

# De Nationale Verjaardags- Kalender

2015—2017  
Eindverslag

**SETUP**

# Inhoud

---

<b>Inleiding</b>	<b>2</b>
<b>Resultaten</b>	<b>3</b>
<b>Deel I: Het bouwen van de verjaardagskalender</b>	<b>5</b>
<b>Deel II: Journalistieke publicaties</b>	<b>9</b>
<b>Deel III: De theatervoorstelling</b>	<b>11</b>
<b>Deel IV: Lezingen, voorstellingen en workshops</b>	<b>14</b>
<b>Conclusie</b>	<b>16</b>

## **Bijlagen**

1. Serverontwerp
2. Correspondentie CPB
3. Contracten hackathons
4. Verzamelde data
5. Publicatie 'Negermuziek ja, lesbische boeken nee'
6. Overzicht hobby's
7. Publicatie 'Cyber-onvoldoende'
8. Alle eigen publicaties NVK
9. Alle spin-off publicaties NVK
10. Foto's van de NVK-voorstelling

# Inleiding

---

## Tussen eind 2015 en start 2017 draaide bij SETUP het project De Nationale Verjaardagskalender.

De aanleiding was de sterke opkomst van ‘data brokers’, bedrijven die zich specialiseren in het verzamelen en verkopen van grote hoeveelheden persoonlijke informatie. Soms gaat dat om ‘harde’ data – zoals een geboortedatum – maar het kan ook om afgeleiden gaan; deze persoon vindt X leuk, dus zal dan wel karaktereigenschap Y hebben. De wereld van big data is enorm, maar ook sterk verborgen. Voor SETUP reden om zelf te onderzoeken hoe ze in de praktijk werkt.

In De Nationale Verjaardagskalender was het doel om een verjaardagskalender van alle Nederlanders te bouwen, op basis van openbare en/of legaal te verkrijgen bronnen. Enerzijds met informatie als naam, adres en geboortedatum, anderzijds met specifieke cadeausuggesties op basis van profiling. Naast een technisch en deels juridisch onderzoek (‘kan/mag dit?’), was het project ook vooral een maatschappelijk onderzoek. Welke ethische kwesties kom je tegen? Welke reacties ontstaan er vanuit bedrijven en publiek? De verhalen en ontdekkingen die we deden, vormden de basis voor een reeks journalistieke publicaties, een theatervoorstelling en diverse presentaties en workshops. Hiermee wilden we het publieke debat rondom big data toegankelijker maken, zodat we kritischer en genuanceerder over dit onderwerp kunnen praten.

# Resultaten

---

**Uiteindelijk hebben we binnen dit project de volgende onderdelen gerealiseerd. In de opeenvolgende hoofdstukken lichten we deze nader toe.**

## **HACKATHONS**

4 hackathons waarin we datapunten vonden van 7.8 miljoen Nederlanders. Hierbinnen waren er 780.000 terug te voeren op een uniek persoon met Naw-gegevens en voldoende data om profiling mee te doen (en een cadeausuggestie te genereren).

## **SOFTWARE & SECURITY**

In de periode van deze hackathons is er zelf software ontwikkeld om de verschillende databronnen te koppelen en doorzoekbaar te maken. Deze software en data is enkel opgeslagen op een gehardende server met onder andere encryptie, boot via een USB-stick en niet verbonden met het internet. Al deze maatregelen zijn gecontroleerd door [Madison Gurkha](#), toonaangevend op het gebied van IT security in Nederland. Juridisch hebben we in samenwerking met de [Privacy Company](#) het hele proces afgestemd met het College ter Bescherming van Persoonsgegevens.

## **PUBLICATIES**

Over dit proces schreven we 5 inhoudelijke artikelen, die onder andere zijn verschenen op *Vice*, in het *VNG Magazine* en op *TMI.News*. Andere artikelen en ook 4 kortere blogs verschenen op de website van SETUP zelf. Daarnaast werd er in diverse media over dit project geschreven (15 artikelen in bladen als bijvoorbeeld *De Nieuwe Revu*, waarvan ook 4 in internationale media, waaronder de Duitse *Die Zeit* en *Der Spiegel*). Ook waren we in tv-programma's als *EenVandaag* te zien.

## **VOORSTELLING**

Om de inzichten uit dit project ook voor een breder publiek invoelbaar te maken, maakten we samen met theatermakers Margit Odems en Ingrid Govers een interactieve voorstelling. In deze voorstelling stapt de bezoeker in de rol van het profilingsalgoritme, om aan de hand van de data-persona zelf toepasselijke verjaardagscadeaus te verzinnen. Deze voorstelling is 11 keer gespeeld op verschillende plekken in Nederland.

## POCKETEDITIE

De interactieve voorstelling is ook vertaald is naar pocketeditie, die in de vorm van een debatspel een soortgelijke ervaring overbrengt. Deze vorm is goedkoper en beter overdraagbaar naar onder andere scholen en bibliotheekmedewerkers.

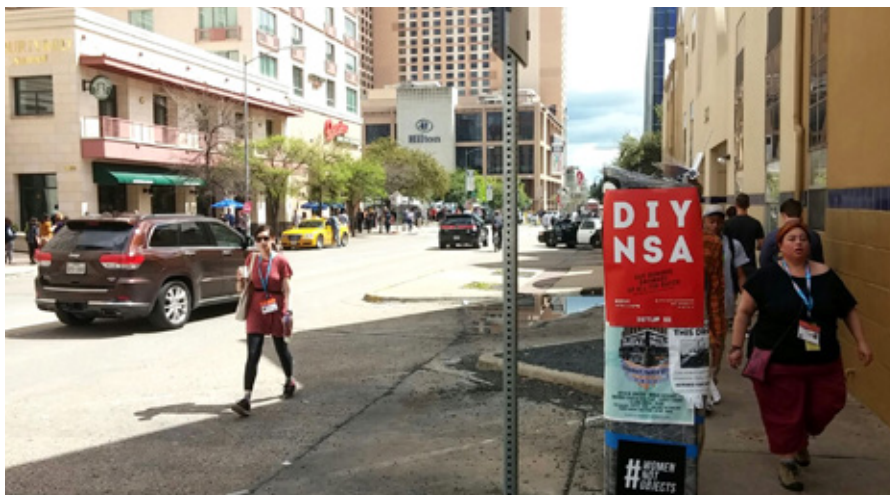
## PRESENTATIES

We gaven 11 grote presentaties over De Nationale Verjaardagskalender, met als hoogtepunt onze talk op het internationale festival *South by Southwest (SXSW)* in Austin, Texas. Ook spraken we op onder andere BNN-radio over dit project.

Foto uit een  
hackathon.



Een SETUP-poster in  
het straatbeeld van  
Austin, Texas.



# Deel I

---

## Het bouwen van de verjaardagskalender

### UITGANGSPUNTEN

We begonnen het project met de doelstelling dat 20 hackers, data-experts en journalisten mee zouden doen in 2 hackathons. Met hen wilden we van een vierde van alle Nederlanders een datapunt vinden en van 1000 Nederlanders echt – eng – veel data verzamelen. Een centraal uitgangspunt was ‘we adviseren niet dat *iedereen* zelf een database van alle Nederlanders gaat bouwen, maar als je het wilt doen, zijn dit de veiligste en meest ethische methoden’. Enerzijds zat dit in de veiligheid, die we waarborgden door een serverplan te laten ontwikkelen en uitvoeren (door het bedrijf *Rootless*). Vervolgens werd deze in een onafhankelijke audit gekeurd, samen met de meer globale werkwijze rond het werken met de data (door *Madison Gurkha*). Naast de technische kant wilden we ook garanderen dat we altijd binnen de kaders van de Wet Bescherming Persoonsgegevens zouden werken en eventuele ethische kwesties goed konden onderbouwen. Hiervoor schakelden we de hulp in van de *Privacy Company*. Zij adviseerden ons in de correspondentie met College Bescherming Persoonsgegevens (inmiddels [Autoriteit Persoonsgegevens](#)) en het opstellen van contracten voor de deelnemers van de hackathons. Bij de hackathons waren ook ervaren privacy officers aanwezig om toe te zien op de naleving van de regels en te adviseren bij twijfelgevallen.

Zie Bijlage 1 voor een overzicht van dit ontwerp.

Zie Bijlage 2/3 voor deze correspondentie en contracten.

Een ander belangrijk uitgangspunt van De Nationale Verjaardagskalender is dat deze volledig is gebouwd met openbaar toegankelijke data (al dan niet met scraping verkregen), en/of legaal verkregen datasets (bijvoorbeeld tegen betaling).

Dit project is neergezet binnen de journalistieke en artistieke uitzondering op de Wet Bescherming Persoonsgegevens. We hebben echter wel vastgelegd dat we ons houden aan gegevensminimalisatie; we verzamelden alleen gegevens voor ons doel (cadeausuggesties genereren) en vermeden bijzondere (gevoelige) persoonsgegevens (politieke voorkeur, etniciteit of gegevens van minderjarigen).

### HACKATHONS

Met de toezegging van het CPB, een technisch veilige server en de aanwezigheid van privacy officers begonnen we aan de hackathons om onze verjaardagskalender te bouwen. In totaal hebben we vier hackathons georganiseerd. Elke sessie bestond uit een volledige zaterdag, waar per keer zo'n 10 tot 20 ontwikkelaars aanwezig waren. Tussen alle sessies zaten enkele weken om nieuwe data te integreren in de bestaande database. De eerste hackathon stond vooral in het teken van het verzamelen van data, de tweede in het koppelen van datasets en de laatste twee in het ontwikkelen van een profilingsalgoritme dat cadeausuggesties kon genereren.

Zie Bijlage 4 voor een overzicht van deze data.

Tijdens deze hackathons kwamen we een aantal interessante zaken tegen. Zo ontdekten we dat het Googlen van een term als ‘ledenadministratie’ en ‘filetype:xls’ verrassend veel

directe databronnen opleverde. Een deel van deze ledenadministraties was zelfs gevoelig. Zo vonden we bijvoorbeeld zo'n administratie van een dagbestedingscentrum, waar naast voor- en achternamen ook wachtwoorden van verschillende leden in stonden.

Ook bleken sportclubs een data-goudmijn te zijn. Veel voetbalclubs gebruiken hetzelfde websysteem om hun leden te beheren, waardoor we in totaal zo'n 70.000 voetballers met voor- en achternaam konden vinden. Hiervan bleek uiteindelijk 60% minderjarig, dus zij zijn uiteindelijk niet toegevoegd aan onze kalender.

Onze hackathons hebben uiteindelijk ook geleid tot concrete aanpassingen. Zo vonden we bij de *Nederlandse Beroepsorganisatie van Accountants (NBA)* een overzicht van alle Nederlandse accountants met voornaam, achternaam, adres en geboortedatum. Naar aanleiding van een interview over dit project op BNN-radio – waarin we dit voorbeeld noemden – namen zij contact met ons op. Ze verzochten ons om de data uit onze kalender te halen, maar hebben ook hun eigen register aangepast. Zo zijn voornamen aangepast naar enkel initialen en is de geboortedatum aangepast naar geboortejaar.

Een andere interessante casus die we tegenkwamen was het forum [50plusser.nl](https://50plusser.nl). Op deze website gaven deelnemers verrassend veel informatie weg, waaronder biografie, hobby's, lievelingsmuziek, favoriete boeken en vakantiebestemmingen. Dat gaf ons een schat aan informatie om cadeausuggesties te genereren. Echter, toen we deze data in detail gingen bekijken, bleek dat er ook heel veel gevoelige informatie in zat. Een antwoord als 'sinds mijn hernia ben ik arbeidsongeschikt' onder het kopje 'werk', dat is een medisch gegeven. Andere gegevens zaten meer in grijs gebied; een antwoord als 'lesbische boeken' onder hobby's, zegt dat ons iets over seksuele voorkeur? Met onze privacy officers hebben we voor al deze gegevens een inschatting gemaakt, waarmee we een checklist van 'verboden woorden' samenstelden.

Zie Bijlage 5 voor een argumentatie achter deze keuzes.

## SOFTWARE

Om de data te verwerken, hebben we zelf software ontwikkeld. Het voornaamste doel hiervan was het koppelen van verschillende datasets, zodat er een completer beeld per persoon ontstaat. Dit is voor ons ook één van de belangrijkste lessen uit De Nationale Verjaardagskalender; kleine, losse datasets lijken heel onschuldig, omdat ze weinig informatie weggeven. Het wordt echter een probleem wanneer iemand die verschillende sets aan elkaar koppelt; opeens ontstaat er dan een bijzonder nauwkeurig profiel met allerlei nieuwe af te leiden informatie.

Foto uit een hackathon.



De software was ook nodig om de datasets te koppelen aan 'standaarddata'. Hiermee kunnen nieuwe datasets worden gekoppeld en gevalideerd. We gebruikten hier twee hoofdzakelijke bronnen voor: de namen-databank van het Meertens Instituut en een database van alle Nederlandse postcode met daaraan gekoppelde straten, steden, gemeenten en provincies. Met deze data konden we valideren dat bepaalde namen in bepaalde regio's voorkomen; dat geeft een 'schoon' datapunt om om de koppeling te maken, omdat het meer zekerheid biedt dat het om dezelfde persoon gaat.

Een ander bijkomstige voordeel van de database van het [Meertens Instituut](#) is dat we hiermee ook een schatting konden maken over de compleetheid van onze datasets. Voor onze presentatie op TEDx Utrecht konden we bijvoorbeeld uitrekenen dat we 25 van de 695 Nederlanders met de voornaam 'Ted' in onze database hadden.

Aanvullend is er ook software geschreven om de database te doorzoeken. Deze software kwam van pas bij onder andere presentaties, waar we bezoekers konden laten controleren of zij zelf in De Nationale Verjaardagskalender stonden. Ook bij lezingen keken we op deze manier live of we aanwezig uit de zaal hadden. Hiermee werden de presentaties over dit project voorzien van een spannend live element, dat sterk tot de verbeelding sprak.

Tot slot hebben we software ontwikkeld om specifieke profilingsvragen te beantwoorden. De database van [50plusser.nl](#) was hier een waardevolle bron voor. Met de software kunnen we interesses per persoon opvragen, maar andersom ook personen met specifieke interesses vinden. Zo weten we bijvoorbeeld dat we 7 personen hebben die van dansen houden en in Almere wonen. Met deze data ontstaat een inzicht van verschillende soorten hobby's per groep.

Zie Bijlage 6 voor een overzicht van deze hobby's.

Om de mogelijkheden te ontdekken, hebben we zelf een aantal zoekvragen opgesteld. Eén daarvan was "Wat was de soundtrack bij je eerste tongzoen?". Om deze vraag te beantwoorden, hebben we als eerst een database van alle historische top-40-lijsten ingevoerd. Deze dataset is vervolgens doorzoekbaar op basis van geboortjaar en laat zien welke hits er in de zomer van een 11e tot 16e levensjaar in de top 40 stonden. Met een andere dataset konden we namelijk kijken wat historisch de gemiddelde leeftijd voor deze eerste zoen was. Samen ontstaat er een indicatie van welke hits er op de achtergrond in de campingdisco werden gedraaid.

Deze meer ludieke analyses zijn aannames; deels onderbouwde *mogelijke* scenario's. Dat lijkt in eerste instantie minder bruikbare data, maar het tegendeel is waar. Veel databrokers maken namelijk ook dit type 'afgeleide data', waarin het gaat om (subjectieve) analyses van meer 'harde data'. Die afgeleide data wordt soms ook weer verkocht als nieuwe data. Dat kan bijvoorbeeld gaan om 'de mate waarin persoon X gevoelig is voor autoriteit'; niet direct meetbaar, maar vanuit een bepaalde analyse het meest aannemelijke antwoord.

## DE AUDIT

Het sluitstuk voor de ontwikkeling van De Nationale Verjaardagskalender was uiteindelijk een technische audit door [Madison Gurkha](#). In dit rapport is gekeken of ons beveiligingsconcept en de uitvoering daarvan adequaat genoeg waren. Het antwoord was zowel 'ja' als 'nee'.

'Ja', de gekozen beveiliging heeft in een specifiek geval nut:

*"De gekozen verharding van het systeem heeft op dit moment maar in specifieke gevallen nut. De Boot USB-stick in combinatie met de passphrase hebben alleen nut indien ze gescheiden opgeslagen zijn en het systeem niet aan staat."*



Maar dus ook 'nee'; er zijn allerlei andere scenario's denkbaar waarin dit niet het geval is. Deze afweging hebben we echter bewust gemaakt, omdat we de server alleen gebruikten in een lokaal netwerk tijdens de hackathons. We hebben maatregelen genomen voor scenario's als 'de server wordt gestolen', maar niet voor schadelijk gedrag van de hackathon-deelnemers die *werkten* met de database. Dit deel hebben we ondervangen door hen een contract te laten tekenen.

Het rapport van de audit kunnen we niet met derden delen. Volgens Madison Gurkha zou dat inzicht geven in hun werkwijze, waarmee het kwaadwillenden ook inzichten zou geven om zwakke plekken te benutten. Wel hebben we een reflectie op de audit geschreven.

Zie Bijlage 7.

## CONCLUSIES EN REFLECTIE

De Nationale Verjaardagskalender bleek een bron van interessante verhalen en ontdekkingen. De tweeledigheid – kan het technisch, kan het ethisch – zorgde voor een constante spanning en dialoog tussen SETUP, de ontwikkelaars, de databezitters (*Funda*, *NBA*, et cetera) en de waakhonden (*CBP*, *Madison Gurkha*, et cetera). Het concept van de verjaardagskalender bleek ook een goed frame, omdat het ons de mogelijkheid gaf om de werking van profiling heel tastbaar te maken. Zo kan de data van Funda bijvoorbeeld worden gebruikt om in te schatten welk huisdier bij iemand past; heeft hij/zij ruimte voor een geit of kippen, of blijft het bij een cavia? Dit soort vragen laat zien dat profiling uiteindelijk geen harde wetenschap is, maar sterk afhankelijk van de gecombineerde datasets en gekozen interpretaties. Kortom, big data is geen exacte wetenschap, er zitten allerlei subjectieve aannames in. Daarnaast bleek het concept van de verjaardagskalender ook goed in te spelen op het tech-startup-gevoel; een aantal keer werd ons in de media gevraagd of dit een echte, serieuze dienst was. Handig, zo vonden sommigen. Die verwarring spreekt boekdelen. Voor ons is dit een kritisch project met een knipoog, voor bepaalde bedrijven is dit dagelijkse realiteit. Zelfs wij waren bijvoorbeeld verbaasd toen bleek dat we legaal een grote dataset van 'kamperende Christenen in Nederland' konden kopen.

De Nationale Verjaardagskalender heeft tijdens de ontwikkeling ook gezorgd voor meer bewustzijn bij bedrijven, zoals het NBA dat haar registers heeft aangepast, of Funda die telefonisch informeerde of zij ook in onze database zouden komen. Ook in de correspondentie met het CPB ontstonden voor beide partijen interessante nieuwe inzichten. Tot slot hebben we het gevoel dat – als we nog meer hackathons zouden doen – we nog meer van dit soort verhalen zouden ontdekken. Bij SETUP kijken we dan ook of we in de toekomst vaker dit soort hackathons kunnen doen, waar we zowel technisch als ethisch onderzoeken wat er kan/mag en zo nieuwe kwesties aankaarten.

Een tegenvaller in dit project was dat het ontwikkelen van de verjaardagskalender niet parallel maar sequentieel liep aan de journalistieke publicaties, theatervoorstelling, workshops en lezingen. We hadden gehoopt dat de theatermakers en journalisten samen met de hackers aan de slag konden gaan om hun onderzoek te stimuleren en vormen. De kloof tussen deze groepen bleek uiteindelijk te groot om echt een sterke dialoog te creëren. In plaats daarvan werkten zij vaker met tussentijdse, afgeronde resultaten. SETUP heeft in dit proces daarom sterker opgetreden om de dialoog te stimuleren, maar dat gebeurde wel in aparte sessies.

Een andere moeilijkheid zat in het tussentijds concreet maken van resultaten. Het nieuws dat SETUP alle Nederlanders wilde vangen in een database werd snel opgepikt door de media. Het bleek daarna niet eenvoudig om die spanningsboog vast te houden, vooral omdat we in het begin relatief weinig konden laten zien. We konden immers externen geen toegang tot de database geven, of de server meenemen naar andere locaties. Dit hebben we opgevangen door relatief snel met eigen blogs onze bevindingen te delen. Ook vonden we trucs om toch 'live' te kunnen werken. Tijdens onze presentatie op *SXSW* belden we bijvoorbeeld live naar ons eigen kantoor, om telefonisch een naam uit het publiek door de database te halen en de informatie weer terug te spelen.

# Deel II

---

## Journalistieke publicaties

### UITGANGSPUNTEN

We hadden als doel om minimaal 3 kleine en 3 grote artikelen vanuit het journalistieke onderzoek in de media te krijgen. Daarnaast spraken we de ambitie uit om deze publicaties te richten op plekken waar het genuanceerde debat over big data mist.

In ons aanvankelijke plan zou er een journalist meelopen tijdens de verschillende hackathons, en dit onderzoeken verder vormgeven met experts. Hier hadden we Douwe Schmidt voor geselecteerd. Tijdens het proces lukte het Douwe echter niet om effectief vorm te geven aan dit onderzoek, waardoor de kloof met de hackers te groot bleef. Om snel te schakelen, hebben we vanaf de tweede hackathon andere journalisten uitgenodigd om te schrijven over dit project. Dat leidde tot onder andere een uitvoerig stuk in *De Nieuwe Revu* en *Het Parool*. Het echte journalistieke onderzoek konden we echter pas na de reeks hackathons oppakken.

### EIGEN PUBLICATIES

De eerste publicatie is geschreven door [Anne Schepers](#) (*De Correspondent*, *EenVandaag*, *DeBalie*). Zij richtte zich op gemeenteambtenaren, na onze ontdekking dat veel gemeenten hun geboortes openbaar publiceren. Dit stuk is verschenen in het *VNG magazine* (Vereniging van Nederlandse Gemeenten).

Een tweede publicatie werd geschreven door [Remmelt Heltzel](#) (*BNN-VARA*, *NPO Radio 2*) en verscheen op *TMI.news* (een nieuwe en veel bekroonde journalistieke startup). Ook werkten we met TMI aan een videoserie over data en profiling, waarin we de kennis uit *De Nationale Verjaardagskalender* haalden. Aan deze series werkten ook Maurits Martijn (*De Correspondent*) en Esther Keymolen mee. Deze serie richtte zich specifiek op jongeren.

De derde publicatie kwam van [Xander Bouwman](#) (*Bits of Freedom*, *Taskforce Cyber*) en haakte in op de audit van Madison Gurkha. In dit stuk gaat hij in op onze aanpak en de risico's die we wel en niet hebben ondervangen. Dit stuk lieten we samenvallen met de ontvellingen van Zembla over de 'data-kraamkamer' van de Belastingdienst. De conclusie; wij volgden in ons project eigenlijk precies dezelfde redenering als de Belastingdienst had. Dit stuk was bedoeld door het Madison Gurkha magazine (voor professionals in de ICT-sector), maar vanwege de prangende actualiteit hebben we deze zelf gepubliceerd.

Daarnaast zijn er gedurende dit project meerdere eigen kleine publicaties en blogs verschenen. Onder andere over onze correspondentie met het CPB, een verslag van onze presentatie op SXSW en een inhoudelijk stuk over onze werkwijze met de vijftigplusser.nl-dataset.

Na al deze publicaties bleef bij ons een behoefte om nog verder in te zoomen op de achtergrond van dit project. Veel media – met name na onze presentatie op SXSW – vonden nieuwswaarde in het feit dat we een eigen database bouwden (“**Wir bauen uns eine NSA**”, *die Zeit*). We hebben daarom naar een nieuwe angle gezocht om de achtergrond opnieuw te presenteren. Dit deden we met socioloog **Ruben Jacobs**, die een stuk heeft geschreven voor *Vice* en het Volkskrant-magazine *Vonk*. Het *Vice*-stuk gaat in op gedrag en hoe dit door de tijd heen veranderd. Ruben gebruikt het oude Hyves-archief om te laten zien hoe we ons anno 2010 publiekelijk presenteerden. Daarin zien we dat de norm inmiddels al verschoven is.

Zie Bijlage 8 voor een overzicht van deze publicaties.

Naast deze eigen publicaties zijn er ook in de media diverse artikelen over dit project verschenen. Zie bijlage 9 voor een overzicht.

## CONCLUSIES EN REFLECTIE

Bij het schrijven van de publicaties merkten we dat dit journalistieke proces niet parallel aan de ontwikkeling van het technische deel kon lopen. Er zat hier een te grote kloof tussen de hackers en journalisten, waardoor de laatste groep afwachtender op resultaten werd. Ook kreeg dit project eigenlijk – deels onbedoeld – al te vroeg veel media-aandacht, simpelweg door het feit dat we hiermee aan de slag gingen. Daar zijn we blij mee – met name als het om kanalen als *EenVandaag* gaat, – maar dat maakte het lastig om het achterliggende onderzoek naar de voorgrond te schuiven. Op dat moment stond er namelijk nog geen journalistiek team klaar. Het kostte daardoor wat meer tijd om later in het project een aantal nieuwe haakjes te vinden om de media opnieuw mee te bereiken (zoals de artikelen van Ruben Jacobs). In vervolprojecten zorgen we daarom dat we journalistiek team eerder aanhaken en voorbereiden.

Artikel in het AD.



# Deel III

---

## De theatervoorstelling

### UITGANGSPUNTEN

Het derde onderdeel van De Nationale Verjaardagskalender was de ontwikkeling van een theatrale ervaring. Hiermee wilden we naast het inhoudelijke verhaal ook een ervaring bieden die op een meer persoonlijk en emotioneel niveau zit. Ook zagen we een uitdaging in de vertaalslag van deze thematiek naar een artistiek werk dat een ander publiek aanspreekt. De installatie moest per definitie geschikt zijn om rond te reizen; zo konden we ook naar plekken toe waar dit debat minder sterk leeft (zoals festivals en openbare bibliotheken door heel Nederland).

### DE BETROKKEN MAKERS

De ontwikkeling, productie en uitvoering lag in handen van theatermakers Margit Odems en Ingrid Govers en technisch studio oneseconds.

We kozen voor [Margit Odems](#) omdat zij in haar voorstellingen probeert om de realiteit op een theatrale manier te verbeelden. Publiek is in haar werk altijd een onmisbaar deel van de voorstelling. Margit onderzoekt vaak het spanningsveld tussen (de identiteit van) het individu versus de massa. Dan gaat het vaker over de mens die zich in allerlei systemen bevindt, maar waar hij niet zomaar uitstapt. Die insteek past naadloos bij onze database van alle Nederlanders. Daarnaast sprak ook de lichtheid en toegankelijkheid van Margit's voorstellingen ons aan.

Margit is vervolgens gekoppeld aan kunstenaar en theatervormgever [Ingrid Govers](#) (1984). Zij studeerde theatervormgeving aan de HKU en was als theatermaker verbonden aan theatergroep *Krabben*. Hierin onderzocht ze hoe je alledaagse handelingen theatraal kunt maken, waardoor ze opnieuw een ervaring worden. Deze insteek was voor ons nodig om een abstract begrip als big data te verbeelden.

Voor de productie van de installatie werkten we met [oneseconds](#). Zij boden de technische en productionele vaardigheden om de ideeën van Margit en Ingrid te realiseren. Naast deze vaardigheden was oneseconds ook thuis in het verhaal over big data en profiling, waarmee we die kennis in de uitwerking konden borgen.

In maart 2016 startte dit team met de ontwikkeling. In de conceptfase hebben verschillende experts uit de hackathons en het journalistieke onderzoek bijgedragen. Eind mei 2016 is de voorstelling in première gegaan in *Bibliotheek Utrecht*, en reisde daarna onder andere naar *Hacking Habitat*, *Betweter Festival* en de *Kunsthof*.

## DE KANS DAT JIJ ER BENT IS ONS BEKEND

**“Hij wacht op jou. Voor hem geen verrassing, dus ongeduldig is hij niet. Verheugd, dat is juist uitgedrukt. Hij is niet zomaar iemand. Hij is iemand die we eigenlijk allemaal wel zouden willen zijn. Zoveel kennis, zoveel vermogen tot waarnemen, zoveel waarheid. Hij is echt en jij gaat hem ontmoeten.”**

In deze voorstelling ga je als bezoeker *Big-data* ontmoeten. Bezoekers registreren zich aan het begin van de voorstelling en stappen in de rol van data-analist. Vervolgens ontmoeten zij Big-data – in de vorm van een fysieke installatie – en zijn twee assistenten in de vorm van twee acteurs. Gezamenlijk halen zij – letterlijk – datapunten van een specifiek persoon uit de big data van de verjaardagskalender. Langzaam ontstaat er zo een persoonsprofiel en ga ze met elkaar in gesprek wat een passend cadeau voor deze persoon zou zijn. Aan het einde van de voorstelling komt het publiek er achter dat ook hun manier van profiling (hun redenering voor wat een goed cadeau zou zijn) aan hun eigen dataprofiel in de kalender is toegevoegd. Ofwel, de data die ze bij de intake hebben afgestaan.

De voorstelling maakt het proces van data-analyse en profiling tastbaar en analoog. Het vormt voor bezoekers een beeld van deze technologische concepten, maar laat de bezoekers tegelijkertijd ervaren dat deze wiskundige werkelijkheid niet alleen rationeel, berekenbaar en onafhankelijk is. Ons eigen wereldbeeld bepaalt voor een groot hoe we data-analyses maken en interpreteren.

De theatrale ervaring had precies het effect dat we voor ogen hadden. De imposante installatie, de simpele verbeelding van het technologische proces en de begeleiding van de acteurs zorgt ervoor dat de bezoekers met elkaar in gesprek gaan. Daarnaast werkte de combinatie tussen vorm en spel heel goed. Waar de installatie heel gestileerd en bijna klinisch de vorm van big data neerzet, is het de Brabantse gezelligheid van de acteurs die in een Lingo-achtige setting de bezoekers in het gesprek meenemen. Die twee kanten (het technologische en wiskundige tegenover het menselijke en subjectieve) laat precies zien waar de spanning zit.

Ook de keuze om heel stapsgewijs meer en meer data aan een naam te koppelen, werkt heel goed. Dat begint met simpele keuzes (is het een man of een vrouw?), maar wordt gaandeweg complexer. De theatermakers hebben per keer ook profielen uitgekozen die aansloten bij de steden waarin de voorstelling werd gespeeld.

Zie Bijlage 10 voor foto- en videomateriaal.

Publiek tijdens de voorstelling.



## CONCLUSIES EN REFLECTIE

De voorstelling is voor ons een onmisbaar deel van De Nationale Verjaardagskalender geworden. De theatermakers zijn erin geslaagd om de verschillende verhaallijnen uit het project te laten samenkomen. Ze maken big data tastbaar voor een publiek dat niet per se aanslaat op rationele argumenten, door vorm en spel te combineren en de werking van big data invoelbaar te maken.

De ruimtelijke installatie draagt hieraan bij; deze trekt voldoende bekijks en roept vragen op. Het nadeel is echter dat de benodigde ruimte niet op elke plek beschikbaar is. Vooral kleinere bibliotheken hebben moeite om dit object te plaatsen. Om hier op in te spelen, hebben we een werkvorm ontwikkeld die de ervaring van de installatie vertaald naar een pocketeditie. Aan de hand van kaarten lopen deelnemers samen met een begeleider hetzelfde profilingsproces door. Door meer spelelementen toe te voegen, kan het spel ook strategisch worden gespeeld. Net als de installatie brengt de pocketeditie een gesprek op gang. Voor bibliotheken blijkt dit een heel goede manier om debat over technologie aan te bieden, zonder gelijk technisch heel vaardig te moeten zijn.

Foto bij artikel AD.



De installatie bij Hacking Habitat.



# Deel IV

---

## Lezingen, voorstellingen en workshops

**Naast de publicaties in de media hadden we als doel om in 5 tot 12 bibliotheken een evenement aan te bieden. Daarnaast wilden we 5 presentaties geven.**

In ons projectplan gingen we uit van een totaalpakket (voorstelling, lezing en workshops). In de praktijk bleek dat de meeste bibliotheken liever geen programma's afnamen, maar enkel losse onderdelen. We hebben daardoor uiteindelijk ook meer losse activiteiten gedaan, verspreid over meerdere locaties en gelegenheden.

Uiteindelijk hebben we de voorstelling in 6 bibliotheken gespeeld en 5 keer op een andere locatie. Op elke speeldag waren er meerdere rondes. Daarnaast hebben we 11 lezingen gegeven over dit project. Hoogtepunten waren onze lezing op SXSW (Texas), de voorstelling tijdens de Mega-Coderdojo in Utrecht en de speelrondes op Digital Days tijdens de Museumnacht. Deze lezingen, voorstellingen en workshops zijn nog steeds deel van ons aanbod. Hoogstwaarschijnlijk worden er in de toekomst dus nog meer edities gedraaid.

### HACKATHONS

28 FEB 2015	Nationale Verjaardagskalender Hackathon #1
23 MEI 2015	Nationale Verjaardagskalender Hackathon #2
31 OKT 2015	Nationale Verjaardagskalender Hackathon #3
28 NOV 2015	Nationale Verjaardagskalender Hackathon #4
30 JAN 2016	Nationale Verjaardagskalender Hackathon #5

### VOORSTELLINGEN

25 MEI 2016	De kans dat jij er bent is ons bekend - try outs @ Bibliotheek Utrecht
27 MEI 2016	De kans dat jij er bent is ons bekend - try outs @ Bibliotheek Utrecht
29 MEI 2016	De kans dat jij er bent is ons bekend @ Bibliotheek Utrecht
4 JUN 2016	De kans dat jij er bent is ons bekend @ Hacking Habitat
26 JUN 2016	De kans dat jij er bent is ons bekend @ De Tweede Verdieping
3 SEP 2016	De kans dat jij er bent is ons bekend @ Vleuterweide
30 SEP 2016	De kans dat jij er bent is ons bekend @ Betweterfestival
15 OKT 2016	De kans dat jij er bent is ons bekend @ Coderdojo XL
3 MRT 2017	De kans dat jij er bent is ons bekend @ DigitalDays
4 MRT 2017	De kans dat jij er bent is ons bekend @ DigitalDays
4 MRT 2017	De kans dat jij er bent is ons bekend @ Museumnacht

## LEZINGEN

10 NOV 2015	G4 CIO congres, Utrecht
11 MRT 2016	DIY-NSA op South-by-Southwest Texas
27 JAN 2016	TEDx Utrecht
6 JULI 2016	Border Sessions, Den Haag
10 NOV 2016	KNVI Congres, Nieuwegein
9 DEC 2016	Human Rights XL (Art. 1 Midden Nederland), Het Huis Utrecht
30 JAN 2017	SURF Privacy Seminar, Utrecht
14 MRT 2017	Nationale Bibliotheekcongres, Assen
18 APR 2017	Next Learning Congres, Den Bosch
6 JUN 2017	Bildungacademie, Amsterdam
14 JUN 2017	SURF Academy, Amsterdam

In ons aanvankelijke plan gingen we uit van combinaties bestaande uit een lezing, voorstelling en workshops. In praktijk bleek de workshop minder goed te werken, omdat veel bibliotheken het programma wel interessant vonden, maar de workshop technisch te hoog gegrepen. Met name omdat het eigen personeel van bibliotheken hierin ook niet kon ondersteunen. Ook werden we door meer gangbare podia (festivals en dergelijke) gevraagd om de voorstelling te spelen, waar minder ruimte was voor een workshopprogramma.

Het lagere aantal workshops hebben we ondervangen door de pocketeditie van de voorstellingen te ontwikkelen. Deze heeft een workshopvorm, maar richt zich meer op debat dan technische ontwikkeling. Daardoor is deze makkelijker te begeleiden door extern personeel (bijvoorbeeld medewerkers van bibliotheken). Trainen van dit personeel om beter dit soort debatten te begeleiden is deel van ons aanbod.

Presentatie op  
TEDxUtrecht





# Conclusie

---

We kijken terug op een geslaagd project. We zijn trots op de nieuwe verhalen, aandacht en impact die we hebben gerealiseerd. Van concrete resultaten bij bijvoorbeeld de NBA tot de nieuwe inzichten bij reguliere bibliotheekbezoekers. In elk aspect van dit langlopende project hebben we big data en profiling concreet en daardoor bespreekbaar gemaakt. We merken tevens hoe De Nationale Verjaardagskalender vaak op andere podia als voorbeeld wordt genoemd als het over profiling gaat.

Hoewel het project uiteindelijk langer duurde dan we aanvankelijk hadden gepland, heeft deze tijd zich wel terugbetaald in de uiteindelijke kwaliteit en de gerealiseerde resultaten. We willen u dan ook bedanken voor niet alleen de financiële steun, maar ook het geduld, waaronder we dit project konden realiseren.

**SETUP**