

Datum: 15 februari 2024



Google in de publieke sector

We onderzochten publieke sector websites op externe trackers.

Door Joran van Apeldoorn

Je komt niet meer om het internet heen. Een afspraak maken bij de dokter, een wijziging van je woonsituatie doorgeven aan de gemeente, het schoolrapport van je kinderen bekijken: het gebeurt allemaal online. Maar wie kijkt er - naast jijzelf - mee op al deze websites? We onderzochten 32.843 websites van de overheid, de zorg en het onderwijs. Hieruit blijkt dat Google op ontzettend veel websites van de Nederlandse publieke sector persoonsgegevens in handen krijgt.

Bij 84% van de onderzochte websites wordt Google op de hoogte gesteld van je bezoek. Dit percentage is het hoogst in het onderwijs, met 90%, direct gevolgd door de zorg met 83%. Bij de overheid is het iets beter gesteld met een - nog steeds schandalig hoog - percentage van 48%.

Bekeek je op de website van je huisarts een pagina over een soa test? Dan weet Google dat waarschijnlijk. Vraag je bijstand aan bij je gemeente? Of gebruikt je kind online lesmateriaal voor kinderen met dyslexie? Google krijgt het allemaal mee. De tools van Google - denk aan Google Maps, Google Forms en Google Analytics - zijn gratis en gemakkelijk te gebruiken voor de ontwikkelaars van websites, maar maken een heftige inbreuk op jouw privacy. Zeker in de publieke sector zou men het goede voorbeeld moeten geven en privacyvriendelijke alternatieven gebruiken.

Gepubliceerd op 15 februari 2024

WAT IS 'TRACKING'?

Tech bedrijven zoals Google verzamelen veel data over je, en dat doen ze niet voor niks. Google verdient haar geld door een gedetailleerd profiel over je op te bouwen, en dit te gebruiken om je advertenties te laten zien. Hoe meer ze over je weten, hoe preciezer ze advertenties op je kunnen afstemmen, en hoe makkelijker ze je geld kunnen laten uitgeven of je mening kunnen beïnvloeden.

Google volgt je dan ook niet alleen op hun eigen websites, maar volgt het liefst je algehele digitale gedrag. Daarom biedt Google diensten aan die ontwikkelaars van websites kunnen gebruiken bij het bouwen van een website. Zo worden die websites afhankelijk van Google, en zal jouw surfgedrag op zulke sites door Google gevolgd kunnen worden.

Omdat Google je op veel verschillende manieren volgt is het belangrijk wat we bedoelen als we het over 'tracking' hebben. We kijken in dit onderzoek naar twee definities van tracking:

1. *Contact met Googlediensten.* Dit is de breedst mogelijke definitie van een 'tracker'. In dit geval wordt bij een bezoek aan een website ook contact gelegd met diensten van Google. Mogelijk gebeurt dit om data te delen, maar mogelijk ook voor andere doeleinden. Dit kan al zo simpel zijn als het ophalen van een lettertype van Google's servers (Google Fonts). Toch krijgt Google bij zulk 'simpel' contact al inzicht in welke websites je bezoekt.
2. *Expliciete Google tracking.* Dit zijn stukjes programmacode van Google op een website die expliciet contact leggen met Google om data over jou te delen. Zo stuurt Google Analytics bijvoorbeeld allemaal gegevens over je surfgedrag naar Google. Denk aan: hoe lang je een pagina bezoekt, welke pagina's je in welke volgorde bezoekt, op wat voor apparaat je dat doet, etc. Ook zal bijvoorbeeld een YouTube filmpje op een pagina aan Google vertellen wat het onderwerp van het filmpje is en hoelang je ernaar kijkt.

Aangezien expliciete trackers contact leggen met Googlediensten vallen trackers uit de tweede categorie ook in de eerste categorie, maar andersom is dit dus niet zo.

ONDERZOEKSOPZET

Het onderzoek bestond grofweg uit twee delen:

1. Het creëren van een overzicht van websites uit het onderwijs, de zorg, en van de overheid; oftewel: het verzamelen van de te onderzoeken domeinnamen
2. Het identificeren van het netwerkverkeer en de tracking bij het bezoeken van die websites.

1. HET VERZAMELEN VAN DE DOMEINNAMEN

De te onderzoeken domeinnamen binnen de drie sectoren zijn op verschillende manieren in kaart gebracht:

- **Overheid.** De overheid houdt geen volledig en up-to-date overzicht bij van alle websites waar zij verantwoordelijk voor is. Wel is er een [register van overheidsorganisaties](#) waarin sommige organisaties hun website vermelden. Voor de overheidsorganisaties uit dit register die geen website vermelden hebben we het KvK-nummer bij de Kamer van Koophandel opgezocht om vaak alsnog hun website te vinden.
- **Onderwijs.** Dienst Uitvoering Onderwijs (DUO) publiceert op hun website [Open onderwijsdata](#): openbare data en publicaties over bekostigd onderwijs in Nederland. De domeinnamen van onderwijsinstellingen zijn hierin opgenomen.
- **Zorg.** Voor zover wij hebben kunnen achterhalen, biedt de overheid geen totaaloverzicht van zorginstellingen met hun contactgegevens. De Patiëntenfederatie Nederland houdt wel [Zorgkaart Nederland](#) bij, een overzicht van alle zorgaanbieders. Op de pagina van iedere zorgaanbieder zijn ook de adres en website gegevens te vinden die we door middel van web-scraping hebben kunnen downloaden.

In het onderzoek ligt de focus op de homepages van organisaties. Campagne websites en andere extra websites zijn buiten beschouwing gelaten. Het verkrijgen van een goed overzicht van deze andere domeinen is namelijk lastig. Daarbij is het lastig om vast te stellen of een website nog actief in gebruik is als dit geen homepage is.

2. HET IDENTIFICEREN VAN TRACKING EN MONITOREN VAN NETWERKVERKEER

Voor het identificeren van trackers bestaan er al meerdere programma's. Wij maakten gebruik van de code die ten grondslag ligt aan [Blacklight](#) van The Markup. Deze tool kun je zelf ook gebruiken om websites te scannen op expliciete trackers.

Blacklight is gebouwd om expliciete trackers te detecteren. De tool bezoekt een paar willekeurige pagina's op een te testen website om normaal gebruik te simuleren. Ondertussen wordt het netwerkverkeer dat bij dat bezoek plaatsvind vergeleken met een lange lijst van bekende expliciete trackers, om zo te detecteren of die trackers aanwezig zijn op de site.

Blacklight houdt ook al het andere netwerkverkeer bij dat plaatsvindt wanneer je een pagina bezoekt. We hebben dit verkeer vergeleken met bekende domeinnamen van Google om te zien welke sites bij jouw bezoek automatisch een connectie met Google leggen, zelfs als dit geen expliciete tracker is. Blacklight accepteert geen cookie banners of privacy statements, dus als er een connectie wordt gelegd met Google, dan is dit zonder dat er toestemming is gegeven via een cookie banner.

RESULTATEN

GEVONDEN DOMEINEN

Uit de drie bronnen komen 71.723 organisaties voort. Niet al deze organisaties hebben een webadres opgegeven of een werkende website. Daarnaast delen veel organisaties een websites, denk aan apotheken die deel uitmaken van dezelfde keten. Uiteindelijk hebben we hierdoor 32.847 unieke domeinnamen kunnen onderzoeken.

Het overheidsorganisaties register bevat ook veel organisaties die vanuit burgers waarschijnlijk minder als overheid gezien worden. Zo staan er ook honderden 'organisaties met overheidsbemoediening' (vooral stichtingen met overheidssubsidies waaraan voorwaarde kleven) en 'regionale samenwerkingsverbanden' in het register. We splitsen daarom de categorie 'overheid' in 'primair' - de rijksoverheid, gemeenten, provincies, de rechtspraak, enz. - en 'secundair' voor de rest. De categorie 'overheid primair' komt zo beter overeen met wat burgers zien als 'de overheid'.

Hieronder een overzicht van het aantal organisaties en websites in elke sector:

	onderwijs	overheid primair	overheid secundair	zorg
Aantal organisaties	8940	692	1930	60161
Aantal met webadres	7666	637	1187	57249
Aantal websites online	7071	592	1112	53881
Aantal unieke domeinen	6189	524	1094	25127

GOOGLE ZIT OVERAL

Bij de overheid hebben bijna de helft, 48%, van de websites Googlediensten in gebruik. Er is een groot verschil zit tussen de overheden die we als 'primair' hebben gelabeld, en de andere overheidsorganisaties: de secundaire categorie doet het nog slechter met 71%. De situatie is nog veel slechter in de zorg, waar meer dan 4 op de 5 websites Google gebruikt. In het onderwijs is dit zelfs 9 op de 10 websites.

Als we de strengere definitie van 'tracker' aanhouden, en alleen kijken naar expliciete trackers, dan maakt nog steeds meer dan de helft van de websites gebruik van Google diensten die tracken. Bij de overheid is dit met 35% iets meer dan 1 op de 3.

	onderwijs	overheid primair	overheid secundair	zorg
% websites met Googlediensten	90%	48%	71%	83%
% websites met Google-trackers	53%	35%	59%	52%

Het is soms lastig de directe impact van Google's surveillance te overzien. We zijn geneigd het als een persoonlijk probleem te zien: jouw data wordt verzameld door Google. Hoewel dit inderdaad een enorme inbreuk op jouw privacy is, wordt het grotere probleem pas duidelijk als je uitzoomt: Google zit overal. In de wachtkamers van de dokter, klaslokalen en gemeentehuizen. Voor Google is jouw individuele data niet veel waard, maar de gecombineerde data van alle Nederlanders is een goudmijn.

Google heeft zich in de publieke sector weten in te graven, om nog maar niet te spreken van het commerciële internet. Zo verzamelt ze zeer persoonlijke gegevens over alle Nederlanders, en dit doet ze niet voor niks. In het informatie tijdperk is kennis macht, en Google weet het meeste van iedereen.

HOE GROTER DE SPELER IN DE ZORG, DES TE MEER GOOGLE-TRACKERS

Zorgaanbieders die deel uitmaken van een keten, hebben Google vaker op hun website. Dit verschil wordt duidelijk als we naar de organisatie zelf kijken in plaats van naar de websites. Honderden verschillende apotheken binnen één grote keten - en daardoor met één website - worden dan namelijk afzonderlijk van elkaar gezien.

Het is zorgelijk dat grote spelers in de zorg slechter scoren op het gebied van privacy dan hun kleinere concurrenten. Aanbieders zoals BENU-apotheken en Service Apotheek hebben honderden apotheken en bereiken zo duizenden mensen. Deze partijen hebben juist de middelen om het goede voorbeeld te geven.

	onderwijs	overheid primair	overheid secundair	zorg
% organisaties met Googlediensten	90%	44%	70%	88%
% organisaties met Google-trackers	56%	32%	58%	65%

ANDERE TRACKERS

Naast Google trackers, zijn er nog tal van andere trackers die schadelijk zijn voor je privacy. Ook deze trackers onderzochten we. Geen enkele kwam ook maar in de buurt van Google's wijde verspreiding. Dat wil echter niet zeggen dat deze trackers minder kwaad doen, ze zijn in sommige gevallen namelijk wel een stuk direc-

ter schadelijk voor je privacy. Het voorkomen van de Meta pixel (vroeger ook wel bekend als: Facebook pixel) is bijvoorbeeld opvallend in de publieke sector. Deze tracker heeft als enige doel om een beter profiel van je op te bouwen voor Facebook- en Instagram advertenties. Iets meer dan 5% van de websites in de zorg en 7% van de websites in het onderwijs gebruikt deze tracker. Op websites van overheidsorganisaties vonden we gelukkig slechts een handvol Meta pixels¹. Waarom Veiligheidsregio Zeeland en de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) het een goed idee vonden om bij elk site bezoek Meta te betrekken blijft echter wel een mysterie.

In de zorg valt ook op dat het bedrijf **Pharmeon** 8% van de websites beheert, inclusief een eigen Google Analytics alternatief, en zo ook gegevens verzamelt van miljoenen Nederlanders. Verder zijn er trackers die 'slechts' op enkele honderden websites actief zijn, maar wel erg veel data verzamelen. Neem bijvoorbeeld Hotjar, aanwezig op 527 (1.6%) van de onderzochte websites. Deze tracker kan precies bijhouden hoe je je muis beweegt op de site en waar je klikt.

Gelukkig is er ook nog positief nieuws. De Rijksoverheid heeft een eigen alternatief voor Google Analytics, dat gratis te gebruiken is voor de websites van de Rijksoverheid en beter omgaat met de privacy van bezoekers. 114 van de gescande websites maakt hier op dit moment al gebruik van. Jammer genoeg is deze service nog niet beschikbaar voor lokale overheden.

¹We vonden er 3 binnen de categorie 'overheid primair' en 77 binnen de categorie 'overheid secundair'. Aangezien deze tweede categorie zowel duidelijke overheden zoals veiligheidsregio Zeeland, als ook bijvoorbeeld Stichting DE KOM (een stadstheater) omvat, houden we het op 'een handvol' overheidswebsites.

CONCLUSIE

Kortom, de bevindingen van dit onderzoek onderstrepen de dringende noodzaak voor de publieke sector om het goede voorbeeld te geven en te kiezen voor privacyvriendelijke alternatieven. Nu we steeds meer in een online wereld leven, is het noodzakelijk dat er verantwoorde maatregelen worden genomen om de privacy van burgers te beschermen. Juist als het gaat om het gebruik van essentiële (publieke) diensten. Dit onderzoek roept dan ook op tot een herevaluatie van de tools en diensten die worden gebruikt door websites van de overheid, de gezondheidszorg en het onderwijs, waarbij wordt gepleit voor oplossingen die privacy vooropstelt.

Specifiek in de publieke sector mogen we ook inzet van de overheid verwachten. [Basisbeveiliging.nl](#) is recent in samenwerking met ministerie voor Binnenlandse Zaken en Koninkrijksrelaties begonnen met het scannen op tracking cookies. De staatssecretaris voor Digitale zaken heeft reeds een wens uitgesproken om tracking van overheidswebsites te halen². Als het aan Bits of Freedom ligt, onderneemt de regering ook actie in andere sectoren waar steeds meer zaken digitaal worden geregeld.

CONTACT

Vragen over het onderzoek kunnen gericht worden aan Joran van Apeldoorn. Hij is te bereiken op [06 2087 7802](tel:0620877802) of joran@bitsoffreedom.nl

DOWNLOAD DE DATA

Je vindt [hier de volledige dataset](#) als CSV-bestand. De data is eind 2023 verzameld. De rijen corresponderen met organisaties en de kolommen zijn als volgt:

Kolomnaam	Beschrijving
sector	De sector waarin de organisatie actief is
categorie	Een subcategorie, bijvoorbeeld “basisschool” of “gemeente”
naam	De organisatie naam
telefoon	Waar beschikbaar het telefoonnummer
email	Waar beschikbaar het e-mailadres
denominatie	Voor onderwijsinstellingen de denominatie van de instelling
kvk	Waar beschikbaar het KVK-nummer
source	Bron waaruit de gegevens zijn opgehaald
from_kvkk	Voor overheidsorganisaties, of het webadres uit de KVK komt (in tegenstelling tot het overheidsregister)

²[Kamerbrief cookies en online tracking, 5 september 2023](#)

Kolomnaam	Beschrijving
url	Het opgegeven webadres
domain	De domeinnaam afgeleid van dat webadres
reachable	Of dit domein tijdens de test bereikbaar was, overige velden zijn alleen ingevuld als dit zo was
redirect_url	Het webadres waar na mogelijke doorverwijzingen op uitgekomen werd
redirect_domain	De domeinnaam waarnaar wordt doorverwezen
google	Of er contact wordt gelegd met Googlediensten
google_example	Een van de mogelijk meerdere domeinnamen van Google waarmee contact is gelegd
google_tracks	Of er Google-trackers zijn gedetecteerd
google_track_example	Een domein behorende aan die trackers
adres	Het straatadres
lat	Breedtegraad van het opgegeven adres
lon	Hoogtegraad van het opgegeven adres

**Bits of Freedom komt op voor jouw
vrijheid en privacy op internet.**

Deze grondrechten zijn onmisbaar voor je ontwikkeling, voor technologische innovatie en voor de rechtsstaat. Maar die vrijheid is niet vanzelfsprekend. Je gegevens worden opgeslagen en geanalyseerd. Je internetverkeer wordt afgeknepen en geblokkeerd.

Bits of Freedom zorgt ervoor dat
jouw internet jouw zaak blijft.

Bits of Freedom
bitsoffreedom.nl
🐦 [@bitsoffreedom](https://twitter.com/bitsoffreedom)
Prinseneiland 97HS
1013 LN Amsterdam

Contactpersoon:
Martijn de Heer
+31 6 1557 7255
martijn@bitsoffreedom.nl

ED16 7F84 5739 203A 9F29
976C 6289 AAB1 70D3 2D7D
(bitsoffreedom.nl/openpgp)

**BITS OF
FREEDOM**
Voor jouw internetvrijheid