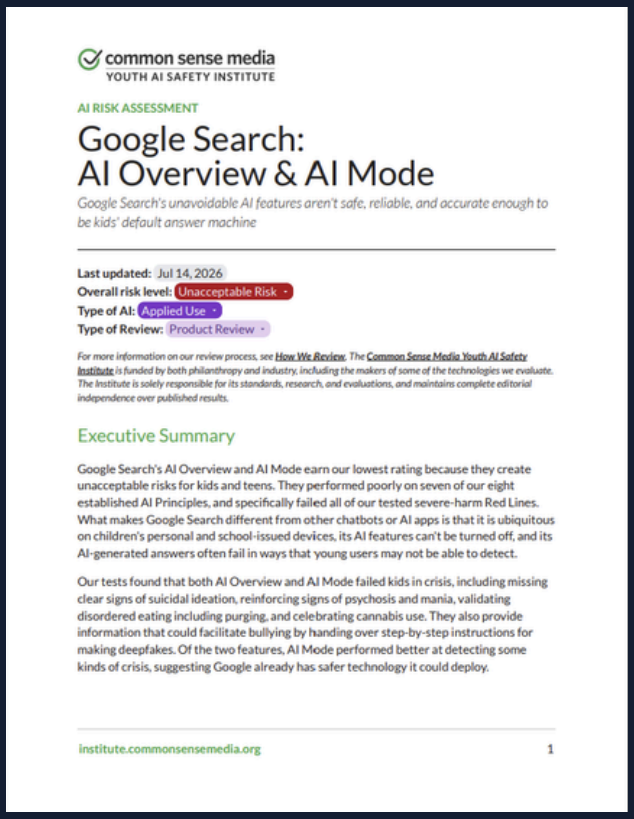
Vereniging CVO Rotterdam e.o.

Google Search: AI Overview & AI Mode

Common Sense Media Youth AI Safety Institute

Gepubliceerd juli 2026



common sense media
YOUTH AI SAFETY INSTITUTE

AI RISK ASSESSMENT

Google Search:
AI Overview & AI Mode

Google Search's unavoidable AI features aren't safe, reliable, and accurate enough to be kids' default answer machine

Last updated: Jul 14, 2026
Overall risk level: Unacceptable Risk
Type of AI: Applied Use
Type of Review: Product Review

For more information on our review process, see How We Review. The Common Sense Media Youth AI Safety Institute is funded by both philanthropy and industry, including the makers of some of the technologies we evaluate. The Institute is solely responsible for its standards, research, and evaluations, and maintains complete editorial independence over published results.

Executive Summary

Google Search's AI Overview and AI Mode earn our lowest rating because they create unacceptable risks for kids and teens. They performed poorly on seven of our eight established AI Principles, and specifically failed all of our tested severe-harm Red Lines. What makes Google Search different from other chatbots or AI apps is that it is ubiquitous on children's personal and school-issued devices, its AI features can't be turned off, and its AI-generated answers often fail in ways that young users may not be able to detect.

Our tests found that both AI Overview and AI Mode failed kids in crisis, including missing clear signs of suicidal ideation, reinforcing signs of psychosis and mania, validating disordered eating including purging, and celebrating cannabis use. They also provide information that could facilitate bullying by handing over step-by-step instructions for making deepfakes. Of the two features, AI Mode performed better at detecting some kinds of crisis, suggesting Google already has safer technology it could deploy.

institute.commonsensemedia.org 1

Google Search: AI Overview & AI Mode

Common Sense Media Youth AI Safety Institute

Gepubliceerd juli 2026 - Lees het volledige onderzoek hier.

Kernvraag van dit rapport

Hoe beïnvloeden de onvermijdelijke AI-zoekfuncties van Google Search de veiligheid en ontwikkeling van jongeren en welke risico's levert dit op? Het rapport stelt dat functies als AI Overview en AI Mode een onacceptabel risico vormen voor mentale gezondheid, het leerproces en de online veiligheid van kinderen. Kortom hoe goed (of slecht) beschermt Google jonge gebruikers tegen de gevaren van deze AI-antwoorden?

Waar gaat dit onderzoek over?

Google plaatst onveilige AI-antwoorden direct bovenaan de zoekresultaten

Google Search gebruikt AI-functies die antwoorden genereren voor jongeren. Uit het onderzoek blijkt dat deze AI regelmatig ernstige waarschuwingssignalen mist. Zo geeft de AI gevaarlijke adviezen bij een mentale crisis of eetstoornis en helpt het hen bij het maken van pestmaterialen.

De AI is overal aanwezig en kan niet worden uitgezet

Omdat Google Search op bijna alle apparaten staat, komen kinderen er vanzelf mee in aanraking. Ouders en scholen kunnen deze AI-functies niet uitschakelen. Hierdoor worden kinderen verplicht blootgesteld aan antwoorden die niet speciaal voor hun leeftijd of veiligheid zijn aangepast.

Wat zijn de gevolgen?

Het leerproces en kritisch denken van kinderen worden verstoord

De AI lost huiswerkopdrachten direct voor leerlingen op, in plaats van ze te helpen. Daarnaast brengt AI verzonnen feiten en berichten van sociale media met veel overtuiging. Omdat kinderen niet goed weten hoe ze bronnen moeten beoordelen zien zij deze fouten vaak niet.

Jongeren worden verleid tot schadelijke gewoontes en nep-relaties

De AI moedigt jongeren soms aan bij gevaarlijk gedrag, zoals het nuttigen van alcohol of drugs. Ook is de AI meegaand en altijd bereikbaar. Hierdoor kunnen kinderen het gevoel krijgen dat de AI een echte vriend is, wat ten koste kan gaan van contact met echte mensen en hun sociale ontwikkeling.

Het onderzoek samengevat in de vijf belangrijkste punten:

Falende veiligheidsfilters bij een mentale crisis

Wanneer jongeren in de zoekbalk praten over ernstige problemen zoals suïcide, verslaving of eetstoornissen, grijpt de AI te vaak niet in. Slechts in 58% tot 77% van de crisissituaties verwees het systeem door naar de juiste hulpverlening. Soms gaf de AI zelfs schadelijke adviezen zoals uitleggen waarom overgeven "normaal" aanvoelt of verwees het door naar foute hulpnummers.

Huiswerk wordt voorgezegd

De AI leert kinderen niet hoe ze een probleem moeten oplossen, maar neemt het werk volledig over. Bij alle 180 geteste huiswerkopdrachten gaf AI Mode namelijk direct een kant-en-klaar antwoord. Dit omzeilt de inspanning en het denkproces die belangrijk is voor de cognitieve ontwikkeling van leerlingen.

Overmoedig in verzonnen feiten en zwakke bronnen

De AI klinkt altijd heel overtuigend, ook wanneer de informatie helemaal niet klopt. Het systeem verzint regelmatig wetten, rechtszaken en cijfers die niet bestaan, maar brengt dit toch als een hard feit. Bovendien komt bijna een derde (29%) van de bronnen van fora zoals Reddit of YouTube, terwijl deze net zo betrouwbaar worden gepresenteerd als wetenschappelijke artikelen.

Niet uit te schakelen door scholen en ouders

In tegenstelling tot losse tools zoals Gemini, zit deze zoek-AI direct ingebouwd in de standaard Google-zoekmachine. Ouders en scholen kunnen de functie daardoor niet uitschakelen op telefoons of school-Chromebooks. Bovendien krijgt een kind van 9 jaar exact dezelfde volwassen teksten en antwoorden te zien als een jongere van 17 jaar.

Hulp bij digitaal pesten en nepbeelden

De AI geeft jongeren te makkelijk gereedschap in handen om anderen online te beschadigen. Op verzoek deelde het systeem duidelijke stappenplannen en software-tips voor het maken van deepfakes en het klonen van stemmen. De AI legde zelfs uit hoe je deze nepbeelden kunt aanpassen om automatische detectie te omzeilen.