress has been made, personalization is still lacking in many chatbots.

7.6.5 Challenges Identified

- **Lack of Personalization** : Most chatbots provide generic responses rather than adapting to user needs.
- **Limited Accessibility** : Few chatbots support multimodal interactions, speech input, or sign language.
- **Low Engagement Features** : Very few chatbots include gamification, habit tracking, or long-term user retention methods.
- **Poor Handling of Natural Language Variations** : Many chatbots fail when users phrase questions differently from what was expected.

7.6.6 Recommendations for Improving UX & Engagement

TAB. 7.12 : Recommendations for Improvement

Issue	Recommendation
Lack of personalization	Implement AI-driven adaptive responses based on user history.
Low accessibility	Introduce voice interaction and sign language features.
Poor engagement	Use gamification (rewards, progress tracking).
NLP limitations	Improve chatbot training with diverse datasets for better understanding.

7.6.7 Conclusion

While healthcare chatbots have improved over the years, their UX and engagement still need significant enhancement. Future developments should focus on **personalization, accessibility, and interactive engagement** to maximize their impact.

Chapitre 8

References

1. Allouch, M., Azaria, A., Azoulay, R.: Conversational agents : goals, technologies, vision and challenges. *Sensors* 21(24), 8448 (2021)
2. Xu, L., Sanders, L., Li, K., Chow, J.C.: Chatbot for health care and oncology applications using artificial intelligence and machine learning : systematic review. *JMIR Cancer* 7(4), e27850 (2021)
3. Abd-Alrazaq, A.A., Rababeh, A., Alajlani, M., Bewick, B.M., Househ, M.: Effectiveness and safety of using chatbots to improve mental health : systematic review and meta-analysis. *J. Med. Internet Res.* 22(7), e16021 (2020)
4. Chew, H.S.J., Achananuparp, P.: Perceptions and needs of artificial intelligence in health care to increase adoption : scoping review. *J. Med. Internet Res.* 24(1), e32939 (2022)
5. Restrepo EGY, Baldassarre M, Boticario JG : Accessibility, biases and ethics in chatbots and intelligent agents for education. In : EDU-LEARN19 Proceedings, pp. 8824–8833. IATED (2019)
6. Suhaili, S.M., Salim, N., Jambli, M.N.: Service chatbots : a systematic review. *Expert Syst. Appl.* 184, 115461 (2021)
7. Okonkwo, C.W., Ade-Ibijola, A.: Chatbots applications in education : A systematic review. *Comput. Educ. Artif. Intell.* 2, 100033 (2021)
8. Polignano, M., Narducci, F., Iovine, A., Musto, C., De Gemmis, M., Semeraro, G.: HealthAssistantBot : a personal health assistant for the Italian Language. *IEEE Access* 8, 107479–107497 (2020). <https://doi.org/10.1109/>

9. Gupta, J., Singh, V., Kumar, I.: Florence-a health care chatbot. In : 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, pp. 504–508. IEEE (2021)
10. Walss, M., Anzengruber, F., Arafa, A., Djamei, V., Navarini, A.A.: Implementing medical chatbots : an application on hidradenitis suppurativa. *Dermatology* 237(5), 712–718 (2021)
11. Chen, J.H., Agbodike, O., Kuo, W.L., Wang, L., Huang, C.H., Shen, Y.S., Chen, B.H.: Online textual symptomatic assessment chatbot based on QA weighted scoring for female breast cancer prescreening. *Appl. Sci.* 11(11), 5079 (2021)
12. Ireland, D., Bradford, D., Szepe, E., Lynch, E., Martyn, M., Hansen, D., Gaff, C.: Introducing Edna : a trainee chatbot designed to support communication about additional (secondary) genomic findings. *Patient Educ. Couns.* 104(4), 739–749 (2021)
13. Siglen, E., Vetti, H.H., Lunde, A.B.F., Hatlebrekke, T.A., Strømsvik, N., Hamang, A., et al.: Ask Rosa-the making of a digital

Conclusion générale

Les chatbots médicaux et les applications de vérification des symptômes (CSC) représentent une avancée significative dans le domaine des soins de santé numériques. Ils offrent des opportunités prometteuses pour améliorer l'accès aux soins, fournir des conseils médicaux préliminaires et guider les utilisateurs vers des professionnels de santé lorsque nécessaire. Cependant, malgré leur potentiel, ces outils présentent des limites importantes qui entravent leur efficacité et leur adoption généralisée.

Les principales limites identifiées incluent un **manque de précision diagnostique**, une **personnalisation insuffisante**, des **problèmes d'accessibilité et d'inclusivité**, ainsi que des **préoccupations concernant la confidentialité des données**. De plus, les chatbots actuels manquent souvent de fluidité conversationnelle et d'empathie, ce qui réduit l'engagement des utilisateurs et leur confiance dans ces outils.

Pour maximiser leur impact, les développeurs doivent se concentrer sur plusieurs axes d'amélioration :

- **Améliorer la précision diagnostique** en intégrant des algorithmes d'IA plus sophistiqués et en assurant une supervision humaine pour les cas complexes.
- **Renforcer la personnalisation** en adaptant les recommandations aux besoins individuels des utilisateurs et en permettant un suivi détaillé des symptômes.
- **Étendre l'accessibilité** en ajoutant des fonctionnalités multilingues, une saisie vocale et un support pour les utilisateurs handicapés.
- **Accroître la transparence** en expliquant clairement les processus de diagnostic et en citant des sources médicales fiables.
- **Améliorer l'engagement** grâce à des fonctionnalités interactives telles que la gamification, les rappels et le suivi des progrès.

En somme, bien que les chatbots médicaux aient le potentiel de transformer les soins de santé, leur succès dépendra de leur capacité à surmonter ces défis. En combinant des technologies avancées, une conception centrée sur l'utilisateur et une approche inclusive, ces outils peuvent devenir des alliés indispensables pour les patients et les professionnels de santé, tout en améliorant l'efficacité et la qualité des soins de santé à l'échelle mondiale.