

# VAN OPNAME TOT ALGORITME

## Oral history in het AI-tijdperk

Afspelen. Luisteren. Pauzeren. Typen. Afspelen. Luisteren. Pauzeren. Typen. Zo ging dat 'vroeger'. Het transcriberen – spraak omzetten naar tekst – van oral history-interviews ging pakweg vijf jaar geleden nog tergend langzaam. Maar dat verandert. En niet alleen het transcriberen. In het tijdperk van kunstmatige intelligentie maakt oral history een evolutie door.

Lian van der Zon, Erik van Vliet en Else Gootjes, Erfgoed Gelderland

► Oral history legt persoonlijke herinneringen vast door middel van interviews. Het gaat om verhalen, emoties en stemmen die bewaard moeten blijven. Kunstmatige intelligentie (AI) is een verzamelnaam voor systemen die zoeken naar patronen en processen versnellen. Wat gebeurt er als de twee werelden elkaar ontmoeten? En wat heeft het erfgoedveld daaraan? Dit artikel verkent de kansen en de vragen die ontstaan wanneer menselijke verhalen en slimme systemen samenkomen.

### De stappen van een oral historian

Waar kan AI oral history bij ondersteunen? Om daar een beeld van te vormen volgen we de stappen die een *oral historian* doorgaans zet. Dat begint vaak met het doen van historisch onderzoek. Dat helpt om in kaart te brengen welke verhalen al worden verteld en welke nog niet. Wiens stem missen we? En welk perspectief kunnen we toevoegen aan de geschiedschrijving door middel van oral history? Op basis daarvan gaat de *oral historian* op zoek naar mogelijke interviewkandidaten: mensen die het verleden van dichtbij en zélf hebben meegemaakt. Uit de eerste hand dus.

Als er geschikte vertellers zijn gevonden, maakt de *oral historian* een lijst met vragen of thema's. Het interview afnemen en opnemen gebeurt idealiter één op één, in een veilige set-

ting, zodat de geïnterviewde vrijuit kan vertellen over wat hij of zij heeft meegemaakt. Eenmaal thuis met de opname gaat ze het interview transcriberen. Dit transcript gebruikt ze als basis voor een tentoonstelling, artikel, documentaire of een ander publieksproduct. Daarna gaat het transcript samen met de opname, toestemmingsformulieren, extra metadata en eventuele andere relevante documenten of afbeeldingen naar een archief. Daar wordt het veilig bewaard voor de toekomst. Onze *oral historian* heeft geschiedenis geschreven!

### AI maakt het verschil

Voor vrijwel elk van deze stappen zijn er AI-tools die het werk kunnen vergemakkelijken of versnellen. Daar is ook vraag naar. Uit de reactie,



Het handmatig transcriberen van één uur interviewopname duurt ongeveer tussen de zes en acht uur.

Het afnemen en vastleggen van een interview blijft mensenwerk (Nemesia Production).

**'Voor vrijwel elk van deze stappen zijn er AI-tools die het werk kunnen vergemakkelijken of versnellen. Daar is ook vraag naar.'**

**'In onze gesprekken met diverse experts horen we uitdagingen die specifiek of extra relevant zijn voor oral history.'**

ties op onze vragenlijst blijkt dat alle respondenten AI willen inzetten bij ten minste één onderdeel uit het oral history-proces. Hoe werkt dat dan? Een voorbeeld: bij het vooronderzoek kan de *oral historian* met één druk op de knop relevante wetenschappelijke artikelen vinden en samenvatten. Op basis van die historische informatie maakt AI daarna automatisch een vragenlijst en doet het suggesties voor mogelijke interviewkandidaten.

Dat is al ontzettend handig, maar waar AI momenteel de meeste impact maakt, is bij het transcriberen. Daan Raven van WO2Net | oorslogsbronnen.nl vertelt: 'De inzet van AI heeft geleid tot een enorme versnelling en kwaliteitsverbetering bij het automatiseren van trans-

cripties, wat bij het Indisch Herinneringscentrum 80% van het werk scheelde.' Het transcriberen van vijf minuten opname kostte eerder namelijk al gauw een uur handmatig typewerk. Met een AI-tool is zo'n klus binnen enkele minuten geklaard. Volgens journalist Laurens Vreekamp is dat nog niet foutloos, maar: 'Als er weinig ruis is en als het een bekende taal is, dan is het goed te doen.'

Ook in de latere stadia van het proces kan AI een verschil maken. De *oral historian* kan een korte, publieksvriendelijke tekst laten maken van het interview, vanuit het perspectief van de

geïnterviewde. Als er meerdere interviews zijn, kan AI ze in mum van tijd analyseren en thema's extraheren. Het kan specifieke fragmenten in interviews met hetzelfde onderwerp opsporen en aan elkaar koppelen. Als de *oral historian* het interview duurzaam wil bewaren en toegankelijk wil maken voor hergebruik, hoeft zij niet zelf de metadata te maken. AI koppelt trefwoorden uit een thesaurus aan het interview op basis van het transcript en schrijft een korte samenvatting. Zo is het te verbinden aan andere interviews en objecten binnen een archief of museumcollectie. Fantastisch, zou je denken. Maar: *what's the catch?*





Een interview wordt vastgelegd op camera, zodat met de AI-tool Descript een publiekversie gemaakt kan worden (Nemesia Production).

### Uitdagingen en risico's

Dat er uitdagingen en risico's verbonden zijn aan AI, is algemeen bekend. Denk aan privacy-vraagstukken, onjuistheden in antwoorden en de wildgroei aan veelal betaalde tools. Laurens Vreekamp legt uit dat chatbots, zoals ChatGPT, getraind zijn om positief en behulpzaam te zijn. Daardoor zullen ze niet snel 'nee' verkopen en verzinnen ze dingen als ze bepaalde informatie niet hebben. Dat is een ontwerpkeuze van de makers. Daarnaast 'hallucineren' chatbots: ze geven antwoorden die grammaticaal kloppen en logisch lijken, maar feitelijk onjuist zijn. Het zijn taalmodellen, geen waarheidsmodellen: goed in taal, niet in feiten. Voor oral history gelden deze algemene risico's natuurlijk ook. In onze gesprekken met diverse experts horen we daarnaast uitdagingen die specifiek of extra relevant zijn voor oral history. Eén cruciale is: transcriptienaauwkeurigheid. Die is van veel factoren afhankelijk en verschilt sterk per programma. Hoewel de programma's steeds beter worden, zijn automatische transcripties nog verre van perfect. Dit heeft verschillende oorzaken. Laurens licht toe: 'We gebruiken veel Engelse termen en dat maakt het voor zo'n tool meteen ingewikkeld.' Ook met dialecten, mompelen en omgevingsgeluiden kunnen de meeste programma's nog niet goed omgaan.

Een andere uitdaging is emotie- en contextblindheid. Bij oral history gaat het niet alleen om wát er wordt gezegd. Hóe het wordt gezegd en wat *niet* wordt gezegd, is net zo belangrijk. Het gaat dus ook over toon, emoties, stiltes en lichaamstaal. AI's onvermogen om deze aspecten te herkennen en te vertalen is op dit moment nog groot. Door de culturele en sociale context die ermee samenhangt, kan er

**'De ontwikkelingen gaan ontzettend snel. Hoe zal de relatie tussen AI en oral history zich de komende jaren ontfouwen?'**

bovendien snel *bias* ontstaan: AI gebruikt vooroordelen en aannames die in data zitten en die mee bepalen hoe AI iets begrijpt en weergeeft. Hierdoor ligt stereotypering op de loer, niet alleen bij de interpretatie van interviews, maar in alle stappen van een oral history-project. Laurens legt uit: 'Trainingsdata zijn gebaseerd op digitale informatie uit archieven en teksten van het web; ze komen uit de maatschappij. De maatschappij is ongelijk, dus daarom zijn ook de data ongelijk.'

### Ethische code

Een voorbeeld. Stel: onze *oral historian* gebruikt een chatbot bij het opsporen van interviewkandidaten uit ondervertegenwoordigde gemeenschappen. Dan moet zij zich heel bewust zijn dat de suggesties van de chatbot vooral het dominante narratief zullen versterken. De werkelijk ondervertegenwoordigde of zelfs onzichtbare stemmen kent AI pas wanneer die worden toegevoegd aan de data waaruit AI put. Marijn Braam van het Knooppunt Spreekende Geschiedenis ziet in die uitdaging ook een kans: 'Het zou mooi zijn als we door middel van oral history tot minder *bias* kunnen komen.' Hoe? Door nieuwe perspectieven te verzamelen met oral history en die toe te voegen aan de data die AI gebruikt. Maar daar raken we meteen aan een nieuwe uitdaging: het gebruik van oral history-materiaal als *data*.

Als de *oral historian* haar werk goed doet, laat zij de mensen die ze interviewt een formulier ondertekenen. Zo geeft de geïnterviewde toestemming voor onder andere opname en afgebakend gebruik. Zou het gebruik van AI daarin moeten worden opgenomen? En in hoeverre is het verantwoord om levensverhalen 'bloot te stellen' aan AI-systemen? En daarmee aan de commerciële bedrijven die daar vaak achter



Een interview ten huize van professor Wilhelmus Johannes Maria Antonius (Willem) Asselbergs (pseudoniem Anton van Duinkerken) (2 januari 1903 – 27 juli 1968) (links) door een reporter van de B.R.T.

zitten? Marijn Braam stelt dat het belangrijk is om daar ethisch mee om te gaan: ‘Samen met de Vrije Universiteit Amsterdam ontwikkelen we een ethische code voor het gebruik en deponeren van oral history-interviews. Daarin is AI meegenomen.’ De code wordt gemaakt met kleine gemeenschappen en initiatieven in het achterhoofd en komt op de website van Spreekende Geschiedenis.

### Toekomstperspectieven

De ontwikkelingen gaan ontzettend snel. Hoe zal de relatie tussen AI en oral history zich de komende jaren ontploegen? Jacoline Zilver-schoon, directeur van het Flipje en Streekmuseum in Tiel, schetst een toekomstbeeld: ‘Over vijf tot tien jaar gebeurt transcriptie automatisch. Vrijwilligers focussen op het gesprek en AI helpt bij het vinden van thematische verbanden tussen honderden verhalen.’ Voor sommige organisaties is dit al realiteit. Voor oorlogsbronnen.nl bijvoorbeeld zoekt AI in interviews naar fragmenten over hetzelfde onderwerp. Zo kun je een reeks persoonlijke ervaringen over één thema of gebeurtenis achter elkaar beluisteren. Daan van WO2Net | oorlogsbronnen.nl ziet nog veel meer mogelijkheden: ‘Waar we nu heel afhankelijk zijn van metadata, zal het veel meer gaan om de inhoud die daarin zit. Over vijf jaar kun je zeggen: ik vind dit archief, dat is getranscribeerd. Daar vind ik die tekst in, daar hoort dit trefwoord bij, daar vind ik nog die foto’s bij over hetzelfde onderwerp – niet alleen op basis van metadata, maar ook op analyse van beeld. Die ontwikkeling zie je nu al, maar zal veel harder gaan.’

Het zijn veelbelovende perspectieven, maar de *oral historian* moet alert blijven op valkuilen en risico’s bij het gebruik van AI. Automatische transcripten zal ze blijven valideren. AI-onder-

## ‘Samen met de Vrije Universiteit Amsterdam ontwikkelen we een ethische code voor het gebruik en deponeren van oral history-interviews. Daarin is AI meegenomen.’

zoek en -analyse zal ze blijven onderwerpen aan haar eigen, menselijke controle. En als ze zich zorgen maakt over bijvoorbeeld privacy, dan heeft journalist Laurens Vreekamp daar in de nabije toekomst een oplossing voor. Want de afhankelijkheid van grote multinationals, die een verdienmodel zien in onze persoonlijke data, zal volgens hem afnemen. Laurens stelt dat er een verschuiving gaat plaatsvinden naar gedecentraliseerde, open source AI-modellen. ‘We gaan dan niet meer naar chatGPT of NotebookLM. Maar met de technologie die open source voorhanden komt, bouwen organisaties zelf hun modellen. Die modellen kunnen we specifiek trainen op eigen onderwerpen en dialecten. Zo kunnen we ons losweken van Big Tech. ‘Dat zie ik wel in vijf jaar gebeuren,’ aldus Laurens.

### Conclusie en aanbevelingen

Eigenlijk is het geschetste toekomstbeeld er al. In theorie kan kunstmatige intelligentie het overgrote deel van het oral history-proces uit handen nemen. Of de *oral historian* daar tenminste drastisch in ontlasten. Welke stappen en voorzorgsmaatregelen moeten we nu nemen om de verdere ontwikkeling in goede banen te leiden? We formuleren, samen met de geïnterviewden in dit artikel, drie aandachtspunten:

### Weet wat je doet.

Om AI in te kunnen zetten moet je globaal weten hoe het werkt. Welke onderdelen van een oral history-project kun je door AI laten doen? Welke moet je stevig controleren? En welke kun je beter zelf in de hand houden? Die informatie verkrijgt je bijvoorbeeld door zelf te experimen-

teren met verschillende programma’s. Kijk eens wat er gebeurt met een (test)opname als je AI erop loslaat. Probeer verschillende programma’s en kies wat bij jou past. Als erfgoedveld moeten we ook meer kennis opbouwen over AI. Dat kan door samen te werken met experts, door codes als die van Spreekende Geschiedenis te ontwikkelen en door trainingen voor *oral historians* te verrijken met tips om AI verantwoord te gebruiken.

### Verlies de menselijke maat niet uit het oog.

Oral history is mensenwerk, per definitie. Dat zal het altijd blijven. AI kan processen versnellen en vergemakkelijken, maar nooit de menselijke maat vervangen. Zorg daarom dat de output van AI altijd wordt gecontroleerd door mensen van vlees en bloed. Blijf kritisch, evalueer je AI-gebruik periodiek en blij trouw aan het verhaal van de geïnterviewde. Wees je bewust van je eigen *bias* en die van AI.

### Zie het als een hulpmiddel, niet als oplossing.

Momenteel kan AI al veel, maar er is nog steeds handmatige correctie en evaluatie nodig. De verschillende stappen van oral history kosten dankzij AI een stukje minder moeite, maar vereisen nog steeds tijd en aandacht. Dat merkten we ook bij het maken van dit artikel. AI hielp ons goed op weg door de interviews te transcriberen en automatisch thema’s aan te wijzen. Maar voor het schrijven van een helder en menselijk verhaal waren overleg, herschrijven, controle en verdiepend onderzoek onmisbaar. ◀

*De auteurs van dit artikel interviewden experts en verzamelden reacties via een digitale vragenlijst. Bij het schrijven gebruikten we ChatGPT, NotebookLM, Whisper, Microsoft Teams transcriptiefuncties. En eigen common sense.*

### Aan de slag met AI en oral history?

Onderstaand stappenplan vraagt maar een beperkte technische kennis.

1. Voorbereiden: ChatGPT voor planvorming, thema’s, interviewvragen
2. Transcriberen: aTrain voor het omzetten van audio naar tekst
3. Verwerken: NotebookLM voor het samenvatten en interpreteren van interviews
4. Publiceren: Descript voor het maken van publiekvideo’s
5. Archiveren: Annif voor metadatering