https://synthid.com/

Mit dem **SynthID-Detector** kannst du überprüfen, ob Inhalte (wie Bilder, Videos oder Audio-Dateien) von künstlicher Intelligenz erstellt oder bearbeitet wurden.

Hier sind die wichtigsten Funktionen und Möglichkeiten im Überblick:

unsichtbare Wasserzeichen erkennen
SynthID bettet digitale, für das menschliche Auge unsichtbare Wasserzeichen direkt in die Pixel von Medien (oder in Audiosignale) ein, während diese von Google-KI-Modellen generiert werden.
Der Detektor kann diese Signale selbst dann noch aufspüren, wenn das Medium bearbeitet, zugeschnitten oder komprimiert wurde.
Herkunft überprüfen
Du erhältst dadurch einen Hinweis darauf, ob ein bestimmtes Medium von einer KI erzeugt wurde.
Unterstützung bei C2PA / Content Credentials
Neben den Google-eigenen SynthID-Wasserzeichen unterstützt das System oft auch offene Standards wie C2PA-Zertifikate. Damit lässt sich die Bearbeitungshistorie (wer hat das Bild erstellt, welche Software wurde genutzt) nachvollziehen, sofern entsprechende fälschungssichere Metadaten in der Datei hinterlegt sind.

Wichtige Hinweise zur Anwendung:
Keine absolute Garantie
Kein Detektor arbeitet zu 100 % fehlerfrei. Nutze die Ergebnisse daher am besten als einen von mehreren Hinweisen und vergleiche sie mit anderen Quellen, wenn es auf Echtheit ankommt.
Einschränkungen bei bestimmten Inhalten:
Bei sehr stark komprimierten Dateien, stark verpixelten Bildern oder Inhalten mit extrem vielen Textzeilen (wie Screenshots von Code) kann es vorkommen, dass keine verlässliche Prüfung möglich ist.

Ausführliche Informationen und offizielle Details dazu findest du auch direkt auf der Hilfeseite unter support.google.com/gemini?p=verify_ai.
Zudem kannst du unter synthid.com nachsehen, welche weiteren Organisationen und Partner diese Technologie unterstützen.

Frage: werden auch Texte überprüft ?
Ja, auch Texte können grundsätzlich mit SynthID versehen und überprüft werden (*SynthID Text*)

Hier ist, wie das bei Texten funktioniert und wo die Unterschiede zu Bildern oder Videos liegen:

Wie das Wasserzeichen funktioniert:
Da Texte keine Pixel oder Audio-Wellenformen haben, nutzt SynthID hier einen statistischen Trick bei der Wortwahl (Token-Auswahl). Ein Sprachmodell wählt bei der Generierung Wörter aus, bei denen es oft mehrere fast gleichwertige Optionen gibt. SynthID steuert diese Auswahl unauffällig so, dass ein bestimmtes statistisches Muster entsteht, das für den Menschen normal und lesbar ist.
Ein spezieller Detektor kann dieses Muster im Nachhinein wiedererkennen.

Wann es funktioniert
Google setzt diese Technologie unter anderem bei Texten ein, die über die Gemini-App oder entsprechende Web-Oberflächen generiert wurden.
Wichtige Einschränkungen bei Text
Umschreibungen und Übersetzungen:
Wenn ein KI-Text stark umformuliert, zusammengefasst oder in eine andere Sprache übersetzt wird, bricht das Wasserzeichen-Muster oft ein, sodass der Detektor es nicht mehr sicher nachweisen kann.

Copyright 2026 , LastminuteITconsulting