01 SCHAT

02 HERLEID

03 TOETS

04 BESLUIT

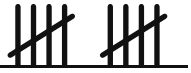
HOE BETROUWBAAR IS DIT CIJFER?

# “62% VAN DE ONDERZOCHTE NEDERLANDSE BOEKEN GAAT OVER EEN MAN.” GETELD MET BEHULP VAN EEN TAALMODEL.

## MAG JE DAT OVERNEMEN?

Een les in drie dingen: de betrouwbaarheid van een cijfer toetsen, kritisch blijven op wat AI oplevert, en zien waarom controle door mensen nodig is.

TIEN GEVALLEN, TWEE BEOORDELAARS



VOOR WIE

Studenten in het hbo en op de universiteit die met cijfers werken, vanaf het tweede jaar. Journalistiek is hier het voorbeeld

DUUR

60 minuten, in tweetallen

NODIG

Een cijfer dat met behulp van AI is geteld, met een openbare verantwoording en openbare gegevens, en het toetsblad. Geen voorkennis van statistiek of programmeren

01

## SCHAT

Hoe betrouwbaar lijkt het cijfer je, op een schaal van 0 tot 10? Schrijf het op, met één reden.

02

## HERLEID

Lees de pagina waarop staat hoe er geteld is, de verantwoording, met vijf vragen. Waar komt het vandaan? Wat is er geteld, en wat niet? Wie of wat gaf het oordeel? Wat staat er over de marge? Wat is door een mens nagekeken?

03

## TOETS

Neem met z'n tweeën dezelfde tien gevallen. Oordeel ieder apart, zonder het oordeel van het model te zien. Leg er daarna pas het model naast.

04

## WEEG

Reken uit hoe vaak jullie het onderling eens zijn, en hoe vaak met het model. Bij elk verschil: lag het aan het model (M), aan een van jullie (J), aan een onduidelijke regel (R) of aan de bron (B)?

05

## BESLUIT

Zou je dit cijfer overnemen? Schrijf de zin op zoals jij hem zou publiceren, met de kanttekening die erbij hoort.

OM TE BESPREKEN

## ALS TWEE MENSEN HET ONDERLING NIET EENS WORDEN, WAT ZEGT EEN VERSCHIL MET HET MODEL DAN NOG? EN HOEVEEL OVEREENSTEMMING IS GENOEG OM EEN CIJFER TE PUBLICEREN?

TOETSBLAD

Vul bij elk geval drie oordelen in: dat van jou, dat van je medestudent en dat van het model. Alleen als die verschillen, zet je in de laatste kolom waar het aan lag.

| NR | GEVAL | MIJN OORDEEL | MEDESTUDENT | MODEL | VERSCHIL DOOR |
|---|---|---|---|---|---|
| *vb.* | *titel van het boek* | *vrouw* | *vrouw* | *man* | *M* |
| 1 | | | | | |
| 2 | | | | | |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |
| 7 | | | | | |
| 8 | | | | | |
| 9 | | | | | |
| 10 | | | | | |

Verschil door: **M** het model had het mis · **J** een van jullie had het mis · **R** de regel is onduidelijk · **B** de bron geeft geen uitsluitsel

**SCHATTING VOORAF** _____ VAN 10 **ONDERLING EENS** _____ VAN 10 **EENS MET MODEL** _____ VAN 10

**De klas telt op.** Tien gevallen zeggen weinig. Tel daarom de uitkomsten van de hele klas bij elkaar op: tien tweetallen is honderd gevallen, en daar kun je iets over zeggen.


Vrij te gebruiken en te bewerken, met vermelding van The Missing Measure (CC BY 4.0).

WAAROM DEZE LES

## EERST TOETSEN, DAN GEBRUIKEN

Steeds meer cijfers zijn met behulp van een taalmodel geteld, in het nieuws, in beleid en in onderzoek. Wie zo'n cijfer gebruikt, moet kunnen nagaan hoe betrouwbaar het is. De journalistiek is in deze les het voorbeeld, omdat dat vak het heeft vastgelegd: het landelijk opleidingsprofiel Journalistiek (2023) rekent het beoordelen van bronnen en datasets en het verifiëren van informatie tot de kern, en de modelrichtlijn van de NVJ zegt dat wat AI oplevert altijd op juistheid wordt gecontroleerd en nooit de enige basis voor een feit mag zijn. Voor andere vakken geldt hetzelfde.

WAT ONDERZOEK LAAT ZIEN

## NAKIJKEN GAAT VAAK MIS

In een experiment met 2.784 deelnemers die voorstellen van AI moesten nakijken, werden fouten waarvoor je de regels moest kennen in ongeveer een derde van de gevallen opgemerkt. Wie AI meer vertrouwde, liet meer fouten staan. Daarom oordelen studenten in deze les eerst zelf, en zien ze het oordeel van het model pas daarna.

Beck e.a., Bias in the Loop, 2025, voorpublicatie

NA DEZE LES

## WAT EEN STUDENT KAN

Na deze les kan een student een cijfer herleiden tot bron, methode en marge, een steekproef nalopen voordat het oordeel van het model zichtbaar is, en zeggen met welke kanttekening het cijfer te publiceren is. Dat is te zien aan de zin die elke student aan het eind opschrijft.

VERLOOP IN MINUTEN, ALS RICHTLIJN

**10** Schatten, en het cijfer herleiden met de vijf vragen

**25** Tien gevallen toetsen, ieder apart, het model pas daarna

**15** Overeenstemming uitrekenen en de verschillen wegen

**10** De klas telt alles bij elkaar op, besluit en bespreekt

PRAKTISCH

## ELK CIJFER MET EEN OPEN VERANTWOORDING

**Past bij** onderzoeksvaardigheden, methodenonderwijs of datajournalistiek, en bij elk vak waarin studenten cijfers overnemen. De les staat op zichzelf.

**Vooraf** Kies een telling waarvan de verantwoording en de gegevens openbaar zijn, en tien gevallen daaruit, ongeveer een half uur werk. Bij Flaptekst staat elk boek met het oordeel erbij, bij Dutch Angle staat bij elk oordeel de zin waar het op rust. AI is het brede woord, een taalmodel is het soort AI dat hier het leeswerk deed.

**Gastdocent** Een gastdocent van The Missing Measure kan de les ook geven, als losse les of op maat voor de opleiding. Prijs op aanvraag, via astrid@missingmeasure.com.

**Heb je de les gegeven?** We horen graag hoe het was, en feedback is welkom, via astrid@missingmeasure.com. We zijn ook benieuwd naar de totalen van de klas, voor wie ze wil meesturen; die worden niet zonder toestemming gepubliceerd.


Vrij te gebruiken en te bewerken, met vermelding van The Missing Measure (CC BY 4.0).