

Interview **Natali Helberger**

Steeds vaker bepalen onzichtbare algoritmen, en niet de redacties, welke berichten we lezen.

In de hoogtijdagen van de papieren krant stond de redacteur centraal. Die bepaalde wat de lezer voorgeschiedt kreeg: hij of zij was 'de poortwachter' en organiseerde de toegang tot het nieuws en de achtergronden.

In het digitale tijdperk is het steeds vaker een wiskundige formule die bepaalt wat de lezer op zijn scherm vindt. Nu 30% van de volwassenen in de VS het nieuws primair via Facebook 'consumeert', is de poortwachter niet langer de nieuwschef van de redactie. Sterker nog: de poortwachter is niet eens meer een mens. In feite is degene die de nieuwsvoorziening regisseert een algoritme. Het zijn nu de onderliggende computerprocessen van Facebook, en ook van The Huffington Post en BuzzFeed, die dicteren welke nieuwsitems naar boven komen drijven op de 'newsfeed' — en in welke volgorde.

Deze selectie en verspreiding van nieuws lijken wetenschappelijk en objectief, maar zijn het niet. Algoritmen zijn verre van neutraal. De wiskundigen die ze bouwen, steken er bewust en onbewust allerlei vooronderstellingen in. Wat een 'trending topic' op Twitter wordt, is afhankelijk van de manier waarop wiskundige modellen de stroom aan berichten wegen en ordenen. Het algoritme bepaalt dus of de prikkelende foto's van Kim Kardashian belangrijker zijn dan het slepende conflict in het Midden-Oosten: welkom in de enerverende, verwarrende wereld van de journalistiek anno 2014.

Kranten zijn momenteel vooral bezig met de vraag hoe ze de overgang naar een digitaal businessmodel moeten overleven — niet met hoe zij hun gedragscodes moeten aanpassen aan de nieuwe spelregels. Het debat over de journalistieke ethiek in het algoritmetijdperk wordt nog maar mondjesmaat gevoerd.

Een van de mensen die dit debat wél aangaan is de Duitse Natali Helberger, hoogleraar informatierecht

Beïnvloeding Algoritmen lijken neutraal, maar zijn dat beslist niet. Ze berusten bewust en onbewust op allerlei aannames

aan de Universiteit van Amsterdam. Helberger is vooral geïntrigeerd door de invloed van 'personalization', 'profiling' en 'targeting' op ons: de moderne digitaal-nieuwsconsumenten.

Dankzij cookies en andere technieken weten de media — en ook adverteerders en andere derden — meer dan ooit over de lezer. Door het instellen van persoonlijke voorkeuren op sites zoals die van de Frankfurter Allgemeine Zeitung ('Mein FAZ.net) en ook die van het FD ('Mijn nieuws'), werkt de consument hieraan mee.

'Dit zijn geweldige nieuwe mogelijkheden', zegt Helberger. 'Er is al zo veel informatie beschikbaar, dan is een soort filter handig en goed. Maar er schuilen ook gevaren in personalisatie. Zo ontstaan er mogelijkheden om de lezer te manipuleren voor doeleinden van de adverteerder, en op termijn zelfs door de politiek.'

In de praktijk is er geen bewijs voor dat de politiek ons daadwerkelijk manipuleert via de werking van algoritmen, maar theoretisch is het wel degelijk mogelijk, terwijl we dit zelf niet zouden merken. De media moeten daarom open en eerlijk praten over de neveneffecten van de nieuwe digitale technologieën, zegt Helberger.

'Er bestaat zoiets als een vertrouwensrelatie tussen de media en het publiek. Ik zou bijvoorbeeld niet willen dat mijn favoriete nieuws-sites mijn leesgedrag delen met adverteerders, zodat die dat kunnen gebruiken om mij te beïnvloeden via de



FOTO: PETER BOER


'Personaliseren sites leidt tot manipulatie'

selectie van de redactionele inhoud die op mijn scherm verschijnt.'

Hoe meer informatie over de lezer, hoe beter, zou je denken. De journalistiek brengt wat de lezer wil. 'Vroeger dachten media in termen van doelgroepen: huisvrouwen bijvoorbeeld, of beleggers, of fans van oude auto's. Nu zoomen ze in op het individu. Elke keer dat je op het internet surft, of iets bestelt op Amazon.com of iets op Facebook zet, wordt een nieuw stukje aan jouw profiel toegevoegd. Het is een lerend systeem. Jouw profiel wordt steeds gedetailleerder.'

Waarin schuilt het gevaar? 'Het is goed dat de media het publiek willen bedienen. Maar waar ligt de grens? In hoeverre moeten de media zich aan het dictaat van de gebruiker aanpassen? En hoeveel zouden de media van de gebruiker moeten weten? Kennis is macht: de macht om je te kunnen beïnvloeden en zelfs te manipuleren. En die machtsverhouding verschuift in de richting van degene die de controle heeft over de algoritmen.'


Heeft u een voorbeeld? 'Kijk naar The Guardian, dat 18 cookies gebruikt en op sommige delen van hun site er wel 38 achterlaat op je computer. Of kijk naar het FD: jullie delen informatie met Google Analytics.'

 *Vervolg op pagina 5*

The Age of Algorithms
Outlook Weekend onderzoekt in 2015 de kansen, risico's en dilemma's van het nieuwe datatijdperk. In welke sectoren zijn algoritmen en big data al niet meer weg te denken? Het komende jaar richten we ons op onder meer de volgende sectoren:

- **Financiële dienstverlening** Banken kennen de transacties van hun klanten, maar die willen dat ze ervan afblijven. De handelsvloer kan al niet meer zonder algoritmen.
- **Gezondheidszorg** Gepersonaliseerde zorg en preventie worden beloofd. Data-analyse moet die leveren, maar patiënten vertrouwen het model nog niet.
- **Handhaving** Computers kunnen profielen creëren van mogelijke overtreders. Dat maakt het werk van fraudejagers efficiënter. Maar het individu kan onterecht bespied worden.
- **Energie** Wiskundigen helpen de energiebedrijven om zelf in gas en licht te handelen. Ze maken zelfs weersvoorspellingen. En ze weten hoe laat de consument thuiskomt.
- **Retail** Big data verbeteren logistiek en marketing. Maar wat merkt de klant hiervan?

The Age of Algorithms

 Vervolg van pagina 3

'Het is goed dat mediabedrijven de lezer willen kennen, maar het delen van deze informatie met externen is riskant. Het opent de mogelijkheden dat derden — Google Analytics, Amazon.com, of de marketingafdeling — zich gaan bemoeien met de inhoud, met beslissingen die eigenlijk de redactie toebehoren.

Als de adverteerder aan mijn profiel ziet dat ik bijvoorbeeld van reizen houd, zou de druk kunnen toenemen om mij meer van dit soort verhalen aan te bieden. De integriteit van de redactionele beslissing kan dus in gevaar komen door gedetailleerde kennis van individuele lezers.

Nu al houden mediaconcerns hun "best gelezen" verhalen nauwkeurig bij. Dit beïnvloedt dan op zijn beurt weer de nieuwsketuzes die de redacties daarna maken. Vroeger was dat puur een redactioneel besluit, nu wordt die beslissing gestuurd door het systeem.'

U vreest manipulatie door adverteerders. Maar er staan toch 'Chinese muren' tussen redactie en commercie. Is dit niet voldoende?
'Adverteerders zijn de grote stuwende krachten achter de trend van personalisatie en het vergaren van steeds meer informatie over de gebruiker. De gedragscodes van redacties maar ook de regelgeving en het toezicht op de media stammen uit de tijd van vóór de algoritmejournalistiek. Momenteel vallen profiling en targeting onder privacywetgeving, terwijl dat eigenlijk



Natali Helberger, hoogleraar informatierecht aan de UvA.

FOTO: PETER BOER

ook een onderwerp voor de mediawetgever zou moeten zijn.'

Waar schort het aan?

'De regelgeving rond bijvoorbeeld minderjarigen is niet van deze tijd. Profiling van kinderen gebeurt nu al via cookies en andere methodes. Terwijl minderjarigen zich juist in een kwetsbare fase van hun ontwikkeling bevinden. Je moet niet willen dat ze in een virtueel profiel zitten opgesloten dankzij de informatie die Google Analytics over ze heeft.

Ik vind trouwens dat je niemand, ook niet volwassenen, moet "profilen" als ze gewoon op nieuwspagina's zitten of items over meer gevoelige onderwer-

pen zoals politiek of religie lezen. Geef mensen hier de ruimte om hun ideeën te vormen en testen zonder dat zij gevolgd en gemonitord worden. Vooral het bijhouden van de politieke en religieuze interesses van gebruikers vind ik te ver gaan, terwijl dat voor sportitems misschien niet geldt.

Misschien moeten we constateren dat het acceptabel is voor mediabedrijven om hun gebruikers te leren kennen, maar niet dat bepaalde informatie gedeeld wordt met de Amerikaanse inlichtingendien-

Monitoren 'Vooral het bijhouden van de politieke en religieuze interesses van lezers gaat te ver'

sten, Google Analytics of Adidas.

Het punt is dat de media hier een debat over moeten voeren, maar dat ontbreekt momenteel.'

In uw werk waarschuwt u voor het gevaar van politieke manipulatie via algoritmen. Gebeurt dat al?

'Afgelopen juni werd bekend dat Facebook een experiment heeft uitgevoerd waarbij de stemming van gebruikers beïnvloed werd door middel van de selectie van berichten op hun tijdslijn. Als de politiek greep krijgt op de media, kan dat theoretisch dus ook tot politieke beïnvloeding leiden. Hoe? Stel dat er in de aanloop tot verkiezingen bewust steeds meer verhalen op mijn tijdslijn verschijnen over pensioenen of andere issues waar ik me normaal

niet zo voor interesseer. Dat is mogelijk door onzichtbare wijzigingen in algoritmen.'

Is de cookieregelgeving contraproductief? In de zin dat mensen toch overal ja op zeggen om zo snel door te kunnen surfen?

'Als je elke cookiehandleiding zou moeten lezen voor elke site die je op een dag bezoekt, ben je langer bezig met het doorspitten van de kleine lettertjes dan met het eigenlijke internetgebruik. Het is te gemakkelijk om te zeggen dat consumenten "informed consent" (weloverwogen toestemming, red.) hebben gegeven als ze ja tegen cookies zeggen. De media moeten er zelf over nadenken of cookies thuishoren op nieuwssites. Nu schuiven ze de verantwoordelijkheid af. "Take it or leave it, do or die." Het is tijd dat de journalistieke gedragscodes hierop aangepast worden en ook op andere punten. En dat er hierover een debat wordt gevoerd, ook intern.

Vergeet niet dat het om meer gaat dan cookies. Cookies zijn tenminste gereguleerd. Door "cannvassing", oftewel "digital fingerprinting", kan jouw digitale gedrag ook worden gemonitord.

De media moeten dus beslissen wat fair is ten opzichte van de consument en wat niet. En wat niet fair is, moeten ze simpelweg niet doen.'

 **Ronald van de Krol** is redacteur van Het Financieele Dagblad.



Web scraping

Web scraping is een andere term voor data-extractie van het wereldwijde web. 'Scraper'-programma's gaan internet-sites af, slaan gegevens op (ongeacht het formaat van de oorspronkelijke data) en laten daar analyses op los. Op deze manier kunnen emailadressen, telefoonnummers en andere data worden vergaard. Ook onderzoeksjournalisten kunnen trouwens scraping inzetten om gevoelige informatie uit ruwe data te halen.

