**Moteurs de diagnostic médical assistés par IA pour la téléconsultation grand public**

Les **moteurs de diagnostic médical** assistés par intelligence artificielle (IA) visent à reproduire l’interrogatoire progressif d’un professionnel de santé en téléconsultation, notamment pour des **pathologies bénignes de médecine générale**. Ces systèmes interactifs procèdent par étapes : ils réalisent d’abord un **triage** des urgences (détection de symptômes graves), génèrent ensuite des **hypothèses diagnostiques** plausibles, posent des **questions ciblées** pour affiner ces hypothèses, et aboutissent enfin à un **diagnostic différentiel probable** ou à la recommandation de consulter un médecin humain. De nombreuses approches ont été explorées pour concevoir de tels moteurs, allant des **arbres de décision médicaux** aux **réseaux bayésiens**, en passant par les **chatbots NLP** et plus récemment les **modèles de langage (LLM) hybrides combinés à des structures de raisonnement médical**. Nous présentons ci-dessous un panorama des travaux scientifiques, projets open source, prototypes et solutions techniques réutilisables dans ce domaine, en insistant sur les sources ouvertes ou librement accessibles.

**Approches par arbres de décision et systèmes à base de règles**

Les premières solutions d’**aide au diagnostic** reposaient souvent sur des arbres de décision ou des règles expertes codifiant le savoir médical. Historiquement, des systèmes experts comme **MYCIN** (diagnostic infectieux) ou **Internist-1/QMR** (Quick Medical Reference) utilisaient des bases de connaissances structurées. De nos jours, des outils modernes perpétuent cette approche : par exemple, la plateforme française **MedVir** implémente un moteur de raisonnement fondé sur plus de 30 ans de connaissances cliniques formalisées. MedVir se présente comme un *symptom checker* intelligent capable d’**orientation et tri** (détection des cas graves) et d’**aide au diagnostic** en s’appuyant sur la *fuzzy logic* pour ne négliger aucun « signal faible »[medvir.fr](https://medvir.fr/#:~:text=La%20rigueur%20scientifique%20au%20c%C5%93ur,de%20la%20d%C3%A9marche). Conçue par un médecin urgentiste, cette IA embarque des protocoles de tri et des arbres décisionnels validés, intégrables en SaaS dans des outils de télémédecine[medvir.fr](https://medvir.fr/#:~:text=Tri%20et%20orientation%20avec%20medvir,PREDIAG)[medvir.fr](https://medvir.fr/#:~:text=medvir%20est%20issu%20de%2034,faible%2C%20pour%20un%20r%C3%A9sultat%20rigoureux). De même, de nombreux **protocoles de régulation médicale** (par exemple pour les services d’urgence ou les plateformes de conseils infirmiers) peuvent être implémentés via des moteurs de règles if-then. Des frameworks open source comme **OpenCDS** (système de décision clinique) ou des outils de chatbot modulaires (p. ex. **Rasa** NLP) permettent de créer des arbres de questions médicales personnalisés. Par exemple, le prototype *DiagnoBot* utilise Rasa pour guider le patient à travers un arbre de symptômes et prédire la maladie probable en sortie[aisel.aisnet.org](https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1281&context=hicss-57#:~:text=DiagnoBot%20was%20designed%20to%20search,Rasa%20NLP%20and%20machine%20learning). Ces approches basées sur des règles présentent l’avantage d’être **transparentes** (chaque décision est justifiable par une règle clinique) et alignées sur les **protocoles médicaux officiels**, ce qui est précieux pour le triage sécurisé des patients.

**Diagnostics probabilistes et réseaux bayésiens**

Pour dépasser la rigidité des règles figées, des approches probabilistes ont été explorées, notamment via les **réseaux bayésiens**. Un réseau bayésien médical représente les maladies et symptômes comme des nœuds reliés par des probabilités conditionnelles, ce qui permet de *calculer la probabilité a posteriori* des diagnostics en fonction des symptômes observés. Cette technique a l’avantage d’intégrer l’incertitude et de **fournir un diagnostic différentiel pondéré** plutôt qu’une réponse unique. Un travail notable est celui de Guan & Baral (2021), qui proposent un moteur de diagnostic basé sur le fameux réseau bayésien QMR (tiré de la base de connaissances INTERNIST)[arxiv.org](https://arxiv.org/abs/2110.08393" \l ":~:text=,new%20changes%20without%20any%20additional" \t "_blank). Leur système réalise l’inférence bayésienne pour estimer les maladies probables et utilise la **théorie de la décision bayésienne expérimentale** pour choisir les questions optimales à poser au patient (phase d’**interrogation adaptive**)[arxiv.org](https://arxiv.org/abs/2110.08393" \l ":~:text=Bayesian%20inference%20in%20the%20inference,dialogue%20datasets%2C%20Muzhi%20and%20Dxy" \t "_blank). En optimisant la stratégie de questions sur plusieurs étapes, ils obtiennent un système **interprétable, sans entraînement coûteux, s’adaptant facilement aux nouvelles données**, tout en égalant ou dépassant l’état de l’art sur des jeux de cas simulés[arxiv.org](https://arxiv.org/abs/2110.08393#:~:text=the%20inquiry%20phase,dialogue%20datasets%2C%20Muzhi%20and%20Dxy). Dans la même veine, Jones *et al.* (2022) ont développé **MidasMed**, un assistant diagnostique bayésien générique, conçu pour remédier aux faiblesses traditionnelles des réseaux bayésiens médicaux[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=Online%20AI%20symptom%20checkers%20and,Our). Leur prototype, combinant réseau bayésien amélioré et heuristiques médicales, a été comparé à des médecins humains et à six *symptom checkers* en ligne (Ada, Babylon, Buoy, Isabel, Symptomate, WebMD). Les résultats montrent que MidasMed a fourni le bon diagnostic en tête de liste dans 93 % des cas testés, surpassant ainsi nettement les autres outils d’auto-diagnostic[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=accuracy%20by%20addressing%20key%20weaknesses,of%20cases). Ces travaux suggèrent que les modèles bayésiens bien conçus peuvent atteindre une **excellente précision diagnostique** tout en restant robustes aux données manquantes et transparents dans leur raisonnement[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=Online%20AI%20symptom%20checkers%20and,2014%3B%20Richens%20et%20al).

En pratique, plusieurs solutions de triage utilisent des modèles probabilistes. Par exemple, le moteur d’inférence de la société **Infermedica** (disponible via API) pose dynamiquement des questions sur les symptômes additionnels en utilisant un algorithme probabiliste propriétaire afin de converger vers le bon diagnostic ou le bon conseil d’orientation[pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov/articles/PMC3270234/#:~:text=Differential%20Diagnosis%20Generators%3A%20an%20Evaluation,suggest%20a%20list%20of)[infermedica.com](https://infermedica.com/inference-engine#:~:text=Our%20AI%20asks%20the%20right,health%2C%20and%20give%20appropriate%20recommendations). De son côté, la plateforme **Isabel Healthcare** (outil clinique existant depuis 2001) s’appuie sur l’IA et des bases de cas cliniques pour fournir aux professionnels et au public une liste différentiée de diagnostics probables en fonction des symptômes saisis[symptomchecker.isabelhealthcare.com](https://symptomchecker.isabelhealthcare.com/#:~:text=Isabel%20,doctors%20use%20and%20trust)[symptomchecker.isabelhealthcare.com](https://symptomchecker.isabelhealthcare.com/#:~:text=1,like%20without%20the%20endless%20questions). Bien que ces solutions commerciales ne soient pas open source, elles illustrent la **fiabilité accrue** obtenue grâce à des algorithmes de diagnostic probabilistes entraînés sur de vastes données cliniques.

**Apprentissage automatique et interrogatoire adaptatif**

*Figure – Exemple de génération de cas patients synthétiques avec symptômes et diagnostic différentiel simulé via une plateforme de téléconsultation. Un moteur de diagnostic peut s’entraîner sur de tels cas en vue d’apprendre à poser des questions pertinentes et émettre des hypothèses de plus en plus précises*[*github.com*](https://github.com/mila-iqia/ddxplus#:~:text=domain)*.*

Les approches plus récentes exploitent le **machine learning** pour apprendre automatiquement le diagnostic à partir de données. En particulier, la **recherche de politiques optimales de questionnement** peut être formulée comme un problème de décision séquentielle. Des techniques d’**apprentissage par renforcement (RL)** ont ainsi été appliquées pour entraîner un agent à interroger le patient (ou un simulateur de patient) de manière à maximiser la précision finale du diagnostic. Zhong *et al.* (2022) ont proposé une architecture de **reinforcement learning hiérarchique** pour le diagnostic médical automatisé[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=,policy%20consists%20of%20amastermodel%20that)[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=is%20responsible%20for%20triggering%20a,up%20researches). Plutôt qu’une politique « plate » traitant tous les symptômes de la même façon, leur agent RL comporte deux niveaux : un **modèle maître** qui décide quelle sous-politique activer (quel groupe de symptômes explorer), et des **modèles subordonnés** spécialisés (p. ex. questionner les symptômes respiratoires vs. neurologiques) suivis d’un classifieur de maladie[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=integrate%20a%20hierarchical%20policy%20structure,We). Cette décomposition hiérarchique permet de gérer un très grand nombre de symptômes/maladies de manière efficace. En entraînant et testant ce système sur des ensembles de cas cliniques simulés et réels, ils montrent une **meilleure précision diagnostique et une meilleure exploration des symptômes** que les approches RL précédentes[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=is%20capable%20to%20deal%20with,this%20http%20URL). Fait important, les auteurs ont rendu leur code et leurs données librement accessibles pour stimuler les recherches ultérieures[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=symptom%20recall%20in%20disease%20diagnosis,this%20http%20URL). D’autres travaux en RL (e.g. Xia *et al.*, 2020) ont utilisé des techniques de **Deep Q-learning** pour apprendre à sélectionner successivement les questions optimales, tandis que des approches supervisées classiques ont traité le diagnostic comme un problème de **classification multi-classe** (prédire la maladie à partir de l’ensemble des symptômes saisis), parfois au moyen de réseaux de neurones profonds. Néanmoins, ces approches purement apprenantes ont souvent besoin de **données d’entraînement massives**. Pour y remédier, la création de **jeux de données synthétiques** a été explorée : par exemple, le dataset **DDXPlus** (NeurIPS 2022) génère des *cas patients artificiels* en combinant une base de connaissances médicale avec un simulateur open source de dossiers de synthèse (Synthea), fournissant pour chaque cas une liste de symptômes et d’antécédents cohérents ainsi qu’un diagnostic différentiel de référence[github.com](https://github.com/mila-iqia/ddxplus#:~:text=domain). De tels corpus permettent d’entraîner et d’évaluer des agents de diagnostic automatisé (*Automatic Diagnosis, AD*), ainsi que des modules de détection de symptômes pertinents (*Automatic Symptom Detection, ASD*) de manière reproductible.

**Chatbots médicaux interactifs (NLP) et moteurs de règles hybrides**

Avec les avancées en **Traitement du Langage Naturel (NLP)**, de nombreux projets ont cherché à bâtir des **chatbots médicaux** capables de dialoguer avec le patient en langage libre tout en suivant un **fil directeur médical**. Un exemple notable est **Babylon Health**, dont le chatbot de symptom checking était déployé commercialement pour des téléconsultations : il combinait une interface de chat interactive à un moteur de règles et de probabilités en arrière-plan pour poser des questions additionnelles et fournir un triage (auto-évaluation)[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=accuracy%20by%20addressing%20key%20weaknesses,of%20cases). D’autres startups comme **Ada Health, Buoy, Ubie, Symptoma** etc. ont développé des assistants virtuels similaires, orientés grand public. Sur le plan open source, des développeurs ont publié des prototypes de chatbots médicaux éducatifs – par exemple, le projet **Symptom-Checker-Chatbot** (Python/PyTorch) simule une conversation avec un patient pour l’aiguiller et peut être étendu ou modifié librement (licence MIT)[github.com](https://github.com/imonishkumar/Symptom-Checker-Chatbot" \l ":~:text=A%20Symptom%20Checker%20Chatbot%3A%20Get,related%20queries" \t "_blank). Ce type de chatbot utilise souvent un **pipeline NLP** classique : reconnaissance d’entités médicales dans le texte du patient (symptômes, durée, antécédents…), correspondance avec une base de connaissances (par ex. dictionnaire de symptômes) puis suivi d’un arbre de décision ou d’un modèle de classification pour proposer la prochaine question ou un diagnostic.

Une approche intéressante consiste à **hybrider NLP et règles expertes**. Par exemple, *DiagnoBot* (Université d’Hawaï, 2024) combine l’analyse NLP du discours du patient avec un **système de prédiction de maladies** : les symptômes extraits alimentent un algorithme de machine learning qui propose un diagnostic, tout en permettant au chatbot de rechercher des informations sur la maladie suggérée afin de poser des questions de clarification[aisel.aisnet.org](https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1281&context=hicss-57#:~:text=DiagnoBot%20was%20designed%20to%20search,Rasa%20NLP%20and%20machine%20learning). De même, des bibliothèques comme **John Snow Labs** (Spark NLP Healthcare) offrent des composants prêts à l’emploi (reconnaissance d’entités cliniques, extraction de symptômes) qui peuvent être intégrés dans un chatbot. Un exemple dans un contexte clinique est le projet **DXplain** (Université de Harvard), un ancien moteur de diagnostic différentiel qui a été adapté avec une interface conversationnelle pour l’enseignement médical : l’utilisateur saisit les symptômes en langage naturel et le système renvoie une liste de diagnostics envisageables avec explications. En somme, les chatbots médicaux interactifs constituent une **interface utilisateur naturelle** pour les moteurs de diagnostic, mais leur efficacité dépend fortement du **coeur de raisonnement médical** (règles, algorithmes) qui les pilote.

**Modèles de langage (LLM) et hybridation IA + raisonnement médical**

Dernière évolution en date, l’essor des **grands modèles de langage** (LLM) offre de nouvelles possibilités pour les moteurs de diagnostic. Des modèles type GPT entraînés sur de larges corpus médicaux peuvent comprendre des descriptions de symptômes complexes et même **fournir des conseils ou un diagnostic probable en langage naturel**. Toutefois, les LLM généralistes ont tendance à manquer de structure dans le raisonnement clinique et peuvent produire des réponses peu fiables sans garde-fous. Pour remédier à cela, plusieurs projets hybrides combinent la **puissance des LLM** avec des **connaissances médicales structurées** ou des mécanismes de contrôle.

Un exemple marquant est le système **AMIE (Articulate Medical Intelligence Explorer)** développé par Google Research en 2023-2024. AMIE est une IA de diagnostic conversationnel basée sur un LLM spécialisé, entraînée dans un *environnement simulé en auto-jeu* (self-play) où elle converse avec des cas patients virtuels et reçoit des feedbacks automatisés pour affiner ses compétences[buzz-esante.fr](https://buzz-esante.fr/amie-nouvelle-ia-de-google-au-service-du-diagnostic-medical/#:~:text=Con%C3%A7u%20pour%20am%C3%A9liorer%20les%20dialogues,diverses%20conditions%20m%C3%A9dicales%20et%20contextes)[buzz-esante.fr](https://buzz-esante.fr/amie-nouvelle-ia-de-google-au-service-du-diagnostic-medical/#:~:text=Le%20moteur%20de%20raisonnement%20m%C3%A9dical,confirmer%20ou%20infirmer%20ces%20hypoth%C3%A8ses). Ce système comporte deux modules principaux : un **moteur de raisonnement diagnostique** qui analyse symptômes/antécédents et génère des hypothèses (s’appuyant sur des techniques d’apprentissage automatique et de NLP avancées), et un **module de dialogue** qui permet de poser des questions de façon naturelle et d’expliquer le raisonnement[buzz-esante.fr](https://buzz-esante.fr/amie-nouvelle-ia-de-google-au-service-du-diagnostic-medical/#:~:text=AMIE%20se%20compose%20de%20deux,un%20syst%C3%A8me%20de%20dialogue%20conversationnel)[buzz-esante.fr](https://buzz-esante.fr/amie-nouvelle-ia-de-google-au-service-du-diagnostic-medical/#:~:text=Le%20syst%C3%A8me%20de%20dialogue%20conversationnel,en%20posant%20des%20questions%20pertinentes). Dans une étude publiée dans *Nature* (2025), AMIE a été comparé en double aveugle à de vrais médecins généralistes sur 159 cas simulés de consultation textuelle. Les résultats montrent qu’**AMIE a atteint une exactitude diagnostique supérieure à celle des médecins**, et a été jugé meilleur sur 30 des 32 critères d’évaluation clinique par des experts (incluant la pertinence des questions, l’exhaustivité de l’anamnèse, la communication, l’empathie, etc.)[nature.com](https://www.nature.com/articles/s41586-025-08866-7?error=cookies_not_supported&code=6cd2194a-dc5e-44bc-b878-12b69d7420f6#:~:text=empathy,several%20limitations%20and%20should%20be)[nature.com](https://www.nature.com/articles/s41586-025-08866-7?error=cookies_not_supported&code=6cd2194a-dc5e-44bc-b878-12b69d7420f6#:~:text=validated%20patient,this%20is%20unfamiliar%20in%20clinical). Il s’agit d’un **jalon important** vers une IA capable de dialogues diagnostiques de haut niveau, même si des recherches supplémentaires sont nécessaires avant une application réelle à grande échelle[nature.com](https://www.nature.com/articles/s41586-025-08866-7?error=cookies_not_supported&code=6cd2194a-dc5e-44bc-b878-12b69d7420f6#:~:text=demonstrated%20greater%20diagnostic%20accuracy%20and,milestone%20towards%20conversational%20diagnostic%20AI).

Parallèlement, la communauté open source s’est mobilisée pour adapter des LLM à la médecine. Par exemple, le modèle **ChatDoctor** (2023) est un LLM *open-source* dérivé de LLaMA, affiné sur des corpus de conversations médicales doctorant-patient[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Overview). Il a été entraîné sur des paires de questions patients/réponses de médecins (issues de forums comme HealthCareMagic et iCliniq), ainsi que sur des informations médicales structurées (listes de maladies, symptômes, traitements). ChatDoctor démontre une capacité à **comprendre les préoccupations du patient** en langage courant et à fournir des réponses alignées sur celles d’un vrai médecin[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Patient%20question%20from%20icliniq%3A%20Hi,Please%20help)[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Answer%20form%20ChatDoctor%3A%20Hi%2C%20Thank,Best%20wishes%2C%20Chat%20Doctor). Par exemple, face à une description de céphalée unilatérale avec rhume, ChatDoctor a correctement évoqué une sinusite bénigne plutôt qu’une tumeur cérébrale, rassurant le patient et conseillant les examens appropriés[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Real%20doctor%20answer%20from%20icliniq%3A,brain%20and%20CT%20paranasal%20sinuses)[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Answer%20form%20ChatDoctor%3A%20Hi%2C%20Thank,Best%20wishes%2C%20Chat%20Doctor). Ce modèle (disponible sous licence open source) illustre comment un LLM généraliste peut être spécialisé via un **fine-tuning sur des données médicales** pour devenir un assistant virtuel fiable.

Plusieurs équipes académiques travaillent également sur l’**intégration explicite de connaissances cliniques dans les LLM**. John Snow Labs, par exemple, a annoncé un *Medical LLM Reasoner* combinant une architecture de chaîne de raisonnement (**Chain-of-Thought**) à des **arbres de décision médicaux intégrés** dans les poids du modèle[johnsnowlabs.com](https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/#:~:text=3.%20Chain,helps%20the%20model%20avoid%20the). Concrètement, ils incorporent des algorithmes de décision médicale (guidelines, arbres de symptômes) au cœur du réseau de neurones, afin que le LLM suive des **chemins diagnostiques conformes aux protocoles cliniques** plutôt que de se fier à des associations de surface[johnsnowlabs.com](https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/#:~:text=4,optimized%20specifically%20for%20handling%20uncertainty)[johnsnowlabs.com](https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/#:~:text=These%20technical%20innovations%20create%20models,to%20diagnosis%20to%20treatment%20recommendations). Ce type d’hybridation (modèle neuronique + règles expertes) vise à éviter les « raccourcis » de raisonnement et à améliorer l’interprétabilité des réponses de l’IA.

Enfin, il est notable que la frontière entre IA propriétaire et IA ouverte s’estompe dans le domaine médical. En 2025, une étude financée par les NIH (Harvard Medical School) a comparé la performance diagnostique de GPT-4 (modèle propriétaire) à celle d’un **modèle open-source** de dernière génération sur des cas cliniques complexes. Le constat est que le modèle open-source – une version fine-tunée de LLaMA nommée *Llama 3.1 (405B)* – a réussi à **égaler GPT-4** dans la génération de diagnostics différentiels pour des cas difficiles nécessitant un raisonnement clinique poussé[dbmi.hms.harvard.edu](https://dbmi.hms.harvard.edu/news/open-source-ai-matches-top-proprietary-llm-solving-tough-medical-cases#:~:text=Has%20open)[dbmi.hms.harvard.edu](https://dbmi.hms.harvard.edu/news/open-source-ai-matches-top-proprietary-llm-solving-tough-medical-cases#:~:text=Hospital). Cela suggère que les **LLM open source récents comblent rapidement l’écart** avec les modèles fermés, ouvrant la voie à des outils de diagnostic IA déployables localement dans les hôpitaux, avec des coûts réduits et un contrôle accru par la communauté médicale[nature.com](https://www.nature.com/articles/s41746-025-01488-3?error=cookies_not_supported&code=94bc4fb4-7306-4fa5-bf32-1347d27a2aff#:~:text=This%20study%20evaluated%20the%20diagnostic,LLMs%20as%20decision%20support%20tools). Des projets comme **Meditron** (EPFL, 2023) fournissent d’ores et déjà une suite de modèles de langage médicaux open source (avec jusqu’à 2k de stars GitHub) pouvant être utilisés comme base pour des chatbots de diagnostic ou d’information santé[jamanetwork.com](https://jamanetwork.com/journals/jama-health-forum/fullarticle/2831206#:~:text=Frontier%20Open,on%20complex%20diagnostic%20challenge%20cases)[github.com](https://github.com/epfLLM/meditron#:~:text=Meditron%20is%20a%20suite%20of,symptoms%2C%20cause).

**Conclusion**

En conclusion, la recherche et l’innovation abondent pour construire des moteurs de diagnostic médical assisté par IA adaptés à la téléconsultation grand public. Les **arbres de décision expert** apportent la fiabilité des protocoles cliniques établis, les **modèles bayésiens** offrent un cadre probabiliste rigoureux pour le diagnostic différentiel, les techniques d’**apprentissage automatique** (RL et autres) apprennent à interroger efficacement le patient, et les **LLM hybrides** promettent des dialogues plus naturels sans sacrifier la précision médicale. De nombreux **travaux académiques** récents fournissent des bases théoriques et parfois du code ouvert (publications en libre accès citées ci-dessus), tandis que des **projets open source** et des **APIs disponibles** (Infermedica, Rasa, ChatDoctor, Meditron, etc.) permettent de réutiliser des composants pour bâtir son propre système. L’enjeu reste de combiner ces avancées de manière sûre et éthique : en veillant à une **exclusion fiable des urgences** (triage robuste aux “red flags”), à la **pertinence clinique** des questions posées et des réponses fournies, et à l’**explicabilité** du raisonnement pour instaurer la confiance des utilisateurs. Avec des solutions bien documentées et librement accessibles, les développeurs de plateformes de téléconsultation disposent désormais d’une **boîte à outils riche** pour intégrer une IA d’aide au diagnostic visant à améliorer l’orientation et la prise en charge des patients pour les petits soucis du quotidien.

**Sources :** Publications et ressources sélectionnées sur les systèmes de diagnostic médical assisté par IA (arbres décisionnels, réseaux bayésiens, apprentissage automatique, chatbots NLP, LLM spécialisés), incluant des projets open source et études récentes sur leur performance : Guan & Baral 2021[arxiv.org](https://arxiv.org/abs/2110.08393#:~:text=,new%20changes%20without%20any%20additional)[arxiv.org](https://arxiv.org/abs/2110.08393#:~:text=Bayesian%20inference%20in%20the%20inference,dialogue%20datasets%2C%20Muzhi%20and%20Dxy); Jones *et al.* 2022[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=Online%20AI%20symptom%20checkers%20and,Our)[frontiersin.org](https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.727486/full#:~:text=accuracy%20by%20addressing%20key%20weaknesses,of%20cases); Zhong *et al.* 2023[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=integrate%20a%20hierarchical%20policy%20structure,We)[arxiv.org](https://arxiv.org/abs/2004.14254#:~:text=is%20capable%20to%20deal%20with,this%20http%20URL); Google Research AMIE 2024[buzz-esante.fr](https://buzz-esante.fr/amie-nouvelle-ia-de-google-au-service-du-diagnostic-medical/#:~:text=AMIE%20se%20compose%20de%20deux,un%20syst%C3%A8me%20de%20dialogue%20conversationnel)[nature.com](https://www.nature.com/articles/s41586-025-08866-7?error=cookies_not_supported&code=6cd2194a-dc5e-44bc-b878-12b69d7420f6#:~:text=empathy,several%20limitations%20and%20should%20be); John Snow Labs 2023[johnsnowlabs.com](https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/#:~:text=4,optimized%20specifically%20for%20handling%20uncertainty); Kim *et al.* 2025[nature.com](https://www.nature.com/articles/s41746-025-01488-3?error=cookies_not_supported&code=94bc4fb4-7306-4fa5-bf32-1347d27a2aff#:~:text=This%20study%20evaluated%20the%20diagnostic,LLMs%20as%20decision%20support%20tools); Harvard Medical School 2025[dbmi.hms.harvard.edu](https://dbmi.hms.harvard.edu/news/open-source-ai-matches-top-proprietary-llm-solving-tough-medical-cases#:~:text=The%20answer%20appears%20to%20be,and%20Brigham%20and%20Women%E2%80%99s%20Hospital); ChatDoctor 2023[github.com](https://github.com/Kent0n-Li/ChatDoctor#:~:text=Answer%20form%20ChatDoctor%3A%20Hi%2C%20Thank,Best%20wishes%2C%20Chat%20Doctor); MedVir 2022[medvir.fr](https://medvir.fr/#:~:text=La%20rigueur%20scientifique%20au%20c%C5%93ur,de%20la%20d%C3%A9marche); DDXPlus dataset 2023[github.com](https://github.com/mila-iqia/ddxplus#:~:text=domain), etc.