



**vakblad voor
informatieprofessionals**



Datahuis Wadden

Zee van informatie
op één plek online

**Galmt het hard
in de echokamer?**

Deel 3 van Achter
het nepnieuws

'Review' ASReview

Supersnel vinden
in grote datasets

**Twintig jaar
spraakherkenning**

Hoe onderzoek
in AV-archieven
eenvoudiger werd

2021 | 05



**LOST IN
DIGITIZATION
BRUG BOUWEN
OVER DE
DIGITALE
KLOOF**

COLOFON

IP is samen met Informatieprofessional.nl het onafhankelijke platform voor de informatiespecialist van vandaag en morgen. De lezers werken in de informatie-, bibliotheek-, archief- en erfgoedwereld.

ISSN: 1385-5328

IP is een uitgave (25ste jaargang) van Uitgeverij IP
Charlotte van Pallandtlaan 18,
2772 TR Voorburg
tel. 06-223 75 75 9
www.informatieprofessional.nl

redactieadres

IP, Charlotte van Pallandtlaan 18,
2772 TR Voorburg,
tel. 06-223 75 75 9,
e-mail redactie@informatieprofessional.nl.

redactie

Marjo Bakker, Maarten Brinkerink,
Wilbert Helmus, Dafne Jansen,
Vincent M.A. Janssen,
Leen Liefsoens, Cindy Lammers
(hoofdredacteur), Matthijs van
Otegem, Mirjam Raaphorst,
Eric Sieverts en Daniël de Vette

vormgeving

Eric van den Berg,
egfvdberg@upcmail.nl,
Tom van Staveren,
graphicisland@upcmail.nl.

medewerkers aan dit nummer

Wouter Brongseest, Joris van
Dierendonck, Anne van den Dool,
Rob Feenstra, Frank Huysmans,
Marja Kingma, Anneke de Maat,
Marijn Mostart, Tineke van
Ommeren, Roeland Ordelman,
John van de Pas,
Astrid van Wesenbeeck

abbonementen

Voor abonnementsprijzen
en andere informatie zie
Informatieprofessional.nl.

advertentieverkoop

Voor informatie over adverteren
in blad of op de site:
Rajin Roopram, e-mail rajin.roopram
@informatieprofessional.nl,
tel. 06-15201724.

Het verlenen van toestemming
tot publicatie in dit tijdschrift
strekt zich tevens uit tot het
in enigerlei vorm elektronisch
beschikbaar stellen.

INHOUD



10

Waddeninfo op één plek

De Wadden zijn werelderfgoed dat beschermd moet worden. Daarvoor zijn goede data nodig, dus kwamen er het Datahuis Wadden en aanverwante diensten. Inkijkje in een lijvig project.

16

Achter het nepnieuws #3

In een serie van zes artikelen duikt Vincent M.A. Janssen, complotdenker-in-opleiding, in de wereld van misinformatie, fake news en aanverwante begrippen. Deel 3: de echokamer.



18

Systematic reviewing

De open source-tool ASReview maakt het mogelijk om in reusachtige datasets snel relevante informatie te vinden. In 2019 won het een Victorine; hoe staat het er anno 2021 mee?



RUBRIEKEN

04

Data in beeld
Acid Clouds

09

IP Lingo
Fieldlabber

21

Column Frank Huysmans
De tao van e-books

35

Buitenland
British Library of the North



26

Spraakherkenning

Automatische spraakherkenning inzetten om AV-archieven beter doorzoekbaar te maken, daar wordt al twintig jaar aan gewerkt. Er zijn tussen veel stappen gezet. Roeland Ordeman blikt terug.

WIM

Het is een spontane ontmoeting; een oudere heer die de deur voor me openhoudt en vriendelijk mompelend iets zegt over 'galant zijn'. Wim, aangenaam. Voor we het weten hebben we een leuk gesprek. Op straat, zonder mondkapje, van mens tot mens omdat het kan. Ik moet zijn leeftijd raden en hij geeft een voorzetje door gelijk maar die van mij te schatten. Tien jaar jonger – ik mag hem natuurlijk meteen. Zelf blij ik nog royaler met mijn 75. Wims gulle lach als reactie weerkaatst tussen de oude stadsmuren, in zijn ogen een sprinkel: 88, mevrouwje!

Hij ziet er vitaal uit, praat open en vol vuur. Uit zijn zak haalt hij trots zijn 'smarttelefoon', klikt op



REDACTIONEEL | Cindy Lammers

een groen icoontje: 'Hiermee klets ik elke dag met mijn maat. Kijk, net afgesproken om te gaan visser.' Meerdere keren hoor ik hem praten over zijn 'compiejoeter' thuis en over 'bifi' (lees: wifi), maar dat zal verder worst zijn. Wim redt zich wel. En dat is meer dan knap voor iemand die ruim acht decennia heeft moeten hink-stap-springen van radio naar tv naar computer naar smart device naar Internet of Things.

Wim wil en kan mee, maar dat is helaas niet iedereen gegund, zo blijkt ook uit een artikel in dit blad. De digitale kloof is groeiende. Laten we in onze digitaliseringsdrang niet vergeten ook bruggenbouwer te zijn.

<

34

De KB gaat FAIR

Een nationale bibliotheek die al het tekstuele erfgoed van Nederland digitaliseert en beschikbaar stelt kan niet zonder FAIR. De eerste stappen zijn in januari dit jaar gezet.



22

Digitale samenleving

Leven en werken in een digitale wereld heeft veel voordelen, maar leidt tegelijk tot een digitale kloof tussen zij die wel en niet meekunnen. Dat moet anders, betoogt John van de Pas.

36

Gelezen

Studie- en leesvoer voor IP'ers

38

Gespot

Handgeschreven bronnen

40

Verenigingen

Beroepsverenigingsnieuws

42

Lifehacking

Informele ontmoetingen

43

Favorieten

MarieCiliène Ogvil





Meer over
Acid Clouds
lees je op
pagina 6.



DONKERE DATAWOLKEN

Een goed deel van de kleine tweehonderd is intussen vastgelegd. Moderne gebouwen vaak, strak, met harde hoeken, glimmend in fullcolour, in schemertijd of volledig donker. Het gaat om datacenters in Nederland. Met het blote oog niet meteen te herkennen als zodanig, maar ze bestaan. Zo ook dit DataCenter Fryslân 2 in Leeuwarden. De foto maakt deel uit van het project *Acid Clouds* van Niels Schrader en Roel Backaert. Een zoektocht in woord en beeld naar de impact van de digitale wereld op onze fysieke leefomgeving. Daar mogen we best wat van vinden, vinden ze.



Vervolg
van
pagina
4 - 5



DATACENTERS IN BEELD: NO CLOUD NINE

De cloud, oneindig en onverzadigbaar in digitale termen, heeft wel degelijk een 'begrenzing'. 'Een fysieke manifestatie die lawaai uitstoot en ruimte en energie inneemt: het datacenter', aldus de initiatiefnemers van onderzoeksproject *Acid Clouds*.

Ze staan soms op onverwachte plekken. Een blok midden in een woonwijk, zoals Dataplace Arnhem. Of een hoge toren langs de weg – Cellnex heeft op 24 plaatsen in Nederland de KPN-torens omgebouwd tot datacenters. Of ze huizen juist op een logische locatie, zoals Equinix AM4 Data Center op het Science Park Amsterdam, 'met een ontwerp dat doet denken aan de iconische harde schijf van Philippe Starck', aldus Niels Schrader, designer (met voorliefde voor technologie) en hoofd van de afdeling Grafisch ontwerp aan de Haagse Koninklijke Academie van Beeldende Kunsten. Samen met fotograaf Roel Backaert bracht hij een goed aantal van de kleine tweehonderd Nederlandse datacenters in beeld voor hun onderzoeksproject *Acid Clouds*.

Meer bewustwording

Data zijn het nieuwe goud, weten beiden, maar we mogen (moeten) kritisch zijn over de toenemende dataficering, vinden ze. Vanwege het slurpen van energie, al tonen datacenters verbetering op dit vlak, maar evengoed interessant is hoe het narratief van 'the cloud' is ont-

staan en wordt ingezet, zegt Schrader op waag.nl. 'Je ziet al wolkjes in de eerste schetsen van ARPANET. Nu wordt het als concept strategisch ingezet om het "schone", "zachte" en "onzichtbare" gevoel te creëren. Dit in combinatie met de opgeruimde, klinische en strakke inrichting van de datacenters zelf, moet het datacenters als schone en noodzakelijke virtuele omgevingen neerzetten, terwijl het grote blokken zijn die in onze publieke ruimte staan.' Hij hoopt op meer bewustwording van het feit dat datacenters de controlekamers van het internet zijn. Zwaarbeveiligd en ontoegankelijk, wat ook vragen oproept over eigenaarschap en transparantie. 'Hoezo mag ik er niet naar binnen? Het zijn toch ook mijn data? Willen we weer grip krijgen op wat datacenters wel en niet bijdragen aan maatschappelijke waarden en belangen, dan moeten de burgers worden betrokken bij de discussie en de besluitvorming hierover.'

Openlucht tentoonstelling

Het project *Acid Clouds* heeft in principe geen einddatum; 'misschien wel als we Nederland compleet in kaart hebben gebracht', reageert Schrader desgevraagd. Over een jaar zal een boek verschijnen met alle foto's en wat aanvullende inhoudelijke reflecties. 'We zijn hierover in gesprek met uitgeverij nai010 in Rotterdam. Verder zijn we op dit moment bezig om samen met Waag voor het najaar een tentoonstelling in de openbare ruimte in te plannen. Daarin worden de foto's van de datacenters in de vorm van grote billboards als openluchttentoonstelling neergezet. Binnenkort is er hopelijk ook nog meer te zien op acidclouds.org.' <

Equinix AM4 Data Center in Amsterdam.



Roel Backaert

ACTIEPLAN 'INVESTEREN

De rijksoverheid heeft onlangs het actieplan Open op Orde gelanceerd. Hiermee wil ze haar informatiehuishouding verbeteren en transparanter maken.

Het actieplan werkt het onderdeel informatiehuishouding uit van de kabinetsreactie op het verslag van de Parlementaire ondervragingscommissie Kinderopvangtoeslag (POK). Informatie is een wezenlijke grondstof voor de werking van de rijksoverheid en vitaal voor verantwoording in een democratie, zo vindt de overheid zelf. Uit vele rapporten en onderzoeken, waaronder het rapport *Ongekend onrecht* naar aanleiding van de POK, blijkt dat verbetering van de informatiehuishouding noodzakelijk is.



OPEN OP ORDE OVERHEID: IN DE INFORMATIEPROFESSIONAL'

Het actieplan bestaat uit een centraal – rijksbreed – deel, waaronder het opstellen van kaders, maar ieder departement is ook ver-

antwoordelijk voor de eigen informatiehuishouding. Daarom maken alle rijksorganisaties voor 1 juli 2021 tevens een eigen actieplan.



In hoofdlijnen gaat het om vier belangrijke thema's:

- > investeren in de informatieprofessional;
- > meer inzicht in hoeveel en welke informatie de rijksoverheid heeft;
- > betere en gebruiksvriendelijker ICT-systemen;
- > bijhouden en laten zien dat het Rijk aan de wettelijke en maatschappelijke eisen voldoet.

Meer informatieprofessionals

Het actieplan wil binnen twee jaar de meest urgente knelpunten aanpakken. Daarna volgt een programma om 'structureel en blijvend de informatiehuishouding in de volle breedte te verbeteren'. Iedereen bij de rijksoverheid zal hier wat van gaan merken, waar je ook werkt binnen het Rijk, zo stelt het actieplan. Er komen trainingen, er wordt sterker gestuurd en er komen informatieprofessionals bij. 'Belangrijk is dat iedere rijksambtenaar zich verantwoordelijk voelt voor het goed omgaan met informatie en dat het Rijk merkbare verbeteringen laat zien', aldus het plan.

OFFENSIEF TEGEN ONTLEZING ONDER JONGEREN MET 'READIFICATION'

De Leescoalitie en Ubisoft willen de ontlezing onder jongeren aanpakken met 'readification': een nieuwe vorm van leesbevordering waarbij populaire games de basis vormen voor nieuwe verhalen. Ronald Giphart en Margje Woodrow hebben voor dit doel Verborgen Verhalen geschreven die zich afspelen in de wereld van de game Assassin's Creed Valhalla.

De samenwerking tussen De Leescoalitie en gameontwikkelaar Ubisoft komt op een belangrijk moment, stellen beide partijen: het leesplezier onder jongeren in Nederland gaat achteruit. Zo vindt bijna de helft van de Nederlandse 15-jarigen lezen een tijdsverspilling, en 60 procent leest alleen als het echt moet. De Leescoalitie wil deze trend ombuigen door aan te sluiten bij de belevingswereld van jongeren, en dat zou kunnen via gaming.

'Unieke leeservaring' door extra's

Ter inspiratie voor het schrijven van de Verborgen Verhalen lieten Giphart en Woodrow zich onderdompelen in de Assassin's Creed Valhalla-game met hulp van YouTuber Jordan 'JorRaptor' van Andel. Op zijn YouTube-kanaal met meer dan 650.000 abonnees neemt hij wekelijks gamers over de hele wereld mee langs de virtuele werelden uit verschillende games. Het leverde de verhalen *Het grootse*

feest voor Bent (Giphart) en *Alf* (Woodrow) op, die inmiddels te lezen zijn via de Ubisoft Special-app. Lezers van de Verborgen Verhalen krijgen een 'unieke leeservaring' door extra visuals, videobeelden en geluidseffecten. Daarnaast is het mogelijk om tijdens het lezen meer uitleg en verdieping te krijgen over historische termen en gebeurtenissen uit het verhaal.

Pilot met Valhalla

De publicatie van de eerste Verborgen Verhalen vormt de pilot van deze samenwerking met de gamewereld, waarbij Stichting Lezen onderzoekt wat het effect is op zowel het lezen zelf als het imago van lezen. Afhankelijk van de uitkomsten van dit onderzoek wordt gekeken of bij een eventuele verdere uitrol in de toekomst gamers ook kunnen worden beloond na het lezen van verhalen.

Save the date!

De VOGIN-IP-lezing 2021 vindt definitief plaats op 21 oktober. Locatie: de Openbare Bibliotheek Amsterdam (OBA). De eerste spreker in inmiddels ook bekend: militair, academicus en softwarespecialist Jelle van Haaster. Check voor updates over het evenement: vugin-ip-lezing.net en informatie-professional.nl.

ONDERZOEK NFI EN UVA NAAR HERKENNEN DEEP- FAKES EN VERBORGEN BERICHTEN VAN CRIMINELEN

Het Nederlands Forensisch Instituut (NFI) en de Universiteit van Amsterdam (UvA) gaan samen onderzoek doen naar computermodellen om deepfakes en verborgen berichten te helpen herkennen. Dat onderzoek zal plaatsvinden in een onderzoekslab bij het ICAI.

Het gezamenlijke onderzoek richt zich op de toepassing van artificial intelligence (AI) in forensische bewijsvoering, zoals het ontwikkelen van computermodellen voor het herkennen van deepfakes en het doen van onderzoek naar de toepassing van AI bij het herkennen van verborgen berichten (steganography). Of de toepassing van AI bij spraakherkenning of het uitlezen van data uit auto's. De samenwerking wordt de komende tijd verder uitgebouwd naar een zelfstandig forensisch onderzoekslab binnen het ICAI - Innovation Center for Artificial Intelligence.

Deepfakes in kinderporno

Criminelen maken steeds vaker gebruik van deepfakes, signaleert prof. dr. Zeno Geradts, onderzoeker bij het NFI, bijzonder hoogleraar Forensic Data Science aan de UvA en één van de initiatiefnemers van het lab. 'Het wordt bijvoorbeeld toegepast om volwassenen in kin-

derpornofilms onherkenbaar te maken. Of om meningen te beïnvloeden met door de computer gecreëerde personen. Het is bijna onmogelijk om echte en deepfakevideo's met het blote oog van elkaar te onderscheiden.' De huidige computermodellen kunnen deepfake in ongeveer acht van de tien video's herkennen, weet hij. 'Dat is nog altijd in twee van de tien video's niet, terwijl je eigenlijk zou willen dat minstens 99,5 procent van de deepfakes eruit wordt gehaald.' 'Er is meer onderzoek nodig om deepfakevideo's te kunnen onderscheiden van echte video's', vult Marcel Worring, hoogleraar Multimedia Analytics aan de UvA aan. 'Negentig procent van alle investeringen in deepfakes gaat naar het verbeteren van de techniek, slechts 10 procent naar onderzoek om deepfakes te herkennen en dan wordt er heel beperkt gekeken naar de forensische bewijswaarde. Dat is echt veel te weinig.'

Herkennen verborgen boodschappen

Er is naast meer onderzoek naar deepfakes ook meer onderzoek nodig naar steganography: het verbergen van boodschappen in foto's en video's. Criminelen kunnen hiervan gebruik maken bijvoorbeeld over wanneer en waar een lading drugs aankomt. 'Met AI kun je computers leren om dit soort verborgen berichten te herkennen door ze te trainen op bepaalde afwijkingen', zegt Geradts. Een derde project gaat zich richten op het verbeteren van de bewijsvoering bij spraakherkenning door het combineren van de herkenning met bijvoorbeeld locatiegegevens op de mobiele telefoon. Het NFI en de UvA richten zich verder op het ontwikkelen van tools om nog beter informatie op telefoons te kunnen doorzoeken. Daarnaast wordt een promovendus aangesteld die in samenwerking met de politie onderzoek gaat doen naar informatie die auto's kunnen opleveren.

Heb je nieuws te melden of een tip of trend gespot die je graag wilt delen met collega-IP'ers? Laat het weten via redactie@informatieprofessional.nl.

FONTYS MORAL

Technologie heeft eigenlijk altijd wel een ethische lading en dit kan soms verregaande consequenties hebben voor consumenten, burgers of organisaties.

Daarom is Fontys Hogescholen per 1 juni het lectoraat Moral Design Strategy gestart onder leiding van lector Bart Wernaart.



Onze samenleving is getransformeerd van een *information society* naar een *smart society*, waarin mens en techniek altijd en structureel met elkaar interacteren. In toenemende mate vertrouwen we daarbij een deel van de ethische besluitvorming



Bart Wernaart

START NIEUW LECTORAAT DESIGN STRATEGY

toe aan slimme technologie. Dat leidt tot allerlei maatschappelijke vraagstukken en uitdagingen. Hoe zeker moet een algoritme zijn van zijn zaak om een burger van fraude te beschuldigen? Wie mag bepalen hoe een algoritme op een socialemidiaplatform omgaat met desinformatie? Het lectoraat gaat aan de slag om dit soort vraagstukken op te pakken.

Ontwerp van technologie

‘Het gaat om meer dan alleen kennis opdoen over ethische vraagstukken’, zegt lector Bart Wernaart. ‘We willen de

en gebruiken als input voor toekomstig ontwerp van technologie.’

Sociaal innovatief leerproces

Wernaart start tegelijk met het nieuwe lectoraat als lector. De afgelopen jaren was hij werkzaam als jurist en docent-onderzoeker Recht en Ethiek aan de Fontys Hogeschool Economie en Communicatie. Wernaart ziet het nieuwe lectoraat als een manier om onderwijs en onderzoek dichter bij elkaar te brengen, waarbij studenten een actieve participerende rol hebben in het doen van betekenisvol onderzoek. ‘We willen enerzijds met onze voeten in de klei staan en bijvoorbeeld met mobiele moral labs de wijken in om te achterhalen welke morele verwachtingen burgers hebben van een voorspellend algoritme dat de overheid gebruikt. Denk daarbij aan *preventive policing*, of een algoritme dat risico op fraude moet detecteren. Anderzijds willen we meedoen op hoogwetenschappelijk niveau en daar tot nieuwe modellen en aanlegroutes komen, gebaseerd op onze praktische expertise.’

InnovatieHub Digitale Transformatie op High Tech Campus Eindhoven

Er komt met ingang van aankomend studiejaar een Fontys InnovatieHub Digitale Transformatie op de High Tech Campus Eindhoven. Studenten, docenten, onderzoekers en organisaties zullen hierin samen en interdisciplinair aan de slag gaan met actuele vraagstukken op het gebied van digitale transformatie. Van nieuwe verdienmodellen en toepassingen voor organisatie en mens tot de ethische aspecten ervan. Het sluit volgens de hogeschool naadloos aan bij het lectoraat Moral Design Strategy. Fontys wil de komende jaren investeren in praktijkgericht onderzoek op het snijvlak van ethiek en technologie. Dat past ook bij de doelstellingen van de hogeschool om een kennisinstelling te zijn met impact in de regio. Het lectoraat zal mede input leveren aan de digitaletransformatievraagstukken binnen de InnovatieHub Digitale Transformatie.

moraliteit van het individu “vangen” en ten goede gebruiken in het ontwerpproces van producten en diensten in zowel de private als publieke sector. Het gaat er vervolgens ook om dit door te vertalen naar strategie. Wat betekent dit bijvoorbeeld voor de missie en visie van organisaties? Of hoe verhoudt zich dit tot decentraal overheidsbeleid? En hoe kan een organisatie uiteindelijk op een transparante manier vaststellen dat het morele ontwerp ook daadwerkelijk datgene doet zoals dat oorspronkelijk bedoeld was? Het technologische ontwerp leidt tot belangrijke uitdagingen op het gebied van morele autoriteit, ethische besluitvorming en morele strategievorming. Door middel van gedegen onderzoek willen we dit beter in kaart brengen



Rob Feenstra

Projectleider/consultant bij de Universitaire Bibliotheken Leiden; heeft als aandachtsgebieden bibliotheeksystemen en de digitale bibliotheek

Fieldlabber

Fieldlabber. Spreek het woord hardop vijf keer achter elkaar uit en ga eens na op welke manier het zich in je hoofd nestelt. In mijn geval werd de ‘a’ gevoelsmatig een ‘e’. Dat heeft te maken met de betekenis van fieldlabber: iemand die meedoet aan een fieldlab-evenement. Bij een aantal van die evenementen namen de fieldlabbers flink wat drank tot zich. Associatief dacht ik daarbij aan fieldlabbers, een nog niet bestaand woord dat zeker een kans verdient. Fieldlabber bestaat pas een paar maanden en toch heeft het opmerkelijk genoeg al meerdere betekenissen. Een fieldlabber kan namelijk ook de organisator van een fieldlab zijn.

Voor fieldlab bestaat een goede Nederlandse vertaling, namelijk proeftuin. Daarop voortbordurend zou je een fieldlabber een proeftuinier kunnen noemen. Maar eigenlijk is een vertaling helemaal niet nodig, want ondanks de Engelse klank komt fieldlabber alleen voor in het Nederlands en vrijwel steeds in verband met coronagerelateerde experimenten. Het past in een rijtje Engels klinkende uitdrukkingen die alleen in het Nederlands bestaan, zoals realitysoap, non-food en touringcar.

Ook wanneer je op fieldlab zoekt, kom je vooral Nederlandse websites tegen. Dat komt door de typisch Nederlandse gewoonte om woorden aan elkaar te plakken. Op Google Trends is te zien dat in de rest van de wereld vrijwel iedereen field lab schrijft (zie afbeelding). De blauwe uithaal naar boven aan het einde van de grafiek is dan ook volledig te danken aan Nederlandse webpagina's.

Ik zocht op de website van de rijksoverheid naar een goede omschrijving van fieldlab en vond dat er in ieder geval een duidelijk onderscheid is tussen fieldlabs en pilots voor testevenementen. Toen ik vervolgens keek wat de overheid beschouwt als een pilot voor een testevenement, zag ik dat daar ook fieldlab-evenementen toe gerekend worden. De overheid is er dus ook nog niet helemaal uit.

Het is hoe dan ook te hopen dat iedereen binnenkort gevaccineerd is en dat de lelijke fieldlab-samenstellingen (ook het werkwoord fieldlabben is al gesignaleerd) net zo snel uit het Nederlands verdwijnen als ze gekomen zijn.



Alle Waddeninformatie

Werelderfgoed dat beschermd moet worden, dat is het Waddengebied. Daarvoor zijn goede data, (onderzoeks)publicaties en literatuur nodig, en helpt het niet als informatie over allerlei plekken is verspreid. Dus kwamen er het Datahuis Wadden en een heleboel aanverwante diensten. Inkijkje in een lijvig project.



Anneke de Maat

Freelance redacteur
en tekstschrijver

De veerboot Schulpengat, die dagelijks tussen Den Helder en Texel vaart, vervoert niet alleen mensen maar ook meetapparatuur. Het Koninklijk Nederlands Instituut voor Onderzoek der Zee (NIOZ) registreert daarmee bij elke overtocht onder andere de hoeveelheid zand, slib, zout en plankton in het water. Dit soort data is onontbeerlijk voor het beheren en beschermen van de Wadden. Het NIOZ is een van de vele organisaties die Waddendata verzamelen en gebruiken voor beheer of onderzoek. Hoe regel je dat al die data en ook onderzoekspublicaties over de Wadden vindbaar en bruikbaar zijn voor iedereen, zodat

online vindbaar op één plek

Zee, zand en zeehond in data gevat

dit unieke natuurgebied optimaal kan worden beschermd?

Daarvoor gingen het ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV) en Rijkswaterstaat (RWS) de samenwerking aan met regionaal historisch centrum Tresoar in Leeuwarden om het Datahuis Wadden te ontwikkelen. Dit Datahuis heeft de ambitie om de beste beschikbare informatie over de Wadden op één plek online vindbaar te maken. Het wil toegang bieden tot data, literatuur en kaarten en inzicht geven in de ontwikkelingen in het Waddengebied.

Oprichting Basismonitoring

‘Zo’n Datahuis opzetten is een enorme

uitdaging’, vertelt Marjan Datema, beleidsmaker bij het ministerie van LNV. ‘De samenwerking met een instelling als Tresoar, die heel andere kennis en expertise heeft, is leuk en inspirerend. Om het beheersbaar te houden ontwikkelen we met kleine stappen.’ Sinds 2016 zet Datema zich samen met Rick Hoeksema, projectmanager bij RWS, in voor samenwerking tussen de Waddenbeheerorganisaties in de Basismonitoring Wadden. Hierin participeren vanaf 2018 het ministerie van LNV, RWS, Coalitie Wadden Natuurlijk, Staatsbosbeheer, de Wadden-academie,

‘De overheid verzamelde enorm veel data over de Wadden, maar kon niets zeggen over de toestand van het Waddengebied’

de Waddenprovincies, de NAM en Tresoar. ‘De Algemene Rekenkamer had gesignaleerd dat de overheid enorm veel data over de Wadden verzamelde via rapportages die door de Europese Unie verplicht zijn, maar dat ze nog steeds niet kon zeggen hoe de toestand van het Waddengebied was’, blikt Datema terug. ‘De overheid wilde daarom dat de Waddenbeheerders zouden gaan samenwerken alsof ze één beheerder zijn, ook wat data en informatie betreft. Hiervoor hebben we de Basismonitoring Wadden opgericht.’

Informatie vindbaar én begrijpelijk

Datema en Hoeksema bedachten vervolgens dat ze nog een stap verder wilden gaan. ‘Door alle gesprekken die we voor de Samenwerkingsovereenkomst Basismonitoring met beheerders voerden, kwamen we op het idee van een Datahuis Wadden’, vertelt Datema. ‘We zijn als overheid verplicht alle data die we met publiek geld verzamelen openbaar te maken, maar deze data waren niet allemaal toegankelijk en vindbaar voor niet-specialisten. Dat is heel jammer, want dergelijke data kunnen ook nuttig zijn voor bijvoorbeeld onderzoekers, studenten, scholieren, gemeentes, provincies, waterschappen, het rijk zelf en bedrijven die een vergunning willen aanvragen. We hebben toen de ambitie geformuleerd dat al die data ook voor deze groepen makkelijk te vinden en te gebruiken moeten zijn.’ Daarbij ging het niet alleen om alle Waddeninformatie op één plek, maar ook dat die op een begrijpelijke, publieksvriendelijke manier ontsloten zou worden, benadrukt Datema. ‘Wij zijn geen datadeskundigen, dus moesten we de expertise voor goede ontsluiting van elders halen. We hebben daar Tresoar voor gevraagd. We kenden Hans Laagland van onze samenwerking in Waddenzee.nl. Je kunt bij Tresoar je familiegeschiedenis uitzoeken; met dezelfde principes blijkt je ook op een publieksvriendelijke manier informatie over de Wadden te kunnen ontsluiten.’

Ontsluiten van brongegevens

Het Datahuis Wadden omvat data over zowel het land- als zeegebied van de Wadden, van socio-economie tot



bodem- en waterecologie. Het kent inmiddels een grote veelheid aan taken en activiteiten, vertelt informatieprofessional Hans Laagland van Tresoar. ‘We zorgen ten eerste voor het verzamelen en ontsluiten van de brongegevens van de Basismonitoring Wadden voor iedereen die er interesse in heeft: van beheerders en bestuurders tot burgers, scholieren, studenten en bedrijven.’

De monitoringdata zijn belangrijk voor de evaluatie van beheer- en beleidsdoelen van allerlei professionele partijen, licht Laagland toe. ‘Bijvoorbeeld voor organisaties als de Waddenvereniging en de provincies. De bodem, het water, de ligging van eilanden en platen, de soorten en aantallen vogels en vissen veranderen voortdurend. Het Datahuis Wadden wil bestuurders, beleidsmakers en beheerders in staat stellen tijdig ontwikkelingen waar te nemen en waar nodig bij te sturen. Als

een verandering bedreigend is voor de natuur of de bewoners, willen we dit kunnen ombuigen. Het is daarom van belang dat bestuurders, beheerders en gebruikers van de Wadden allen over dezelfde betrouwbare informatie kunnen beschikken. De Waddenviewer zorgt daarvoor met een kaartviewer waarin de brongegevens gekoppeld zijn aan locaties in het Waddengebied. Je kunt de onderliggende datasets downloaden. Ook hebben we alvast een historische kaart van Tresoar toegevoegd. De data zijn voor leken niet altijd makkelijk te interpreteren, daarom helpt het Datahuis eveneens om er begrijpelijke informatie van te maken.’

Bronhouder blijft eigenaar

Het Datahuis Wadden maakt dat Waddendata via één plek online vindbaar zijn, maar vindbaar betekent niet perse via één zoekbalk, vervolgt Laagland. ‘Het Datahuis kent veel verschillende

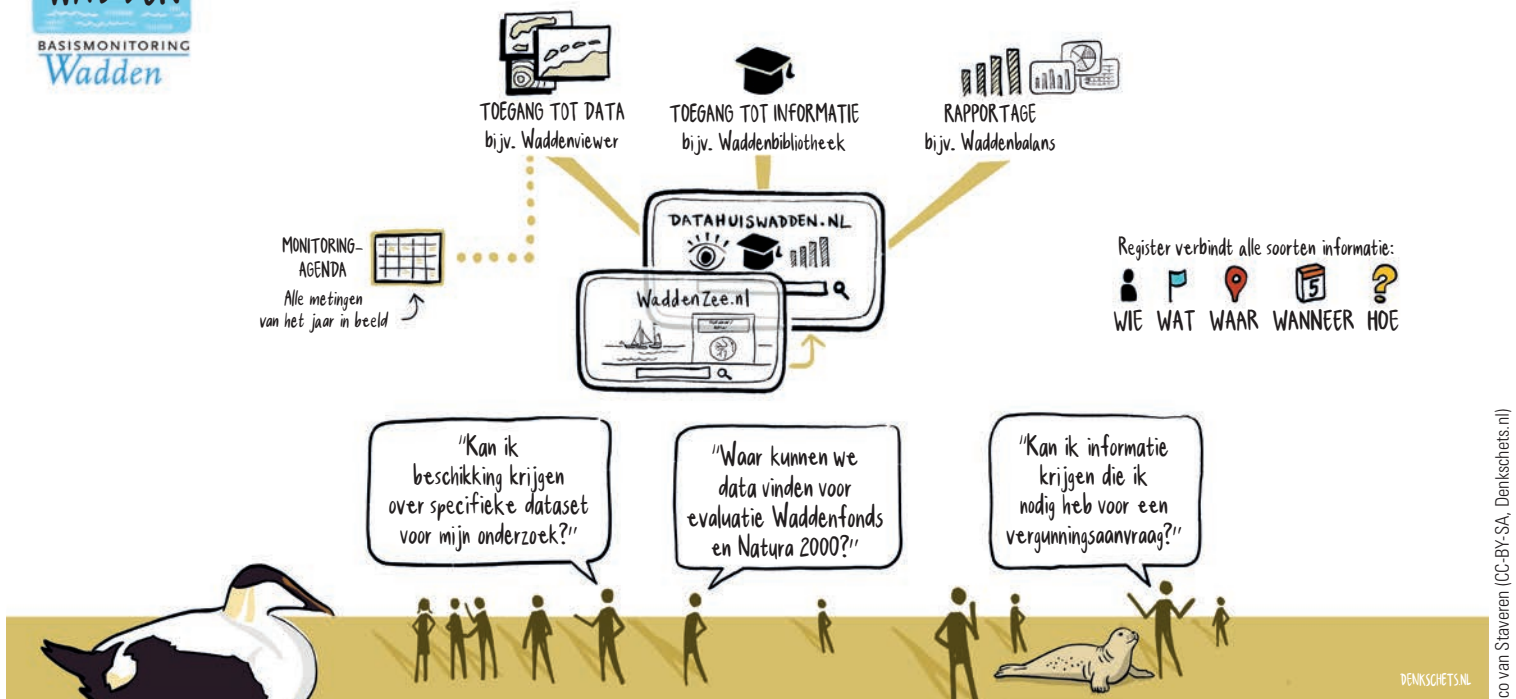
‘Je kunt bij Tresoar je familie-geschiedenis uitzoeken. Met dezelfde principes blijf je ook informatie over de Wadden te kunnen ontsluiten’

soorten data. Daarom willen we de datasets via een register op één plek vindbaar maken, het zogeheten Waddenregister. Vervolgens maken we data die op elkaar lijken nader toegankelijk in aparte subsites (of portalen), zoals ruimtelijke data via een Waddenviewer, bibliografische data via de Waddenbibliotheek of sociaal-economische data via de Waddenbalans. Het Datahuis Wadden concentreert zich op *toegang* tot data: de data zelf blijven zo veel mogelijk bij de bronhouders. De data van de overheid zijn open access en herbruikbaar voor iedereen, maar die van sommige andere partijen niet. Bij data uit wetenschappelijk onderzoek is er soms nog een spanningsveld.’

WaddendataproTOCOL

De bronhouders zijn als eigenaar van de data verantwoordelijk voor de kwaliteit, vertelt Laagland. ‘Wij ondersteunen hen bij het duurzaam

TOEGANGSPORTALEN



houdbaar, vindbaar, doorzoekbaar en toonbaar maken van de data. Om de datakwaliteit te verbeteren willen we een Waddendataprotoocol opstellen. Als een partij vanuit een project data in het Datahuis Wadden wil plaatsen, maken we graag zo vroeg mogelijk in het traject afspraken over een datamanagementplan bij het project, zodat we ze in het Datahuis kunnen opnemen. Bijvoorbeeld een goede beschrijving van de data met metadatering. We willen dat steeds meer mensen de data gaan gebruiken en we willen zo vraaggestuurd mogelijk werken.'

De Waddenacademie (zie ook kader) heeft net een open dataset aan het Datahuis Wadden toegevoegd: de Waddenbalans, vertelt Klaas Deen, secretaris van de Waddenacademie, enthousiast. 'We zagen dat er al veel data over de natuur worden gepubliceerd, maar weinig sociaal-demografische data. In samenwerking met Telos hebben we daarom de Waddenbalans gemaakt. Dit digitale rapport geeft een beeld van de ontwikkeling per gemeente op de Waddeneilanden en langs de Waddenkust vanuit sociaal-cultureel, ecologisch en economisch duurzaam-

heidsperspectief. Dit is voor het eerst gedaan. Het is nu mogelijk gemeentes op indicatoren met elkaar te vergelijken. Het is onze taak om ervoor te zorgen dat zoiets tot stand komt, daarna moet een andere partij het overnemen. In dit geval de Waddenprovincies.'

Internationale Waddenbibliotheek

Op initiatief van de Waddenacademie en Tresoar is onlangs een nieuw onderdeel van het Datahuis Wadden gerealiseerd: de internationale Waddenbibliotheek. Deze bibliotheek wil digitaal toegang bieden tot allerlei soorten bronnen over het Waddengebied: van wetenschappelijke artikelen en boeken tot (historische) kaarten. De eerste stap was het digitaal ontsluiten van (historische) literatuur. Deen: 'We misten een plek waar alle literatuur over de Wadden vindbaar is. We wisten dat er veel literatuur over de Wadden bestaat, maar het was heel lastig uit te vinden wat zich waar bevond. De catalogi van de collecties van de verschillende kennisinstellingen waren niet met elkaar verbonden. Oude boeken die alleen in gedrukte vorm bestonden, waren nog maar op een paar plekken beschikbaar.

'We willen bestuurders, beleids-makers en beheerders in staat stellen tijdig ontwikkelingen waar te nemen en waar nodig bij te sturen'

We wilden die oude literatuur digitaliseren en ter beschikking stellen.'

Tresoar beheert de erfenis van de voormalige Universiteit van Franeker, waaronder de collecties van de vroegere universiteitsbibliotheek. Volgens Laagland is dat een van de redenen waarom Tresoar wilde meedoen aan het opzetten van de Waddenbibliotheek. 'We hebben van oudsher een rol in wetenschappelijke informatievoorziening. Via onze participatie in het Datahuis Wadden geven we hier nu een hedendaagse invulling aan. We verzamelen en ontsluiten niet alleen literatuur, maar maken ook data toegankelijk voor een breed publiek. We hebben

Waddenacademie: kennis over Wadden vergroten

Omdat de kennis over de Wadden versnipperd en slecht ontsloten was, is in 2008 de Waddenacademie opgericht 'om de kennis over het Waddengebied te vergroten'. Dat doet de organisatie door te agenderen, programmeren, informeren en monitoren. Ze heeft vijf wetenschappers met een deelaanstelling in dienst met kennis op het gebied van geowetenschappen, ecologie, economie, cultuurhistorie, klimaat en natuurrecht.



vanuit de historie een collectie over de Wadden in huis en willen zorgen voor een goede infrastructuur voor wetenschappelijk onderzoek in Friesland.’

Digitaliseren historische literatuur

‘We wilden niet zelf een academische bibliotheek beginnen,’ vervolgt Deen, ‘maar ervoor zorgen dat de standaardliteratuur over het Waddengebied die nog niet digitaal beschikbaar was, gedigitaliseerd zou worden en zo voor iedereen beschikbaar zou komen. Een goed voorbeeld is het standaardwerk *Ecology of the Wadden Sea* van Wim Wolff uit 1983. Dat boek was nog maar op een paar plekken te vinden. We wilden dat iedereen het vanachter zijn of haar computer zou kunnen inzien; de auteur en uitgever hebben hier belangeloos aan meegewerkt. Daarnaast willen we zogenaamde grijze literatuur, zoals oude onderzoeksrapporten van Rijkswaterstaat, digitaliseren. Doordat de Waddenbibliotheek toegang biedt tot historische literatuur kun je huidige ontwikkelingen in hun historische context plaatsen.’

De wetenschappers van de Waddenacademie maakten de selectie van de te digitaliseren literatuur, vervolgt Deen. ‘Om het open access beschikbaar te kunnen stellen, moet in verband met auteursrecht de maker minstens zevenig jaar geleden overleden zijn. Met het digitaliseren van grijze literatuur moeten we nog starten. We merken wel dat

daar steeds meer behoefte aan is. Als iemand met pensioen gaat, wil men de oude rapporten veiligstellen.’ Alle gedigitaliseerde historische literatuur is vindbaar gemaakt via de online bibliotheekcatalogus WorldCat, vult Laagland aan. ‘Moderne literatuur stellen we beschikbaar via een virtuele bibliotheek. Om daar in te kunnen heb je een bibliotheekpas van Tresoar nodig.’

Overzicht van internationaal onderzoek

De tweede stap voor de Waddenbibliotheek was het ontwikkelen van het Waddenportaal. De Waddenacademie publiceerde begin 2020 het onderzoeksrapport *De kennishuishouding van het Waddengebied*. Voor dit rapport maakte een onderzoeksbureau een overzicht van het wetenschappelijke onderzoek naar het Waddengebied, met aandacht voor belangrijke onderwerpen van onderzoek en welke instellingen zich inzetten voor Waddenonderzoek. Hierop voortbouwend is het Waddenportaal ontstaan: een overzicht van onderzoekspublicaties over de Wadden. Deze publicaties zijn geselecteerd uit een verzameling publicaties van onderzoeksplatform Dimensions van Digital Science. Hierbij is in het Nederlands, Fries, Engels, Duits en Deens gezocht op publicaties met betrekking tot de Waddenzee, Waddeneilanden en het Nederlandse, Duitse en Deense Waddengebied.

‘Via crowdsourcing en met behulp van artificial intelligence trainen we het zoekstelsel voor onderzoekspublicaties over de Wadden’

Via het publiek toegankelijke Waddenportaal kan iedereen nu eenvoudig overzichten verkrijgen van onderzoekspublicaties over de Wadden, vervolgt Laagland. ‘Je kunt zoeken op onderwerp, auteur, jaartal, titel en kennisinstelling. Je krijgt ook visualisaties van aantallen publicaties per jaar, per land of thema (klimaat, ecologie, economie) van de Trilaterale Kennisagenda van de drie Waddenzeelanden. Je kunt zien in welk land naar welk thema veel onderzoek is gedaan. Via het dashboard krijg je per thema inzicht in de stand van zaken van het onderzoek.’

‘We vragen gebruikers om aan te geven of ze de publicaties in het overzicht wel of niet relevant genoeg vinden’, zegt Laagland. ‘Via deze crowdsourcing en met behulp van artificial intelligence trainen we het zoekstelsel. Het portaal krijgt jaarlijks een update met nieuwe publicaties uit Dimensions. We breiden de Waddenbibliotheek de komende jaren verder uit en dragen zo bij aan een kwalitatieve en stabiele stroom van informatie. Dat is belangrijk voor het uiteindelijke doel: het beheer en behoud van het Wadden Werelderfgoed.’

Waddenregister voor beschikbare datasets

Datema van LNV is tevreden over wat al is bereikt en is enthousiast over hoe het Datahuis Wadden zich verder ontwikkelt. ‘We gaan nu het Waddenregister maken, waarmee je kunt zien welke datasets in het Datahuis Wadden aanwezig zijn. Je krijgt een link naar de dataset en de contactgegevens van de bronhouder. En we gaan de Waddenviewer verder verbeteren. Tot slot werken we steeds meer samen, zoals met het Informatiehuis Marien.’ <

Monitoringagenda

‘We publiceren elk jaar de Monitoringagenda’, vertelt Hans Laagland (Tresoar). ‘Dit is een overzicht van de lopende en geplande monitoringactiviteiten in het Waddengebied voor een kalenderjaar. Die agenda is publiek toegankelijk. Zo weten alle partijen van elkaar wanneer er wordt gemeten en wanneer de data beschikbaar komen. Als er ergens een schip data gaat verzamelen, kan een andere organisatie bijvoorbeeld vragen of ze mee mag varen om ook data te vergaren.’

CURSUSAGENDA

GO opleidingen

2021

SEPTEMBER

08/09	ONLINE PRESENTEREN	ONLINE
09/09	MAAK HET VINDBAAR	GAAT DOOR ONLINE
09/09	LEERGANG INFORMATIEADVISEUR	GAAT DOOR
10/09	FUNCTIONEEL BEHEER	
15/09	DATA STEWARD	GAAT DOOR
15/09	DATA-ARCHITECTUUR & DATAMANAGEMENT	
16/09	COMMUNICEREN EN BEÏNVLOEDEN IN EEN VERANDERENDE OMGEVING	
20/09	KLANTGERICHTE INFORMATIEVAARDIGHEDEN	
23/09	CONTENT CURATOR	ONLINE
24/09	AAN DE SLAG ALS COMMUNITY LIBRARIAN	
24/09	EMBEDDED LIBRARIAN	
28/09	KWALITEITSZORG	GAAT DOOR
29/09	ADVISEREN ALS SAMENSPEL	*NIEUW*
01/10	ZOEKSTRATEGIEËN VOOR GESTRUCTUREERDE DATABANKEN	

DE ECHOKAMER



Vijf maanden geleden likete ik voor het eerst een post over radioactieve mondkapjes. Sindsdien ben ik steeds dieper in een echokamer vol nepnieuws gewandeld. Galmt het er hard? En hoe geluidsdicht is die? Voor deze derde editie van 'Achter het nepnieuws' leg ik mijn oor te luister in mijn eigen echokamer.



Vincent M.A. Janssen
Redacteur van IP
en complotdenker-in-opleiding

DEEL 1: NEPNIEUWSINFLATIE

DEEL 2: HET MEDIADIEET

DEEL 3: DE ECHOKAMER

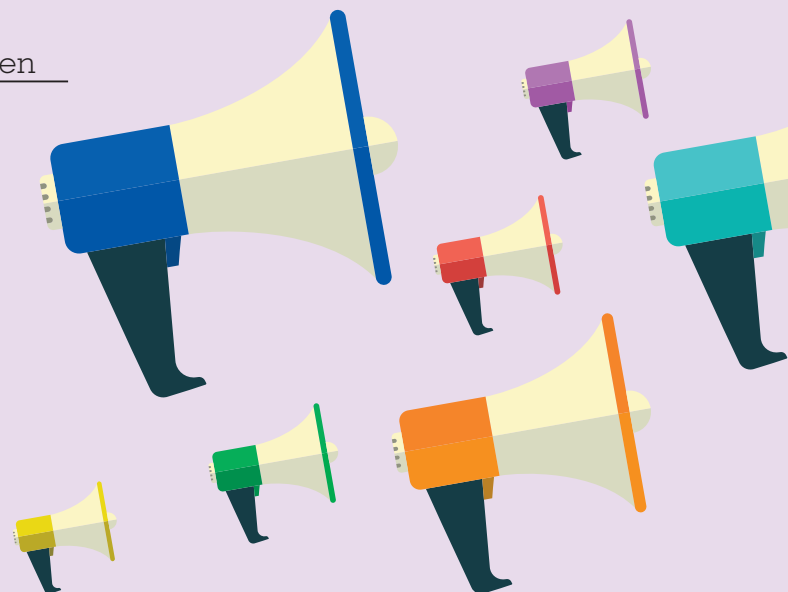
Deel 4: DE GELOOFSCRISIS

Deel 5: PLEISTERS EN REMEDIES

Deel 6: EN NU?

Deel 4 zal verschijnen in IP #7 van dit jaar.

'Er is nog steeds veel onenigheid in de echokamer. In mijn bubbel zie ik bijvoorbeeld verdeeldheid over het nieuws dat Sigrid Kaag eigenlijk een man is. Of een demon'



Mijn aandacht is uiterst kostbaar. Daarom doen techbedrijven er alles aan om mij aan hun platformen te kluisteren. Ik zoek een keer naar een video over Herbalife en ik kijk de rest van het jaar naar aanbevolen filmpjes over coronagenezende lavendelshakes en hoe ik mijn eigen baas kan worden in een piramideschema. Het resultaat is een gepersonaliseerde feed van informatie, ofwel een filterbubbel. Handig voor techbedrijven die zo meer advertenties verkopen, onhandig voor mijn beïnvloedbare brein. Die personalisatie vindt ook plaats bij thema's als politiek, corona en klimaatverandering. Door daarnaast lid te worden van allerlei nepnieuwsgroepen zorg ik ervoor dat geen alternatief feit aan mijn aandacht ontsnapt. Door mijn experiment ben ik in een echokamer terechtgekomen. Hoe ziet die er vanbinnen uit?

Blinde oelewapper

Als ik mijn medemillennials op Reddit moet geloven, zitten de echokamers vooral gevuld met klimaatsceptische boomers en 'Karens' (een stereotyperende benaming voor bevooroordeelde, racistische en anti-vax-vrouwen). Nepnieuws, kom ik achter. Socioloog Jaron Harambam zei al jaren geleden dat nepnieuwslezers en complotdenkers geen specifieke demografie kennen. Ook in mijn eigen echokamer zit een gemêleerd gezelschap, van alle leeftijden en achtergronden. Al zijn er bovengemiddeld veel life coaches en multi-levelmarketingverkopers van essentiële oliën actief. Zijn er dan overeenkomsten te vinden in het opleidingsniveau? Nee, dat varieert namelijk van de universiteit (waar je blijkbaar gebrainwasht wordt tot D66'er) tot de *Red Pill University*, een videokanaal vol complotten, vernoemd naar de rode pil uit *The Matrix*.

Wat veel echokamergenoten delen, is hun oprechte zorgen over de toekomst, hun betrokkenheid bij maatschappelijke problemen en hun onderzoekende geest. Alleen slaat wat ze lezen soms als een tang op een varken. Zo zou een potje gemalen kurkuma per dag tegen kanker helpen en is er bewijs dat Finland niet bestaat. Uiteraard zijn velen van hen even sceptisch over mijn scepsis. Sommigen hebben zelfs medelijden met me omdat ik een 'product van het systeem', een 'oelewapper' en een 'blinde' ben. Niets persoonlijks, verzekeren ze me.



Illusie van bewijs

Zoals ik in het vorige deel van deze serie schreef, schiet de bewijsvoering voor veel nepnieuwsberichten tekort. In mijn echokamer en in veel commentaarsecties is er soms zelfs een *feedback loop* van bronnen. Om bijvoorbeeld te bewijzen dat 'zielskracht' ziektes kan genezen, wordt verwezen naar andere artikelen op dezelfde nepnieuwsite. Soms van dezelfde auteur. In die artikelen zijn vervolgens ook geen geloofwaardige bronnen te vinden. Deze *loop* – die versterkt wordt door de aanbevelingen van algoritmes – creëert een illusie van bewijs, dat ondanks een gebrek aan feiten geloofwaardig is. Oftewel een echo.

Maar zijn echokamers geluidsdicht? Als ik de onderzoeken van *Zondag met Lubach* en de *NOS* moet geloven, krijg ik niets meer van de mainstream media mee zodra ik in de 'fabeltjes-fuik' kuier. Omdat ik inmiddels 'wakker' ben en weet dat 'de publieke omroep corrupte staatspropaganda is met een beeldfrequentie die ervoor zorgt dat kijkers sneller gehoorzamen', zoek ik verder. Maar goed ook, want mijn aannames over echokamers houden wetenschappelijk gezien nauwelijks stand.

Wanden niet geluidsdicht

Een team van Duitse en Noorse wetenschappers zocht het afgelopen jaar uit of 'echokamer' wel een geschikte term is. Zij kwamen erachter dat haar wanden allesbehalve geluidswerend zijn; er dringen nog steeds veel mainstream nieuws en tegenstrijdige opvattingen de kamer binnen. De vaccinatievrijelaars in mijn omgeving gebruiken bijvoorbeeld artikelen van *NU.nl*, zoals over AstraZeneca, in hun argumentatie. De effecten van echo's zijn volgens de onderzoekers daardoor minder ingrijpend dan we denken. Zo is er nog steeds veel onenigheid in de echokamer. In mijn bubbel zie ik bijvoorbeeld verdeeldheid over het nieuws dat Sigrid Kaag eigenlijk een man is. Of een demon. Die verdeeldheid is niet vreemd, gezien de schaal en de diversiteit van de nepnieuws wereld. Online wordt er regelmatig gesproken over de nepnieuwsberg. Het nepnieuws dat de meesten van ons kennen, zoals over Big Pharma, de Rothschilds en klimaatverandering, is slechts het topje. Daaronder schuilt een hele wereld van vergaande theorieën en intenties die variëren van bewuste desinformatiecampagnes van autoritaire regimes tot grootschalige complotten om met sojabonen de wereld te ontdoen van mannelijkheid.

'Ook mensen die niet het eelt van hun duimen scrollen op Facebook, kunnen ontvangers en delers van nepnieuws zijn'



'Lezers van zowel nep- als echt nieuws beschuldigen elkaar van gevangenschap in de echokamer om hun eigen gelijk te verdedigen'



Rekenfout

Die diversiteit en het gebrek aan consensus betekenen overigens niet dat corrigerende feedback in de echokamer altijd met open armen wordt ontvangen. De reactie die ik het vaakst zie op mijn weerwoord is 'ga zelf maar op onderzoek uit'. Met alle plezier, maar als dat onderzoek het tegendeel bewijst willen ze het liever niet horen. Zo lees ik op de website van platteaardegoeroe Frans Heslinga dat onze planeet met 66.666 miljard per uur om de zon draait. Dat moet volgens hem wel iets te maken hebben met de duivel. Genoeg volgers wijzen hem erop dat het getal niet klopt, het is een rekenfout. Deze reacties worden echter genegeerd of weerlegd met *whataboutisme*. Het Duitse en Noorse onderzoek laat ook zien dat echokamers en sociale media niet leidend zijn voor het ontwikkelen van sterke overtuigingen. Ook mensen die niet het eelt van hun duimen scrollen op Facebook, kunnen ontvangers en delers van nepnieuws zijn. Recentelijk hoorde ik op een vrijdagdag de claim dat waterzuivering via planten moet plaatsvinden. Dat haalde volgens een kennis – die zelf niet actief is op sociale media – alle chemicaliën waarmee de overheid ons wil vergiftigen uit het drinkwater. Een snelle factcheck op mijn telefoon toont aan dat de overheid ons water allang zuivert met deze methode.

Othering

Het idee dat echokamers en algoritmes onze opvattingen bepalen, ondermijnt de complexiteit van de wereld achter fake news. Het is een luie verklaring, waarin causatie en correlatie door elkaar lopen. Uit onderzoek van de Universiteit van Oxford blijkt bijvoorbeeld dat sociale media en echokamers niet de oorzaak zijn van politieke polarisatie; iets wat wel vaak gedacht wordt. Net als de term 'nepnieuws' lijkt ook de 'echokamer' eerder een verwijt. Het is een manier van *othering*; lezers van zowel nep- als echt nieuws beschuldigen elkaar van gevangenschap in de echokamer om hun eigen gelijk te verdedigen.

Ironisch genoeg zijn deze overgesimplificeerde verklaringen zelf gaan echoën in de media; ze geven simpele antwoorden op moeilijke vragen. Als de mainstream media hier de mist ingaan, waarover dan nog meer? Is dit het begin van mijn geloofscrisis? In het volgende deel van deze serie laat ik zien hoe mensen in nepnieuws tuinen zodra de twijfel toeslaat. <

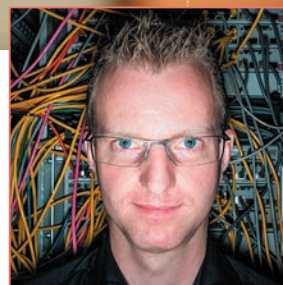
De **juiste** informatie

Wanneer stop je met grasduinen in die eindeloze lijsten met bronnen die naar aanleiding van je zoekacties zijn komen bovendrijven? Die vraag herkennen Rens van de Schoot, Jan de Boer, Felix Weijdema en Bianca Kramer. Al in 2019 oogstten zij lof met ASReview, afkorting voor Active learning for Systematic Reviewing, waarvoor zij een Victorine wonnen. Met deze open source-tool, gebaseerd op machine learning, maakten zij het mogelijk om in reusachtige datasets snel relevante papers te vinden. Hoe staat het er anno 2021 mee?



Anne van den Dool

Tekstschrijver, auteur
en cultureel journalist



Rens van de Schoot

‘Veel mensen denken dat open source nooit commercieel rendabel kan zijn. Integendeel: het kan bij de ontwikkeling van een product juist waardeverhogend werken’



Jan de Boer

Rens van de Schoot heeft als professor Statistics for Small Data aan de Universiteit Utrecht dagelijks met deze zoekuitdaging te maken. Ook in coronatijd werd de kwestie weer extra relevant toen onderzoekers zo snel mogelijk de juiste informatie over dit virus boven tafel wilden krijgen. ‘Stel je voor dat je een zoekterm invoert op Google’, legt hij uit. ‘Dan krijg je tienduizenden hits. Wanneer je een klassiek literatuuronderzoek uitvoert, bekijk je alleen de eerste pagina. Bij een systematische review bekijk je ze allemaal.’

in een vingerknip



Felix Weijndema



Bianca Kramer

Je beantwoordt je onderzoeksvraag dus niet door zelf onderzoek uit te voeren, maar door systematisch eerdere onderzoeken te doorploegen, vult Jan de Boer aan. Hij is bij de Universiteitsbibliotheek Utrecht aanspreekpunt voor de faculteit Sociale Wetenschappen, waarvan ook Van de Schoot deel uitmaakt. 'Ik ondersteun onderzoekers bij het uitvoeren van hun systematische onderzoek. Daar komt heel wat bij kijken. Wie op deze manier te werk gaat, moet zijn werkwijze uitgebreid documenteren. Een onderzoek dat je uitvoert volgens de principes van systematic review moet

'Wanneer heb je genoeg artikelen gelezen om je een volledig beeld te vormen en voldoende vertrouwen om te stoppen? Het is een bijna filosofisch vraagstuk'

reproduceerbaar en transparant zijn. Ook daarin onderscheidt het zich van literatuuronderzoek. Dit is wetenschap ten top.'

Dat laatste aspect is belangrijk, want wanneer je niet duidelijk laat zien hoe je een onderzoek precies hebt uitgevoerd, is het niet reproduceerbaar. Terwijl het juist interessant kan zijn om een proef na een bepaalde periode nog eens te herhalen. 'Daarom heeft het ook weinig zin alleen de resultaten van je onderzoek te delen', legt De Boer uit. 'Eenzelfde onderzoek kan binnen heel korte tijd andere uitkomsten geven. Vergelijk het met het delen van de uitkomsten van een vragenlijst zonder te vermelden wie hebben meegedaan en welke vragen precies zijn gesteld. Zo'n uitvraging valt niet goed te herhalen.'

Breed toepasbaar

Felix Weijndema vervult dezelfde taken als De Boer, maar voor de faculteit Diergeneeskunde. Eerder hielp hij onderzoekers van het UMC Utrecht bij systematic reviews. Vanuit die ervaring weet hij dat ook in universitair medische centra grote behoefte is aan een manier om veel informatie razendsnel te doorzoeken en te sorteren op relevantie. 'Ook daar zoekt men vaak zo snel mogelijk naar de juiste kennis. Het laat zien hoe breed het automatiseren van systematic review toepasbaar is.'

Datzelfde zag Bianca Kramer in haar tijd bij het UMC, waarin zij ondersteuning bij deze werkwijze bood. Nu houdt zij zich vooral bezig met universiteitbreed beleid omtrent open science, bijvoorbeeld in de vorm van een open source-code. 'Veel mensen denken dat open source nooit commercieel rendabel kan zijn. Integendeel: openheid van zaken geven kan bij de ontwikkeling van een product juist waardeverhogend werken.'

Actief leren

Met ASReview (asreview.nl) lees je de hits die je gevonden hebt niet willekeurig, zoals bij een klassieke systematic review, maar geordend van meest naar minst relevant, legt Van de Schoot uit. Dankzij *active learning* wordt ASReview steeds slimmer. 'Met iedere beslissing of een paper echt relevant is of niet, begrijpt de tool beter wat een gebruiker precies nodig heeft. Alleen zo konden we bijvoorbeeld de onderzoekers die zich bezighielden met de bestrijding van het coronavirus helpen honderdduizenden artikelen te doorploegen en aan hun de teksten voorschotelen die het meest relevant waren voor hun zoektocht naar het effect van een bepaald geneesmiddel.' Op die manier informatie zo goed mogelijk rangschikken is slechts een deel van het verhaal, benadrukt De Boer. 'Een uitdaging is: wanneer heb je genoeg artikelen gelezen om je een volledig beeld te vormen? Wanneer heb je voldoende vertrouwen om te zeggen: nu stop ik? Het is een bijna filosofisch vraagstuk, maar we houden ons er wel degelijk mee bezig. Onze laatste berekeningen zeggen dat je met onze tool soms al na vijf procent van de totale hoeveelheid teksten kunt stoppen met lezen. Dat zorgt voor een enorme winst in tijd en energie.'

Het past niet altijd helemaal in het plaatje van de onderzoeker, die vaak het liefst alles zelf bekijkt en in de hand heeft. 'Toch zullen ze moeten toegeven dat op een gegeven moment de moeheid toeslaat', aldus Van de Schoot. 'Om nog niet te spreken over de tijdsdruk die op een onderzoek kan liggen, zoals in het geval van de bestrijding van Covid-19.'

Includeren en excluseren

Hoe gaat zo'n screening van al die informatie precies in z'n werk? 'Wij kijken vooral naar titels en samen-

vattingen van artikelen’, legt Kramer uit. ‘Op basis daarvan includeren en excluseren we. We zouden graag met de complete teksten werken, maar dat is op dit moment lastig. Sommige vragen kun je op basis van een titel en abstract niet beantwoorden, bijvoorbeeld als het gaat om de precieze methodologie. Het is nog even toekomstmuziek – niet vanwege de techniek, maar vanwege de toegang. Wetenschappelijke teksten staan verspreid over platforms van vele uitgevers, gevangen in verschillende formaten. Ook zijn ze niet altijd vrij beschikbaar: zelfs als je er, zoals de Universiteitsbibliotheek Utrecht, voor betaalt om de content in te zien, is het niet altijd mogelijk deze in bulk te downloaden.’

Van de Schoot heeft goede hoop dat die barricades ooit worden neergehaald, maar wanneer is nog maar de vraag. ‘Waarschijnlijk worden op dat moment alleen nieuwe publicaties op die manier beschikbaar gesteld. Oude publicaties blijven onbereikbaar, en ook die heb je natuurlijk nodig voor gedegen onderzoek.’

Menselijke vooroordelen

De ASReview-tool helpt niet alleen snel relevante papers te vinden, maar neemt ook een deel van de vooroordelen van het menselijke zoekproces weg. ‘We krijgen wekelijks mailtjes waarom je in onze tool geen namen van auteurs en tijdschriften te zien krijgt’, vertelt Van de Schoot. ‘Daar hebben we goed over nagedacht. Hoe graag je het ook wilt ontkennen, als mens ben je toch geneigd op basis van de naam van een onderzoeker of tijdschrift een bron extra te vertrouwen of juist te wantrouwen. Met deze tool selecteren we de teksten automatisch op relevantie en kwaliteit – op basis van de inhoud, en niet op basis van de maker of de bron. Zo hopen we tevens bij te dragen aan een democratischer onderzoekslandschap, waarin ook minder bekende maar zeker niet minder betrouwbare auteurs en tijdschriften aan bod komen.’

FAIR te werk

Toch loopt het team nog wel eens tegen bepaalde zaken aan. ‘Ik was zo naïef te denken dat we samenvattingen van teksten vrijelijk konden delen’, herinnert Van de Schoot zich. ‘Dat bleek een illusie: ook die vallen vaak onder het auteursrecht van de uitgever. ‘De abstracts kunnen niet zomaar worden gedeeld’, bevestigt Kramer. ‘We zouden graag willen dat uitgevers die teksten openbaar toegankelijk maken, zodat onderzoeksresultaten beter te hergebruiken zijn. Daarom startten we het initiatief voor het open delen van abstracts: I4OA.org.’ Die open manier van werken wordt steeds gangbaarder. De FAIR-principes (Findable, Accessible, Interoperable en Reusable) krijgen steeds meer voet aan de grond, met name in empirische studies. ‘Je kunt tegenwoordig bijna niet meer in een tijdschrift publiceren zonder de data te delen waarmee je het hele onderzoeksproces kunt nagaan’, aldus Van de Schoot. ‘Toch wordt het delen van de set van gescreende papers met alle daarbij horende beslissingen amper gedaan.’ Samen met zijn collega’s besloot De Boer systematic reviews gepubliceerd door Utrechtse onderzoekers te analyseren. ‘Er zit verbetering in, maar de bewoordingen blijven vaag. “Ik heb met die-en-die woorden gezocht”, dat soort werk. Gelukkig worden de exacte zoekacties steeds vaker als bijlage bijgevoegd, zodat je deze zelf kunt herhalen, maar de beschikbaarheid van datasets met titels, abstract en inclusie- en exclusiebeslissingen voor elk record was op een hand te tellen. Er kan dus nog een flinke slag worden geslagen.’

Responsible AI

Eigenlijk is het gek, vindt Van de Schoot: na de affaire rondom Diederik Stapel had hij wel verwacht dat de hele onderzoeksweld, inclusief die van het systematic reviewen, zich anders ging gedragen. ‘Ik beschouwde dat als een enorme wake-upcall wat betreft het foutief reproducen van data’, vertelt hij. ‘Ik hoop dat onze tool en open werkwijze hierin

‘Een onderzoek dat je uitvoert volgens de principes van systematic review moet reproduceerbaar en transparant zijn’



‘Met iedere beslissing of een paper echt relevant is of niet, begrijpt de tool beter wat een gebruiker precies nodig heeft’

meer kunnen betekenen.’

Dat houdt in: ook de machinematige beslissingen die de tool neemt inzichtelijk maken. ‘Niet alleen de tool zelf is open source, maar ook alle informatie tijdens het hele proces in interactie met de machine, zodat alles volledig reproduceerbaar is’, aldus Van de Schoot. ‘Daarmee nemen we een voorschot op de nieuwste ontwikkelingen op het gebied van responsible AI. Extra prettig als je weet op basis waarvan besluitvorming heeft plaatsgevonden.’

ASReview is daarmee meer dan een tool: het is een onderzoeksproject. ‘We kijken continu: hoe werkt ASReview in welke situatie?’, aldus De Boer. ‘We nemen alles onder de loep: van domme, snelle modellen die alleen naar losse woorden kijken tot deep learning-modellen die ook de context meenemen in hun zoektocht.’

Community

Er ontstaat een heuse community rondom het project. ‘Iedereen mag met ons meedenken’, aldus Van de Schoot. ‘We hebben een discussieplatform waarop je je expertise kunt delen. Het liefst zouden we dit doel collectief blijven doorontwikkelen, zoals we nu al doen. Op die manier kunnen we de vragen die we nu aan elkaar stellen in een groter verband aankaarten. We kunnen nog zo veel van elkaar leren.’

Wie nu denkt dat de tool alleen nuttig is voor *diehard* onderzoekers komt bedrogen uit, weet Van de Schoot. ‘ASReview is breed toepasbaar: de tool wordt al gebruikt voor verslagen, wetsteksten, krantenartikelen en e-mailverkeer. Het is een mooie optie voor iedereen die een grote hoeveelheid tekst wil doorploegen.’ <

ASReview won in 2019 de initiatiefprijs bij de Victorines, prijzen bedoeld om publicaties en vernieuwing in het informatievak te stimuleren. Inzenden voor de Victorines 2021 kan nog tot 1 september dit jaar. Meer informatie vind je op knvi.nl/victorine.

De tao van e-books



‘Nu moest hij iets verzinnen om de honderdduizend dollar aan computertijd die hem in de schoot was geworpen nuttig te besteden’

Frank Huysmans

Bijzonder hoogleraar bibliotheekwetenschap aan de UvA en zelfstandig onderzoeker en adviseur bij WareKennis

Het moet goed toeven zijn geweest aan de Amerikaanse universiteiten op de drempel van de jaren zeventig van de vorige eeuw. Zeker voor degenen die zich met de jonge ‘informatica’, de wiskunde en de logica achter de opkomende computertechnologie bezighielden. Recent las ik over de congressen van de ‘human potential movement’ aan de Californische kust. In Esalen, een plek waar bronnen uit het gebergte in Big Sur in zee uitkomen, trachtte men westerse wetenschap en oosterse wijsheid te verenigen. De zon scheen, de zee ruiste, het warme natuurlijke bronwater verzachtte huid en spieren, terwijl de wind de lange manen deed wapperen. En als je in gezelschap van

grote denkers maar diep genoeg nadacht over zenuwstelsel en bewustzijn, informatie- en chaostheorie, cybernetica, wiskunde en logica, kwam je vanzelf bij tao en Boeddha uit. Desnoods met een beetje *acid*. Met bewustzijnsverruiming zou er vrede tussen de mensen komen. Eerst Vietnam, dan de wereld. *Imagine*.

In dit academische klimaat verliet op Independence Day, 4 juli 1971, in Urbana (Illinois) een jonge computerwetenschapper een kruidenierszaak. Zijn broers beste vriend had net een account voor hem aangeemaakt op de mainframecomputer van de plaatselijke universiteit. En nu moest Michael Hart iets verzinnen om de honderdduizend dollar aan computertijd die hem in de schoot was geworpen nuttig te besteden. In het rekencentrum ontdekte hij dat de kruidenier een facsimile van de handgeschreven Onafhankelijkheidsverklaring in zijn tas had gestopt. Op slag wist

Hart wat zijn missie zou gaan worden. Het mainframe maakte deel uit van ARPANET, het Amerikaanse computernetwerk dat kort daarvoor met zijn Europese evenknie EARN/BITNET was verbonden. Op dit embryonale internet konden rechten-vrije boeken en teksten zonder extra kosten worden gekopieerd en gedeeld! Hart typte de verklaring over en zette die op het net. Dat het bestand maar zes keer werd gedownload door de circa honderd *internauts* deerde hem niet. Het e-book was er. *And it was here to stay*.

Tot 1993 schoot het niet echt op met het overtypen van boeken en teksten. Hart zon op een manier om vrijwilligers te kunnen mobiliseren voor het omzet- en correctiewerk. Een bevriende programmeur hielp met het opzetten van de benodigde list- en mailservers. Zo vond Hart ook het *crowdsourcen* uit. Dankzij het vrijwilligerscorps, scanners met tekstherkenning en slimme arbeidsdeling bij het scannen en correctielezen ging het ineens een stuk sneller.

In zijn kelder werkte Hart vanachter een oude computer veertig jaar onbetaald en onverstoort door aan zijn project. Aan giften accepteerde hij alleen wat hij nodig had om in zijn basisbehoeften te kunnen voorzien. Hij streefde een zo groot mogelijke mate van autonomie na en die gunde hij anderen ook.

Toch was altruïsme niet wat hem dreef. Dat was het verdedigen van het publieke domein tegen de oprukkende commercialisering: *‘I am doing Project Gutenberg for the most selfish of reasons – because I want a world that has Project Gutenberg in it.’*

<

De downside van datagedrevenheid

Verbinding zoeken

Leven en werken in een digitale wereld heeft veel voordelen, maar kan tegelijk een barrière worden die toegang weigert aan mensen die niet over de competenties beschikken om met de toenemende digitalisering om te gaan. Dat moet anders kunnen, betoogt John van de Pas.



John van de Pas

Bestuurslid van VOGIN en hoofddocent Informatie-management, onderzoeker bij het lectoraat Ethiek & Technologie en het lectoraat Maatschappelijke Veiligheid van Hogeschool Saxion

De coronacrisis heeft denken over data hoog op de agenda van elke organisatie gezet. Bij de overheid en aanverwante sectoren (zoals het hoger onderwijs) is thuiswerken de norm geworden. Teams en andere elektronische werkomgevingen zijn voor veel

voor de digitale kloof

organisaties het enige lijntje dat verbinding tussen collega's en klanten mogelijk maakt. Het belang van betrouwbare informatiediensten is duidelijk. Werken in lockdown is zonder digitale netwerkinfrastructuur onmogelijk.

Virtuele organisatie

Buiten een digitale omgeving werken was al uitdagend in precorona-tijd. Wie geen toegang had tot de kennis, vaardigheden en attitudes die noodzakelijk zijn om met een virtuele organisatie te werken, had al een uitdaging te pakken. Tot in de jaren negentig van de vorige eeuw was het bij de meeste organisaties gebruikelijk om een balie met een medewerker erachter als frontoffice te hebben. Daarmee was direct toegankelijke communicatie beschikbaar die voor enige vorm van overleg en uitleg op menselijk niveau kon

zorgen.

Sindsdien zijn organisaties en hun werknemers meedogenloos meegesleurd in de efficiencyverhoging van digitale communicatiemiddelen. Baliemedewerkers zijn vervangen door callcenters. Het invullen van een formulier aan een balie, waar een al dan niet welwillende medewerker helpt met het juiste vinkje in het juiste vakje zetten, is getransformeerd in een digitaal formulier met een juridisch dichtgetimmerde online toelichting. Vanuit

'Organisaties en hun werknemers zijn meedogenloos meegesleurd in de efficiencyverhoging van digitale communicatiemiddelen'

het perspectief van de organisatie geeft dat grote voordelen: het zorgt voor processen die formeel correct georganiseerd zijn. De virtuele organisatie is in veel gevallen een feit.

Voldoende leesvaardigheid

Tegelijk groeit het besef dat het optimaliseren van de organisatie door middel van ICT niet alleen andere vaardigheden van medewerkers verlangt, maar ook eisen stelt aan de gebruiker die de interactie aangaat met de organisatie. Dat begint met correct kunnen omgaan met tekst: communicatie met organisaties hangt af van de mogelijkheid om te lezen welke informatie je als gebruiker moet invoeren om een dienst geleverd te krijgen. Zonder voldoende leesvaardigheid lukt het simpelweg niet om te begrijpen wat je op

een website moet doen en grijp je als gebruiker mis.

Dat misgrijpen is nog geen probleem in een omgeving waarin mensen beschikbaar zijn die je kunnen helpen als je er zelf niet uitkomt. Als je laagdrempelig kunt langslopen bij een baliemedewerker, is de kans op fout ingevulde formulieren of onbedoelde vergissingen

een stuk kleiner dan in een virtuele omgeving. Teksten op websites als enige interface tussen gebruiker en organisatie bieden nu eenmaal een beperkte mogelijkheid tot aansluiten bij het niveau van de individuele klant. Waar een welwillende ambtenaar achter een loket moeiteloos met passende uitleg kan schakelen tussen – bijvoorbeeld – een jurist van provinciale staten en de functioneel-analfabete medewerker van een sociale werkplaats, is de laatste in virtuele organisaties zonder hulp reddeloos verloren.

Organisatie- versus gebruikersperspectief

Bellen met het callcenter biedt in veel gevallen weinig soelaas. Ook daar is vrij strikt omschreven welke vragen beantwoord kunnen worden en welke niet. De formeel-juridisch gescipte antwoorden in de database die een callcentermedewerker opleest, zijn opgezet vanuit organisatieperspectief om correct te kunnen argumenteren ten behoeve van juridisch sluitende besluitvorming. Niet om contextueel te redeneren gericht op interpretatie, waarbij wordt uitgegaan van de situatie van de individuele gebruiker, tenzij je het geluk hebt een helpdesk te treffen waar deskundige medewerkers de vrijheid wordt gegund om vanuit klantperspectief te helpen. *Twenty-first century skills*, geroemd en verguisd, vormen het toegangskaartje tot de hedendaagse samenleving.

Digitale kloof

Wie niet beschikt over voldoende talent of kennis om te voldoen aan de toegangseisen die informatietechnologie stelt, blijft achter aan de oevers van de ‘digitale kloof’. Daarmee wordt de scheiding bedoeld tussen mensen die de vruchten van voortschrijdende digitalisering plukken en mensen die vooral worden geconfronteerd met hun onvermogen om deel te nemen aan een wereld die in toenemende mate toegang tot dienstverlening exclusief laat afhangen van digitale vaardigheden. Informatietechnologie, in principe bedoeld om zo veel mogelijk maatwerkdienstverlening op het niveau van de gebruiker af te leveren, wordt daarmee

tegelijkertijd uitsluitingsgrond voor degenen die niet slim genoeg zijn om daarin mee te kunnen. Doorgaans meewarig bekeken door hen die het allemaal wel beheersen. Met naar schatting 2,5 miljoen mensen die moeite hebben met lezen en schrijven en met de digitaliseringsgolf van 2020-2021 lijkt de stelling gerechtvaardigd dat de actuele ontwikkelingen niet leiden tot verkleining van die kloof ...

Kafkaëske systemen en filterbubbels

Het concept digitale kloof is een passend concept, omdat het naast een vaardighedenkloof (bedienen ICT) en een financiële kloof (ICT-hardware kunnen betalen) ook een intellectuele



kloof (snappen wat van je verwacht wordt) benoemt. Dat stelt het idee van ‘samenleving’ ter discussie. Welke impact de versnelde doorvoering van de digitale samenleving – met haar digitale organisaties, *digital twins* en thuiswerken als norm – heeft op dat *samen*, brengen organisaties als bibliotheken en Stichting Lezen en Schrijven onder de aandacht. Ook de Nationale Ombudsman waarschuwt al jaren voor toenemende complexiteit in de communicatie tussen burger en overheid.

‘Communicatie met organisaties hangt af van de mogelijkheid om te lezen welke informatie je als gebruiker moet invoeren om een dienst geleverd te krijgen’

‘Wie niet in staat is om toegang te krijgen tot de digitale wereld is letterlijk “af” en doet niet meer mee’

En dat leidt bijna logischerwijs tot een gebrek aan probleemoplossend vermogen ‘als het ingewikkeld wordt’ en als standaardprocedures problemen niet oplossen maar verergeren. Het wemelt van de publicaties die beschrijven hoe open en toegankelijke informatienetwerken net zo makkelijk kunnen worden getransformeerd in digitale kooien die gebruikers opsluiten in kafkaëske systemen en filterbubbels.

In een digitale wereld kan een digitale omgeving dus een barrière worden die toegang weigert aan degene die niet beschikt over de competenties om ermee te werken. De voordelen van de samensmelting van informatiesystemen tot vernetwerkte digitaliseringsecosystemen zijn evident: goedkopere en snellere dienstverlening, 24/7 beschikbaar, tegen lagere kosten voor organisaties en gebruikers. Tegelijk wordt door grote groepen in de samenleving een hoge prijs betaald voor het verder optimaliseren en het efficiënter en globaler maken van dienstverlening voor degenen die kunnen lezen en schrijven met ICT. Wie niet in staat is om toegang te krijgen tot de digitale wereld, is letterlijk ‘af’ en doet niet meer mee.

Vertrouwen

Dat zou anders moeten kunnen, al is het niet eenvoudig. Daarvoor is de vervlechting tussen complexiteit en digitalisering te hardnekkig en levert het verder digitaliseren voor alle betrokkenen simpelweg te veel op. Tegelijk is duidelijk dat het invoeren van de geoptimaliseerde digitale dienstverlening neveneffecten met zich meebrengt die onvermijdelijk lijken. Bovendien blijkt het keer op keer buitengewoon complex om ongewenste negatieve uitkomsten van procesoptimalisatie door digitalisering te corrigeren. Vaak blijkt het voor organisaties feitelijk onmogelijk om de gevolgen van foutieve beslissingen voor gebruikers ongedaan te maken, laat staan een incorrecte beslissing terug te draaien. Dat zet een maatschappelijke waarde als ‘vertrouwen in eerlijke beslissingen’ onder druk.

Maatschappelijke onrust

Organisaties investeren fors in nieuwe informatiesystemen. Bij het aanbesteden

den daarvan is de focus bijna exclusief gericht op efficiëntie en effectiviteit – de kernwaarden die met ICT worden nagestreefd. De angel zit in de *andere* waarden, die niet noodzakelijk te koppelen zijn aan de twee kernwaarden van ICT. Zorgvuldigheid moet in sommige gevallen wijken voor snelheid. En ‘beslissingen terugdraaien’ is geen feature die standaard in procesoptimalisatiesystemen wordt ingebouwd omdat een dergelijke functionaliteit te zelden nodig is.

In het grotere plaatje is dat wellicht begrijpelijk. Dat een beperkt aantal beslissingen desastreus uitpakt voor degenen die de gevolgen ervan moeten ondergaan, terwijl 99,9 procent van de besluiten correct is, is vanuit rationeel oogpunt een statistisch detail. Maar als die 0,1 procent foute beslissingen tot onaanvaardbare hardheid leidt die niet gecorrigeerd kan worden, tast precies dat kleine detail de geloofwaardigheid aan van de beslissingen die worden genomen. Dat is een sociaal-maatschappelijke waarde die wellicht in aantallen kan worden gerelativeerd, maar in de alledaagse realiteit van het menselijk samenleven tot grote maatschappelijke onrust kan leiden.

Leven in een alternatieve wereld

In 2017 schreef de Franse filosoof-antropoloog Bruno Latour het essay *Waar kunnen we landen? – Politieke oriëntatie in het Nieuwe Klimaatregime*, waarin hij de geculmineerde woede van de Trumpstemmers en de ‘fact free politics’ bespreekt. Hij stelt dat we er niet uitkomen als we ons rationele beeld vasthouden tegenover een irrationeel verklaarde massa. ‘De “rationele geesten” geloven nog steeds dat feiten op eigen kracht standhouden, zonder een gedeelde wereld, zonder instituties, zonder openbaar leven, en dat heel dat lieve volkje alleen maar hoeft te worden teruggejaagd naar een ouderwetse schoolklas, met een schoolbord en schriftelijke overhoringen, waarna het verstand eindelijk zal zegevieren.’ Voor de digitaliseringsgolf geldt in dat verband iets vergelijkbaars. Latour parafraserend: degenen die geloven in de eindeloze zegeningen van verdere digitalisering, zitten óók

‘Latour parafraserend: degenen die geloven in de eindeloze zegeningen van verdere digitalisering, zitten óók vast in de strikken van de des-informatie’

‘Zonder nieuwe vormen van digitalisering leidt doorgaan op de huidige weg slechts tot verdere verdieping en verbreding van de kloof tussen digitale *haves* en *have-nots*’

vast in de strikken van de des-informatie. Zij zien niet dat het geen zin heeft verantwoordigd te zijn over mensen ‘die geloven in alternatieve feiten’, terwijl zijzelf daadwerkelijk leven in een alternatieve wereld, een wereld waarin digitalisering alleen maar voordelen oplevert, terwijl dat in de wereld aan de andere kant van de digitale kloof niet het geval is.

Veiligheidskleppen aanbrengen

Aangezien de inrichting van informatiesystemen de grenzen stelt en tegelijk de voorwaarden schept voor mogelijke beslissingen die een organisatie kan nemen, verdient de maatschappelijke discussie over datagedrevenheid een vertaling naar de besluitvormingsprocessen in organisaties bij het inrichten van die systemen. Repareren achteraf blijkt keer op keer veel te kostbaar en wordt veelal op economische gronden afgeschoten. Om de hoogst noodzakelijke veiligheidskleppen aan te brengen in de digitale informatiesystemen die het maatschappelijk leven in groeiende mate bepalen, is het essentieel om debat te voeren over de maatschappelijke waarden die geborgd moeten worden. Zonder nieuwe vormen van digitalisering leidt doorgaan op de huidige weg slechts tot verdere verdieping en verbreding van de kloof tussen digitale *haves* en *have-nots*. Waarbij – als we niet oppassen – de digitale erosie van maatschappelijke waarden de laatste restjes menselijke interactie wegslijpt.

Daarom neemt VOGIN samen met NEMO Kennislink het initiatief om, naast de cursussen voor informatiepecialisten, een bredere scope te kiezen. Waar de toepassing van informatietechnologie nog steeds (sinds 1977) de kernfocus is, zoekt VOGIN ook partijen op die beslissen over de inrichting van informatiesystemen. Dat doet ze met een reeks publicaties, lezingen en debatten, om antwoord te vinden op de vraag: *Anders digitaliseren – hoe dan?* <



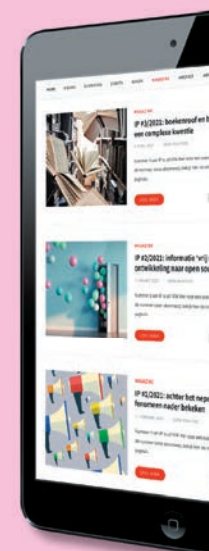
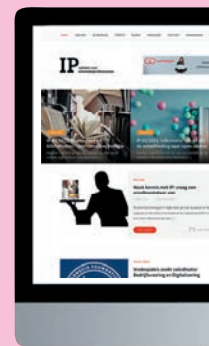
vakblad voor
informatieprofessionals

IP kun je
ook
**online
lezen**

**Vanaf
jaargang
2011
zijn alle
bijdragen
full-text
en als pdf
beschikbaar.***

* De laatste negen verschenen edities van IP zitten achter een ‘slotje’. Abonnees via Aboland hebben inloggegevens toegestuurd gekregen. Abonnees via de KNVI loggen in via mijn.knvi.nl. Vragen? Mail naar administratie@informatieprofessional.nl.

www.informatieprofessional.nl



Onderzoek in AV-archieven

Twintig jaar spraakherkenning in Nederland

De wens om automatische spraakherkenning in te zetten om audiovisuele archieven beter doorzoekbaar te maken is niet van gisteren. Samen met archieven in Nederland werken computer- en geesteswetenschappers al zo'n twintig jaar hieraan. De recente aandacht voor AI heeft deze wens extra op de agenda gezet. Want naast dat spraakherkenning op zichzelf al interessant is, kan het ook als motor dienen voor andere 'intelligente' toepassingen. Een beknopt verslag van hoe het er in Nederland voorstaat.

'Niet alleen voor onderzoekers en journalisten zijn technieken om te zoeken en relaties te leggen in AV-collecties interessant, maar ook voor het algemeen publiek'

Nederlands audiovisueel erfgoed is een schat voor onderzoekers. De 'stoffige' archieven van weleer hebben de afgelopen jaren belangrijke stappen gezet naar digitale depots om erfgoed te bewaren en beter toegankelijk te maken. Naast digitalisering en mogelijkheden om content online beschikbaar te stellen, is het kunnen zoeken in de data een belangrijke voorwaarde: er zijn metadata nodig om gebruikers te kunnen verwijzen naar items die ze interessant zouden kunnen vinden.



Roeland Ordelman

Hoofdonderzoeker taal- en spraaktechnologie
aan de Universiteit Twente en technology
innovation manager bij Beeld en Geluid

Het creëren van deze metadata was altijd de verantwoordelijkheid van archivariissen of documentalisten. Maar dit veelal handmatige proces staat onder druk door de toestroom van digitale data door digitalisering en de creatie van nieuwe, digital born content. Daar komt bij dat gebruikers behoefte hebben aan tijd-gecodeerde metadata om ook *binnen* collectie-items te kunnen zoeken. Dit geldt met name in het audiovisuele domein waar je niet zomaar met ‘control-F’ door tekst kunt springen, maar een item zult moeten bekijken en beluisteren om te vinden waar je naar op zoek bent. Doordat dit soort uitgebreide, beschrijvende metadata meestal maar beperkt aanwezig zijn, is de schat aan informatie in de digitale archieven ook maar beperkt toegankelijk. Sterker nog, vaak is de aanwezigheid van pareltjes volledig onbekend.

Kunstmatige intelligentie

Er wordt al decennia gekeken naar kunstmatige intelligentie als hulpmiddel om content automatisch door computers van metadata te voorzien. Een voorbeeld hiervan in het audiovisuele domein is automatische spraakherkenning. Al in 2001 nam het Nederlands Instituut voor Beeld en Geluid samen met de Universiteit Twente deel in een

Europees onderzoeksproject (ECHO) dat onderzocht of automatische spraakherkenning zou kunnen worden ingezet voor het doorzoekbaar maken van het archief. Hoewel de techniek toen nog niet zo ver was als vandaag de dag, bleek al wel dat het automatisch laten uitschrijven van het gesproken woord in audio en video enorm behulpzaam kan zijn voor het ontsluiten van archiefmateriaal. Want ook al maakt de spraakherkenning fouten en is een automatische transcriptie vaak niet goed leesbaar voor een mens, het gaat erom dat de belangrijke inhoudswoorden goed worden herkend, zodat je met behulp van een zoekmachine via die woorden interessante fragmenten kunt vinden. Met andere woorden: als de computer de data maar kan ‘lezen’.

Interesse vanuit onderzoek

Een doelgroep die al vroeg grote interesse had in het gebruik van spraakherkenning, waren geesteswetenschappelijke onderzoekers die met ‘gesproken getuigenissen’ (*oral history*) werkten. Een belangrijke aanleiding voor deze interesse was het MALACH-project. Doel van dit iconische project dat startte in 2001, was het preserven van getuigenissen van overlevenden van de holocaust en het breed toe-

‘Gebruikers die geen expliciete informatie-behoefte hebben over een onderwerp, typen vaak als eerste hun eigen naam in, of die van bekende personen’

gankelijk maken van deze getuigenissen met behulp van ‘moderne’ digitale zoek- en analysetechnieken. Vanwege de enorme hoeveelheid van circa 50.000 opgenomen interviews waren die automatische technieken voor zoeken in gesproken documenten (*spoken document retrieval*) onontbeerlijk.

Vanwege de grote verscheidenheid van de spraak in de holocaustgetuigenissen bleek het echter nog een hele toer om de technologie over de volle breedte van de collectie op een acceptabel kwaliteitsniveau te krijgen. Denk aan diverse talen, dialecten en etnolecten die je kon tegenkomen, spraak van soms zeer geëmotioneerde mensen op leeftijd, en specifieke termen, namen en plaatsen gerelateerd aan de oorlog. Delen van de collectie zijn daarom jarenlang gebruikt voor onderzoek in de spraaktechnologie om manieren te vinden om de kwaliteit te verbeteren. In Nederland spitste zich dat toe op het gebruik van Nederlandse spraakherkenning voor oral history-onderzoek in het CHoral-project, dat een vervolg heeft gekregen in diverse projecten en samenwerkingsverbanden. Een voorbeeld hiervan is het Buchenwald-project waarin in samenwerking met het NIOD getuigenissen van overlevenden van kamp Buchenwald online toegankelijk werden gemaakt (zie figuur 1).

Transcripties oplijnen

Naast het automatisch uitschrijven van het gesproken woord, kan spraakherkenning ook op een andere manier behulpzaam zijn bij het onderzoek en de ontsluiting van interviewdata. Het komt regelmatig voor dat onderzoekers de interviews waarmee ze onderzoek doen zelf handmatig transcriberen. Deze ‘transcripties’ voorzien ze vervolgens weer van commentaar gerelateerd aan het specifieke onderzoek dat ze doen. De onderzoekers voegen aan deze commentaren codes toe die het mogelijk maken om interviewfragmenten met eenzelfde code met elkaar te vergelijken. Het zou heel handig zijn als nieuwe onderzoekers op basis van die codes door het materiaal zouden kunnen scrollen, maar om dat mogelijk te maken zouden de commentaren ook van tijdcodes moeten worden voorzien.



Figuur 1. Zoeken in het gesproken woord van getuigenissen van overlevenden van kamp Buchenwald met behulp van spraakherkenning. Via de zoekterm ‘fiets’ springt een gebruiker naar het fragment waarin dhr. De Vries in dit voorbeeld praat over hoe hij op de fiets van Hilversum naar Amsterdam ging om in het sportfondsenbad te zwemmen.

Figuur 2. Schermfoto van een deel van de alignment service van de Radboud Universiteit Nijmegen waar een gebruiker een transcriptie en een audiofile kan uploaden die het systeem vervolgens omzet in een bestand met tijdcodes.

Dit is met de hand onbegonnen werk, maar voor een spraakherkenner relatief eenvoudig. De spraakherkenner 'weet' als het ware al wat er werd gezegd en hoeft alleen nog maar aan te geven op welke plek in de spraak elk woord werd uitgesproken. Zolang het verschil tussen de transcriptie en wat letterlijk werd gezegd niet te groot wordt, kan de spraakherkenner hier goed mee overweg en tot op de milliseconde nauwkeurige tijdcodes opleveren. Deze spraakherkenningstechniek wordt oplijnen (*alignment*) genoemd. De Radboud Universiteit Nijmegen heeft hier samen met onderzoekers handige tools voor ontwikkeld (zie figuur 2).

Diverse onderzoeksdomeinen

Naast historici en sociale wetenschappers die met interviewdata werken, ontstond de laatste jaren ook vanuit andere onderzoeksdomeinen interesse in het gebruik van spraakherkenning. Bijvoorbeeld door mediawetenschappers, om te onderzoeken hoe de verslaglegging van een bepaald onderwerp in nieuws- en actualiteitenrubrieken van de publieke omroep zich heeft ontwikkeld. Of hoe de verslaglegging zich verhoudt tot andere media, zoals kranten of sociale media, maar ook hoe een individueel programma zich door de tijd heeft ontwikkeld. Ook voor (on-

derzoeks)journalisten zijn dit soort data-analyses die gebruikmaken van automatische spraakherkenning interessant. Tijdens de verkiezingen van 2021 gebruikten journalisten spraakherkenning om te achterhalen welke onderwerpen aan bod kwamen in talkshows tijdens de verkiezingen. Verder werden er wordclouds gemaakt die mooi lieten zien welke woorden politici het meest gebruikten (zie figuur 3).

Thesaurus persoons- of plaatsnamen

Bovenstaande voorbeelden geven aan hoe spraakherkenning van nut kan zijn voor audiovisuele archieven. Niet alleen om te zoeken naar woorden als 'fiets' of 'vluchtelingen crisis', maar ook als basis voor verdere of diepere analyse. Die analyse kan bestaan uit eenvoudig woordjes tellen en het resultaat hiervan op een mooie manier laten zien, maar kan ook complexere vormen aannemen.

Om verbindingen te leggen tussen verschillende bronnen binnen een audiovisueel archief of tussen meerdere archieven (denk aan het voorbeeld van televisie en kranten hierboven), kan een thesaurus van persoons- of plaatsnamen als verbindende factor worden gebruikt. Het herkennen van personen of plaatsen in de automatische transcripties (*named entity extraction*) is voor dit aan elkaar relateren van da-

'Met het inzetten van spraakherkenning als CLARIAH-dienst starten we dit jaar de eerste testen'

tabronnen (linked data) een belangrijk hulpmiddel. Andere voorbeelden van technieken waarvoor spraaktranscripties kunnen worden gebruikt, zijn het automatisch segmenteren op basis van (de verandering in) onderwerpen waarover gesproken wordt of het inschatten van het sentiment (bijvoorbeeld positief, neutraal of negatief) van een bron (zie figuur 4).

Stimuleren van serendipiteit

Niet alleen voor onderzoekers en journalisten zijn technieken om te zoeken en relaties te leggen interessant. Ook voor het 'algemeen publiek' kunnen dit soort technieken helpen bij het toegankelijk maken van collecties. We weten na decennia van onderzoek best veel over het gedrag van dit type gebruikers als het gaat om toegang tot audiovisuele content. Bijvoorbeeld dat gebruikers niet altijd precies weten waar ze naar op zoek zijn. Dat ze vaak helemaal niet zo'n beeld hebben bij wat er in AV-collecties van erfgoedinstellingen zoal te vinden is. En ook dat ze snel afgeleid of weer vertrokken zijn. Als ze dan gaan zoeken, is het fijn als ze vanuit de zoekresultaten direct naar een fragment in de data kunnen springen (*jump-in point*) dat relateert aan de zoekvraag. En dat ze kunnen grasduinen, 'exploratief' kunnen zoeken, om al browsend op pareltjes in de collectie te stuiten. Het 'stimuleren van serendipiteit' wordt dat wel genoemd, het toevallig ergens tegenaanlopen. Namen of plaatsen als startpunt nemen ligt hierbij voor de hand. Want gebruikers die geen expliciete informatiebehoefte hebben over een onderwerp, typen vaak als eerste hun eigen naam in, of die van bekende personen, of plaatsen waar ze vandaan komen of geboren zijn.

Digitale infrastructuur

Automatische spraakherkenning biedt dus diverse mogelijkheden om audiovisuele bronnen toegankelijk te maken. Hoe kun je als archief deze techniek praktisch inzetten? Het korte antwoord is dat allereerst de spraakherkenning bij de bronbestanden van een archief moet kunnen om de audio op te halen, te verwerken en de resultaten op te leveren aan een bestaand archi-



Figuur 3.
Wordclouds van de spraak van Gertjan Segers (rechts) en Jesse Klaver (boven) in een selectie van gesprekken die op de publieke radio/tv van 1 februari t/m 14 maart 2021 werden uitgezonden.



vering- of metadateringsproces. Verder kunnen de spraaktranscripts als metadata door een aanwezige zoekmachine worden geïndexeerd zodat gebruikers erin kunnen zoeken. Wil je gebruikers direct naar een relevant fragment binnen een item kunnen sturen in plaats van naar het hele AV-document, dan moet de zoekmachine er rekening mee houden dat de metadata tijd-gecodeerd is. Daarbij moet de software die wordt gebruikt om bestanden af te spelen (*play-out*) kunnen omgaan met verwijzingen naar een specifieke plek in het bestand.

Achter het korte antwoord zit in de praktijk vaak een wereld van nieuwe vragen. Het beschikbaar maken van bronbestanden aan een spraakherkenningsproces kan al een lastige horde zijn. Bijvoorbeeld omdat het vanwege ingewikkelde digitale werkprocessen niet zo eenvoudig is om bronbestanden uit het archief te trekken om aan een spraakherkenner te voeren. Ook privacygevoeligheid of auteursrechten kunnen de beschikbaarstelling aan een spraakherkenner in de weg staan, vooral omdat spraakherkenning door-

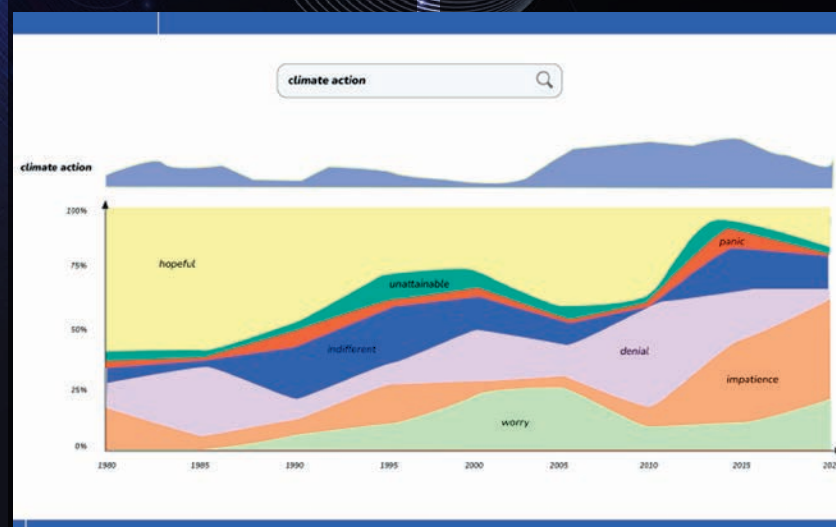
Om onderzoekers spraakherkenning te kunnen laten verbeteren, is het belangrijk om te weten wanneer spraakherkenning het goed doet en wanneer niet'

gaans als online dienst wordt geleverd. De bestanden moeten daarvoor 'naar buiten' of, nog spannender, naar de cloud. Denk bijvoorbeeld aan de spraakherkenningsdiensten die partijen zoals Google en Microsoft bieden. Vanuit archiefperspectief noopt dit volkomen terecht tot voorzichtigheid en

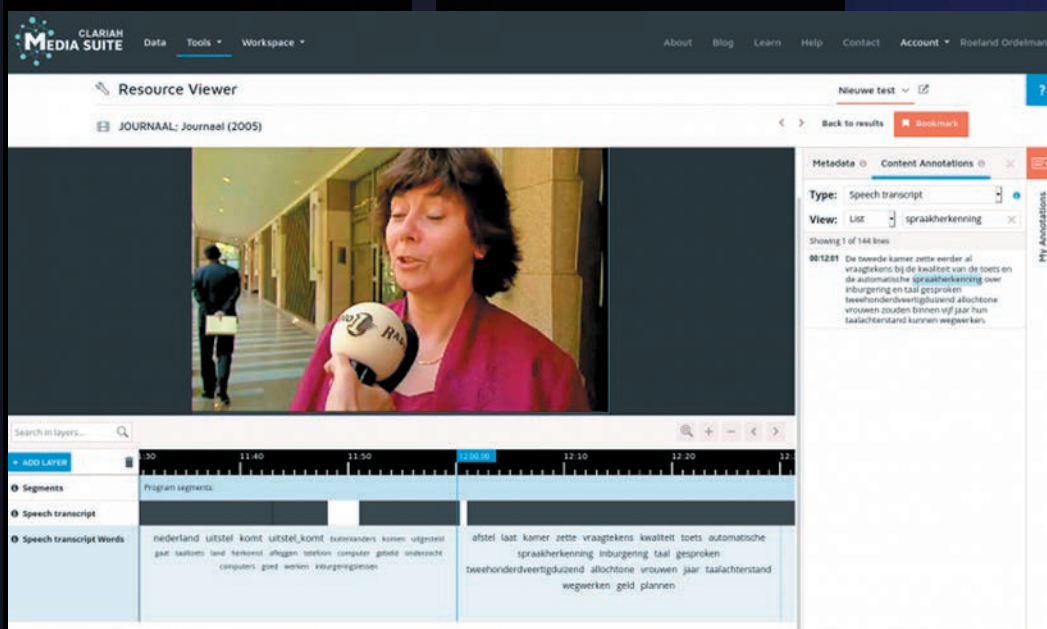
aarzeling om groen licht te geven voor het versturen van bronbestanden. Ten slotte spelen ook vragen met betrekking tot de spraakherkenning zelf, met name: wat kost het en is de kwaliteit goed genoeg?

Zo veel mogelijk AV-collecties aanbieden

Zoals hierboven reeds aangestipt, was er vanuit het onderzoeksveld al jaren veel vraag naar spraakherkenning en de toepassing ervan in de praktijk van audiovisuele archieven waar de voor onderzoekers interessante data veelal liggen opgeslagen. Een aantal jaar geleden is daarom in het CLARIAH-project gestart met het ontwikkelen van een spraakherkenningsdienst voor onderzoeks- en erfgoedinstellingen. CLARIAH bouwt aan een landelijke digitale onderzoeksinfrastructuur. Kort gezegd moet deze infrastructuur het voor onderzoekers eenvoudiger maken om vanuit hun bureaustoel met een paar drukken op de knop online databestanden te raadplegen en te analyseren. Met het beschikbaar stellen van spraakherkenning binnen deze infrastructuur kunnen AV-collecties door onderzoekers veel beter worden benut. De CLARIAH Media Suite (zie figuur 5) is de portal binnen de CLARIAH-infrastructuur die zich expliciet richt op onderzoek met mediabronnen, met name audiovisuele media. Via de



Figuur 4. Impressie van sentimentanalyse op basis van spraak voor het onderwerp 'klimaatactie' door de tijd heen, waaraan gewerkt wordt in het Newsslider-project, een samenwerking tussen Beeld en Geluid, IDfuse, Cream on Chrome en Fontys Hogeschool (Stimuleringsfonds Creatieve Industrie).



Figuur 5. Schermafbeelding van de CLARIAH Media Suite, waar zoeken op 'spraakherkenning' een item uit het *NOS Journaal* uit 2005 oplevert waarin je direct kunt springen naar Rita Verdonk die spreekt over de kwaliteit van de inburgeringstoets die gebruikmaakt van automatische spraakherkenning.

Media Suite kunnen onderzoekers de collecties van Beeld en Geluid en partners (denk bijvoorbeeld aan Tweede Kamer-data) raadplegen. Ook worden stapsgewijs collecties van bijvoorbeeld Eye Filminstituut, DANS en Meertens Instituut beschikbaar gemaakt. Het ultieme doel is om via de Media Suite zo veel mogelijk audiovisuele collecties aan te bieden: van nationale archieven tot regionale en lokale archieven, en zelfs microarchieven die 'in de kast' liggen bij kleine instellingen, project-organisaties of personen.

Spraakherkenning als CLARIAH-dienst

Om deze collecties van spraakherkenning te kunnen voorzien, is de afgelopen jaren gewerkt aan een dienst die binnen CLARIAH kan worden gebruikt voor het efficiënt verwerken van (soms grootschalige) AV-bronnen. Deze dienst haalt audiobestanden op en laat daar vervolgens meerdere spraakherkenners tegelijk op los. Dit gebeurt doorgaans op een relatief krachtige computer die supersnel kan rekenen. Om te beginnen is de dienst getest op een groot deel van de collectie van Beeld en Geluid, honderdduizenden uren aan data, gebruikmakend van een rekencluster van SURF, de

organisatie die voor onderzoek en onderwijs digitale diensten beschikbaar stelt. Maar inmiddels draait de dienst ook op een lokale server bij Beeld en Geluid. Dat is wat handzamer en betekent dat de dienst ook op kleinere collecties kan worden ingezet, en ook bij andere CLARIAH-partners met AV-collecties kan draaien. Met het inzetten van spraakherkenning als CLARIAH-dienst starten we dit jaar de eerste testen.

Open source-modellen

De spraakherkenner die bij Beeld en Geluid en CLARIAH wordt gebruikt, is gebaseerd op de open source-spraakherkenner Kaldi. Deze maakt gebruik van de nieuwste technieken op het gebied van machine learning (deep learning). Wetenschappers aan de Universiteit Twente en de Radboud Universiteit hebben met behulp van een grote hoeveelheid Nederlandse spraak- en tekstdata Nederlandse modellen getraind. Deze zijn open source beschikbaar gesteld onder de vlag Kaldi-NL. Om het gebruik van Nederlandse spraakherkenning te bevorderen, hebben spraakonderzoekers van Nederlandse universiteiten en kennisinstellingen zich verzameld onder de vlag van

de Stichting Open Spraaktechnologie. Deze stelt Kaldi-NL beschikbaar en denkt mee over de verdere ontwikkeling van spraakherkenning in samenwerking met universiteiten en diverse private partijen.

Inclusieve spraakherkenning

Eén van de onderwerpen waar de stichting zich hard voor maakt, is de kwaliteit van de spraakherkenning, die doorgaans wordt uitgedrukt in Word Error Rate (WER): het aantal fout herkende woorden gedeeld door het totale aantal woorden. Fout herkend kan betekenen dat een verkeerd woord werd herkend (substitutie), een woord foutief werd ingevoegd (insertie) of juist foutief werd weggelaten (deletie). Kenmerkend voor AV-collecties is dat de data heel divers zijn: de kwaliteit van de brondata wisselt (denk aan oude opnames), er komen heel veel typen spraak voorbij (verschillende sprekers, leeftijden, dialecten) in diverse contexten (interview, redevoering, vergadering) en opgenomen onder verschillende omstandigheden (in een studio, een kerk, op straat, in een Zoom-meeting). Om onderzoekers spraakherkenning te kunnen laten verbeteren, is het belangrijk om te weten wanneer spraakherkenning het goed doet en wanneer niet. We weten bijvoorbeeld al wel dat spraakherkenners het moeilijk hebben met spraak van kinderen, dialect-spraak of spraak van Nederlanders met een migratieachtergrond. Omdat in AV-collecties dit type spraak ook voorkomt en allerlei publieke diensten worden ontwikkeld waar spraakherkenning een rol in speelt (denk aan Google Home), is het belangrijk dat onderzoekers werken aan wat wel inclusieve spraakherkenning wordt genoemd (zie figuur 6).

Woorden (her)kennen

Als we het over AV-archieven hebben, geldt die inclusiviteit ook voor domeinen. Hierboven kwam al even ter sprake dat in oral history-interviews rond de Tweede Wereldoorlog veel woorden worden gebruikt die specifiek met de oorlog te maken hebben. De spraakherkenner moet deze woorden kennen om ze te kunnen herkennen.

'We weten al dat spraakherkenners het moeilijk hebben met spraak van kinderen, dialectspraak of spraak van Nederlanders met een migratieachtergrond'

Dit speelt ook bij bijvoorbeeld stadsarchieven waar personen, straatnamen of gebouwen uit de stad voorkomen in de spraak, maar mogelijk niet bekend zijn bij de spraakherkenner. Een ander voorbeeld is collecties waar veel medische termen in voorkomen. In het Homo Medicinalis (HOMED)-project onderzoeken we hoe we spraakherkenning zo eenvoudig mogelijk kunnen aanpassen aan dit soort specifieke domeinen in CLARIAH en daarbuiten, zodat spraakherkenning in de volle breedte van de bijzondere collecties in archieven succesvol kan worden ingezet.

Benchmarking

Omdat er naast Kaldi-NL andere spraakherkenners voor het Nederlands bestaan, zou het ook mooi zijn om verschillende spraakherkenners te kunnen vergelijken zodat partijen kunnen kiezen welke spraakherkenner het beste past bij hun wensen. Om de kwaliteit van spraakherkenning op een gestructureerde manier te kunnen testen, is de Stichting Open Spraaktechnologie met een aantal mediapartijen, waaronder Beeld en Geluid, NPO en RTL, een project gestart rond benchmarking. Het doel is om het voor partijen eenvoudiger te maken om aan de hand van voorbeelden uit hun eigen datacollectie een beeld te krijgen of de kwaliteit van spraakherkenning of van verschillende spraakherkenners voldoende is voor hun doeleinden. Een partij die geïnteresseerd is in automatische ondertiteling, stelt uiteraard andere eisen aan de kwaliteit dan een partij die vooral wil kunnen zoeken (zie figuur 7).

Vorderingen

We zijn in Nederland dus al best lang bezig met het onderwerp spraakherkenning in audiovisuele archieven. Enerzijds zien we daarbij dat het in gebruik nemen van nieuwe technologieën in het archiefdomein een ingewikkeld proces is. Het vergt niet alleen de inrichting van nieuwe digitale processen die als ‘complex’ worden ervaren, maar ook een mindset: de wil om mee te bewegen met veranderende behoeften en wensen van gebruikers van archieven, en met de veranderende

‘De recente aandacht voor AI zal helpen om samen met computerwetenschappers in Nederland onderzoek te doen naar de verdere verbetering van spraakherkenning’



Figuur 6. In het project *Kinderen in Gesprek met Media*, mogelijk gemaakt door SIDN fonds en CLICKNL, wordt onderzocht hoe kinderen op een verantwoorde manier met een robot kunnen communiceren, bijvoorbeeld om interessante content in het archief te ontdekken door een gesprek aan te gaan met de robot (Furhat).

maatschappelijke en technologische context. Technologie niet zien als gevaar of iets engs wat je misschien liever uitbesteedt, maar als kans om op andere manieren met AV-bronnen bezig te gaan.

Anderzijds zien we ook dat er aan een aantal basisvoorwaarden moet worden voldaan om het gebruik van spraakherkenning bij audiovisuele ar-

chieven te helpen realiseren. Daar zijn de afgelopen jaren belangrijke stappen in gezet: de kwaliteit is sterk verbeterd, het is makkelijker geworden om zelf spraakherkenning te gebruiken en er is ondertussen ruime ervaring met het gebruik van spraakherkenning om archieven doorzoekbaar te maken. De recente aandacht voor kunstmatige intelligentie zal ook helpen om samen met computerwetenschappers in Nederland onderzoek te doen naar de verdere verbetering van spraakherkenning. Dus archieven, laat die stem maar horen!

Roeland Ordelman promoveerde in 2003 aan de Universiteit Twente op de inzet van spraakherkenning voor het ontsluiten van audiovisuele archieven. Sindsdien staat zijn werk in het teken van het gebruik van informatietechnologie om audiovisuele archieven beter toegankelijk te maken. Hij combineert onderzoek en praktijk bij respectievelijk Universiteit Twente en Beeld en Geluid, waar hij onder meer het R&D ontwikkelteam (O-team), LABS en CLARIAH Media Suite coördineert. In 2000 schreef hij voor IP een artikel over spraakherkenning, het begin van wat hij hierboven heeft samengevat over twintig jaar onderzoek. <

Categorie	WER
Nieuws	18.15
Opinie	26.13
Radio	31.03
Sport	34.36
Samenleving	36.74
Kennis	42.36
Expressie	44.61
Amusement	46.88
Kinderprogrammering	62.15

Figuur 7. Score van de Kaldi-NL open source-spraakherkenner op een aantal mediacategorieën in termen van Word Error Rate (WER), geëvalueerd in het benchmarkingproject. Hoewel de categorieën niet altijd wat zeggen over het type spraak, geven de scores wel een globaal beeld van de kwaliteitsverschillen voor verschillende typen data.

Erfgoed - advies,
detachering & projecten

Reekx

Cultureel Erfgoed

Cultureel Erfgoed is van iedereen en
moet voor iedereen bereikbaar zijn.
Reekx registreert, ontsluit en
digitaliseert uiteenlopende collecties.



Meer weten?

070 - 300 06 84

info@reekx.nl

www.reekx.nl

**KB
ONDERZOEKS-
KRONIEK**

BUITENLAND

**GESPOT
+
GELEZEN
+
VERENIGINGEN
+
LIFEHACKING
+
FAVORIETEN**

‘Techbedrijf Thales start eigen speakersbureau’

**Het wil de kennis en expertise van zijn specialisten op
het gebied van technologische vraagstukken delen
met organisaties, studenten en consumenten, aldus
een bericht op www.informatieprofessional.nl.**



@informatieprofessional



@IPnieuws



[linkedin.com/company/informatieprofessional](https://www.linkedin.com/company/informatieprofessional)

MEER IMPACT MET FAIR

Een nationale bibliotheek die al het tekstuele erfgoed van Nederland digitaliseert en beschikbaar stelt voor huidige en toekomstige onderzoekers en andere gebruikers; waar zou die zijn zonder FAIR? Precies, nergens. Daarom gaat de KB FAIR.

In de rubriek
'KB Onderzoekskroniek'
beschrijven medewerkers
van de afdeling Onderzoek van
de Koninklijke Bibliotheek
hun resultaten, trends
en vondsten.

Het begon bij een intern manifest dat afgelopen januari is vastgesteld. In dit *KB Manifest* staan maatregelen voor de korte en lange termijn ten behoeve van het Findable, Accessible, Interoperable en Reusable maken van onze collecties en datasets. Er is werk aan de winkel, maar we zijn al wel een goed eind op weg. De gedigitaliseerde collecties zijn bijvoorbeeld al vindbaar voor het brede publiek en voor wetenschappelijk onderzoek via onder andere Delpher.nl, DBNL.org en Dataservices.

Aanbevelingen

Om een beter beeld te krijgen van de 'FAIR-heid' van de collecties van de KB hebben we drie 'zware jongens' onder de collecties beoordeeld aan de hand van de FAIR-principes (zie kader). Deze drie zijn de Digitale Bibliotheek voor de Nederlandse Letteren (DBNL), de gedigitaliseerde historische kranten en het webarchief. Vervolgens hebben we, om tot verbetering te komen, een aantal aanbevelingen genoteerd die dit jaar en in de komende jaren binnen de organisatie zullen worden besproken en naar acties zullen worden vertaald. De aanbevelingen hebben betrekking op bijvoorbeeld eigenaarschap en verantwoordelijkheid, persistent identifiers, linked data en juridische helderheid, maar ook op het nalopen van andersoortige databestanden die we FAIR willen gaan maken.

Open science

Door onze collecties FAIR te maken sluiten we aan bij de internationale transitie naar open science, waarin FAIR een hoofdrol speelt. Eenmaal

FAIR worden collecties gemakkelijker te vinden, te citeren, te gebruiken en te hergebruiken. Omdat de zichtbaarheid zal toenemen, verwachten we dat het gebruik van de collecties en datasets zal toenemen zodat ook de maatschappelijke impact groter wordt. Omdat de KB al jarenlang hofleverancier van data is voor onderzoekers in de geesteswetenschappen, willen we met het FAIR maken van de collecties beter kunnen aansluiten bij de huidige praktijk van de (professionele) onderzoeker. We verwachten dat hoe 'FAIR-der' de collecties, hoe beter bruikbaar ze zijn voor onderzoekers buiten de geesteswetenschappen.

Het FAIR maken leidt ook tot een nieuw perspectief op de relatie tussen collecties en user interfaces, zoals Delpher en DBNL. Momenteel zijn de collecties (de achterkanten) en de user interfaces (de voorkanten) met elkaar verbonden, wat betekent dat we de collecties organiseren op basis van hoe ze volgens de user interface moeten worden getoond. Door de voorkant en de achterkant los te koppelen kunnen we met een collectie meer doen voor verschillende doelgroepen. Zo maken we onze infrastructuur flexibeler, toekomstbestendiger en professioneler, en kunnen we elke eindgebruiker precies bieden wat zij, hij of het (machine) nodig heeft.



Astrid van Wesenbeeck

Open science officer
bij de Koninklijke Bibliotheek

De vijftien principes van FAIR

F – Findable (vindbaar)

- F1** (Meta)data krijgen een wereldwijd unieke en blijvende identifier toegewezen.
- F3** Data worden beschreven met rijke metadata.
- F3** (Meta)data bevatten duidelijk en expliciet de identifier van de data die ze beschrijven.
- F4** (Meta)data worden geregistreerd of geïndexeerd in een doorzoekbare bron.

A – Accessible (toegankelijk)

- A1** (Meta)data kunnen via hun identifier worden opgehaald met behulp van een gestandaardiseerd communicatieprotocol.
- A1.1** Het protocol is open, vrij en universeel implementeerbaar.
- A1.2** Het protocol maakt waar nodig een authenticatie- en autorisatieprocedure mogelijk.
- A2** Metadata zijn toegankelijk, ook als de data niet meer beschikbaar zijn.

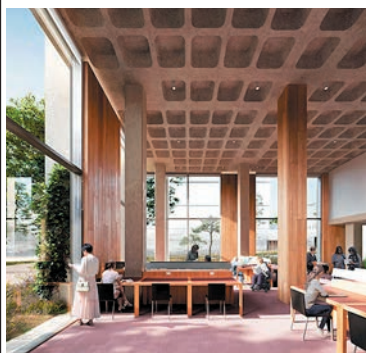
I – Interoperable (uitwisselbaar)

- I1** (Meta)data gebruiken een formele, toegankelijke, gedeelde en breed toepasbare taal voor kennisrepresentatie.
- I2** (Meta)data gebruiken vocabulaires die de FAIR-principes volgen.
- I3** (Meta)data bevatten gekwalificeerde verwijzingen naar andere (meta)data.

R – Reusable (herbruikbaar)

- R1** (Meta)data hebben meerdere nauwkeurige en relevante attributen.
- R1.1** (Meta)data worden vrijgegeven met een duidelijke en toegankelijke datagebruiklicentie.
- R1.2** (Meta)data zijn voorzien van een gedetailleerde herkomstbeschrijving.
- R1.3** (Meta)data voldoen aan domeinrelevante community standards.

Meer over FAIR en over de principes vind je op www.go-fair.org.



Slab/Carmody Groarke

OP WEG NAAR EEN EXTRAVERTE BIEB

Als je gebruikmaakt van het British Library's Document Supply Centre (DSC), dan weet je misschien dat er een vestiging buiten Londen bestaat. In Yorkshire, tussen York en Leeds, iets buiten een dorpje genaamd Boston Spa. Oorspronkelijk een munitiefabriekcomplex uit de Tweede Wereldoorlog, later de locatie van de National Lending Library for Science and Technology. Als deze bibliotheek in 1973 onderdeel wordt van de British Library, blijft ze gevestigd op die plek. Uitbreidingen volgen, en anno nu zitten er afdelingen zoals IT, HR, metadata, 'digitalising' en 'acquisitions' die honderdduizenden nieuwe aanwinsten voor de collecties per jaar verwerken, inclusief alle boeken die ik bestel uit Nederland (2400 per jaar). Ongeveer de helft van de collecties is er opgeslagen, onder meer in twee gigantische magazijnen met een laag zuurstofgehalte, waardoor mensen er maar twintig minuten kunnen verblijven. Zodoende gebeurt het plaatsen en ophalen van stukken geheel automatisch.



Marja Kingma

Curator Nederlandstalige Collecties bij de British Library

Een tienjarenplan moet de 'British Library of the North' helpen boetseren: warmer, gastvrij, modern. Met 2026 als eindstip een blik terug en vooruit.

Een beetje verlept

In 2016 wordt het programma *Boston Spa Renewed* gelanceerd met als doel om het gebouwencomplex te moderniseren en Boston Spa opener en toegankelijker te maken. Nu werkten we als BL altijd al voor de hele wijde wereld, maar dat gebeurde vooral achter de schermen. Dat moet veranderen met de visie voor een 'British Library of the North'.

In 2018 komt er een herziening van het oorspronkelijke plan, omdat besloten wordt het markante hoofdgebouw te behouden en te verbouwen. Dit is waar de meeste medewerkers werken en waar een deel van de uitleencollectie opgeslagen staat. Op dit moment heeft het gebouw een zekere retro-uitstraling, zij het een beetje verlept; het heeft zijn beste tijd zeker gehad. Het eerste idee was daarom om het af te breken, maar daar werd van afgezien omdat bleek dat het verwijderen van 600.000 ton

beton nogal wat CO₂-uitstoot zou opleveren.

Gastvrijere omgeving

Het gebouw zal nu worden verbouwd om een betere fysieke en modernere werkomgeving te bieden. Zo komen medewerkers dicht bij het raam te zitten en ook dicht bij teams waar ze mee samenwerken, zodat ze niet meer tien minuten hoeven te lopen om elkaar te zien. De coronapandemie heeft ontwikkelingen zoals hybride werken en *smart working* versneld; deze nieuwe praktijken zullen in de plannen worden verwerkt. De medewerkers hebben overigens een stem in het ontwikkelen hiervan. Ook bezoekers zullen een veel warmere en gastvrijere omgeving aantreffen. Naast het hoofdgebouw zullen omliggende gebouwen worden verbouwd tot zakenruimtes, werkplaatsen en studio's die aan kleine bedrijven zullen worden verhuurd. Dat zou een diverser pu-

Artist impressions van hoe Boston Spa er in 2026 moet uitzien, met nieuwe publieke ruimtes, een nieuwe entree en een nieuw geautomatiseerd magazijn.

bliek moeten trekken en het aantal bezoekers moeten verhogen. Daartoe zullen er eveneens meer leeszalen komen, en tentoonstellingsruimtes en zalen om schoolklassen te ontvangen, zoals nu al in Londen gebeurt. Verder komt er nog een derde magazijn bij. De bedoeling is een deel van die opslagruimte te delen met andere culturele instellingen.

Nieuwe bibliotheek in Leeds

Om de omschakeling van 'introvert' naar 'extravert' te maken en een British Library of the North te realiseren, is meer nodig dan een bouwplan. Er moet ook worden gewerkt aan een sterke relatie met het stadsbestuur van Leeds en met andere partners. We willen namelijk ook een bibliotheek vestigen in het centrum van Leeds die verbonden is met Boston Spa. Daarvoor zijn bijvoorbeeld betere busverbindingen met Leeds (en York) nodig.

Het is een tienjarenproject dat nu op de helft is. Afgelopen april zijn twee belangrijke stappen gezet: de BL stuurde een aanvraag voor een bouwvergunning in en de aanbestedingsprocedure voor het vinden van een bouwbedrijf is gestart. Intussen wordt gewerkt aan het ontruimen van kantoorruimte en het verhuizen van delen van de collecties. Dat kan betekenen dat je aanvraag bij het DSC iets langer duurt dan normaal. Geduld alsjeblieft en *watch this space!* <

'Van het afbreken van het gebouw werd afgezien omdat het verwijderen van 600.000 ton beton nogal wat CO₂-uitstoot zou opleveren'



Hoe we de macht over onze gegevens kunnen terugkrijgen | Kathalijne Buitenweg | De Bezige Bij | paperback 9789403125213 | € 20,99 | e-book 9789403144610 | € 11,99 | 208 pagina's

DATAMACHT EN TEGENKRACHT

Van Facebook tot de Belastingdienst, van Magister tot werkgevers: bedrijven en overheden willen steeds meer van ons weten. Door het koppelen van data worden we geanalyseerd en raken we gevangen in profielen. Gemeenten voorstellen wie een potentiële fraudeur is, of wie grote kans heeft slachtoffer te worden van huiselijk geweld. Een groeiende groep werknemers wordt aangestuurd via algoritmes, en ons venster op de wereld wordt bepaald door bedrijven die weten wat we willen

zien, horen en lezen. Digitalisering heeft ons leven makkelijker gemaakt, maar ook nieuwe vormen van beïnvloeding en machtsuitoefening gecreëerd. Wat betekent dit voor onze zeggenschap? En wat doet het met de samenleving als we gereduceerd worden tot onze data? In *Datamacht en tegenkracht* roept Kathalijne Buitenweg, voormalig GroenLinks-politicus, alle burgers, werknemers en internetgebruikers op om tegenmacht te organiseren en zo de macht over onze data te heroveren. <

UIT HET TRANSFORMATIE-MOERAS

Menno Lanting heeft al wat boeken op zijn naam staan die lezenswaardig zijn: *Connect!*, *Iedereen CEO*, *De slimme organisatie*, *Lead with a Tweet* en *Olietankers en speedboten*. Zijn laatste boek is al even interessant en snijdt een actueel onderwerp aan: de digitale transformatie. Veel organisaties zijn ermee bezig, maar Lanting zag en ziet dat de meeste na een enthousiast begin vastlopen in 'het transformatie-moeras'. Ondanks alle investeringen, veranderplannen, transformatiemodellen, canvassen en consultancytrajecten lukt het ze niet om resultaten te boeken. De auteur analyseert wat er misgaat en beschrijft aan de hand van voorbeelden uit binnen- en buitenland de vijf kritische faalfactoren die echte transformatie in de weg staan.

Uiteraard schetst hij ook de uitweg: om los te komen uit het moeras moeten organisaties zich omvormen tot een lerende organisatie. Willen ze relevant blijven in het digitale tijdperk, dan is het namelijk niet genoeg om alleen de bestaande processen te digitaliseren. Ze moeten hun cultuur, businessmodel en organisatiemodel aanpassen, niet eenmalig, maar in een continu proces. Zich verbinden aan een ecosysteem en daarbij wezenlijke toegevoegde waarde leveren, dan zijn ze klaar voor de toekomst.

Lanting put zijn kennis en ervaring uit zijn werk als strateeg en leiderschapsexpert, en als associate aan een aantal business schools. Hij onderzoekt doorlopend hoe organisaties van gesloten, hiërarchisch en onpersoonlijk naar open, authentiek en verbonden kunnen transformeren.



De 5 faalfactoren bij de digitalisering van organisaties | Menno Lanting | Business Contact | paperback 9789047014539 | € 24,99 | e-book 9789047014621 | € 14,99 | 240 pagina's | mennolanting.nl <

EN HOE VIND JE ZELF DAT HET GING?



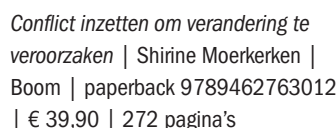
Effectief feedback geven in leersituaties | Ward Grootens | Haystack | paperback 9789461264510 | € 12,50 | 100 pagina's

Vakmanschap begint met goede scholing. Mensen scholen, trainen en oefenen zonder het geven van feedback is zinloos. In dit boek kun je lezen hoe je als waarnemer, collega of leidinggevende in leersituaties omgaat met feedback. Succesvol feedback geven begint met enkele eenvoudige spelregels, weet de auteur die sinds 2008 werkt als zelfstandig trainer en OTO (opleiden, trainen en oefenen)-adviseur in diverse veiligheidsregio's voor brandweer, GHOR, ambulancedienst, meldkamer, bevolkingszorg, politie en ziekenhuizen. Deze spelregels zijn: spreek gewonemensen-

taal, beschrijf alleen wat relevant is, leg blinde vlekken bloot door specifiek te zijn en wees voorbereid. 'Met deze en andere adviezen kun je ervoor zorgen dat anderen je boodschap begrijpen en serieus nemen.'

In zijn trainingen over waarnemen en feedback geven combineert Grootens zijn praktijkervaring als waarnemer met zijn achtergrond in Human Performance Technology en Neuro Linguïstisch Programmeren.

Meer dan ooit is het noodzakelijk dat we ons repertoire vergroten in het functioneel omgaan met conflict en verschillen. Een kind kansen bieden, een metrolijn aanleggen, een pandemie trotseren, terroristische aanslagen voorkomen, een duurzame samenleving creëren; allemaal actuele voorbeelden van vraagstukken die samenwerking van mensen, organisaties en regeringen noodzakelijk maken. En waar mensen samenwerken, is conflict onvermijdelijk – zelfs essentieel. ‘Alle neuzen dezelfde kant op’ klinkt comfortabel, maar daarmee kun je maar één richting uitgaan. Je hebt verschillende perspectieven en belangen nodig om echt in beweging te komen. En voor complexe maatschappelijke

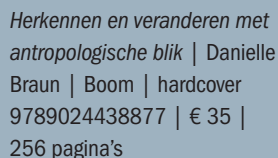


vraagstukken hebben we vooral meervoudig, interdisciplinair en conflicterend handelen nodig.

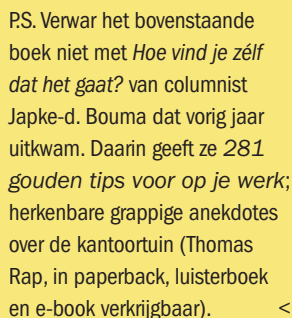
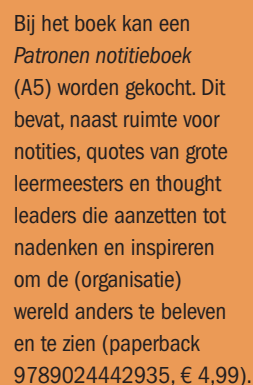
Conflict in samenwerking wordt echter meestal gezien als een probleem. De oplossing ligt dan vaak in het conflict hanteren, beheersen of vermijden.

Interventie- en veranderkundige Shirine Moerkerken pleit ervoor om het conflict juist niet te schuwen en om disfunctioneel conflict functioneel te maken. *Conflict eren* laat zien dat je moed nodig hebt om conflict functioneel te maken, maar meer nog methodische precisie in het uitzoeken van wat er in een specifieke situatie speelt en hoe je kunt werken met wat je aantreft. De auteur biedt praktijkervaringen en ideeën om met conflicteren te gaan oefenen, zodat je conflict en verschil gaat waarderen en effectief kunt inzetten om tot innovatieve oplossingen voor complexe problemen te komen.

De coronacrisis liet afgelopen jaar zien dat patronen die al jaren vastliggen toch ineens kunnen worden doorbroken. Denk aan het thuiswerken. 'Als de reden goed genoeg is, de angst groot genoeg

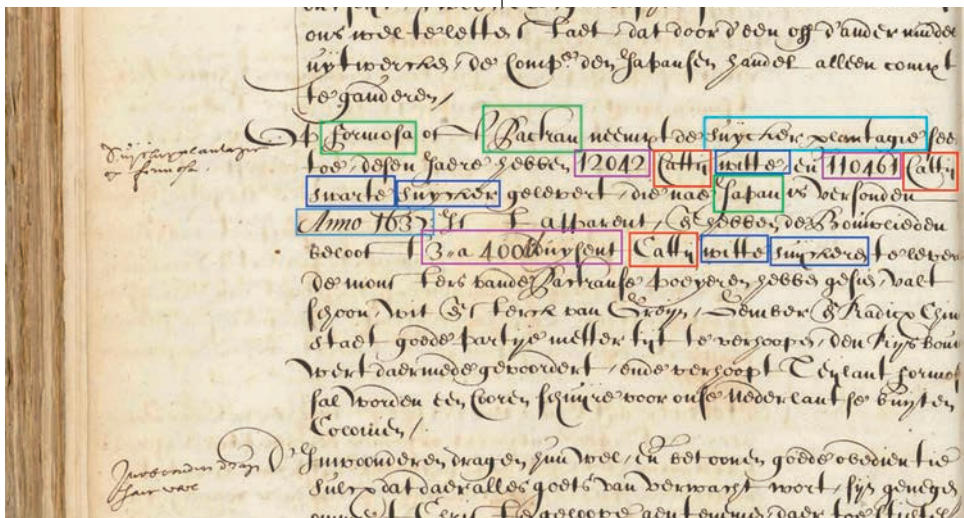


is of de macht het echt wil, kunnen we patronen snel doorbreken', zegt corporate antropoloog Danielle Braun. Ze wordt veel gevraagd door organisaties waar de patronen diep zijn ingesleten en die anders willen gaan werken. Met haar nieuwste boek wil Braun veranderaars, leiders van organisaties, bewegingbouwers, coaches, docenten en studenten (corporate) antropologie helpen patronen te ontrafelen en deze te benutten of te veranderen als dat nodig is. Het gaat over hoe je ze systematisch onderzoekt, hoe belangrijk het is om jezelf daarbij goed te kennen en hoe je patronen in beweging krijgt, consolideert en verstevigt. Een methodisch handboek met mooie beelden en verhalen, voorbeelden en handreikingen uit de praktijk van organisaties en het dagelijks leven.



ZOEKINTRANSCRIPTIES.NL MAAKT IJSBERG ZICHTBAAR

Het heeft veel publiciteit opgeleverd, en dat is mooi voor het vak: de lancering van zoekintranscripties.nl, of in elk geval het project *De ijsberg zichtbaar maken* dat tot deze website heeft geleid. Hiermee kunnen bezoekers handgeschreven historische bronnen – vaak moeilijk leesbaar – snel bekijken en doorzoeken.



Het snel doorzoeken van archiefstukken op zoekintranscripties.nl is mogelijk omdat een grote hoeveelheid handgeschreven bronnen met behulp van transcriptietechniek machineleesbaar is gemaakt en met kunstmatige intelligentie is verrijkt. Dat biedt grote kansen voor historisch onderzoek. Op dit moment zijn 1,5 miljoen bronnen op de website te bekijken. Daarvan zijn er 90.000 verrijkt en doorzoekbaar. Het is een eerste tussen-

tijds resultaat dat nu al aangeboden wordt aan het publiek. De verwachting is dat deze zomer 2 miljoen verrijkte transcripties doorzoekbaar zijn op zoekintranscripties.nl. Het gaat om historische handgeschreven documenten van de Verenigde Oost-Indische Compagnie (VOC) uit de zeventiende en achttiende eeuw van het Nationaal Archief en notariële archieven uit de negentiende eeuw van het Noord-Hollands Archief en andere regionale historische centra.

Zoekintranscripties.nl is qua inhoud en gekozen techniek, waar mogelijk, open source en bruikbaar voor andere erfgoedinstellingen.

Schat aan data

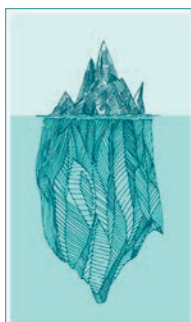
Tijdens het project *De Ijsberg zichtbaar maken* zijn steeds transcripties toegevoegd aan open-datawebsite zenodo.org. Diverse onderzoekers en historici gingen al aan de slag met deze schat aan data. Gerhard de Kok, universitair docent geschiedenis van de Universiteit Leiden, maakte bijvoorbeeld een zoekinterface en deed interessante vondsten in de automatisch getranscribeerde archiefstukken.

Ook schreven het Noord-Hollands Archief en het Nationaal Archief de Innovatie in Archiefonderzoek-prijs uit. Het doel hiervan was studenten en onderzoekers uit te dagen om te laten zien welke nieuwe onderzoeksmogelijkheden deze data bieden. De jury kende de prijs toe aan Saskia Noot, Clare Shutt en Thijs Vorstenburg, studenten van de Universiteit van Amsterdam (UvA). Zij gaven vrouwen en tot slaaf gemaakten in de VOC-testamenten een stem door gebruik te maken van automatische transcripties van testamenten.

Betrokken partijen

In dit project werkt het Nationaal Archief samen met het Noord-Hollands Archief en andere regionale historische centra. Ter gelegenheid van de afsluiting ervan werd op 17 mei het online symposium *De Ijsberg zichtbaar maken* georganiseerd over de resultaten van het project. Dit is terug te kijken op YouTube.

Voor de selectie van leveranciers voor de doorontwikkeling is een aanbestedingsprocedure uitgevoerd. Picturae bleek de beste partij voor de ontwikkeling van de achterkant van de projectwebsite. Dit bedrijf is verantwoordelijk voor de verbetering en modernisering van de transcripties, het toekennen van entiteiten en het ontwikkelen van een transcriptiebeheersysteem. Sioux Technologies is als onderaannemer aangesloten voor de kunstmatige intelligentie; Aincient voor zijn kennis van AI, het erfgoedveld en projectmanagement. Clean Code ontwikkelde de voorkant van de projectwebsite, inclusief de inrichting van de zoekmachine. Van Duizend-vijftig zijn het grafisch ontwerp en coderen. <



Zoeken in transcripties

Op deze website zoekt u in de tekst van handgeschreven documenten. De documenten zijn afkomstig uit de archieven van de Verenigde Oost-Indische Compagnie (VOC) uit de 17e en 18e eeuw en archieven van notarissen uit de 19e eeuw.

[Begin met zoeken](#)

VIJF ADVIEZEN OM TIJDIG OP DIGITALE VERANDERING TE KUNNEN REAGEREN



Wie: Bart Scheffers
Wat: deed tijdens zijn master Information Management aan de Universiteit van Tilburg onderzoek naar de invloed van Covid-19 op de digitale transformatie binnen gemeentelijke organisaties; zijn onderzoek voerde hij uit bij Upstream.

De afgelopen maanden zijn veel organisaties druk bezig geweest met (digitale) verandering, zo ook gemeenten. Het vereiste flexibiliteit om processen op andere (digitale) manieren in te richten. Scheffers voerde voor zijn onderzoek gesprekken met meerdere gemeenten om te kijken hoe ze hun processen inrichten en hun capaciteiten inzetten om flexibel te kunnen zijn. Daaruit konden vijf adviezen worden gedistilleerd.

1

Maak gebruik van autonome crossfunctionele werkgroepen

Autonome werkgroepen helpen de organisatie een verandering door te maken van het werken in verticale silo's naar samenwerken tussen traditionele silo's. Uit het onderzoek is gebleken dat autonome werkgroepen, die vaak veel eigen verantwoordelijkheid krijgen om hun strategie te bepalen, gemeenten helpen flexibeler te worden en sneller te reageren op veranderende digitale vraagstukken.

2

Creëer een cultuur waarin innovatie gestimuleerd wordt

Een innovatieve cultuur is een cultuur waarin medewerkers zich vrij

voelen om te experimenteren met nieuwe (digitale) technieken en vraagstukken. Voor veel gemeenten is het stimuleren van innovatie nog niet een van de kernwaarden, ondanks dat het kan helpen bij het spotten van digitale trends in de omgeving van de gemeente.

3

Zie het totaalpakket aan services verleend aan burgers en organisaties als een portfolio van verschillende diensten

Zodra gemeenten hun dienstverlening zien als een portfolio van verschillende diensten, kunnen medewerkers op afdelingsniveau meer verantwoordelijkheid krijgen en besluiten nemen over de stra-

tegie voor hun betreffende onderdeel van het portfolio. Dit helpt gemeenten om op afdelingsniveau hun diensten te verbeteren en voor hun diensten te kijken waar vraag is naar digitale oplossingen.

4

Participeer in/initieer digitale platforms

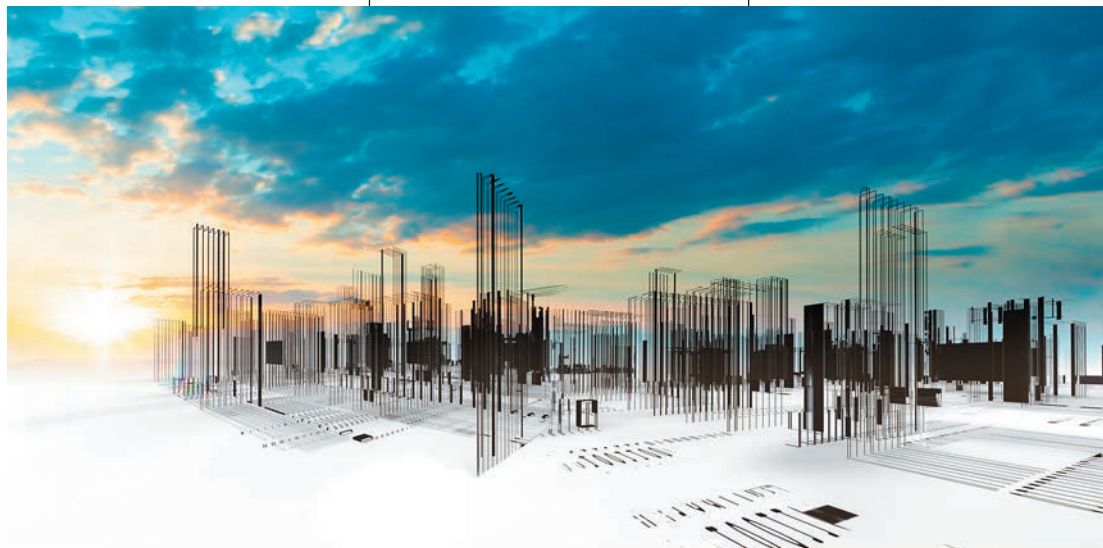
In de huidige samenleving zijn de meeste burgers online op zoek naar oplossingen voor problemen. Gemeenten kunnen hierop inspelen door burgers maar ook organisaties met elkaar in contact te brengen door middel van digitale platforms. Op deze manier verandert de rol van de gemeente in een faciliterende rol, waarin vraag en aanbod rondom zowel com-

merciële als maatschappelijke vraagstukken binnen de gemeente gemakkelijk kunnen worden samengebracht.

5

Participeer in/initieer digitale ecosystemen

Waar steeds meer burgers online antwoorden proberen te vinden voor hun vragen, zijn er ook steeds meer organisaties online actief waarmee kan worden samengewerkt. Gemeenten moeten zich meer bewust zijn dat veel digitale vraagstukken de virtuele grens van de gemeente overschrijden en dat ze met andere gemeenten of gespecialiseerde organisaties de samenwerking moeten opzoeken. <





Oorverdovende stilte (2)

Mijn vorige column ging over een docent die een foto van zichzelf op Instagram zag met een zeer kwetsende titel erboven, vermoedelijk geplaatst door leerlingen. Direct nadat ze hiermee was geconfronteerd, stapte ze vol vertrouwen naar haar leidinggevende. Een illusie armer kwam ze ervandaan. Ze werd wel aangehoord, maar actie bleef uit. Dat verwacht je niet op een school waar de kernwaarden veilig, eerlijk, respect, aandacht en ambitie prominent aanwezig zijn in publicaties en lessen. Als je het onderwijs en de omgang met anderen baseert op deze kernwaarden en op christelijke identiteit, dan zou je je veilig en geborgen moeten voelen. Tot voor kort had ik daar mijn handen voor in het vuur willen steken. Nadat de docent geen gehoor had gekregen bij haar teamleider, klopte ze aan bij andere leidinggevend.

Een enkeling toonde zich verontwaardigd, maar actie bleef uit. Het is echter een illusie om te denken dat het dan wel vanzelf overwaait. Het ongeloof, de verbijstering, de vernedering en de eenzaamheid waarin je alles moet verwerken, schud je niet zomaar van je af. De scherpe kantjes verdwijnen misschien, maar het blijft een 'onvergetelijke' ervaring. Tijdens een bijeenkomst van de sectie waar de betreffende docent deel van uitmaakt, kwam alles weer naar boven toen de vraag werd gesteld of medewerkers zich op deze school voldoende veilig achten. Ze brak en deed haar verhaal. De collega's deelden haar verontwaardiging, maar ook nu bleef een baanbrekend protest uit. Uiteindelijk volgde er wel een goed gesprek met haar teamleider. Ook heeft ze een vaste aanstelling gekregen. Inmiddels heeft ze ervoor gekozen om verder te gaan met haar leven en haar werk op school. De blik op de toekomst gericht. Het gaat haar lukken, maar toch ... Als het eropaan komt, moet je doen waar je voor staat; dan moet je kleur bekennen.

Tineke van Ommeren, voorzitter Beroepsvereniging Mediathecarissen Onderwijs (BMO) en mediathecaris Christelijk College De Noordgouw Heerde



ARCHIEF DAGEN

'Archivaris komt naast de burger te staan'

'Hoe ziet de toekomst van een archiefinstelling eruit?' Die prikkelende vraag stond onlangs centraal tijdens een online editie van de Archiefdagen, georganiseerd door KVAN en de VNG. Verschillende sprekers bespraken de uitdagingen van nu met een blik op de toekomst.

Sigfried Janzing, directeur Streekarchief Midden-Holland en medeauteur van de *Toekomst-verkenning Archiefdiensten* (VNG, 2020), trapte de bijeenkomst af. Met de Archiefwet en de Wet open overheid (Woo) worden we uitgedaagd om onze prioriteiten te herijken, aldus Janzing. De nadruk gaat liggen op 'actief', proactief archiveren en actief openbaar maken naast duurzame toegankelijkheid creëren. Volgens Janzing komt de archivaris naast de burger te staan. Minder een verlengstuk van de overheid, soms zelfs tégen die overheid. Daarvoor moet hij een andere informatiehouding gaan ontwikkelen die bestaat uit nieuwe vaardigheden (andermans taal spreken) en een nieuw frame hanteren (niet vanuit de informatie maar vanuit de burger denken). Janzing ziet een herontwerp van de archieffunctie voor zich richting 2030.

ARCHIEFFUNCTIE ONTWIKKELEN EN POSITIONEREN

Margriet van Gorsel (Stadsarchief Amsterdam) onderschreef het verhaal van Janzing. 'Preserveren is in essentie vindbaar maken en houden. Wij, professionals, moeten die preservering ook naar de bron brengen.' Zij pleit ervoor de archieffunctie integraal verder te ontwikkelen, zowel in de controlketen als in de informatieketen. Bert de Vries (KVAN) vond 2030 veel te ver weg. 'Archiveren zal al veel eerder

automatisch verlopen. Wat is dan de rol van de professional?' Ook voor De Vries staat de burger voorop. 'Om te komen tot een sterke maatschappelijke archieffunctie moeten we onze waarden vóór in de keten positioneren en daar een herkenbaar verhaal vertellen. We hebben als sector veel andere kennis en vaardigheden nodig. KVAN speelt daarop in met het register en de opzet van een programma lifelong learning.'

SAMENWERKING

Ilonka Sas (Gelders Archief) en Joris Uilenreef (Gemeente Utrecht) hamerden allebei op samenwerking tussen alle functionarissen die met informatiebeheer te maken hebben: 'We hebben een gemeenschappelijk belang!' En: 'We moeten als informatiebeheerders het onderwerp zelf hoog op de agenda zetten. Informatie moet je managen, net zo goed als mensen en geld. Gebruik het Strategisch Informatieoverleg (SIO) en wees een geïnformeerde gesprekspartner. Technologie helpt ons bij ons werk, maar houding en gedrag op de werkvloer en de sturing daarop zijn zeker zo belangrijk. De CIO en de gemeentearchivaris moeten elkaar goed kunnen vinden, want zij hebben een gedeeld belang.'

Je kunt deze online sessie terugzien op YouTube via het kanaal 'Archiefdagen', aflevering 'Inspiratiesessie in samenwerking met de VNG'.

In welke tv-serie zitten we straks?

Het is soms bijna niet voor te stellen dat onze kinderen (of hun kleinkinderen) over een jaar of vijftig net zo terugkijken op onze 'moderne' tijd als wij dat doen op de tijd van onze ouders en grootouders. Bij ons in huis staat als herinnering aan 'vroeger' nog steeds een oude bakelieten telefoon met draaischijf. Mijn dochter weet nu in ieder geval nog wat dat is, mocht het een keer in de geschiedenisles op school terugkomen. Een plaatje van bijvoorbeeld de Harvard Mark I, ook wel bekend als de IBM Automatic Sequence Controlled Calculator (ASCC), ziet er bijna romantisch uit.



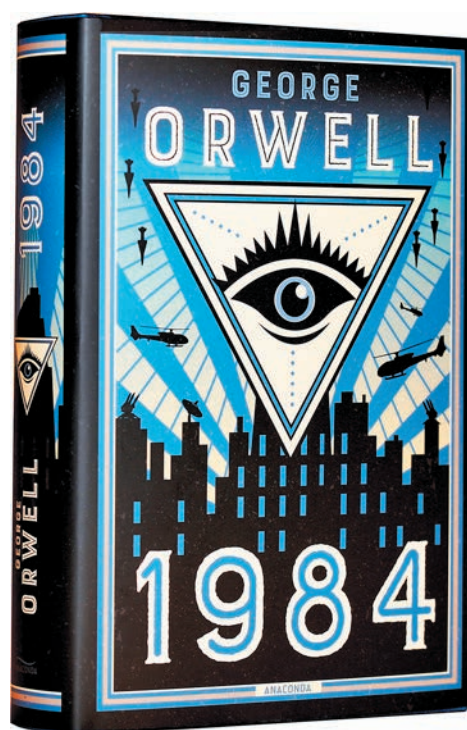
VOORUITDENKENDE (SCRIPT)SCHRIJVERS

Tegenwoordig kijken we met kleine nudging en een gealgoritmiseerd tv-aanbod via Netflix de mooiste series. Ik heb een persoonlijke voorkeur voor sciencefiction, boeken en tv: wegdromen met vooruitdenkende scriptschrijvers. Het is interessant wat deze schrijvende breinen voor ons bedenken. Het valt op dat sommige schrijvers een goed beeld kunnen neerzetten van een mogelijke toekomst. George Orwell zal zichzelf geen sciencefictionschrijver hebben genoemd. Met *Animal Farm* en *1984* uit respectievelijk 1945 en 1949 gaf hij een blik op ontwikkelingen die nu soms realiteit zijn. Het lijkt zelfs af en toe wel dat het gedachtegoed van 'newspeak' nieuw leven wordt ingeblazen.

DYSTOPISCHE SAMENLEVING

Series als *Colony* en *The Handmaid's Tale*, maar ook de Canadese serie *Continuum*, zijn voorbeelden. Vanuit deze series zijn prachtige parallellen te trekken naar wat we nu ervaren. De macht van corporaties die groter is dan die van de overheid, zoals in *Continuum*, leest als een mogelijke storyline van het boek *Society 4.0* van prof. dr. Bob de Wit. In andere series is het beangstigend om te zien hoe een dystopische samenleving eruitziet als alle macht in 'Davos' komt te

liggen. Of als je met een mobiele telefoon in de hand en drones in de lucht altijd iemand kunt volgen. Of als de wereld door vervuiling en chemische stoffen wordt verdeeld in groepen, in het bijzonder de groep vruchtbare vrouwen.



SMART HUMANITY

Ik ben vast niet de enige die een aantal voorbeelden in het hier en nu denkt terug te zien. Bijvoorbeeld als ik in de rij sta te wachten tot ik op basis van een onbetrouwbare test de landsgrens over mag – mits ik een uitnodigingsbrief kan overhandigen. Of als ik lees over het 'teruggeven van vrijheden' door de autoriteiten – zoals minister De Jonge zei op 7 mei. Of als ik op de site van de fabrikant lees welke tests (voor vruchtbaarheid én voor zwangere vrouwen!) er nog moeten worden gedaan voorafgaand aan de formele goedkeuring van de vaccins. Of als ik lees dat er voor de tweede keer (Chinees) ruimtepuin naar beneden stort. De vorige keer werd een kaasmakerij in Afrika geraakt, nu stortte het puin in de Indische Oceaan bij Sri Lanka. Dat lijkt wel heel erg op scènes uit de genoemde boeken en tv-series.

In het kader van Smart Humanity en de rol van al die (informatie)professionals die onze samenleving mede bedenken, maken, beheren en de technologie helpen toepassen, is mijn vraag: in welke tv-serie belanden wij straks als samenleving?

Wouter Brongseest, duovoorzitter KNVI



Mibo



Topia



Gather

GEZELLIGHEID IN CYBERSPACE

Rondhangen op een eiland met **MIBO**

Ook al is er uitzicht op verbetering nu het aantal coronabesmettingen daalt en de vaccinatiegraad stijgt, toch kunnen we de anderhalvemeterafstand nog niet loslaten. Drukbezochte fysieke vrijdagmiddagborrels zitten er daarom nog niet in. De bekende tools Zoom en Teams hebben grappige functies als stickers en virtuele achtergronden toegevoegd om ook voor informele ontmoetingen te kunnen worden ingezet, maar ze halen het niet bij het Nederlandse Mibo. Via getmibo.com kom je terecht op een tropisch eiland in 3D. Je kunt er naar hartenlust ronddwalen, waarbij je vanzelf anderen tegen het lijf loopt. Je lichaam is een zwevende robot met je webcam als hoofd, en je hebt real-time verbinding. Dankzij positionele audio hoor je alleen de mensen die dicht bij je staan. Dit maakt van Mibo de perfecte omgeving om op een persoonlijke manier in kleine groepjes te socializen. Ga met elkaar rond de tafel zitten in de bar om lekker bij te kletsen, speel een potje basketbal, of zwem samen een rondje om het eiland. Ook is er een podium voor een centraal momentje.

Mibo is een verfrissende ervaring voor weer eens een echt goed gesprek met collega's. In de gratis versie kun je met maximaal twaalf personen op je eigen eiland vertoeven. In de

Corona zorgde voor de doorbraak van videovergaderen. Het gebruik van Zoom en Teams verveelvoudigde, maar voor het koffiepraatje of de vrimibo voelen deze tools te zakelijk aan. Informelere alternatieven zijn Mibo, Topia of Gather.

betaalde versie kunnen honderden deelnemers over meerdere (thema)eilanden worden verdeeld.

Creëer je eigen ontmoetingsplek met **TOPIA** of **GATHER**

Met Mibo kom je in de gratis versie standaard terecht op een eiland. Wil je gratis je eigen virtuele ontmoetingsplek creëren, maak dan gebruik van Topia of Gather. Deze platforms zijn het best te beschrijven als een combinatie van de wereld van Zoom en het Nintendo-spel Super Mario Bros. Je avatar loopt in een gameachtige wereld waarin je andere avatars tegenkomt. Via webcamschermpjes zie je de mensen achter de avatars, vergelijkbaar met Zoom of Teams. Deze schermpjes verschijnen of verdwijnen naargelang de bijbehorende avatars bij elkaar komen of van elkaar weglopen.

Bij Topia is je avatar een gekleurd figuurtje. Je kunt Topia leren kennen via topia.io/welcome, de welkomstlobby die iedereen kan betreden. Via topia.io/welcome?createworld=1 kun je je eigen wereld creëren. De mogelijkheden zijn legio. Standaard kom je in een bos- en rotsachtige omgeving terecht, maar je kunt ook een tentenkamp bouwen of zelfs een hele stad. Nodig anderen uit om elkaar in deze wereld te ontmoeten en mee te bouwen. Topia werkt zo leuke creatieve ontmoetingen in de hand.

Bij Gather heb je keuze uit diverse avatars met diverse outfits. Via gather.town/app kun je bij het creëren van een eigen wereld kiezen uit templates. De gratis versie van Gather heeft 24 kantoortemplates zoals een gezellig kantoor (inclusief koffieapparaat voor het spontane koffiepraatje) of een ruimtestationkantoor, dertig sociale templates zoals een eiland of een bibliotheek, drie onderwijsttemplates zoals een auditorium of een klaslokaal, drie conferentietemplates en drie blanco templates. Gather zet zichzelf in de markt als een combinatie van videogesprekken met een 2D kaart. Meerdere mensen kunnen afzonderlijke gesprekken tegelijkertijd voeren, waarbij ze net zo gemakkelijk in en uit die gesprekken kunnen lopen als in het echte leven. Maar een deel van de templates ondersteunt dus ook formelere bijeenkomsten als colleges of congressen. Met de gecreëerde wereld waarin de deelnemers zich bevinden, kun je deze bijeenkomsten echter meer inspirerend maken dan in de doorsnee Teams- of Zoom-bijeenkomsten.

In de gratis versies van zowel Topia als Gather kunnen 25 mensen tegelijk in je gecreëerde wereld deelnemen.

**Leen Liefsoens**

Redacteur van IP
en senior informatiespecialist
bij de bibliotheek van
De Haagse Hogeschool



Wie: Marie Ciliène Ogvil

Wat: informatieprofessional bij de dr. Jules Sedney Bibliotheek van de Centrale Bank van Suriname

FILM

Ik heb een aantal favorieten, maar een film die voor mij niet oud wordt, is *Diary of a Mad Black Woman*. Hierin kiest een vrouw er vanuit kracht voor om haar man bij te staan na een ongeluk, ondanks zijn slechte gedrag naar haar. 'Behandel mensen zoals je zelf behandeld wilt worden.'



APP

Een aanrader vind ik Duolingo, een gebruiksvriendelijke app om op een eenvoudige en vriendelijke manier een andere taal te leren. De app heeft meer dan twintig talen beschikbaar die je in je eigen tempo eigen kunt maken, met oefeningen gebaseerd op jouw leerstijl.

MUZIEK

Ik houd van gospelmuziek ter ondersteuning van mijn geloof, maar ook kompa is een favoriet van mij. Dat is een moderne merengue, een Haïtiaanse muziekstijl, die sinds 1950 wordt gezongen en gedanst.

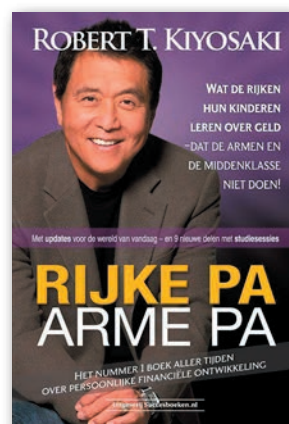


GAME

Ik heb geen favoriete game, maar een woordzoeker vind ik wel fijn om af en toe te doen. Deze puzzel verrijkt je woordenschat en stelt je geduld op de proef, vooral als bepaalde woorden moeilijk te vinden zijn.

WEBSITE

Handig om af en toe spullen bij te kopen is Aliexpress.nl, zeker in de afgelopen tijd – door corona is online shoppen niet meer weg te denken in de wereld. De website is gebruiksvriendelijk en je vindt er van alles, van kleding tot huishoudelijke artikelen.



MUSTREAD

Niet per se een favoriet, maar wel een aanrader: *Rich Dad Poor Dad* van Robert Kiyosaki. Een bestseller, vertaald in meer dan 51 talen. Het boek gaat over persoonlijke financiële ontwikkeling en leert je hoe je hersenen te gebruiken als het gaat om investeren.

Met de kennis van (n)u!

schrijf u in voor onze webinar

www.ingressus.nl/evenementen/

Volg een van onze gratis online kennissessies over het optimaal aanbieden en toegankelijk maken van uw (digitale) collectie & content via de Vakliteratuur app.

De Hoge Raad der Nederlanden ging u al voor in het inrichten van een persoonlijke kennisomgeving.

INGRESSUS

www.ingressus.nl/evenementen/