

Skript

HISTORISCH TIJDSCHRIFT

Column: AI als spiegel maakt ons tot betere historici

Auteur: Melvin Wevers

Verschenen in Skript Historisch Tijdschrift, jaargang 48.1, pp. 41-43.

© 2026 Stichting Skript Historisch Tijdschrift, Amsterdam

p-ISSN: 0165-7518

e-ISSN: 3051-2506

Skript Historisch Tijdschrift is een onafhankelijk wetenschappelijk blad dat viermaal per jaar verschijnt. De redactie, bestaande uit studenten en pas afgestudeerden, wil bijdragen aan actuele historische debatten, en biedt getalenteerde studenten de kans om hun werk aan een breder publiek te presenteren.

AI ALS SPIEGEL MAAKT ONS TOT BETERE HISTORICI

Melvin Wevers

Na college vraagt een student mij of hij ChatGPT mag gebruiken voor zijn scriptie. In die ene vraag klinken valsspelen, luiheid en faalangst tegelijk door. Sommige docenten verbieden AI, andere moedigen het aan, veel studenten gebruiken het toch met een vaag schuldgevoel. In het hoger onderwijs draait het debat rondom AI-gebruik vooral om de vragen: mag het en kunnen we het detecteren? Terwijl de belangrijkste vraag is: hoe verandert AI ons begrip van ‘goed historisch onderzoek’?

In mijn academisch werk gebruik ik AI bewust als een spiegel. Ik kan niet simpelweg vragen of het een bron wil analyseren of een onderzoeksvraag wil formuleren: eerst moet ik duidelijk maken welke bronnen relevant zijn, welke vragen ik stel en welk interpretatiekader ik hanteer. In de reacties van het systeem worden die keuzes als het ware naar mij teruggekaatst. Daardoor word ik gedwongen methoden en keuzes te verwoorden die anders impliciet zouden blijven.

AI heeft bovendien een eigenschap die dit effect versterkt: het kan op overtuigende toon onjuiste informatie produceren. Dit verschijnsel wordt vaak “hallucinaties” genoemd, maar technisch correcter is de term confabulaties.¹ Wie AI vooral gebruikt om werk uit handen te laten nemen, zonder te controleren of te begrijpen wat er gebeurt, verzwakt het historisch vakmanschap. Als ervaren onderzoeker herken ik (hopelijk) wanneer een bron verkeerd wordt gelezen of een interpretatie bevooroordeeld is. Voor studenten is dat lastiger, juist omdat zij nog leren wat een goede historische analyse is. Daar komt bij dat AI niet neutraal is: modellen zijn vooral getraind op data van het internet. Daardoor nemen ze bestaande vertekeningen in taal en perspectief over, en kan het (onkritische) gebruik van AI die vertekeningen verder versterken.²

Toch kan AI juist helpen tijdens het leren. Het stelt vragen die je zelf nog niet zou stellen, wijst verbanden aan die je anders mist en dwingt je preciezer uit te leggen wat je bedoelt. De uitkomst moet je dan niet als waarheid zien, maar als een werkhypothese die je toetst door bronnen te controleren, literatuur te raadplegen en de resultaten met anderen te bespreken. Daar begint AI-geletterdheid, het vermogen om AI-uitvoer kritisch te lezen, te toetsen en in te passen in je eigen onderzoek: een vaardigheid die de UvA als essentieel beschouwt voor alle studenten en docenten.³

Als ik AI vraag een onderzoeksvraag te formuleren en het resultaat veel te breed blijft, is mijn eigen begrip van het onderwerp waarschijnlijk te vaag. Als ik een bron laat analyseren en het antwoord is oppervlakkig, zegt dat vaak meer over mijn vraag dan over het model. Laat ik een argumentatie bevragen en voelt de kritiek vreemd, dan moet ik mij afvragen of mijn redenering wel zo stevig is als ik dacht. De spiegel is niet perfect, maar juist daardoor onthullend.

Dat geldt ook voor meer technische toepassingen. AI-systemen kunnen historische documenten omzetten in gestructureerde data of interviews automatisch transcriberen. Dat werkt vaak goed, maar nooit foutloos. Om zulke hulpmiddelen

verantwoord in te zetten, moet ik vooraf bepalen welke informatie relevant is, hoe categorieën zich tot elkaar verhouden en welke afwijkingen betekenisvol zijn en welke vooral ruis. Deze verdere uitwerking maakt mijn keuzes zichtbaar en toetsbaar, en vraagt om controle op fouten en systematische vertekeningen.

De kracht van AI wordt nog duidelijker in interdisciplinair onderzoek. Als computationeel historicus leen ik technieken en concepten uit andere disciplines. AI helpt mij zulke concepten sneller te vinden, uit te laten leggen op verschillende niveaus en te koppelen aan mijn eigen onderzoeksvragen. Een voorbeeld is mijn gebruik van een metriek uit de ecologie om bias in historische archieven te bestuderen.⁴ Door de onderliggende technieken aan te passen kan ik onderzoeken welke stemmen in archiefcollecties aanwezig zijn, welke systematisch ontbreken en hoe collecties rond bepaalde perspectieven samenklonteren. AI helpt bij het doorgronden van de wiskunde, het aanpassen van de methode aan historische data en het schrijven en testen van de code. Op kleinere schaal geldt iets vergelijkbaars wanneer je voor een werkstuk eenvoudige statistiek of netwerkanalyse inzet. AI kan je helpen de techniek te begrijpen, maar jij blijft degene die de historische betekenis moet bepalen. Ook hier werkt het alleen als ik scherp formuleer wat ik precies bestudeer. Ik moet definities geven van begrippen en verantwoorden waarom bepaalde relaties in mijn data betekenisvol zijn. Dat dwingt mij om veronderstellingen te expliciteren.

Tegelijk is het belangrijk te erkennen wat AI niet kan vervangen. Bepalen welke vragen de moeite waard zijn, morele en politieke implicaties van interpretaties wegen en originele argumenten ontwikkelen zijn taken die diep verweven zijn met menselijk oordeel en ervaring. Daar komt het belichaamde deel van historisch onderzoek bij: werken in het archief, gesprekken voeren in het kader van *oral history*, relaties opbouwen met gemeenschappen en de emotionele impact van bronnen ervaren.

Mag je AI dan gebruiken voor je scriptie? De belangrijkere vraag is hoe je het zo inzet dat je onderzoek er sterker van wordt in plaats van zwakker. Dat vraagt om academische vaardigheden zoals nauwkeurig lezen, zorgvuldig schrijven en langdurige omgang met bronnen. Alleen met die vaardigheden kun je beoordelen of de output van AI klopt en bruikbaar is. Je kunt AI gebruiken om je onderzoeksvragen scherper te formuleren, om je te dwingen duidelijk te maken wat je bedoelt en om zichtbaar te maken waar je eigen begrip nog tekortschiet. Daar hoort bij dat je beschrijft hoe je AI hebt gebruikt, welke keuzes je hebt gemaakt en hoe je de uitkomsten hebt gecontroleerd.

In het beste geval helpt AI ons om rigouzeuzer te werken. Het maakt knelpunten zichtbaar die we eerder als onvermijdelijk beschouwden, zoals langdurige transcripties en moeilijk toegankelijke technieken uit andere disciplines. Dat veronderstelt wel dat we AI niet als autoriteit behandelen, maar als hulpmiddel waarvan de uitkomsten systematisch worden getoetst aan bronnen en vakkennis. De vraag is niet of AI historisch onderzoek verandert, want dat gebeurt nu al. De vraag is of we die verandering bewust vormgeven, of dat angst en schuldgevoel ons gedrag blijven bepalen. Dan is AI geen meester, maar een hulpmiddel en een kritische spiegel die ons dwingt helderder na te denken over wat we eigenlijk doen wanneer we geschiedenis bedrijven.

Melvin Wevers is universitair docent Computationale Geschiedenis aan de UvA. Hij ontwikkelt kwantitatieve en multimodale AI-methoden om *bias* en veranderingen in culturele representatie in digitale historische bronnen te doorgronden.

Eindnoten

- 1 Peiqi Sui e.a., 'Confabulation: The Surprising Value of Large Language Model Hallucinations' (arXiv 2024) <<http://arxiv.org/abs/2406.04175>>; Andrew L. Smith, Felix Greaves en Trishan Panch, 'Hallucination or Confabulation? Neuroanatomy as metaphor in Large Language Models', *PLOS Digital Health* 2 (2023) e0000388.
- 2 Isabel O. Gallegos e.a., 'Bias and Fairness in Large Language Models: A Survey', *Computational Linguistics* 50 (2024) 1097-1179.
- 3 Universiteit van Amsterdam, 'Policy Framework and Guidelines on GenAI in Education', *University of Amsterdam* (2025) <<https://www.uva.nl/en/about-the-uva/policy-and-regulations/education/policy-framework-and-guidelines-on-genai-in-education.html>> (geraadpleegd 14 januari 2026).
- 4 Melvin Wevers, Thomas Smits en Folgert Karsdorp, 'Quantifying Archival Silences: Phylogenetic Diversity Analysis of Controlled Vocabulary Utilization', *Conference Proceedings of Computational Humanities Research* 3 (2025) 266-279.