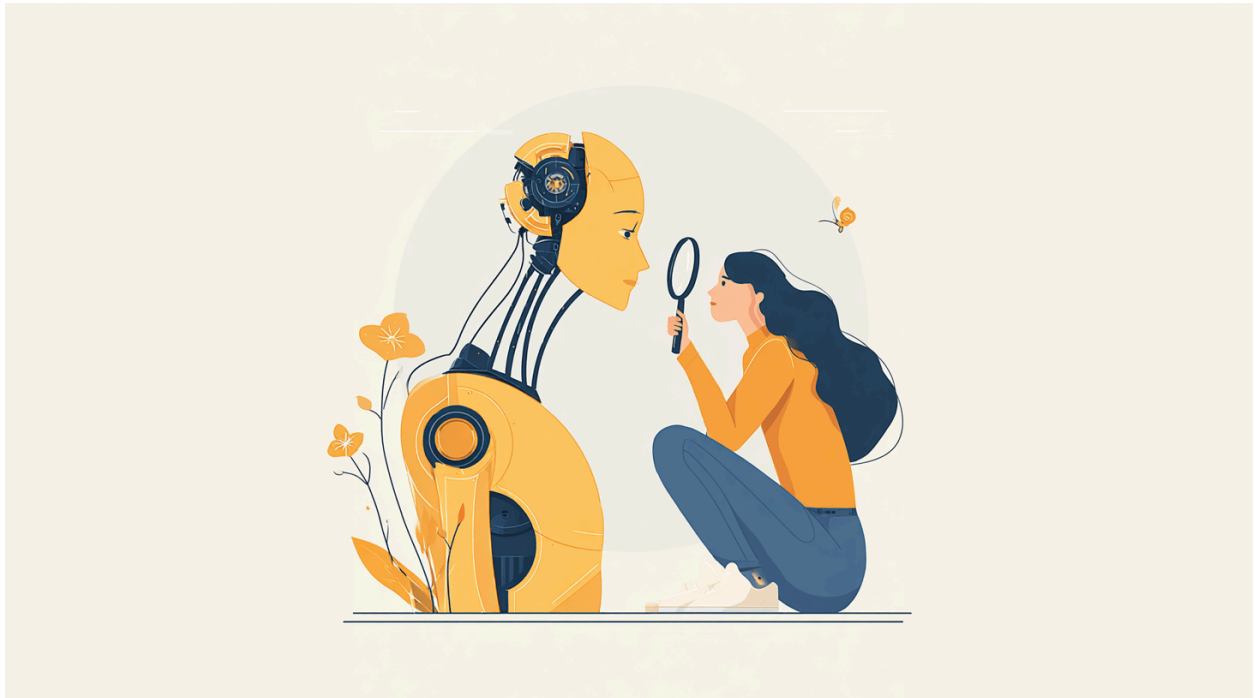


# Précision des détecteurs d'IA [Mise à jour mars 2026]



## Principales conclusions :

*Les recherches universitaires sur la précision de détection de l'IA montrent que les performances rapportées varient généralement entre 83 % et 100 %, la plupart des études faisant état de niveaux de précision compris entre 96 % et 99 %, ce qui indique que les détecteurs IA modernes peuvent atteindre des performances élevées dans des conditions de test contrôlées.*

Les comparaisons entre les différents outils révèlent des variations notables dans la précision de la détection de l'IA, les valeurs rapportées allant de 77 % pour Copyleaks à 98 % pour Originality.ai, tandis que Turnitin AI fait état d'une précision de 92 % et GPTZero d'une précision de 86 % dans l'ensemble de données évalué.

L'analyse de la précision de détection de Turnitin AI montre que le système a détecté les textes générés par l'IA avec une précision de 29 %, contre 83 % pour Originality.ai sur le même ensemble de données, tandis que les deux outils ont affiché une précision relativement élevée pour l'identification des textes rédigés par des humains (respectivement 93 % et 96 %).

La précision de détection varie en fonction du modèle d'IA qui a généré le texte, les systèmes identifiant le contenu provenant de ChatGPT, Grok et Gemini avec une précision de 100 %, tandis que la précision de détection était légèrement inférieure pour GPT-3.5 (99,7 %) et GPT-4 (98,7 %).

Les résultats montrent également que les performances de détection diffèrent selon le type de texte, avec une précision de 98 % pour les textes entièrement générés par l'IA, de 96 % pour les textes rédigés par des humains, de 90 % pour les contenus humains édités par l'IA et de 87 % pour les écrits hybrides IA-humain, ce qui indique que la paternité mixte est plus difficile à classer pour les détecteurs.

L'analyse du taux d'erreur démontre que les systèmes de détection de l'IA peuvent produire des erreurs de classification, avec des taux de faux positifs allant de 2 % à 38 % et des taux de faux négatifs allant de 2 % à 20 %, tandis que les évaluateurs humains ont affiché un taux de faux négatifs de 15 % lors de l'identification de contenus générés par l'IA.

*La comparaison de plusieurs plateformes de détection indique que la précision déclarée pour les principaux outils varie de 77 % à 99 %, GPTZero, Smodin et Hive affichant une précision de 99 %, Turnitin de 98 % et Quillbot AI Detector de 80 %, ce qui illustre des variations significatives de performances entre les systèmes de détection de l'IA.*

*Dans l'ensemble, les résultats montrent que la précision de la détection par IA dépend fortement des ensembles de données, de la méthodologie d'évaluation, de la source du modèle d'IA et du type de texte, ce qui signifie que les valeurs de précision rapportées doivent être interprétées comme des résultats de recherche dépendants du contexte plutôt que comme des garanties de performance universelles.*

La croissance rapide des outils d'IA générative a soulevé une question importante dans les domaines de l'éducation, de l'édition et du contenu numérique : les détecteurs d'IA sont-ils précis lorsqu'ils identifient un texte généré par l'IA ? À mesure que les systèmes de rédaction basés sur l'IA deviennent plus avancés, de nombreuses organisations s'appuient sur des outils de détection pour distinguer le contenu rédigé par des humains de celui généré par des machines.

Dans la pratique, l'évaluation de la précision de la détection de l'IA est plus complexe qu'un simple pourcentage. Les chercheurs évaluent généralement les détecteurs à l'aide de multiples indicateurs, ensembles de données et environnements de test. De ce fait, les résultats rapportés peuvent varier considérablement en fonction de la méthodologie utilisée. Cette variation explique pourquoi les discussions sur la précision des détecteurs d'IA aboutissent souvent à des conclusions différentes.

Les taux de précision des détecteurs d'IA varient considérablement en fonction de l'ensemble de données, du type de texte et de la méthodologie d'évaluation. L'analyse suivante regroupe les résultats rapportés dans différentes études universitaires et tests de référence.

Les valeurs de précision évoquées dans cet article représentent les résultats rapportés par différentes évaluations plutôt qu'un seul benchmark standardisé.

Dans les sections ci-dessous, nous examinons les données empiriques sur la précision de la détection de l'IA, y compris les résultats de recherches universitaires, les comparaisons entre les principaux outils de détection et les taux d'erreur tels que les faux positifs et les faux négatifs. Cette analyse permet de déterminer si les détecteurs d'IA sont précis dans des scénarios réels et comment les performances varient selon les détecteurs, les ensembles de données et les types de texte.

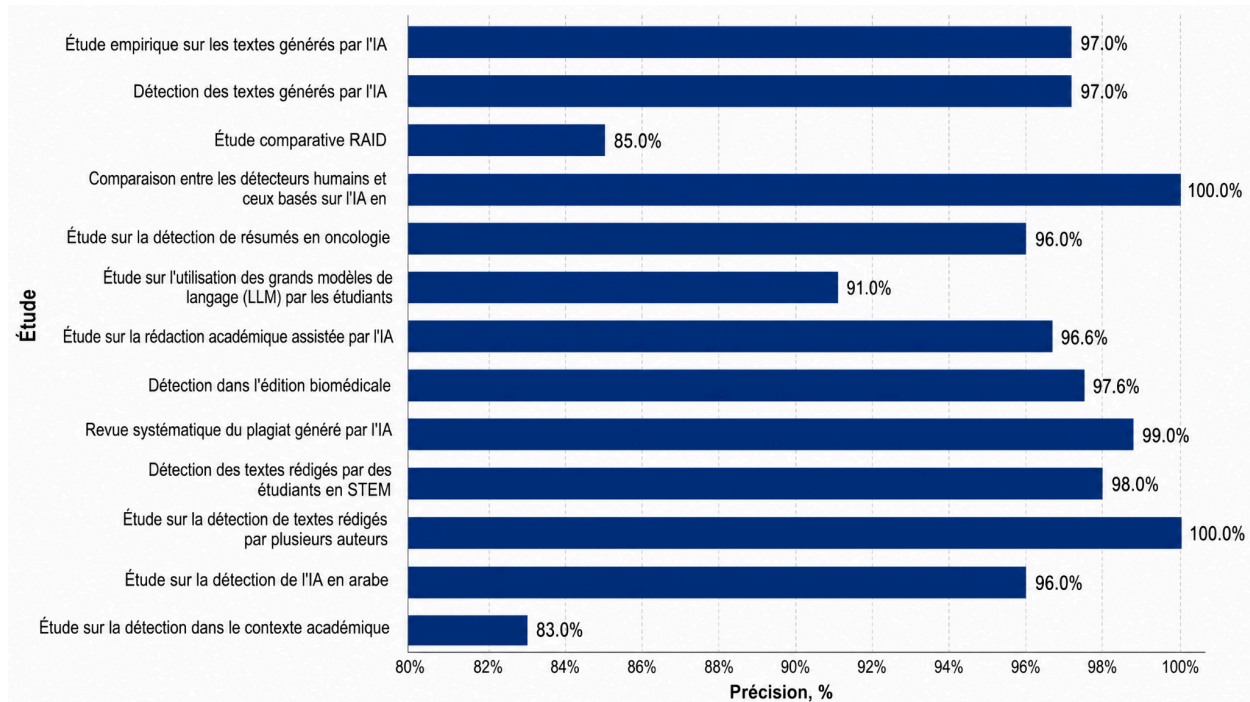
Après avoir abordé le concept général de la précision de la détection de l'IA, il est utile d'examiner les données empiriques issues d'études universitaires qui ont mesuré la capacité des détecteurs d'IA à identifier les textes générés par l'IA.

## Précision de la détection de l'IA dans les études universitaires

Le graphique ci-dessous résume les résultats rapportés dans plusieurs études évaluées par des pairs et des tests de référence évaluant la précision de la détection de l'IA à travers différents ensembles de données et contextes de recherche. Ces études abordent la question courante de savoir si les détecteurs d'IA sont précis en mesurant dans quelle mesure les outils de détection distinguent de manière fiable les textes générés par l'IA de ceux rédigés par des humains.

Les taux de précision des détecteurs d'IA varient considérablement en fonction de l'ensemble de données, du type de texte et de la méthodologie d'évaluation. La comparaison suivante regroupe les résultats rapportés dans différentes études universitaires et tests de référence.

Dans l'ensemble, les résultats fournissent aux lecteurs une vue d'ensemble fondée sur des données, en examinant la précision des détecteurs d'IA et la capacité des systèmes actuels à identifier de manière cohérente les contenus générés par l'IA.



- La précision de détection de l'IA la plus élevée rapportée a atteint 100 % tant dans l'étude « Humans vs AI Detectors in Medical Writing » que dans l'étude « Multi-Authored Text Detection Study ».
- La précision la plus faible dans l'ensemble de données était de 83 % dans l'étude « Academic Context Detection Study », ce qui montre que les résultats peuvent varier considérablement en fonction de l'environnement de test.
- La plupart des études ont rapporté une précision comprise entre 96 % et 99 %, notamment 97,6 % dans l'étude sur la détection dans les publications biomédicales et 98 % dans l'étude sur la détection dans les écrits d'étudiants en STEM.

## Quelle est la précision des détecteurs IA ? Données issues de la recherche universitaire

Étude	Précision, %
Étude empirique sur les outils de détection de textes générés par l'IA	97,0
Détection des textes générés par l'IA (étude portant sur 16 détecteurs)	97,0
Étude comparative RAID	85,0
Comparaison entre les détecteurs humains et ceux basés sur l'IA dans les textes médicaux	100,0
Étude sur la détection de résumés en oncologie	96,0
Étude sur l'utilisation des modèles de langage de grande envergure (LLM) par les étudiants	91,0

---

Étude sur la rédaction académique assistée par l'IA	96,6
Étude sur la détection des publications biomédicales	97,6
Revue systématique des détecteurs de plagiat basés sur l'IA	99,0
Étude sur la détection des écrits d'étudiants en STEM	98,0
Étude sur la détection de textes rédigés par plusieurs auteurs	100,0 %
Étude sur la détection par IA en arabe	96,0 %
Étude sur la détection de contextes académiques	83,0

---

Dans l'ensemble des études analysées, la plupart des valeurs de précision rapportées se situent dans une fourchette de 90 à 100 %, ce qui indique que les outils modernes de détection de l'IA peuvent atteindre des performances relativement élevées dans des environnements de recherche contrôlés. Parallèlement, la variation entre 83 % et 100 % de précision démontre que l'efficacité de ces systèmes dépend fortement de l'ensemble de données, du domaine textuel et de la méthodologie d'évaluation.

Ces résultats aident à répondre à des questions telles que le degré de précision de l'IA dans l'identification de textes générés par l'IA et la fiabilité des détecteurs d'IA dans des scénarios concrets. Si les données montrent que de nombreux systèmes obtiennent de bons résultats lors des tests de référence universitaires, les variations entre les études

suggèrent que les résultats de détection par l'IA doivent être interprétés en fonction du contexte plutôt que considérés comme un score de précision unique et universel.

Si les études universitaires fournissent un aperçu général de la précision de la détection de l'IA, une comparaison plus approfondie des différents outils permet de répondre à des questions pratiques, telles que la précision du détecteur d'IA de Turnitin et la comparaison de ses performances avec celles d'autres systèmes de détection de l'IA.

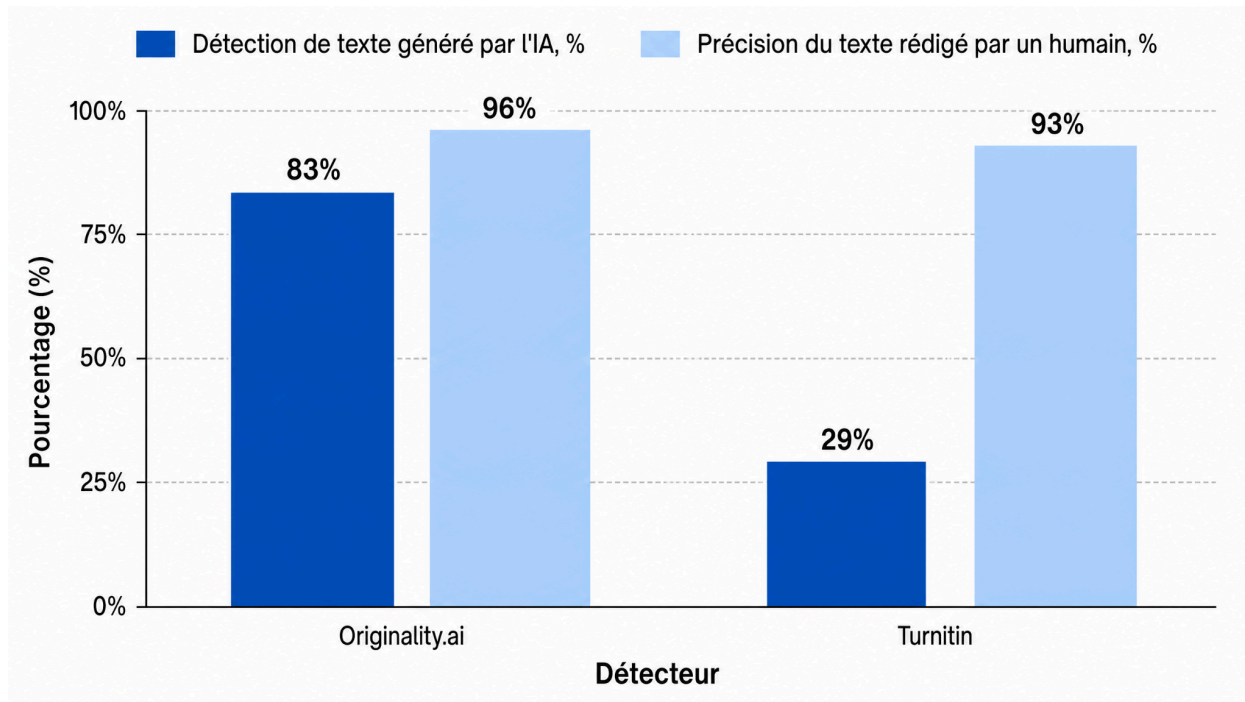
## Précision de la détection de l'IA de Turnitin par rapport à d'autres détecteurs d'IA

Le graphique ci-dessous compare les performances de détection de deux détecteurs d'IA selon deux critères : la capacité à identifier correctement les textes générés par l'IA et la précision de la classification des contenus rédigés par des humains.

Cette comparaison permet de répondre à des questions courantes telles que le degré de précision du détecteur d'IA de Turnitin et la question de savoir si ses performances correspondent à la précision de détection de l'IA de Turnitin rapportée dans les évaluations universitaires. Les résultats proviennent d'une étude qui a testé les deux détecteurs sur des ensembles de données contenant des textes générés par l'IA, rédigés par des humains et hybrides.

Les taux de précision des détecteurs d'IA varient en fonction de l'ensemble de données, du type de texte et de la méthodologie d'évaluation. La comparaison suivante reflète les résultats rapportés dans une étude spécifique plutôt qu'une référence universelle.

## Turnitin par rapport aux autres détecteurs d'IA



- Originality.ai a détecté les textes générés par l'IA avec une précision de 83 %, contre 29 % pour Turnitin dans la même évaluation.
- Turnitin a correctement identifié les textes rédigés par des humains dans 93 % des cas, un taux légèrement inférieur à celui d'Originality.ai (96 %).
- La différence de capacité de détection de l'IA entre les deux systèmes a atteint 54 points de pourcentage (83 % contre 29 %) dans cet ensemble de données.

Quelle est la précision du détecteur d'IA de Turnitin par rapport à d'autres outils ?

Détecteur

Détection de texte par IA, %

Précision du texte humain, %

---

Originality.ai	83 %	96
Turnitin	29 %	93 %

---

Cette comparaison met en évidence à quel point les résultats peuvent varier considérablement d'un outil de détection d'IA à l'autre lors de l'évaluation de textes générés par l'IA. Dans cet ensemble de données, la précision de détection de l'IA rapportée par Turnitin pour l'identification de contenus générés par l'IA était nettement inférieure à celle du système alternatif, même si les deux outils ont démontré une précision relativement élevée lors de la classification de textes rédigés par des humains.

Ces résultats alimentent les débats en cours sur la précision du détecteur d'IA de Turnitin et illustrent pourquoi la précision de détection de l'IA rapportée par Turnitin doit être interprétée dans le contexte des conditions de test spécifiques. Dans l'ensemble, les données suggèrent que les performances de détection dépendent non seulement de l'outil lui-même, mais aussi de l'ensemble de données et de la méthodologie d'évaluation utilisés dans l'étude.

Après avoir examiné la précision de détection de l'IA de Turnitin par rapport à un autre détecteur, la prochaine étape consiste à comparer les performances de plusieurs grands outils de détection de l'IA.

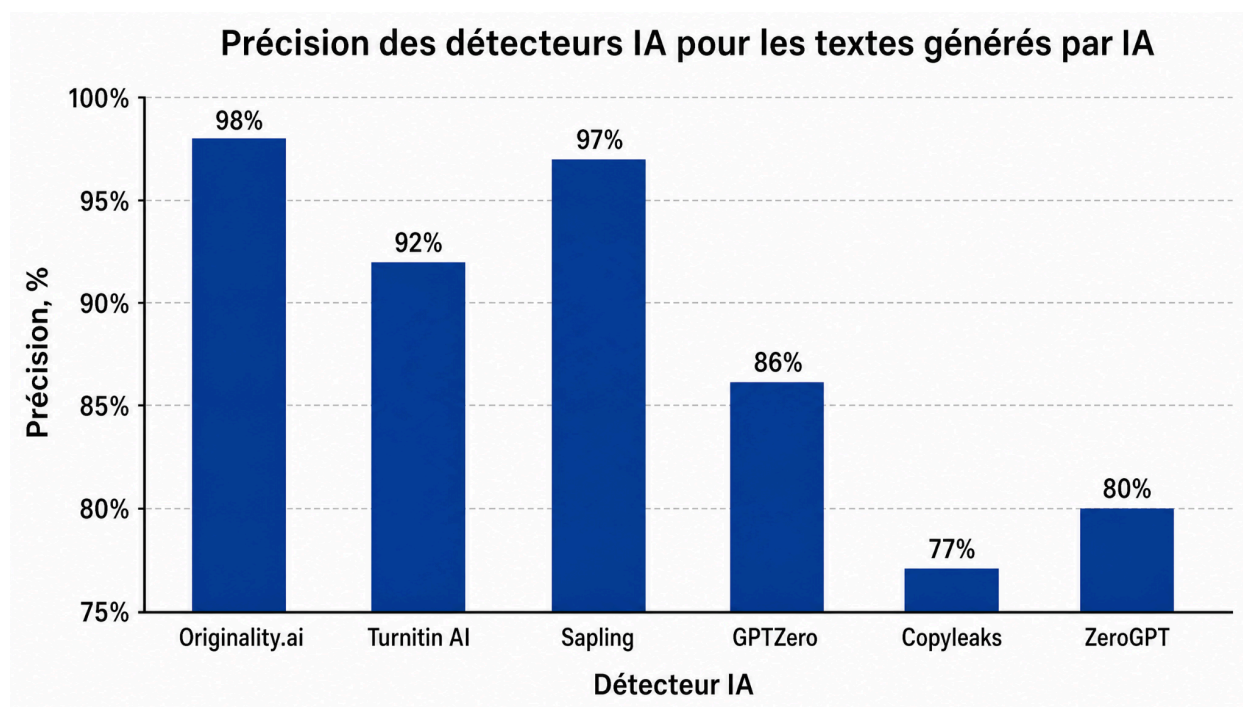
## Comparaison de la précision des principaux détecteurs d'IA

Le graphique ci-dessous compare la précision de détection de l'IA rapportée pour plusieurs outils de détection de l'IA largement utilisés. Cette comparaison permet de répondre à des questions courantes, telles que : quel est le détecteur d'IA le plus précis et quelle est la précision des détecteurs d'IA lorsqu'ils sont évalués sur différentes plateformes ?

Les taux de précision des détecteurs d'IA varient considérablement en fonction de l'ensemble de données, du type de texte et de la méthodologie d'évaluation. La

comparaison suivante regroupe les résultats rapportés dans différentes études universitaires et tests de référence.

En raison de ces différences, le graphique doit être interprété comme un aperçu comparatif plutôt que comme un classement définitif des performances des détecteurs.



- Originality.ai affiche la précision la plus élevée rapportée, à 98 %, suivi de près par Sapling à 97 %.
- Turnitin AI affiche une précision de 92 %, ce qui le place entre les détecteurs les plus performants et les outils les moins performants.
- Copyleaks et ZeroGPT affichent des niveaux de précision plus faibles, à 77 % et 80 %, tandis que GPTZero fait état d'une précision de 86 % dans les résultats évalués.

## Quel est le détecteur d'IA le plus précis ? Comparaison de la précision entre les différents outils

Détecteur IA	Précision, %
Originality.ai	98 %
Turnitin AI	92 %
Sapling	97 %
GPTZero	86 %
Copyleaks	77 %
ZeroGPT	80 %

Cette comparaison montre que la précision de détection de l'IA peut varier considérablement d'un détecteur à l'autre. Alors que certains outils affichent des taux de précision supérieurs à 95 %, d'autres affichent des performances plus proches de 77 à 86 %, selon l'évaluation.

Ces différences aident à expliquer pourquoi des questions telles que « quel est le détecteur d'IA le plus précis ? » restent sujettes à interprétation. Étant donné que les

résultats de détection dépendent des ensembles de données, des méthodes d'évaluation et des types de texte analysés, les valeurs de précision annoncées doivent être considérées comme des indicateurs comparatifs plutôt que comme des garanties de performance fixes.

Au-delà de la comparaison des détecteurs individuels, un autre facteur important influant sur la précision de la détection de l'IA est le type de modèle d'IA qui a généré le texte.

## Précision de la détection de l'IA par source LLM

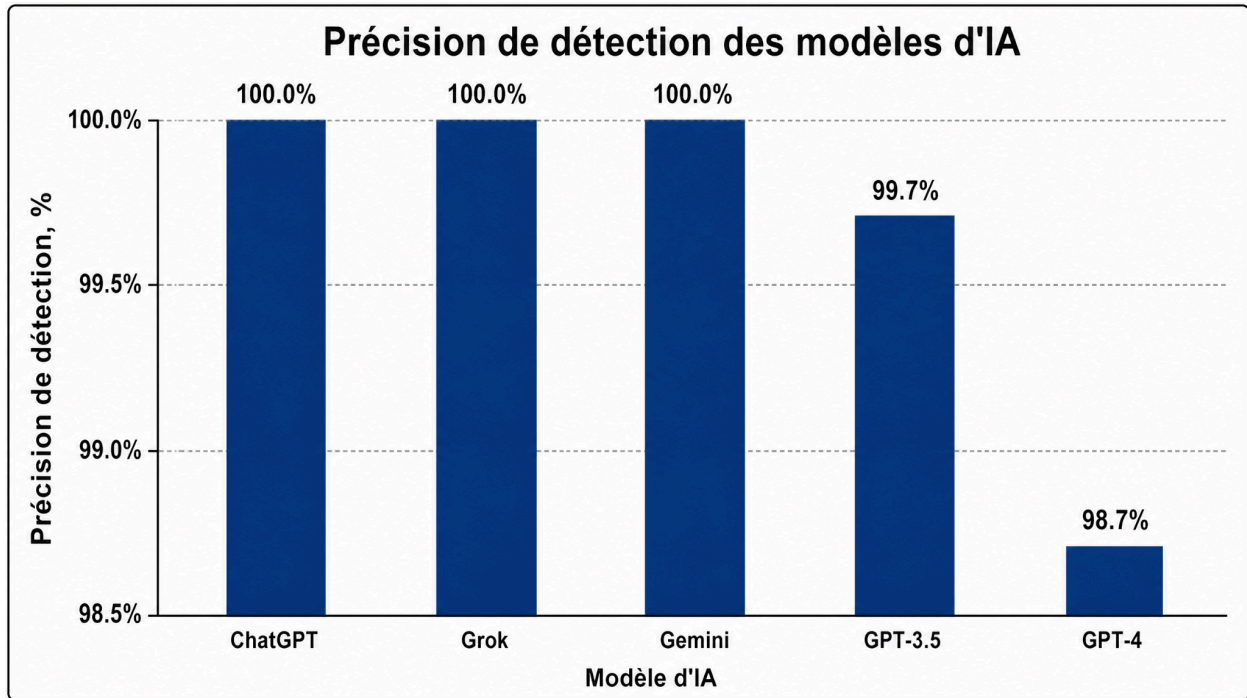
Le graphique ci-dessous montre dans quelle mesure les systèmes de détection basés sur l'IA identifient avec précision les textes générés par différents grands modèles linguistiques (LLM). L'évaluation des performances de détection sur plusieurs modèles permet d'aborder des questions plus générales, telles que le degré de précision de l'IA lorsqu'il s'agit d'identifier des textes générés par l'IA provenant de différentes sources.

La précision rapportée peut varier en fonction de l'ensemble de données, de la structure du texte et de la méthodologie d'évaluation utilisée dans chaque étude. Par conséquent, les performances de détection peuvent différer lors de l'analyse des résultats de différents modèles d'IA.

Les valeurs du graphique représentent les résultats rapportés d'évaluations spécifiques plutôt qu'une référence standardisée unique.

Comprendre ces différences permet d'expliquer pourquoi les débats sur la précision des détecteurs d'IA dépendent souvent du modèle d'IA testé.

## Précision de la détection de l'IA par source LLM



- Les textes générés par l'IA à partir de ChatGPT, Grok et Gemini ont été détectés avec une précision de 100,0 % lors des tests évalués.
- La précision de détection pour GPT-3.5 a atteint 99,7 %, ce qui indique une identification quasi parfaite du contenu généré par l'IA.
- GPT-4 a affiché une précision de détection légèrement inférieure, à 98,7 %, tout en restant au-dessus du seuil des 98 %.

Comment la précision de détection varie selon les différents modèles d'IA

Modèle d'IA

Précision de détection, %

---

---

ChatGPT	100,0 %
Grok	100,0 %
Gemini	100,0 %
GPT-3.5	99,7 %
GPT-4	98,7

---

Les résultats suggèrent que les systèmes de détection de l'IA peuvent atteindre un haut niveau de précision lorsqu'ils identifient du texte généré par les principales plateformes LLM. Dans l'ensemble de données analysé, la précision de détection variait entre 98,7 % et 100 % selon le modèle d'IA.

Ces résultats apportent un éclairage supplémentaire sur des questions telles que la précision des détecteurs d'IA et l'efficacité avec laquelle les systèmes de détection distinguent les contenus générés par l'IA. Si les résultats présentés démontrent de solides capacités de détection sur plusieurs grands modèles de langage (LLM), les variations entre les ensembles de données et les méthodes d'évaluation signifient que les performances réelles peuvent varier en fonction de la source textuelle spécifique analysée.

Outre le modèle d'IA qui génère le texte, un autre facteur influençant la précision de la détection de l'IA est le type de contenu analysé.

## Précision de la détection par type de texte

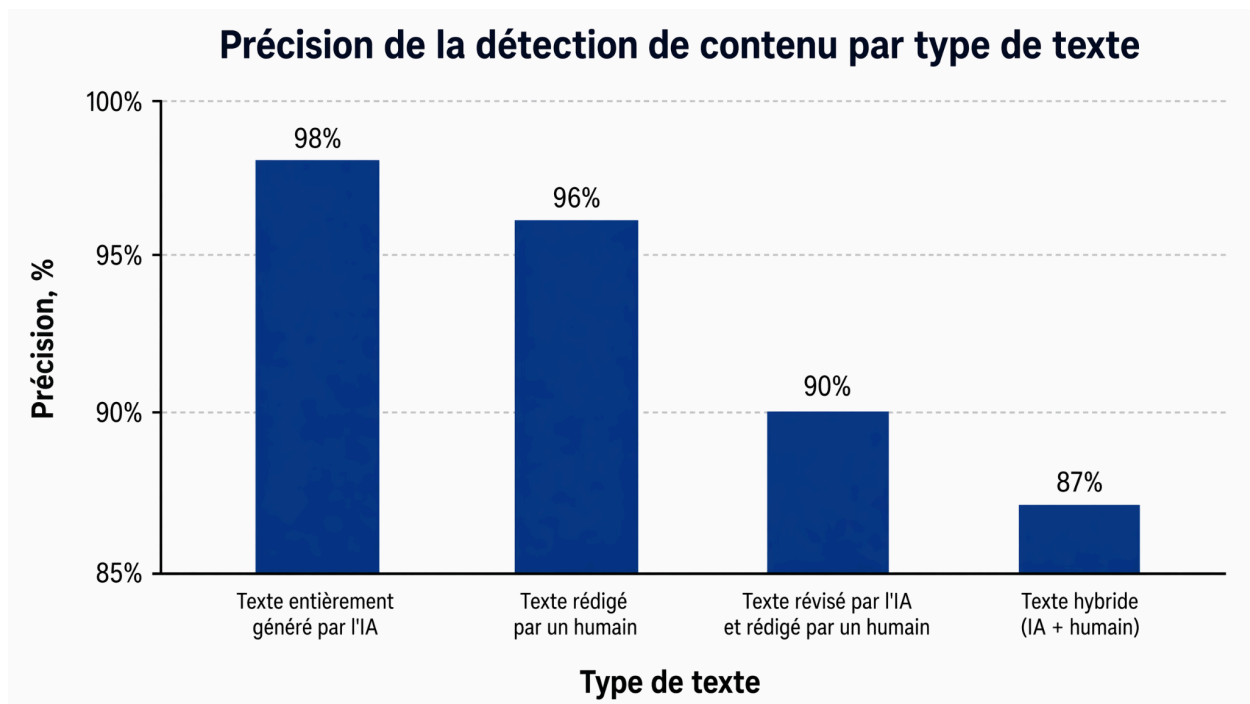
Le graphique ci-dessous compare les performances de détection entre différentes catégories de textes, notamment le contenu entièrement généré par l'IA, le texte rédigé par un humain, les textes édités par l'IA et le contenu hybride (IA-humain).

Ces distinctions sont importantes pour évaluer la précision des détecteurs d'IA, car le niveau d'implication de l'IA dans le texte peut considérablement affecter les résultats de détection. Par exemple, il est généralement plus facile d'identifier un texte entièrement généré par l'IA que de détecter un contenu partiellement édité ou hybride.

Les taux de précision des détecteurs d'IA varient en fonction de l'ensemble de données, de la structure du texte et de la méthodologie d'évaluation. Les valeurs indiquées dans le graphique reflètent les résultats rapportés dans le cadre d'une évaluation spécifique plutôt qu'une référence universelle.

Comprendre ces différences aide également à éclairer les débats actuels sur la fiabilité des détecteurs d'IA lorsqu'ils analysent des textes mixtes ou partiellement assistés par l'IA.

## Précision de la détection par type de texte



- La précision de détection a atteint 98 % pour les textes entièrement générés par l'IA, ce qui représente la meilleure performance parmi les types de textes évalués.
- Les détecteurs d'IA ont correctement classé 96 % des textes rédigés par des humains, ce qui indique un taux de faux positifs relativement faible dans cet ensemble de données.
- La précision est tombée à 90 % pour les textes humains édités par l'IA et à 87 % pour les contenus hybrides IA-humains, ce qui montre que la paternité mixte est plus difficile à détecter.

## Comment le type de texte influe sur la précision de la détection de l'IA

Type de texte	Précision, %
Texte entièrement généré par l'IA	98 %
Texte rédigé par un humain	96
Texte rédigé par un humain et édité par l'IA	90
Texte hybride IA + humain	87

Les résultats montrent que la précision de la détection par l'IA varie en fonction de la manière dont le contenu a été créé. Les systèmes obtiennent de meilleurs résultats lorsqu'ils analysent des textes entièrement générés par l'IA, où la précision de détection atteint 98 %, tandis que les contenus mixtes ou partiellement édités posent davantage de difficultés.

Ces résultats apportent un éclairage supplémentaire sur des questions telles que la précision des détecteurs d'IA et leur fiabilité lors de l'évaluation de textes réels. À mesure que l'utilisation des outils d'édition assistés par l'IA se généralise, la distinction entre les contenus rédigés par des humains, ceux édités par l'IA et les contenus hybrides pourrait devenir un facteur clé influençant les performances des systèmes de détection de l'IA.

Si les indicateurs de précision offrent une vue d'ensemble de la précision de la détection de l'IA, l'évaluation des faux positifs et des faux négatifs permet de répondre à une autre question importante : les détecteurs d'IA peuvent-ils se tromper ?

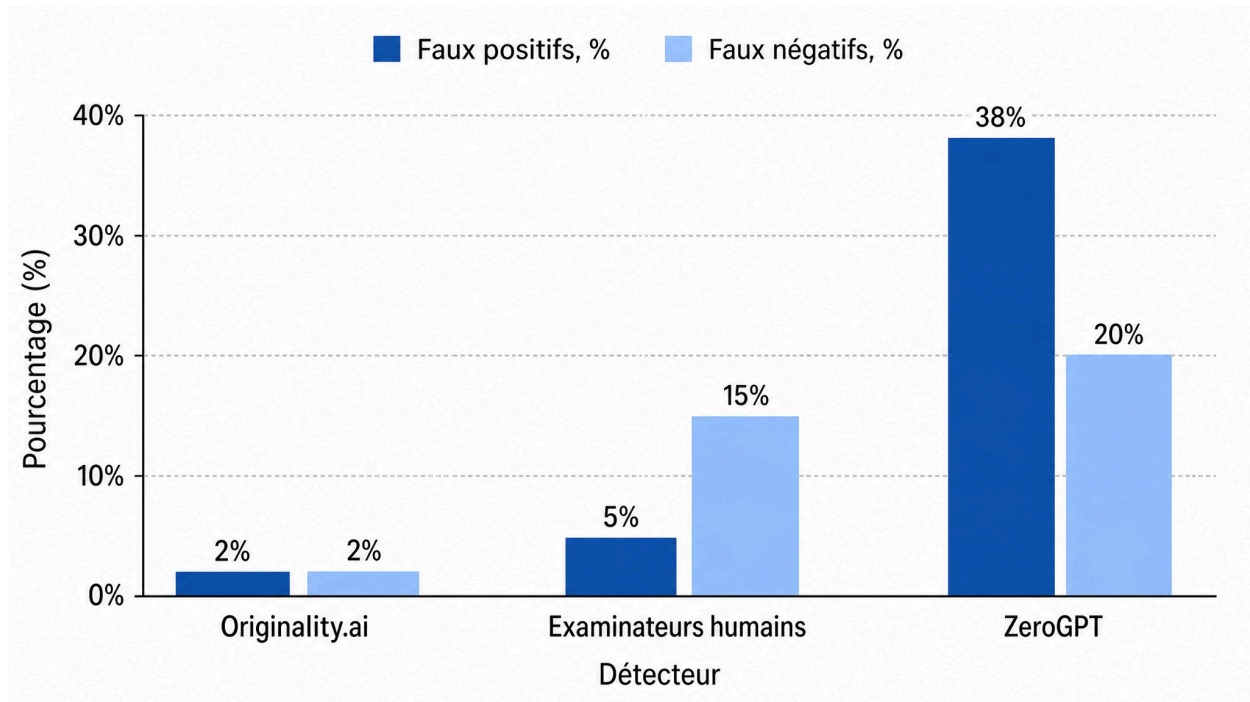
## Taux de faux positifs et de faux négatifs dans la détection de l'IA

Le graphique ci-dessous compare les taux d'erreur de différents évaluateurs, y compris les détecteurs d'IA automatisés et les réviseurs humains. Ces résultats sont généralement rapportés dans des recherches évaluant les méthodologies des articles sur les détecteurs d'IA et les vérificateurs d'IA, où des indicateurs de matrice de confusion tels que les faux positifs et les faux négatifs sont utilisés pour mesurer la fiabilité.

Les faux positifs se produisent lorsque du texte rédigé par un humain est identifié à tort comme généré par l'IA, tandis que les faux négatifs se produisent lorsque du contenu généré par l'IA est classé comme rédigé par un humain. Ces deux types d'erreurs déterminent la fiabilité des détecteurs d'IA dans les applications concrètes.

Les taux de précision des détecteurs d'IA varient en fonction de l'ensemble de données et de la méthode d'évaluation. Les valeurs présentées ci-dessous reflètent les résultats rapportés dans des études spécifiques plutôt qu'une référence standardisée unique.

## Taux de faux positifs et de faux négatifs



- Originality.ai a enregistré les taux d'erreur les plus bas, avec 2 % de faux positifs et 2 % de faux négatifs dans l'ensemble de données évalué.
- Les évaluateurs humains ont affiché un taux de faux négatifs plus élevé, à 15 %, ce qui signifie que le texte généré par l'IA n'a pas été détecté dans 15 % des cas.
- ZeroGPT a affiché les taux d'erreur les plus élevés, avec 38 % de faux positifs et 20 % de faux négatifs lors de la même évaluation.

## Les détecteurs d'IA peuvent-ils se tromper ? Taux d'erreur dans les systèmes de détection d'IA

Détecteur	Faux positifs, %	Faux négatifs, %
Originality.ai	2 %	2
Évaluateurs humains	5	15 %
ZeroGPT	38 %	20

Les données montrent que l'évaluation de la précision de la détection de l'IA ne se limite pas à un simple pourcentage de précision. Les faux positifs et les faux négatifs fournissent des informations supplémentaires sur les performances des systèmes de détection lors de la classification de textes générés par l'IA et de textes rédigés par des humains.

Dans cet ensemble de données, les détecteurs automatisés ont affiché des taux d'erreur allant de 2 % à 38 %, tandis que les évaluateurs humains ont présenté un taux de faux négatifs de 15 %. Ces différences contribuent à expliquer pourquoi les débats sur la possibilité que les détecteurs d'IA se trompent restent d'actualité dans la recherche universitaire et technique.

Dans l'ensemble, les résultats rapportés dans de nombreux articles sur les détecteurs d'IA et les études sur les vérificateurs d'IA suggèrent que tant les outils automatisés que les évaluateurs humains peuvent commettre des erreurs de classification, ce qui renforce l'importance d'interpréter les résultats de détection de l'IA dans le contexte d'ensembles de données et de méthodes d'évaluation spécifiques.

Après avoir examiné la précision de la détection, la variation des ensembles de données et les taux d'erreur, il est utile de comparer globalement les différents outils de détection de l'IA en termes de précision rapportée.

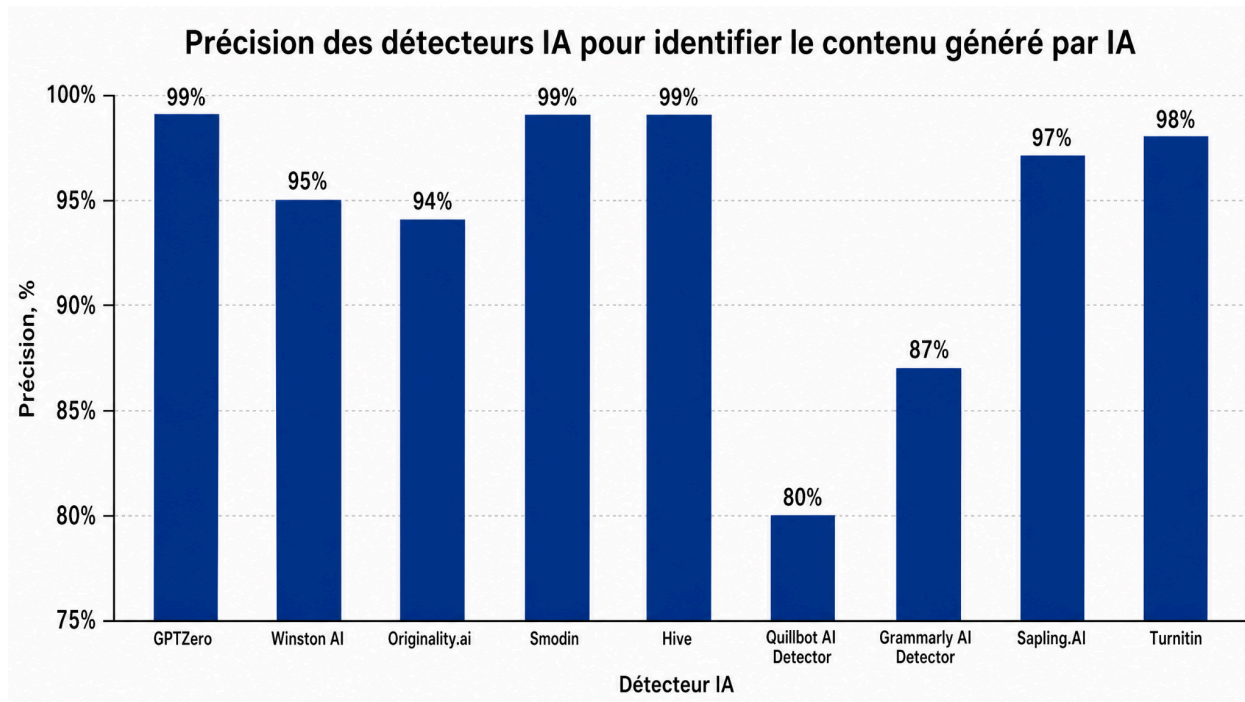
## Les détecteurs de contenu IA les plus précis

Le graphique ci-dessous compare plusieurs outils de détection de l'IA largement utilisés sur la base des valeurs de précision rapportées à partir de différents ensembles de données et évaluations. Cette comparaison permet de répondre à la question courante de savoir quel est le détecteur d'IA le plus précis en résumant les indicateurs de performance rapportés pour différents outils.

Les taux de précision des détecteurs d'IA varient considérablement en fonction de l'ensemble de données, du type de texte et de la méthodologie d'évaluation. La comparaison suivante regroupe les résultats rapportés dans le cadre de différentes évaluations et tests de référence.

En raison de ces différences méthodologiques, les comparaisons doivent être interprétées comme un aperçu général plutôt que comme un classement définitif des performances des détecteurs.

## Les détecteurs de contenu IA les plus précis



- Selon les données disponibles, GPTZero, Smodin et Hive affichent les niveaux de précision les plus élevés, à 99 %.
- Turnitin affiche une précision de 98 %, tandis que Sapling.AI affiche une précision de 97 % avec des taux de faux positifs relativement faibles.
- Certains outils affichent des taux de précision plus faibles, notamment Quillbot AI Detector (80 %) et Grammarly AI Detector (87 %).

Quel est le détecteur d'IA le plus précis ? Comparaison de la précision des principaux outils

Détecteur d'IA

Précision, %

Taux de faux positifs

---

GPTZero	99 %	Faible
Winston AI	95 %	Modéré
Originality.ai	94	Modéré à élevé
Smodin	99 %	Modéré
Hive	99 %	Très faible
Détecteur d'IA Quillbot	80 %	Modéré
Détecteur d'IA Grammarly	87	Élevé
Sapling.AI	97	Faible
Turnitin	98 %	Faible

---

La comparaison montre que la précision de détection de l'IA rapportée varie considérablement d'une plateforme à l'autre. Certains systèmes affichent des niveaux

de précision proches de 99 %, tandis que d'autres se situent plutôt entre 80 et 90 %, selon les conditions d'évaluation.

Ces différences expliquent en partie pourquoi il reste difficile de répondre de manière définitive à des questions telles que « quel est le détecteur d'IA le plus précis ? ». Les performances de détection dépendent non seulement de l'outil lui-même, mais aussi de l'ensemble de données, de la méthodologie de test et du type de contenu analysé.

Par conséquent, les comparaisons entre les détecteurs doivent être interprétées dans le contexte des études publiées plutôt que considérées comme des garanties de performances fixes pour tous les cas d'utilisation.

## Conclusions

- Les données disponibles sur la précision de la détection de l'IA indiquent que les systèmes modernes de détection de l'IA peuvent atteindre des performances relativement élevées dans des environnements de recherche contrôlés. Dans l'ensemble des études universitaires analysées, les valeurs de précision rapportées variaient de 83 % à 100 %, la plupart des études faisant état de résultats compris entre 96 % et 99 %, ce qui suggère que de nombreux détecteurs d'IA sont capables d'identifier de manière fiable les textes générés par l'IA dans des conditions de test spécifiques.
- Les comparaisons entre les différents outils montrent que la précision de la détection de l'IA varie considérablement d'une plateforme à l'autre. Dans les ensembles de données évalués, la précision rapportée allait de 77 % pour Copyleaks à 98 % pour Originality.ai, tandis que Turnitin AI rapportait une précision de 92 % et GPTZero de 86 %, ce qui indique que différents outils peuvent produire des résultats sensiblement différents.
- Une analyse complémentaire montre que la précision de détection de Turnitin AI peut varier en fonction de l'ensemble de données et de la méthode d'évaluation.

Dans l'ensemble de données examiné dans cet article, Turnitin a détecté les textes générés par l'IA avec une précision de 29 %, contre 83 % pour Originality.ai, tandis que les deux outils ont démontré une précision relativement élevée pour l'identification des textes rédigés par des humains (respectivement 93 % et 96 %).

- Les performances de détection varient également en fonction du modèle d'IA ayant généré le texte. D'après les résultats de l'évaluation, les détecteurs d'IA ont identifié avec une précision de 100 % les textes produits par ChatGPT, Grok et Gemini, tandis que des taux de détection légèrement inférieurs ont été enregistrés pour GPT-3.5 (99,7 %) et GPT-4 (98,7 %), ce qui démontre que les résultats de détection peuvent varier en fonction de la source du LLM.
- L'analyse montre en outre que la structure du texte et le type d'auteur influencent les performances de détection. La précision de détection a atteint 98 % pour les textes entièrement générés par l'IA, 96 % pour les textes rédigés par des humains, 90 % pour les contenus humains édités par l'IA et 87 % pour les écrits hybrides IA-humain, ce qui indique que les contenus mixtes ou partiellement assistés par l'IA peuvent être plus difficiles à classer correctement pour les détecteurs.
- L'analyse du taux d'erreur indique que tant les systèmes automatisés que les évaluateurs humains peuvent produire des erreurs de classification. Dans l'ensemble de données analysé, les taux de faux positifs variaient de 2 % à 38 %, tandis que les taux de faux négatifs variaient de 2 % à 20 %, et les évaluateurs humains ont affiché un taux de faux négatifs de 15 %, ce qui illustre que les résultats de détection par l'IA ne sont pas exempts d'erreurs.
- Dans l'ensemble, les résultats suggèrent que la précision de la détection par IA dépend fortement de l'ensemble de données, de la méthodologie d'évaluation, de

la source du modèle d'IA et du type de texte analysé. Par conséquent, les valeurs de précision rapportées doivent être interprétées comme des résultats dépendants du contexte plutôt que comme des garanties de performance universelles lorsqu'il s'agit d'évaluer si les détecteurs d'IA peuvent identifier de manière fiable les contenus générés par l'IA.

## Sources

- Barlow, Written. « Les 9 meilleurs détecteurs d'IA offrant la plus grande précision en 2026 ». Ressources sur la détection de l'IA | GPTZero, 2 janvier 2026, <https://gptzero.me/news/best-ai-detectors/>. Consulté le 16 mars 2026.
- Gillham, Jonathan. « Études sur la précision de la détection de l'IA – Méta-analyse de 13 études – Originality.AI ». Originality.AI, <https://originality.ai/blog/ai-detection-studies-round-up>. Consulté le 16 mars 2026.
- « Nous atteignons une précision de 99 % dans la détection de l'IA : étude d'Originality.AI – Originality.AI. » Originality.AI, <https://originality.ai/blog/ai-accuracy>. Consulté le 16 mars 2026.