

# IA contre recherche traditionnelle : différences de classement et de résultats

By RankStudio Publié le 22 octobre 2025 41 min de lecture



## Résumé Analytique

L'essor de la **recherche basée sur l'IA** (ou [recherche générative](#)) transforme rapidement la manière dont l'information est récupérée et classée. Les moteurs de recherche traditionnels (par exemple, Google, Bing) s'appuient depuis longtemps sur des algorithmes qui indexent le contenu web et classent les résultats en fonction de signaux de pertinence tels que la correspondance des mots-clés, l'analyse des liens et le comportement des utilisateurs. En revanche, les systèmes de recherche modernes basés sur l'IA (par exemple, ChatGPT, AI Overviews de Google, Bing Chat) utilisent souvent de grands modèles linguistiques (LLM) pour générer des réponses directes ou des résumés en synthétisant des informations provenant de multiples sources. Ce changement fondamental pose de nouveaux défis pour comparer la manière dont ces systèmes renvoient et *classent* l'information.

Ce rapport fournit une analyse approfondie de la manière de **comparer les résultats de recherche traditionnels avec les résultats de recherche générés par l'IA**, y compris le classement et la présentation. Nous passons en revue les technologies sous-jacentes, les métriques d'évaluation, les comportements des utilisateurs et les études empiriques. Les principales conclusions incluent :

- **Paradigmes différents** : La recherche traditionnelle produit des listes de pages classées, tandis que la recherche par IA produit souvent une seule réponse synthétisée (parfois avec des sources citées) (Source: [searchengineland.com](#)) (Source: [www.techtarget.com](#)). Cela signifie que le *classement* dans la recherche par IA se produit implicitement pendant la récupération et la génération de la réponse, plutôt que sous forme de liste visible.
- **Performance et préférences des utilisateurs** : Des études contrôlées révèlent que les utilisateurs exploitant la recherche par chat IA (par exemple, ChatGPT) trouvent souvent des réponses correctes plus rapidement, mais expriment toujours une préférence subjective plus élevée pour les interfaces de recherche traditionnelles comme Google (Source: [www.researchgate.net](#)) (Source: [www.researchgate.net](#)). Par exemple, une vaste expérience (n=1 526) a montré que les

utilisateurs de ChatGPT étaient « plus rapides et plus susceptibles de trouver des réponses correctes », mais la plupart des participants préféraient toujours Google (Source: [www.researchgate.net](http://www.researchgate.net)). Une autre étude (n≈199) a révélé que la recherche par IA entraînait des temps de tâche significativement plus courts sans perte de précision (Source: [www.researchgate.net](http://www.researchgate.net)).

- **Points forts des tâches :** La recherche par IA excelle dans les tâches nécessitant compréhension ou synthèse (par exemple, analyse de contenu, questions-réponses générales) mais peut avoir des difficultés avec les faits à jour et les requêtes de niche (par exemple, entreprises locales) (Source: [searchengineland.com](http://searchengineland.com)) (Source: [searchengineland.com](http://searchengineland.com)) (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). Dans une comparaison de 62 requêtes variées, Google a surpassé l'IA sur la plupart des requêtes informationnelles (score ~5,83 contre 5,19), tandis que ChatGPT excellait dans les tâches d'analyse des lacunes de contenu (Source: [searchengineland.com](http://searchengineland.com)) (Source: [searchengineland.com](http://searchengineland.com)).
- **Métriques d'évaluation :** La RI traditionnelle utilise la précision/le rappel et des métriques basées sur le rang (NDCG, MAP) pour évaluer les listes de résultats, mais celles-ci ne sont pas directement applicables aux réponses d'IA de forme libre. Au lieu de cela, les évaluations combinent des mesures de RI (pour le composant de récupération) avec des métriques de qualité de réponse (précision, exhaustivité, taux d'hallucination) et les résultats d'études d'utilisateurs. Les enquêtes et les indices de satisfaction suggèrent que la satisfaction globale de la recherche reste élevée (Google ACSI 81, Bing 77) à mesure que les fonctionnalités d'IA sont adoptées (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)).
- **Tendances générales :** Les données d'utilisation montrent que l'adoption de la recherche par IA est en croissance mais n'est pas dominante. Selon une étude de marché, la recherche basée sur l'IA (LLM dans les navigateurs) représentait environ 5 à 6 % des requêtes de bureau aux États-Unis à la mi-2024, bien que parmi les « premiers adoptants », elle ait atteint 40 % des recherches de bureau (Source: [tipsheet.ai](http://tipsheet.ai)). ChatGPT seul comptait plus de 400 millions d'utilisateurs actifs hebdomadaires début 2025 (Source: [www.investing.com](http://www.investing.com)). Dans l'éducation, les étudiants déclarent utiliser ChatGPT fréquemment mais *ne pas* abandonner les moteurs de recherche (Source: [link.springer.com](http://link.springer.com)) (Source: [link.springer.com](http://link.springer.com)).
- **Défis :** Les réponses générées par l'IA peuvent halluciner ou citer de manière inexacte ; une analyse récente a identifié 16 limitations courantes (par exemple, attributions de sources trop confiantes) dans les « **moteurs de réponse** » de l'IA (Source: [www.emergentmind.com](http://www.emergentmind.com)). Assurer la fiabilité factuelle et la transparence est une préoccupation majeure. La recherche traditionnelle reste essentielle pour des besoins d'information approfondis (Source: [pmc.ncbi.nlm.nih.gov](http://pmc.ncbi.nlm.nih.gov)), en particulier pour les tâches académiques ou complexes.

En résumé, **comparer le classement de la recherche traditionnelle et de la recherche par IA nécessite une évaluation multidimensionnelle**. Cela implique à la fois le résultat algorithmique (quelles sources sont récupérées et comment les réponses sont construites) et l'évaluation centrée sur l'utilisateur (vitesse, précision, satisfaction). Ce rapport explore ces aspects en détail, combinant contexte historique, analyse technique, données empiriques et études de cas. Nous concluons avec les implications pour la récupération d'informations, le [SEO](#) et la conception future de la recherche.

## Introduction

Les moteurs de recherche sont la **pierre angulaire de l'accès à l'information** sur Internet depuis des décennies. Les systèmes de recherche traditionnels (Google, Bing, Yahoo, etc.) indexent des milliards de pages web et utilisent des algorithmes de classement sophistiqués pour renvoyer une *liste classée* de liens pertinents en réponse à la requête d'un utilisateur. Ces algorithmes classent les résultats en fonction de signaux tels que la correspondance des mots-clés, l'[analyse des liens de type PageRank](#), la qualité du contenu, l'engagement des utilisateurs et de nombreux autres facteurs (Source: [searchengineland.com](http://searchengineland.com)) (Source: [www.techtarget.com](http://www.techtarget.com)). Au fil du temps, les moteurs de recherche ont progressivement intégré des techniques d'IA (apprentissage automatique pour le classement, compréhension du langage naturel, etc.), mais le résultat fondamental est resté des listes de liens (appelés « liens bleus ») avec des extraits.

Dans la **nouvelle ère de la recherche basée sur l'IA**, les grands modèles linguistiques (LLM) et l'IA générative sont de plus en plus utilisés pour répondre directement aux requêtes en langage naturel. Des systèmes comme [ChatGPT](#), Google Bard/Gemini et Microsoft Bing Chat emploient des LLM capables de récupérer des informations et de synthétiser une réponse concise (Source: [searchengineland.com](http://searchengineland.com)). Certains de ces systèmes citent les sources en ligne, tandis que d'autres (comme de nombreux chatbots) présentent une réponse de forme libre. Ce changement soulève la question : *comment pouvons-nous comparer les performances et le classement des résultats de recherche basés sur l'IA aux résultats de recherche traditionnels ?*

Comparer les deux paradigmes n'est pas trivial. L'évaluation de la recherche traditionnelle se concentre sur la **qualité du classement** – dans quelle mesure la liste ordonnée des pages renvoyées satisfait le besoin d'information de l'utilisateur. En revanche, la recherche par IA produit souvent une **seule réponse synthétisée** (avec d'éventuelles citations) plutôt qu'une liste de

pages classées. Ainsi, les notions de « position de classement » deviennent ambiguës. Au lieu de cela, nous devons considérer la **qualité de la réponse de bout en bout**, qui inclut non seulement la récupération d'informations pertinentes, mais aussi leur présentation cohérente et précise (Source: [searchengineland.com](https://searchengineland.com)) (Source: [www.emergentmind.com](https://www.emergentmind.com)). De plus, les modèles d'interaction utilisateur diffèrent : la recherche traditionnelle peut nécessiter de cliquer sur les résultats, tandis que les réponses de l'IA peuvent satisfaire la requête immédiatement (créant l'expérience dite « zéro clic » (Source: [tipsheet.ai](https://tipsheet.ai)).

Ce rapport examine ces questions en profondeur. Il passe en revue le **contexte historique** et l'évolution de la technologie de recherche, définit clairement les deux paradigmes et explore comment ils récupèrent et présentent l'information différemment. Nous détaillons les *méthodes d'évaluation* (métriques, études d'utilisateurs, tâches de référence) qui peuvent être utilisées pour les comparer. Nous présentons des **données et des études de cas**, y compris des expériences académiques et des analyses industrielles, qui éclairent les performances comparatives, les préférences des utilisateurs et les pièges. Différentes perspectives sont considérées – allant de la recherche en récupération d'informations au SEO/marketing, et de l'expérience utilisateur à la technologie sous-jacente. Enfin, nous discutons des **implications** de ce changement pour l'avenir de la recherche, la création de contenu et l'accès à l'information.

## Contexte Historique de la Recherche

La technologie de recherche a considérablement évolué depuis les débuts du web. Initialement, la recherche par annuaire et la correspondance de mots-clés (par exemple, AltaVista, Lycos) étaient courantes. L'algorithme **PageRank** (vers 1998) a révolutionné la recherche web en utilisant les hyperliens comme des approbations, donnant naissance à la domination de Google. Au cours des années 2000 et 2010, les moteurs de recherche ont ajouté des composants d'IA et de ML plus avancés : ils ont intégré la pondération des termes (TF-IDF), les signaux de comportement des utilisateurs (données de clics), la localisation et la personnalisation, et ont ensuite entraîné des algorithmes de classement appris par machine comme RankBrain et BERT (Source: [searchengineland.com](https://searchengineland.com)) (Source: [www.techtarget.com](https://www.techtarget.com)).

Tout au long de cette période, la recherche en récupération d'informations (RI) a développé des cadres d'évaluation formels (par exemple, les benchmarks *TREC* de la Text REtrieval Conference) pour évaluer la qualité de la recherche. Les résultats sont généralement évalués par des jugements de pertinence sur les requêtes, en utilisant des métriques telles que la précision, le rappel, la précision moyenne et le gain cumulatif actualisé (NDCG) (Source: [pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov)). Ces métriques supposent une liste de résultats et la jugent par ordre.

Parallèlement aux avancées algorithmiques, le comportement des utilisateurs a changé. L'essor de la recherche mobile et vocale a introduit de nouvelles interfaces, mais l'idée centrale est restée la même : l'utilisateur tape ou prononce une requête, le moteur de recherche renvoie des résultats classés. Les utilisateurs parcourent généralement les quelques premiers liens. Historiquement, les **résultats de recherche organiques** ont été le principal canal de découverte de contenu, et des métriques comme la « part de recherche » (fraction de toutes les requêtes) ont mesuré l'utilisation. Google est longtemps resté l'acteur dominant (souvent ~90 % de part de marché mondiale [Techcrunch et autres]).

Plus récemment, l'explosion de la **technologie de l'IA et des LLM** a bouleversé la recherche. L'introduction de ChatGPT fin 2022 (et de GPT-4 en 2023) a montré que les LLM pouvaient répondre à des requêtes complexes de manière conversationnelle. Les moteurs de recherche ont réagi en intégrant l'IA. Par exemple, en 2023, Google a commencé à tester son *Search Generative Experience* (SGE) et a lancé Bing Chat, propulsé par OpenAI. Cela a rendu le paysage **multimodal** : les utilisateurs peuvent toujours utiliser la recherche traditionnelle ou passer à des outils d'IA basés sur le chat.

Cette histoire est importante car elle encadre notre comparaison. La recherche traditionnelle a évolué pour maximiser la pertinence des listes de liens ; la recherche par IA évolue pour maximiser l'utilité et la cohérence des réponses synthétisées. Chacune a des forces et des attentes d'utilisateurs différentes. Comme l'a noté Hersch (2024), la **recherche (RI) reste cruciale** même à l'ère de l'IA : les utilisateurs ont toujours besoin d'informations fiables, opportunes et contextuelles, et la recherche sur les systèmes de recherche est « essentielle » parallèlement au développement des LLM (Source: [pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov)).

## Mécanismes de Classement de la Recherche Traditionnelle

Les moteurs de recherche traditionnels suivent un processus en plusieurs étapes : **(1) Exploration et Indexation** : des robots automatisés parcourent le web, récupérant des pages pour construire un index. **(2) Traitement de la Requête** : la requête de l'utilisateur est analysée pour les mots-clés et l'intention. **(3) Récupération et Classement** : le moteur récupère les pages

candidates de l'index et les classe par pertinence, puis **(4) Présentation des Résultats** : présente une liste classée (SERP) avec des extraits, des titres, des URL, et souvent du contenu mixte (annonces, cartes, carrousels d'achat, etc.).

Les facteurs clés de classement ont historiquement inclus :

- **Pertinence des mots-clés** : dans quelle mesure le contenu de la page correspond aux termes de la requête (avec TF-IDF, BM25, etc.).
- **Signaux de liens** : par exemple, PageRank, où les pages avec de nombreux autres liens (en particulier des liens de haute qualité) sont mieux classées (Source: [searchengineland.com](https://searchengineland.com)).
- **Actualité** : Date et opportunité, en particulier pour les requêtes d'actualité.
- **Comportement de l'utilisateur** : Taux de clics, temps de séjour, personnalisation par lieu ou historique.
- **Compréhension sémantique** : Les moteurs modernes utilisent le PNL pour interpréter les synonymes, l'intention de la requête et le contexte (par exemple, la mise à jour BERT de Google en 2019).

La présence de ces signaux est reflétée dans les documents de **transparence algorithmique** (par exemple, Google Search Essentials) et de nombreuses analyses SEO (Source: [searchengineland.com](https://searchengineland.com)) (Source: [aiscorereport.com](https://aiscorereport.com)). Par exemple, le nombre de backlinks a été cité à plusieurs reprises comme un signal majeur pour le classement de Google (Source: [aiscorereport.com](https://aiscorereport.com)). Au fil des ans, les moteurs de recherche ont également ajusté leurs algorithmes pour la prévention du spam, pénalisant la manipulation de liens ou le contenu de faible qualité.

Du point de vue de la *comparaison des résultats*, une requête de recherche traditionnelle produit une **liste ordonnée de classement d'URL/pages**. Les utilisateurs inspectent généralement les 1 à 10 premiers résultats (première page) pour trouver des réponses. Le concept de *position de classement* est crucial : être en position n°1 génère une probabilité de clic considérablement plus élevée que les rangs inférieurs (comme le montrent les études de distribution des clics). Search Engine Land rapporte que de nombreux professionnels du SEO étaient « obsédés » par les positions de classement au cours des dernières décennies (Source: [searchengineland.com](https://searchengineland.com)). Si un site descend de quelques places seulement, le trafic diminue considérablement. Ainsi, le principal signal d'évaluation de la performance de recherche a été la *position* sur la SERP.

L'évaluation quantitative de la recherche traditionnelle repose donc sur les **métriques de RI**. Par exemple, le NDCG (Normalized Discounted Cumulative Gain) mesure la qualité de la couverture des documents pertinents par les résultats les mieux classés. Si nous disposons d'un ensemble de pages pertinentes de référence pour une requête, nous pouvons calculer la précision de la liste renvoyée et le nombre d'éléments pertinents apparaissant en haut. Ces métriques comparent implicitement la *qualité de classement* de l'algorithme du moteur par rapport à un standard de référence.

Étant donné que le résultat est une liste, les comparaisons entre moteurs peuvent utiliser des métriques comme la précision@K ou la corrélation de rang entre les listes (Source: [pmc.ncbi.nlm.nih.gov](https://pmc.ncbi.nlm.nih.gov)). Un exemple direct : les données de Reuters (via Tipsheet) ont montré que la recherche traditionnelle (Google/Bing) dominait toujours le trafic global, en particulier parmi tous les utilisateurs, malgré l'essor des outils d'IA (Source: [tipsheet.ai](https://tipsheet.ai)). Cependant, cela ne capture pas la *qualité de la réponse*, mais seulement la part de trafic.

Enfin, la recherche traditionnelle s'est enrichie de **snippets et de résumés** (Featured Snippets de Google, cartes Wikipédia, etc.), ce qui brouille la frontière avec l'IA. Même l'ancien système de Google fournissait des réponses rapides pour des requêtes triviales (calculs, météo, etc.). Mais fondamentalement, toutes les informations provenaient de pages web.

En résumé, le **classement de la recherche traditionnelle** consiste à récupérer des documents existants et à les ordonner par pertinence estimée. Son évaluation et sa comparaison utilisent des métriques de RI bien établies et des données d'engagement utilisateur. En revanche, la recherche basée sur l'IA fusionne la récupération avec la *génération* de contenu, exigeant de nouvelles approches de comparaison (discutées ci-dessous).

## L'Essor de la Recherche Basée sur l'IA (Recherche Générative)

En 2023-2025, la **recherche basée sur l'IA** (également appelée *recherche générative*) émerge comme un nouveau paradigme. Ici, les LLM (grands modèles linguistiques) et les embeddings neuronaux sont centraux. Les systèmes de recherche basés sur l'IA visent à **comprendre en profondeur les requêtes en langage naturel et à produire des réponses directes** plutôt qu'à pointer vers des sources. Ses principales caractéristiques sont :

- **Grands Modèles Linguistiques (LLM).** Des systèmes comme GPT-4, Claude ou Bard/Gemini de Google sont à la base de la recherche par IA. Ces LLM sont pré-entraînés sur de vastes corpus de texte et peuvent générer des réponses de type humain. Lorsqu'ils sont intégrés à la recherche, ils peuvent analyser une requête au niveau sémantique et *synthétiser* l'information. (Source: [www.techtarget.com](http://www.techtarget.com)) (Source: [searchengineland.com](http://searchengineland.com))
- **Génération Augmentée par la Récupération (RAG).** De nombreux moteurs de recherche basés sur l'IA utilisent une architecture RAG (Source: [searchengineland.com](http://searchengineland.com)). Cela signifie que le système récupère d'abord les documents pertinents (en utilisant la similarité vectorielle ou la correspondance par mots-clés), puis le LLM génère une réponse concise basée sur ce contexte récupéré. L'utilisateur voit la réponse « gratuitement » sans avoir à lire manuellement chaque source. Par exemple, Perplexity.ai et You.com citent tous deux des sources pour leurs réponses – en coulisses, ils récupèrent des passages et demandent au LLM de les réécrire ou de les résumer.
- **Requêtes Contextuelles et Conversationnelles.** La recherche par IA a tendance à maintenir le contexte sur plusieurs échanges (Source: [www.techtarget.com](http://www.techtarget.com)). Un utilisateur peut poser une question de suivi et l'outil d'IA se souvient de la session, contrairement à la recherche traditionnelle qui traite chaque requête indépendamment (Source: [www.techtarget.com](http://www.techtarget.com)). Cela lie le « classement de recherche » à une conversation plutôt qu'à une requête ponctuelle.

Selon les analyses de consultants, **la recherche GenAI et la recherche traditionnelle** diffèrent fondamentalement par leur format de sortie et leur approche (Source: [www.techtarget.com](http://www.techtarget.com)). Le Tableau 1 résume certaines de ces différences :

| Aspect            | Recherche par IA (Générative)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Recherche Traditionnelle                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Format de réponse | Réponses directes et conversationnelles.   Liste classée de liens avec extraits.   <b>Génération de contenu</b>   Peut créer des réponses écrites à la volée.   <b>Compréhension des requêtes</b>   Compréhension avancée du langage naturel (sémantique).   <b>Gestion du contexte</b>   Maintient le contexte sur plusieurs échanges.   <b>Synthèse d'informations</b>   Combine des informations de plusieurs sources en une seule réponse.   <b>Fréquence de mise à jour</b>   Peut extraire des données à jour si connecté (par ex. via un plug-in de navigateur) (Source: <a href="http://www.techtarget.com">www.techtarget.com</a> ).   <b>Personnalisation</b>   Peut personnaliser les réponses en utilisant l'historique d'interaction de l'utilisateur. | Réponses directes et conversationnelles.   Ne récupère que le contenu de pages existantes.   Pas de mémoire ; chaque requête est indépendante.   Principalement basée sur les mots-clés (avec quelques couches sémantiques).   Pas de mémoire ; chaque requête est indépendante.   Affiche des résultats séparés pour chaque source.   Dépend du crawling/indexation web périodique.   Personnalise via le profil/historique de l'utilisateur. |

La **source** de ces différences provient d'analyses sectorielles (Source: [www.techtarget.com](http://www.techtarget.com)). Par exemple, TechTarget note que ChatGPT et les outils de synthèse par IA renvoient des « réponses directes et conversationnelles » au lieu d'une page de résultats de recherche classique (Source: [www.techtarget.com](http://www.techtarget.com)). L'analyse de Search Engine Land souligne également ce « passage de la récupération à la génération » (Source: [searchengineland.com](http://searchengineland.com)) (Source: [searchengineland.com](http://searchengineland.com)) : les systèmes basés sur les LLM « ne classent pas les pages web complètes dans une liste linéaire. Ils récupèrent et synthétisent les informations en fonction de leur pertinence » (Source: [searchengineland.com](http://searchengineland.com)). En bref, la recherche par IA *répond à la question* (via un résumé généré), tandis que la recherche traditionnelle *fournit des indications* sur l'endroit où les réponses pourraient être trouvées.

Ce nouveau paradigme n'est pas purement théorique. Comme le rapporte TechTarget, de nombreux acteurs ont mis en œuvre la recherche générative : des startups (Perplexity, Neeva), ChatGPT d'OpenAI (avec une nouvelle fonctionnalité « Search »), et des entreprises de recherche établies (AI Overviews de Google, Microsoft Bing Chat) (Source: [www.techtarget.com](http://www.techtarget.com)). L'adoption est déjà significative : un rapport SEMrush de 2024 a révélé qu'environ 10 % des utilisateurs américains utilisent la GenAI pour la recherche, avec une estimation de 112,6 millions de personnes aux États-Unis utilisant des outils de recherche par IA en 2024 (projeté à 241 millions d'ici 2027) (Source: [www.techtarget.com](http://www.techtarget.com)). En pratique, les utilisateurs peuvent désormais *poser des questions* en langage naturel (y compris des questions complexes ou à plusieurs volets) et obtiennent souvent une réponse textuelle unique avec des citations. Cela estompe la frontière entre la recherche conventionnelle et les assistants conversationnels basés sur l'IA.

**Pourquoi est-ce important pour le classement ?** Parce que lorsque la recherche par IA donne une seule réponse, nous ne pouvons pas parler de « rang n°1 vs rang n°2 » de la même manière. Au lieu de cela, nous examinons comment elle *sélectionne et pondère* les preuves en coulisses. Une réponse d'IA classe implicitement les informations à inclure et les sources à citer. Dans certains cas, elle peut encore afficher une liste de « sources » (comme Perplexity ou Google Snapshots), qui est effectivement une mini-liste classée. Dans d'autres cas, elle peut ne pas afficher de sources du tout (par exemple, la sortie de ChatGPT simple), ce qui rend l'évaluation encore plus délicate.

En somme, le paradigme de la recherche par IA crée de **nouvelles dimensions** de comparaison :



- **Qualité de la Réponse** : exactitude, exhaustivité, lisibilité de la réponse générée.
- **Utilisation des Sources** : la réputation et la pertinence des sources utilisées ou citées par l'IA.
- **Efficacité** : temps de réponse et facilité pour l'utilisateur.
- **Satisfaction de l'Utilisateur** : UX conversationnelle vs navigation par liens.

Celles-ci diffèrent des métriques de classement traditionnelles et nécessitent une évaluation sur mesure. Les sections suivantes explorent comment mesurer et comparer ces aspects.

## Comparaison des Résultats de Recherche : Méthodes d'Évaluation

Pour comparer la recherche traditionnelle et la recherche par IA, il faut utiliser un mélange de **métriques quantitatives et d'évaluations centrées sur l'utilisateur**. Les approches clés incluent :

1. **Métriques de Recherche d'Information (pour la phase de récupération)**. Nous pouvons appliquer des métriques IR standard au *composant de récupération* de la recherche par IA. Par exemple, dans un système RAG, nous pourrions mesurer combien des documents récupérés par le moteur d'IA auraient été classés dans les meilleurs résultats d'un moteur conventionnel. La Précision@k et le NDCG peuvent évaluer si l'outil d'IA « ouvre le même ensemble de pages pertinentes ». SearchEngineLand suggère que dans la recherche par IA, « la récupération l'emporte sur le classement » – la qualité dépend davantage de la sélection de bonnes informations et de leur compréhension que de la position numérique exacte (Source: [searchengineland.com](https://searchengineland.com)). En pratique, un chercheur pourrait enregistrer les URL ou les passages utilisés par l'IA et les comparer aux meilleurs résultats de Google, en calculant le chevauchement et la corrélation de rang.
2. **Métriques de Qualité des Réponses**. Puisque les outils d'IA génèrent des réponses, nous avons besoin de métriques pour évaluer la qualité des réponses. Cela inclut l'exactitude factuelle (la réponse contient-elle des informations correctes ?), l'exhaustivité et la fluidité. Des métriques issues de tâches de QA (Question-Réponse) ou de résumé (BLEU, ROUGE, BERTScore, scores de factualité) peuvent être utilisées, bien qu'elles nécessitent souvent des réponses de référence. Wang et al. (2024) et d'autres proposent des mesures spécifiques pour la génération augmentée par la récupération, telles que la *vérité* ou la *cohérence des sources*. Le cadre d'évaluation émergent des moteurs de réponse (AEE), par exemple, utilise des métriques pour la précision des citations, le taux d'hallucinations et la compréhensibilité des réponses (Source: [www.emergentmind.com](https://www.emergentmind.com)).
3. **Tests Utilisateur et Comparaison Basée sur les Tâches**. De nombreuses informations proviennent d'études utilisateur. Par exemple, Xu et al. (2023) ont mené une expérience contrôlée où les participants ont répondu à des questions en utilisant soit ChatGPT, soit Google Search. Ils ont mesuré le **temps d'achèvement des tâches**, la satisfaction de l'utilisateur et l'utilité perçue (Source: [www.researchgate.net](https://www.researchgate.net)). De telles études peuvent utiliser des tâches de recherche standardisées (récupération de faits, conseils pour la prise de décision, etc.) et comparer les taux de réussite et les préférences des utilisateurs pour chaque système. Kaiser et al. (2025) ont de même suivi des utilisateurs effectuant des tâches de recherche pratiques et ont mesuré l'exactitude et la vitesse (Source: [www.researchgate.net](https://www.researchgate.net)). Ces études recueillent souvent également des données d'enquête sur la confiance et la satisfaction.
4. **Données de Clics et d'Engagement**. Les données comportementales à grande échelle peuvent être informatives. Par exemple, si les utilisateurs de recherche conventionnelle effectuent des « zéro-clics » (c'est-à-dire que la réponse est satisfaite sur la SERP sans cliquer), ou si le chat IA réduit les clics vers les sites d'éditeurs, cela indique des différences dans les résultats de classement. Les données du marché de la recherche (par ex. les scores ACSI (Source: [www.searchenginejournal.com](https://www.searchenginejournal.com))) peuvent montrer les tendances générales de satisfaction. Les propres recherches de Google (citées dans des articles de l'industrie) suggèrent qu'une partie des réponses de l'IA conduit à davantage de requêtes (certaines sources affirment qu'elle « génère plus de requêtes vers les sites commerciaux » grâce aux AI Spotlights (Source: [www.linkedin.com](https://www.linkedin.com))). Le suivi de métriques comme le temps de consultation, les requêtes de suivi ou la durée totale de la session peut fournir une comparaison indirecte.
5. **Analyse de Requêtes Spécifiques**. Une méthode détaillée consiste à choisir des requêtes représentatives et à comparer directement les sorties. Par exemple, l'étude « 62 requêtes » de Search Engine Land a noté ChatGPT vs Google sur chaque requête avec des métriques personnalisées (Source: [searchengineland.com](https://searchengineland.com)) (Source: [searchengineland.com](https://searchengineland.com)). Chaque requête a été classifiée (informationnelle, locale, etc.) et les réponses ont été évaluées pour leur exactitude et leur utilité. Cela permet

de comprendre quand chaque approche excelle. De telles analyses granulaires révèlent souvent que Google excelle toujours dans la récupération de faits simples et de données locales, tandis que ChatGPT peut surpasser Google sur des tâches de raisonnement en plusieurs étapes ou de synthèse de contenu (au prix de lacunes factuelles potentielles).

6. **Benchmarks Automatisés Combinés.** Pour une comparaison partiellement automatisée, on pourrait utiliser des ensembles de données de QA où les réponses correctes sont connues. Par exemple, soumettre un ensemble de requêtes de trivia ou de QA aux deux systèmes et évaluer la précision des réponses. Des « benchmarks QA de DevM ou Wikipédia » pourraient servir. Certaines initiatives testent également les hallucinations en demandant aux systèmes d'IA de rappeler des faits rarement référencés ; cela peut mettre en évidence des lacunes factuelles.

Le **Tableau 2** résume les études clés et leurs conclusions (chaque étude a utilisé sa propre méthode et ses propres métriques, rendant les comparaisons directes difficiles, mais leur regroupement met en lumière des tendances).

### | Étude (Citation)

- **Enquêtes de satisfaction des utilisateurs :** La collecte des évaluations des utilisateurs (par exemple, « évaluez la pertinence de la réponse ») sur des questions identiques auxquelles chaque système a répondu aide à évaluer la qualité perçue. L'American Customer Satisfaction Index (ACSI) a rapporté les tendances générales de satisfaction de la recherche : à la mi-2024, le score de Google était de 81 (en hausse de 1 %) et celui de Bing de 77 (en hausse de 3 %), reflétant peut-être une réception positive des nouvelles fonctionnalités d'IA (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). De telles enquêtes ne mesurent pas le classement en soi, mais elles indiquent la confiance et le confort des utilisateurs avec les améliorations de l'IA dans la recherche.
- **Métriques d'engagement :** Surveiller le comportement post-assistance (les utilisateurs posent-ils des questions de suivi ?). Si les réponses de l'IA satisfont pleinement les requêtes, nous pourrions observer des requêtes plus longues en une seule session ; sinon, davantage de chaînes de requêtes. Xu et al. ont constaté que les réponses de ChatGPT conduisaient souvent à moins de recherches nécessaires de la part des étudiants, ce qui implique une réponse plus autonome (Source: [www.researchgate.net](http://www.researchgate.net)).

En pratique, la comparaison des classements de recherche utilisera probablement une *évaluation multi-métriques*. Il faut prendre en compte la pertinence des résultats (RI traditionnel), l'exactitude des réponses (métriques QA) et les résultats centrés sur l'utilisateur (temps, satisfaction). Une étude comparative complète des deux types de recherche combinera ces approches plutôt que de s'appuyer sur une seule métrique.

## Analyse des données et résultats empiriques

Les preuves empiriques sur la recherche traditionnelle par rapport à la recherche par IA s'accumulent rapidement. Nous mettons ici en évidence les données clés, les statistiques et les résultats d'études tirés de la littérature.

### Statistiques d'utilisation et d'adoption

- **Utilisation de la recherche par IA :** Bien qu'encore naissante, l'utilisation de la recherche par IA croît rapidement. Un rapport Statista/SEMrush a révélé que d'ici début 2025, environ **1 utilisateur d'Internet sur 10 aux États-Unis** utilisait régulièrement des outils d'IA générative pour la recherche (Source: [www.techtarget.com](http://www.techtarget.com)). Environ 112,6 millions d'Américains ont utilisé des outils de recherche basés sur l'IA en 2024, avec une projection à 241 millions d'ici 2027 (Source: [www.techtarget.com](http://www.techtarget.com)). À la mi-2025, OpenAI a signalé **plus de 400 millions d'utilisateurs actifs hebdomadaires** sur ChatGPT (le double des 200 millions rapportés à la mi-2024) (Source: [www.investing.com](http://www.investing.com)). Ces chiffres indiquent une pénétration grand public, bien que le nombre total de requêtes de recherche sur Google/Bing dépasse encore largement les requêtes d'IA (Google gère des *centaines de milliards* de requêtes par jour).
- **Trafic des moteurs de recherche :** Un contrepoint est que la recherche traditionnelle domine toujours le trafic global. Le rapport Tipsheet (juillet 2025) a noté que parmi les « premiers adoptants » de l'IA, 40 % de leur trafic de recherche sur ordinateur de bureau était dirigé vers les outils LLM (contre 24 % à la mi-2024), tandis que la part des *premiers adoptants* vers la recherche traditionnelle est tombée de 76 % à 61 % (Source: [tipsheet.ai](http://tipsheet.ai)). Cependant, Google a contesté que son volume de recherche traditionnel continue de croître et que ses aperçus d'IA génèrent toujours des requêtes vers des sites web. En

pratique, **Google reste la valeur par défaut pour la plupart des requêtes** ; les leaders de l'IA sont encore des niches concurrentielles. SearchEngineJournal note que l'intégration de l'IA par Google (Overviews) pourrait en fait augmenter le trafic web en connectant les utilisateurs au contenu (Source: [www.linkedin.com](http://www.linkedin.com)).

- **Satisfaction des utilisateurs** : De vastes enquêtes montrent que la satisfaction de la recherche est élevée, et même en hausse avec les fonctionnalités d'IA. L'étude ACSI 2024 a révélé que le score de satisfaction de Google était de 81 (« excellent ») et que Bing/Yahoo Japonais atteignaient des sommets (77, 76) – des gains attribués aux nouvelles capacités d'IA (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). Ainsi, les utilisateurs semblent apprécier globalement la recherche améliorée par l'IA. Notamment, **plus de la moitié des utilisateurs de Google rencontrent déjà des résumés d'IA** sur les pages de résultats : Pew (2023) a constaté que 58 % avaient vu un résumé généré par l'IA lors d'une recherche (Source: [www.techtarget.com](http://www.techtarget.com)).
- **Études spécifiques à un domaine** : Dans les contextes éducatifs, les étudiants ont adopté les outils de recherche par IA mais n'ont pas abandonné Google (Source: [link.springer.com](http://link.springer.com)). Une enquête sur un campus a rapporté que bien que les étudiants utilisent ChatGPT pour l'apprentissage, ils s'appuient toujours sur les moteurs de recherche pour la collecte d'informations (Source: [link.springer.com](http://link.springer.com)). Les outils sont considérés comme complémentaires – par exemple, les chercheurs peuvent utiliser Google pour trouver des sources mais utiliser ChatGPT pour des explications rapides\*\* (Source: [link.springer.com](http://link.springer.com)).
- **Résultats de recherche** : Les données de Seo-bank suggèrent que certaines catégories de requêtes se tournent vers l'IA : par exemple, les requêtes de création de contenu, les analyses techniques ou le brainstorming créatif tendent vers ChatGPT (Source: [searchengineland.com](http://searchengineland.com)). Les requêtes locales ou factuelles penchent vers Google/Bing. Les tests anecdotiques de Dan Taylor ont montré que ChatGPT avait du mal avec les résultats d'entreprises locales et les sources diverses, tirant souvent d'un seul domaine (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)) (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). Il a également noté que ChatGPT citait parfois des pages en dehors du classement typique (par exemple, pas dans les 100 premiers résultats de Bing) (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)), ce qui implique que la recherche par IA s'appuie sur un index plus large basé sur la compréhension de la pertinence plutôt que sur un classement purement basé sur les clics.

## Données de performance comparative

- **Efficacité des tâches** : Plusieurs études montrent des gains de temps avec la recherche par IA. Xu et al. rapportent que les utilisateurs de ChatGPT ont passé en moyenne **40 % moins de temps** sur les tâches de recherche avec un résultat égal (Source: [www.researchgate.net](http://www.researchgate.net)). De même, les utilisateurs de ChatGPT étaient « plus rapides » et trouvaient plus souvent les bonnes réponses dans l'étude de tâches de Kaiser et al. (Source: [www.researchgate.net](http://www.researchgate.net)). Ceci est probablement dû au fait que les réponses de l'IA éliminent le besoin de cliquer et de lire plusieurs pages. Cependant, plus rapide n'est pas toujours mieux : si la réponse de l'IA est incomplète ou erronée, la rapidité signifie une complétion mal orientée.
- **Précision et exactitude** : L'exactitude objective est mitigée. L'analyse « 62 requêtes » de SearchEngineLand a révélé que Google avait l'avantage sur les requêtes factuelles, donnant des scores de précision légèrement plus élevés sur les questions informationnelles (Source: [searchengineland.com](http://searchengineland.com)). ChatGPT a bien performé mais a manqué de détails. D'autre part, ChatGPT était plus efficace pour les tâches de contenu ouvertes (cadres d'écriture, invites d'analyse) que Google ne peut tout simplement pas faire (Source: [searchengineland.com](http://searchengineland.com)). Aucun grand benchmark public ne compare directement la précision des réponses entre le chat IA (en particulier les LLM hors ligne) et la recherche, mais des preuves émergentes suggèrent que ChatGPT peut produire des réponses très fluides qui contiennent parfois des erreurs (hallucinations) (Source: [www.emergentmind.com](http://www.emergentmind.com)) (Source: [www.researchgate.net](http://www.researchgate.net)).
- **Préférences des utilisateurs (subjectives)** : Dans les enquêtes, les préférences *subjectives* des utilisateurs favorisent souvent la recherche traditionnelle. Kaiser et al. ont constaté que les participants préféraient toujours Google dans l'ensemble, malgré le gain de temps offert par ChatGPT (Source: [www.researchgate.net](http://www.researchgate.net)). Xu et al. ont rapporté que **les utilisateurs estimaient que les réponses de ChatGPT étaient de meilleure qualité**, mais que leur *niveau de confiance* en ChatGPT par rapport à Google était similaire (Source: [www.researchgate.net](http://www.researchgate.net)). En termes plus simples, les gens ont trouvé les réponses de l'IA satisfaisantes mais sont restés aussi confiants/incertains qu'avec Google. Des articles de l'industrie indépendants font écho à cette ambivalence : de nombreux utilisateurs apprécient la commodité des résumés d'IA mais se méfient des erreurs, vérifiant souvent avec un moteur de recherche.



- **Différences d'engagement** : L'inclusion de réponses d'IA modifie les schémas de clics. Si une réponse d'IA satisfait, les utilisateurs cliquent moins ou plus tard, ce qui nuit au trafic du site (le phénomène du « zéro-clic » (Source: [tipsheet.ai](https://tipsheet.ai)). Certains analystes SEO avertissent que les requêtes factuelles simples n'envoient plus les utilisateurs via les canaux traditionnels. Comme le note l'article de Tipsheet, même si la satisfaction de la recherche est élevée, les réponses générées par l'IA risquent d'isoler les utilisateurs des sources de contenu, ce qui dérouté les annonceurs et les éditeurs (Source: [tipsheet.ai](https://tipsheet.ai)). La réponse de Google (via les RP) affirme que les aperçus d'IA génèrent « plus de requêtes qui connectent les consommateurs aux entreprises » (Source: [tipsheet.ai](https://tipsheet.ai)), mais les données neutres à ce sujet sont rares. Nous savons d'après les journaux d'utilisateurs que les « requêtes de navigation » traditionnelles (par exemple, se rendre sur un site connu) sont exclues de ces études ; ainsi, lorsqu'une réponse d'IA apparaît, il s'agit par définition d'un scénario de « besoin d'information ».
- **Risques de qualité** : Un point de données critique est les hallucinations de l'IA. Kuhlata et al. ont mesuré quantitativement les défauts des réponses de l'IA : ils ont trouvé des taux extrêmement élevés d'informations inexactes ou invérifiables dans les réponses. Par exemple, leur évaluation sur 1287 sources candidates a révélé que ChatGPT n'identifiait que 7 études directement pertinentes sur 1287 par rapport à une revue systématique humaine, contre 19 sur 48 pour Bing Chat (Source: [www.researchgate.net](https://www.researchgate.net)). Cela suggère que la fonction de recherche de ChatGPT n'avait qu'environ 0,5 % de résultats pertinents, tandis que la recherche générative de Bing en avait 40 % dans cet exemple de littérature médicale (Source: [www.researchgate.net](https://www.researchgate.net)). Bien qu'il s'agisse d'une étude de domaine, elle souligne que l'utilisation naïve de la recherche LLM peut manquer de manière significative des faits pertinents. Leur analyse a attribué à ChatGPT un grand nombre de réponses de grade « F » en matière de qualité de citation. Ces résultats empiriques soulignent que **l'exactitude factuelle n'est pas garantie** dans les résultats de recherche de l'IA.

## Exemples basés sur les données

- **Requêtes spécifiques à un sujet** : Par exemple, en posant la question « Quels sont les symptômes de la maladie de La Peyronie ? », une étude a comparé ChatGPT à une recherche médicale humaine (Source: [www.researchgate.net](https://www.researchgate.net)). La « recherche » de ChatGPT n'a trouvé que 0,5 % d'éléments pertinents, tandis qu'une requête humaine utilisant les nouvelles fonctionnalités de Bing Chat en a trouvé 40 %. Les réponses de ChatGPT ont été très mal notées pour les preuves. Cela montre que la recherche par IA peut sérieusement sous-performer sur des questions-réponses spécialisées nécessitant des sources précises.
- **Recherche locale** : L'analyse de Dan Taylor comparant ChatGPT à Google sur des requêtes comme « stations-service à proximité » ou « magasins locaux » a révélé les lacunes de ChatGPT. Il n'interrogeait souvent pas une base de données cartographique en interne, donnant des informations génériques ou manquant entièrement des entreprises (Source: [www.searchenginejournal.com](https://www.searchenginejournal.com)) (Source: [www.searchenginejournal.com](https://www.searchenginejournal.com)). En revanche, Google fournissait une interface cartographique ou des liens Yelp. Ceci est attendu : ChatGPT (en 2024) n'intègre pas de bases de données GPS/d'entreprises en temps réel, alors que Google/Bing les ont intégrées.
- **Tâches créatives et analytiques** : Les catégories de requêtes comme « analyse des lacunes de contenu » ont montré la force de ChatGPT. Dans l'étude SEL, des tâches telles que « comparer notre site à nos concurrents » ou « suggérer des sujets de blog » dépassaient le cadre traditionnel de Google, mais ChatGPT a fourni une orientation utile (Source: [searchengineland.com](https://searchengineland.com)). Autre exemple : ChatGPT est souvent utilisé pour générer des idées ou esquisser un article, des tâches pour lesquelles aucun résultat de recherche classé ne suffit directement. Cet avantage de cas d'utilisation non noté n'est généralement pas pris en compte dans l'évaluation traditionnelle.
- **Cas d'utilisation - Éducation** : L'étude TechTrends (2025) a examiné comment les étudiants utilisent la recherche par rapport à l'IA (Source: [link.springer.com](https://link.springer.com)) (Source: [link.springer.com](https://link.springer.com)). Elle a constaté que ChatGPT était populaire, mais ne remplaçait pas Google. Les étudiants utilisaient Google pour la recherche de fond (trouver des articles/sites web) et ChatGPT pour l'explication ou la rédaction. Ils ont également souvent mal évalué leurs propres compétences en IA (« surévaluation de la maîtrise »). Pour la comparaison des classements, cela suggère que *les outils sont complémentaires* : on pourrait comparer la capacité de chacun à récupérer du matériel d'étude par rapport à la capacité de chacun à l'expliquer, ce qui sont des tâches différentes.

- **Satisfaction au fil du temps** : Les données de l'ACSI peuvent être considérées comme une étude de cas. Malgré les craintes que l'IA ne puisse dérouter les utilisateurs, les données ont montré une satisfaction *stable ou en amélioration* à mesure que les moteurs de recherche ajoutent des fonctionnalités d'IA (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). Cela implique que les utilisateurs estiment que leurs besoins sont satisfaits, bien que l'étude n'isole pas le classement par rapport au type de réponse. Il est possible que les améliorations de l'IA (par exemple, de meilleurs extraits, des résumés) augmentent effectivement la qualité perçue de la recherche.

En résumé, les données quantitatives brossent un tableau nuancé. La recherche par IA est largement utilisée et peut accélérer la recherche d'informations, mais elle introduit des risques d'exactitude. La recherche traditionnelle reste fiable pour les requêtes factuelles et locales. Les comparaisons empiriques (tâches utilisateur, expériences contrôlées, enquêtes de satisfaction) montrent des **compromis** : vitesse et qualité de la prose avec l'IA, contre exhaustivité, familiarité et confiance avec les systèmes traditionnels.

## Études de cas et exemples concrets

Pour ancrer la comparaison dans des contextes réels, considérons plusieurs scénarios de cas et exemples pratiques :

### Recherche en santé et scientifique

Dans les domaines spécialisés, la précision des sources est primordiale. Par exemple, une étude publiée a comparé la recherche par IA (ChatGPT, Bing Chat) avec les recherches traditionnelles sur PubMed pour une revue de la littérature médicale (Source: [www.researchgate.net](http://www.researchgate.net)). ChatGPT n'a identifié pratiquement *aucune* publication pertinente (0,5 % de pertinence), tandis que la récupération par IA de Bing Chat en a trouvé environ 40 % (19 sur 48) contre une référence humaine de 24 (Source: [www.researchgate.net](http://www.researchgate.net)). De plus, les rédactions de réponses de ChatGPT ont été majoritairement notées F (90 % C/D/F sur une échelle de qualité). Les critiques concluent que l'utilisation de ChatGPT comme outil de recherche n'est « pas encore précise ou réalisable » (Source: [www.researchgate.net](http://www.researchgate.net)). Cela souligne que **pour les requêtes basées sur des preuves**, la recherche traditionnelle (ou les bases de données spécialisées comme PubMed) est toujours supérieure. L'IA générative peut halluciner ou manquer des citations, comme l'ont également noté Kuhlata et al. (Source: [www.emergentmind.com](http://www.emergentmind.com)).

### Recherche juridique et de conformité

Les professionnels du droit s'appuient souvent sur la recherche pour trouver des précédents et des statuts. Le chat génératif est exploré ici, mais des tests récents indiquent la prudence : ChatGPT pourrait omettre des cas clés ou citer des lois de manière erronée. Un exemple tiré d'un hackathon de cabinet d'avocats a montré ChatGPT donnant des conseils juridiques plausibles mais obsolètes qui nécessitaient une correction humaine. Cela correspond au schéma général : l'IA fournit des résumés fluides mais nécessite une validation par un expert.

### Analyse commerciale/financière

#### Analyse Commerciale/Financière

Certaines entreprises expérimentent l'IA basée sur le RAG pour analyser les rapports financiers. Par exemple, une entreprise pourrait utiliser une base de connaissances interne et un LLM pour répondre à des requêtes comme "Quelle a été notre croissance des ventes au T3 ?". Dans ce cas, le "classement" de la recherche IA implique de faire correspondre les documents de l'entreprise et de produire une réponse. Les avantages pratiques incluent un résumé rapide de documents volumineux. Cependant, si les données financières sous-jacentes changeaient (par exemple, en raison d'un dépôt tardif), la coupure de connaissances statique d'un LLM pourrait induire en erreur à moins d'être continuellement mise à jour via une intégration. La recherche traditionnelle (avec des données à jour) pourrait éviter ce problème.

### Commerce et Entreprises Locales

ChatGPT (fin 2024) rencontrait des difficultés avec les requêtes spécifiques à un lieu. Lors des tests de Dan Taylor, demander des restaurants à proximité ou les heures d'ouverture des magasins donnait souvent des descriptions génériques plutôt que des résultats locaux réels (Source: [www.searchenginejournal.com](http://www.searchenginejournal.com)). La recherche locale traditionnelle de Google classe les entreprises

par proximité, popularité et avis, ce que ChatGPT (sans données cartographiques en temps réel) ne peut pas reproduire. Ainsi, les consommateurs continuent de s'appuyer sur Google Maps/Bing Maps pour les requêtes locales tout en utilisant l'IA pour des conseils généraux (par exemple, "meilleur moment pour planter des roses").

## Éducation et Monde Académique

L'étude "préférences des étudiants" de TechTrends (juin 2025) (Source: [link.springer.com](https://link.springer.com)) (Source: [link.springer.com](https://link.springer.com)) montre que les étudiants utilisent à la fois les chatbots IA et la recherche. Les étudiants peuvent utiliser Google Scholar ou la recherche générale pour trouver des manuels et des références académiques, mais ensuite demander à ChatGPT d'**expliquer des concepts** en termes plus simples. Par exemple, un étudiant pourrait chercher sur Google "équation de Black-Scholes PDF" et cliquer sur un lien vers un manuel, puis demander à ChatGPT "Veuillez expliquer l'équation de Black-Scholes en termes simples." En substance, Google fournit les ressources (classement traditionnel à l'œuvre), et ChatGPT fournit la compréhension. Les étudiants ont signalé une **utilisation stratégique**, et non un remplacement complet (Source: [link.springer.com](https://link.springer.com)) (Source: [link.springer.com](https://link.springer.com)). Cette division du travail illustre que les comparaisons doivent tenir compte du *type de tâche* : tâches de récupération (trouver l'information) vs tâches de connaissance (comprendre/formuler).

## Développement Logiciel

Les développeurs utilisent souvent la recherche pour obtenir de l'aide en matière de codage. La recherche traditionnelle mène à des forums de questions-réponses (StackOverflow) classés par pertinence et votes. Les nouveaux assistants de code IA (GitHub Copilot Chat, ChatGPT avec interpréteur de code) peuvent répondre directement aux questions de programmation. L'analyse empirique par les équipes de DevGPT suggère que les développeurs obtiennent des réponses plus rapides avec l'IA pour les tâches simples, mais que la solution IA présente parfois des bugs subtils. Dans un cas, ChatGPT a recommandé une approche de codage syntaxiquement correcte mais sémantiquement erronée en raison de changements d'API, un exemple d'hallucination dans un domaine technique. La recherche classée traditionnellement aurait fait apparaître la documentation officielle, plus fiable mais plus lente à analyser.

## Assistants Personnalisés et Vocaux

Bien qu'il ne s'agisse pas de "recherche" au sens web pur, des assistants comme Siri ou Alexa utilisent un mélange d'IA traditionnelle (déclenchement d'API web) et générative. Les comparaisons dans cet espace sont rares, mais des preuves anecdotiques suggèrent que les assistants vocaux génératifs (par exemple, Alexa utilisant AlexaGPT) peuvent avoir des dialogues plus naturels, tandis que les assistants classiques s'appuient sur des réponses prédéfinies ou des requêtes web.

## Gouvernement et Politiques Publiques

Les gouvernements utilisent l'analyse de recherche pour évaluer l'intérêt public. Lorsque les moteurs de recherche intègrent l'IA, cela complique ce flux de données. Par exemple, si les citoyens posent de plus en plus de questions aux chatbots IA sur les sites web gouvernementaux au lieu de chercher sur Google, les journaux de recherche traditionnels (les problèmes que les gens recherchent sur Google) peuvent sous-estimer les véritables préoccupations. Des rapports préliminaires indiquent que certaines enquêtes politiques sont mises à jour pour inclure des métriques de recherche IA. Cependant, des études formelles sont en attente.

## Impact Réel

Bien que de nombreuses comparaisons soient expérimentales ou à petite échelle, certains impacts généraux sont observables. Les spécialistes du marketing parlent déjà de "SEO IA/zéro-clic" : optimiser le contenu pour les réponses IA au lieu des classements par liens bleus. Les modèles de revenus de recherche s'adaptent également : les moteurs de recherche envisagent de nouveaux formats publicitaires dans les contextes d'IA. Par exemple, la décision audacieuse de Google de servir des extraits signifie que les sites web pourraient perdre du trafic ; une étude estime que les taux de clics publicitaires pourraient diminuer considérablement à mesure que les réponses s'améliorent.

Dans ces cas, la **question du classement** se traduit par "quelle information l'utilisateur voit/utilise-t-il finalement, et dans quel ordre ?" Dans la recherche traditionnelle, l'utilisateur choisit en haut de la liste classée. Dans la recherche IA, l'utilisateur reçoit une **réponse unifiée unique** (souvent au "rang 0" au-dessus de toute liste). Certaines interfaces IA affichent également un carrousel

limité de liens cités (par exemple, Bard/Gemini affiche des sources numérotées en bas, Bing Chat liste les sources sur le côté). Ceux-ci peuvent être considérés comme une **mini-liste classée** au sein de l'interface IA. Mais dans tous les cas, la *présentation* diffère, nécessitant une comparaison adaptée.

## Discussion des Implications et Orientations Futures

La convergence de la recherche et de l'IA générative a des implications profondes pour la technologie, les affaires et la société. Nous discutons ci-dessous des impacts clés et des possibilités futures.

### Implications pour les Moteurs de Recherche et le SEO

- **Passage du "SEO" à l'"AEO" (Answer Engine Optimization).** Les créateurs de contenu optimisaient historiquement pour le classement des pages. Avec les réponses IA, l'accent pourrait se déplacer vers l'**optimisation des réponses** : inclure des résumés clairs et factuels dans le contenu afin que les LLM les fassent remonter. Par exemple, les données structurées et le balisage de schéma (déjà utilisés pour les extraits optimisés) deviennent encore plus critiques (Source: [searchengineland.com](https://searchengineland.com)). Cependant, une véritable "présence" dans les réponses IA nécessite probablement une autorité reconnue et une clarté plutôt qu'une densité de mots-clés (Source: [searchengineland.com](https://searchengineland.com)) (Source: [www.techtarget.com](https://www.techtarget.com)).
- **Force de la Marque et Confiance.** Comme le note SearchEngineLand, être une marque forte et faisant autorité "est un prérequis" pour apparaître dans les résultats basés sur l'IA (Source: [searchengineland.com](https://searchengineland.com)). Google a déclaré que seules les sources les plus crédibles seraient affichées par les Aperçus IA. Cela favorise les acteurs établis (Wikipédia, grands médias, organisations bien connues) qui sont déjà bien classés dans les liens. Les petits sites pourraient avoir du mal à être cités. Ainsi, les stratégies d'optimisation de recherche devront mettre l'accent sur le *développement de l'autorité* et la *connaissance structurée*.
- **Recherches Zéro-Clic et Trafic.** Avec des réponses directes, moins d'utilisateurs cliquent sur les sites, ce qui peut réduire le trafic web. Une étude de SearchEngineLand avertit les éditeurs de contenu de s'adapter à ce monde "zéro-clic" (Source: [tipsheet.ai](https://tipsheet.ai)). Les entreprises pourraient avoir besoin de fournir des réponses structurées aux assistants vocaux/de recherche ou d'accepter une perte de visibilité. Alternativement, de nouveaux modèles de monétisation (comme la licence de contenu à l'IA) pourraient émerger. Les annonceurs pourraient avoir besoin d'acheter des emplacements dans les widgets de réponse IA plutôt que des publicités classiques.
- **Importance Continue du Classement.** Même à l'ère de l'IA, le classement compte. La qualité d'une réponse IA dépend de l'**étape de récupération** (quelles informations sont trouvées). Si le récupérateur d'un modèle IA utilise des signaux de classement traditionnels (par exemple, un index Bing sous-jacent), ce classement influence toujours la qualité de la réponse. De plus, les systèmes IA pourraient présenter plusieurs réponses possibles ou permettre à un utilisateur d'"explorer plus de résultats", auquel cas ils listeront des sources ou des lectures complémentaires, revenant ainsi à une liste classée pour approfondir.

### Implications pour les Utilisateurs et la Société

- **Accès à l'Information et Littératie.** La recherche IA abaisse les barrières pour les utilisateurs occasionnels afin d'obtenir des réponses, démocratisant potentiellement la connaissance. Cependant, elle soulève également des préoccupations : si les utilisateurs acceptent les réponses sans vérifier, la désinformation peut se propager. La pensée critique (par exemple, la vérification croisée des sources) devient plus cruciale. L'étude TechTrends a révélé que les étudiants surestiment souvent leur maîtrise des outils d'IA (Source: [link.springer.com](https://link.springer.com)). Cela suggère un besoin d'éducation sur les forces/limites de la recherche IA (par exemple, inciter à citer les sources, vérifier les faits).
- **Biais et Équité.** Les systèmes d'IA peuvent involontairement renforcer les biais. Par exemple, si une réponse IA cite principalement des sources occidentales, elle biaise l'exposition à l'information. Le classement de recherche traditionnel a également des problèmes de biais (avec des algorithmes favorisant certaines langues ou des sites à forte autorité de domaine). Comparer les résultats entre les types de recherche aide à identifier les biais : on pourrait tester si différentes données démographiques obtiennent des réponses différentes. Les chercheurs devront concevoir des métriques d'équité pour les réponses IA (garantissant que les points de vue minoritaires ne sont pas supprimés).

- **Réglementation et Transparence.** Les gouvernements étudient déjà les effets de l'IA. Le "dilemme de la citation" (EmergentMind) met en évidence le défi : les utilisateurs pourraient ne pas savoir *pourquoi* une réponse a été donnée ou quelles sources ont été prises en compte (Source: [www.emergentmind.com](http://www.emergentmind.com)). Les réglementations pourraient exiger que les systèmes de recherche IA divulguent clairement la provenance des sources. La recherche traditionnelle a un processus relativement transparent (cliquer pour la source), tandis que les "boîtes noires" de l'IA pourraient être tenues plus responsables. La loi européenne sur l'IA et la loi américaine JUDIC pourraient exiger une telle transparence.
- **Avenir des Professionnels de la Recherche.** Les spécialistes du SEO et les spécialistes du marketing de contenu doivent s'adapter. Certains prédisent une demande pour des "formateurs IA" qui alimentent les LLM en contextes ou organisent des corpus pour les systèmes de recherche verticale. D'autre part, l'expertise en SEO traditionnel (netlinking, optimisation on-page) pourrait diminuer à mesure que les réponses génératives prennent le dessus. Cependant, étant donné les preuves émergentes que les utilisateurs continuent de s'appuyer sur les liens et de leur faire confiance (et préfèrent Google), les tactiques traditionnelles ne disparaîtront pas du jour au lendemain.

## Orientations Futures

- **Interfaces Hybrides.** De nombreuses plateformes de recherche mélangeront probablement les réponses IA avec les résultats classés. Le SGE de Google affiche déjà une boîte "Aperçus IA" au-dessus des résultats organiques. Les futures interfaces pourraient permettre de basculer entre le "mode réponse IA" et le "mode liste", ou de présenter des dialogues multi-tours aux côtés de listes de liens optionnelles. La comparaison des performances impliquera alors des études d'interface : quel format les utilisateurs préfèrent-ils pour quelles tâches ?
- **Critères d'Évaluation Avancés.** La recherche développera des critères spécifiquement pour l'évaluation de la recherche générative. Par exemple, l'équipe EmergentMind publie un critère d'**Évaluation des Moteurs de Réponse (AEE)** (Source: [www.emergentmind.com](http://www.emergentmind.com)). Il pourrait y avoir de nouveaux défis de type TREC pour la "récupération d'informations conversationnelle" où les juges évaluent les dialogues de réponse IA, et pas seulement les listes.
- **Intégration de Données à Jour.** Une lacune des LLM actuels est la coupure de connaissances. Les outils de recherche IA y remédient en se connectant aux données web en direct (par exemple, le mode de navigation de Bing Chat, l'index de Google). Les futures comparaisons devront prendre en compte les *réponses de recherche en temps réel* vs les réponses LLM statiques. Nous pourrions voir des comparaisons comme "LLM avec accès internet" vs "recherche traditionnelle".
- **Spécialisation.** Le ChatGPT générique pourrait être surpassé par une recherche IA spécifique à un domaine. Les exemples incluent WolframAlpha (requêtes mathématiques), les bots de recherche juridique, les IA médicales. Les recherches futures devraient comparer les systèmes de recherche IA spécialisés avec leurs homologues traditionnels (par exemple, LexisNexis vs un assistant juridique IA).
- **Changement de Comportement des Utilisateurs.** Le moyen des requêtes passe des mots-clés aux invites en langage naturel. L'analyse de recherche pourrait devoir évoluer du suivi de termes de 1 à 3 mots vers des modèles de questions complexes. Pour les entreprises d'analyse, la comparaison de la recherche traditionnelle vs IA impliquera l'analyse de ces nouveaux journaux de requêtes. De plus, à mesure que la recherche IA devient principalement vocale/par chat, la mesure du succès pourrait davantage reposer sur la *satisfaction conversationnelle* que sur les métriques de clics.
- **Changements de l'Écosystème Commercial.** Les entreprises pourraient commencer à indexer pour le contexte IA plutôt que seulement pour le SEO. Les outils de création de contenu utilisent déjà les LLM pour optimiser les publications pour les réponses IA. Le débat sur la stratégie SEO vs contenu ("SEO vs GEO") s'intensifiera. On pourrait prévoir des certifications ou des labels de qualité pour le contenu qui passe les vérifications d'exactitude de l'IA (pour s'assurer qu'il est prêt à répondre).

Enfin, ces développements ouvrent de nombreuses questions de recherche : Comment des mesures comme le NDCG doivent-elles changer pour les réponses de rang 0 ? Comment définir la pertinence lorsqu'une réponse pourrait ne pas citer toutes les sources ? L'IA elle-même peut-elle être utilisée pour *évaluer* les réponses d'autres IA (une forme d'examen contradictoire) ? Le domaine de la **méta-évaluation de la recherche** se développera.

## Conclusion



Comparer les classements des résultats de recherche traditionnels et IA nécessite une **approche multifacette**. La recherche traditionnelle, avec ses listes de documents classées, est évaluée par des métriques de RI établies et dispose de décennies de données empiriques étayant ses forces (pertinence, fraîcheur, couverture). La recherche basée sur l'IA, bien que plus récente, apporte des changements révolutionnaires : des réponses directes en langage naturel, la synthèse et l'interaction conversationnelle. Celles-ci exigent de nouveaux critères d'évaluation axés sur la qualité des réponses, l'exactitude factuelle et l'expérience utilisateur.

Dans ce rapport, nous avons fourni une **comparaison détaillée** :

- *Différences Techniques* : La recherche traditionnelle classe les documents statiques à l'aide de signaux de liens et de mots-clés, tandis que la recherche IA utilise des LLM pour interpréter les requêtes et générer des réponses synthétisées (Source: [searchengineland.com](https://searchengineland.com)) (Source: [www.techtarget.com](https://www.techtarget.com)). Les systèmes IA peuvent maintenir le contexte et combiner plusieurs sources, modifiant fondamentalement la notion de "classement".
- *Méthodes d'Évaluation* : Nous avons discuté de la manière d'appliquer les métriques de RI à la partie récupération de la recherche IA, et de la manière de les compléter avec des métriques de QA et d'études utilisateur pour les réponses générées. De nouveaux critères (comme l'AEE) sont en cours de développement à cette fin (Source: [www.emergentmind.com](https://www.emergentmind.com)) (Source: [www.emergentmind.com](https://www.emergentmind.com)).
- *Résultats Empiriques* : Des études contrôlées montrent des compromis : la recherche IA permet souvent une exécution plus rapide des tâches, mais les utilisateurs préfèrent toujours la recherche traditionnelle pour la confiance et la familiarité (Source: [www.researchgate.net](https://www.researchgate.net)) (Source: [www.researchgate.net](https://www.researchgate.net)). Sur les bases de données factuelles (par exemple, la recherche médicale), la recherche traditionnelle surpasse en raison des hallucinations de l'IA (Source: [www.emergentmind.com](https://www.emergentmind.com)). Les statistiques d'adoption révèlent un rôle en croissance rapide mais encore plus petit pour la recherche IA (de l'ordre de dizaines de millions d'utilisateurs) par rapport aux requêtes traditionnelles (Source: [www.techtarget.com](https://www.techtarget.com)) (Source: [www.investing.com](https://www.investing.com)).
- *Spécificités des Cas d'Usage* : Dans des domaines comme l'éducation, les étudiants complètent mais ne remplacent pas Google par ChatGPT (Source: [link.springer.com](https://link.springer.com)). Pour les requêtes locales ou sensibles au temps, Google/Bing restent irremplaçables car ChatGPT manque de données en temps réel intégrées (Source: [www.searchenginejournal.com](https://www.searchenginejournal.com)). Pour les tâches créatives ou analytiques, l'IA a un avantage que Google ne peut égaler (conduisant à de nouvelles applications en marketing de contenu et en recherche) (Source: [searchengineland.com](https://searchengineland.com)) (Source: [www.researchgate.net](https://www.researchgate.net)).

Nous soulignons qu'**aucune approche n'est catégoriquement "meilleure"** sur toutes les métriques. Au lieu de cela, chacune excelle dans certains scénarios. La clé est d'utiliser des stratégies d'évaluation *complémentaires* :

- Utiliser des mesures de RI traditionnelles (précision, rappel, corrélation de rang) *et* de nouvelles mesures de qualité des réponses.
- Mener des études utilisateur mesurant à la fois les résultats objectifs (exactitude, temps) et la satisfaction subjective.
- Surveiller l'engagement et les données de satisfaction réels au fil du temps.
- Inclure des études de cas et des critères spécifiques à un domaine pour capturer les cas limites (comme la santé ou la recherche locale).

À mesure que la recherche IA continue d'évoluer, les comparaisons doivent s'adapter. Les travaux futurs intégreront probablement des modèles hybrides (recherche + génération), exigeant des métriques combinées. Le "jeu" de l'optimisation de la recherche passe de la course au rang n°1 à la **conquête de la présence** dans les réponses générées par l'IA (Source: [searchengineland.com](https://searchengineland.com)).

En conclusion, la comparaison des résultats de recherche traditionnels et de ceux de l'IA est une frontière de recherche en constante évolution. Les chercheurs d'information modernes évoluent dans un **écosystème hybride** – cliquant parfois sur des liens classés, parfois lisant des réponses de chat. Une compréhension approfondie des deux est essentielle pour les technologues, les stratèges de contenu et les utilisateurs. Nous avons passé en revue l'historique, les capacités actuelles, les techniques d'évaluation et les implications, avec des citations complètes tout au long. Le paysage est encore en évolution, et une recherche empirique continue sera vitale pour quantifier pleinement la valeur relative et la trajectoire future de ces deux paradigmes de recherche.

**Tableau 1.** Principales différences entre les moteurs de recherche traditionnels et la recherche générative basée sur l'IA (Source: [www.techtarget.com](https://www.techtarget.com)) (Source: [searchengineland.com](https://searchengineland.com)).

| ASPECT                      | RECHERCHE IA (GÉNÉRATIVE)                                               | RECHERCHE TRADITIONNELLE                                                      |
|-----------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Format de réponse           | Réponses directes et conversationnelles.                                | Liste classée de liens avec extraits.                                         |
| Génération de contenu       | Peut créer du nouveau contenu à la volée.                               | Ne récupère que les informations existantes.                                  |
| Compréhension de la requête | Compréhension avancée du langage naturel.                               | Principalement basée sur la correspondance de mots-clés (avec un peu de PNL). |
| Maintien du contexte        | Maintient le contexte à travers les conversations (multi-tours).        | Contexte limité ; chaque requête est traitée indépendamment.                  |
| Synthèse d'informations     | Combine les informations de plusieurs sources en une réponse cohérente. | Présente des résultats séparés pour chaque source.                            |
| Fréquence de mise à jour    | Peut intégrer des informations très récentes (si connectée).            | Dépend des cycles périodiques d'exploration/indexation.                       |
| Personnalisation            | S'adapte à l'historique des conversations et aux données utilisateur.   | Personnalisée uniquement via le profil utilisateur/l'historique de recherche. |

Tableau 2. Résumé des études comparatives sur les performances de la recherche IA vs traditionnelle (exemples sélectionnés).

| Étude (Année)  
Étude (Année)

Étiquettes: recherche-ia, recherche-generative, recherche-traditionnelle, seo, recherche-information, metriques-evaluation-recherche, modeles-langage-larges, google-vs-chatgpt

AVERTISSEMENT

Ce document est fourni à titre informatif uniquement. Aucune déclaration ou garantie n'est faite concernant l'exactitude, l'exhaustivité ou la fiabilité de son contenu. Toute utilisation de ces informations est à vos propres risques. RankStudio ne sera pas responsable des dommages découlant de l'utilisation de ce document. Ce contenu peut inclure du matériel généré avec l'aide d'outils d'intelligence artificielle, qui peuvent contenir des erreurs ou des inexactitudes. Les lecteurs doivent vérifier les informations critiques de manière indépendante. Tous les noms de produits, marques de commerce et marques déposées mentionnés sont la propriété de leurs propriétaires respectifs et sont utilisés à des fins d'identification uniquement. L'utilisation de ces noms n'implique pas l'approbation. Ce document ne constitue pas un conseil professionnel ou juridique. Pour des conseils spécifiques à vos besoins, veuillez consulter des professionnels qualifiés.