



Overzicht landelijke databronnen

Relevant in het kader van visitor management

Februari 2020

Aanleiding

Veel plekken en regio's zijn op zoek naar tooling die ze helpen bij het managen van bezoekersstromen. Daarbij is een goed en liefst realtime (of zelfs voorspellend) inzicht in bezoekersomvang en –stromen van cruciaal belang. Veel van de nu beschikbare tools verzamelen hun data door mensen ter plekke de mate van drukte te laten invoeren (user generated content) of door sensoren te plaatsen die drukte meten. In een enkel geval worden deze data aangevuld met data van parkeergarages. Er wordt nog maar beperkt van landelijk beschikbare (big) datasets gebruik gemaakt. Zoals data van Google, mobiele telefoons of van het wegenverkeersnet.

Hoewel deze bronnen in potentie veel nuttige inzichten kunnen bieden, blijkt het in praktijk lastig hier toegang toe te krijgen. In dit document zetten we op een rij in hoeverre het mogelijk is dergelijke datasets te gebruiken en wat daarvoor de randvoorwaarden zijn. Ook gaan we in op de specifieke inzichten die je kunt verkrijgen per dataset, relevant in het kader van visitor management.

Benieuwd op welke manier je naast landelijke databronnen inzicht kan krijgen in dagbezoekers?

Kijk dan ook eens in de [tool 'Vergelijking meetmethoden dagbezoek'](#), het resultaat van een Data en Development Lab project. In dit project zijn in totaal 20 onderzoekstools en –methoden beoordeeld op hun effectiviteit t.a.v. het meten van dagbezoekers. Welke informatie levert elke methode precies op, welke gebruik je in welke situatie en wat zijn voor- en nadelen van elke methode? Dat kan je allemaal vinden in de online publicatie. Van enquêtes tot zelfregistratie en van tellingen en het plaatsen van trackers tot het vaststellen van ratio's, alles komt voorbij. Veel van de onderzoeksmethoden hebben een lokale of regionale scope.

Aanvullend op de informatie uit deze tool richt dit document zich enkel op de (open) databronnen met een landelijke scope.

Overzicht relevante databronnen

Voor visitor management zijn verschillende databronnen in potentie relevant:

1. Mobiele telefoniedata
2. GPS data uit apps (incl. case Smart City Dashboard Eindhoven)
3. Google popular times
4. Google mobility reports
5. Google trend data
6. Apple data over mobiliteit
7. NDW data
8. Floating Car Data
9. Satellietdata

Analysecriteria

Per bron zetten we de volgende aspecten op een rij

1. **Toelichting** > Algemene uitleg van de bron.
2. **Bruikbaarheid** > Waar kan de data voor gebruikt worden?
3. **Toegankelijkheid** > Zijn de data vrij beschikbaar (via een open data platform), of moeten de data aangekocht worden?
4. **Status** > Wat is de laatste stand van zaken ten aanzien van (de ontwikkeling van) de databron?

Daarnaast geven we in het kader snel zicht op de belangrijkste voor- en nadelen van de bron, alsmede waar de data input voor kunnen zijn.

Analyse per databron



1. Mobiele telefoniedata

Toelichting

Tot een jaar geleden was het mogelijk om analyses uit te voeren op data van mobiele telefoonproviders. Zo werd data van Vodafone via Mezuro vermarkt en voerde het CBS analyses uit op T-mobile data. Telefoonproviders kunnen via zendmasten en miljoenen contactmomenten per dag van alle klanten hun locatie bepalen en deze data geanonimiseerd uitgeven voor analyses.

Bruikbaarheid

De data op basis van mobiele telefoons biedt met name inzicht in de aanwezigheid van mensen/ telefoons in een bepaald gebied op een bepaald moment in de tijd. Het zijn momentopnames, waar mooie heatmaps van gemaakt kunnen worden. Daarnaast biedt het ook een indicatie van het aantal Nederlanders op een bepaald moment op een bepaalde plek. Er is voldoende kennis om de data van één provider op te hogen naar totalen. Dat is bij bijvoorbeeld GPS-data lastiger, evenals bij internationale bezoekers. Door de data van meerdere momenten in de tijd te analyseren ontstaan patronen die kunnen helpen bij het beter voorspellen van drukte. Waar het weinig over zegt, is de kenmerken van de telefoongebruikers. Data analisten proberen wel op basis van definities bv een dagbezoeker van een inwoner te onderscheiden. Meer informatie over de bruikbaarheid van mobiele telefoondata voor inzicht in dagbezoek vind je ook [hier](#).

Toegankelijkheid

De Autoriteit Persoonsgegevens houdt het toestaan van gebruik van locatiegegevens van telecombedrijven momenteel tegen. Voor de coronacrisis was het al onderwerp van gesprek tussen de overheid, de autoriteit en het CBS. Tijdens de crisis is gesproken over een noodwet waardoor de locatiegegevens ingezet zouden kunnen worden om het virus te remmen. Die noodwet is er niet gekomen en de data is zelfs voor dat doel niet beschikbaar gekomen. De stap om deze data toegankelijk te krijgen voor analyses binnen ons domein lijkt daarom nog er weg. Het advies van de AP (mei 2020) is [hier](#) te vinden.

Status

De uitbereiding van de – in de volksmond zogeheten – CBS-wet waardoor het CBS data van alle providers zou kunnen ontvangen, vindt veel weerstand. De Autoriteit Persoonsgegevens stelt vragen bij de mate van anonimiteit van de data. De gesprekken tussen de AP, het ministerie van EZK en het CBS lopen nog. Ook een eventuele noodwet in het kader van corona heeft dus nog niet tot beschikbaarheid van mobiele telefoondata geleid. Dit terwijl de potentie van deze databron groot is in relatie tot bezoekersmanagement-doeleinden.

Voordelen:

- Grote groep gebruikers, veel data
- Gegevens over tijd & ruimte, dus herkomst, bewegingen, overnachtingen zijn grotendeels te achterhalen
- Goed zicht op Nederlandse gebruikers

Nadelen:

- Niet vrij beschikbaar, momenteel zelfs helemaal niet beschikbaar
- Locatiebepaling gebeurt via telefoonmasten. Dit is minder nauwkeurig dan locatiebepaling op basis van bijv. GPS. Gebieden beschikbaar in mobiele telefoondata zijn wat grootschaliger dan straatniveau. Is daarnaast ook mede afhankelijk van de dichtheid van telefoonmasten in de betreffende regio.
- Data zijn niet realtime
- Relatief kostbaar
- Onderscheid toerist/ bewoner is lastig te maken

Input voor:

- Voorspelmodellen van drukte
- Bewegingsgedrag van mensen

2. GPS data uit apps

Toelichting

Door het gebruik van een app met locatiegegevens aan, worden de exacte locaties en bewegingen van personen vastgelegd. Voor deze data moeten gebruikers toestemming geven. Er zijn landelijke aanbieders van gps-data van Nederlandse gebruikers verkregen via hun apps. Er zijn apps die hun eigen data vermarkten (zoals Strava), maar er zijn ook aanbieders die de gps data van meerdere apps samenbrengen en vermarkten, zoals Resono (data wordt gebruikt door o.a. Zicht op Data) en Perfect Place (die naar eigen zeggen de data van honderden apps omzet in geaggregeerde en anonieme realtime informatie). Tot slot is er het Nationaal Verplaatsingspanel (Dat.mobility/ Kantar) dat door de koppeling met een online panel meer mogelijkheden biedt voor verdiepende analyses.

Bruikbaarheid

Uitgaande van de partijen die de data delen van meerdere apps, dan betreft het data met een goede dekking over ons land en is de databron dus zowel voor stedelijke als landelijke omgevingen te gebruiken. Het geeft de aanwezigheid van telefoons meer realtime weer dan bv de mobiele telefoondata. Het biedt in potentie dus concrete stuurinformatie. De grote beperking is dat het (vooralsnog) enkel op Nederlandse apps gericht is, internationale bezoekers meet je dus niet. De kenmerken mis je ook, zoals motief van het bezoek. Wel kan je de bezoekduur en de frequentie in kaart brengen. Afhankelijk van de aanbieder zijn er wel mogelijkheden om persoonskenmerken te koppelen aan de data en/of aanvullend te verzamelen. Bij het verplaatsingspanel kan er zelfs aanvullend onderzoek onder de panelleden worden gedaan. Meer informatie over het gebruik van GPS-trackers om dagbezoek te meten is te vinden in de digitale publicatie '[Vergelijking meetmethoden dagbezoek](#)' (zie optie 4 'Indirecte observatie').

Toegankelijkheid

De data kunnen aangekocht worden bij de apps zelf (Strava) of via partijen als Resono, Zicht op Data en Perfect Place. De data zijn niet vrij beschikbaar. De kosten van de dataoplossingen variëren.

Status

De eerste voorbeelden van regio's en steden die met de data werken zijn bekend. Zo heeft Eindhoven de data opgenomen in haar Smart City Dashboard en heeft de provincie Noord-Brabant analyses laten doen door Zicht op Data. Op de Veluwe is gestart met een pilot met Resono-data en ook Rotterdam gebruikt Resonodata in hun druktemonitor (www.spotrotterdam.nl). De mogelijkheden van het Nationaal Verplaatsingspanel worden o.a. bekeken bij de opzet van een nieuw vrijetijdsonderzoek.

Voordelen:

- Op relatief klein niveau (gebiedsdefinities vanaf circa 8 m², ook mogelijk voor bijv. wijken, wegen, fietspaden, gebouwen etc.)
- Gegevens over mate van drukte op een bepaald moment en over bewegingen (aantallen, bezoekfrequentie en herkomst (postcode 4 NL)
- Achtergrondgegevens/ bezoekerskenmerken kunnen, afhankelijk van de gekozen oplossing, worden toegevoegd.

Nadelen:

- Afhankelijk van de dataset kan je met een specifieke doelgroep te maken hebben
- Data zijn niet altijd realtime

Input voor:

- Voorspelmodellen van drukte
- Bewegingsgedrag van mensen

CASE: Resono data in Smart City Dashboard Eindhoven

Vanaf november 2020 maakt dit dashboard gebruik van locatie data op basis van gps. Het betreft Resono data van ruim 1,5 miljoen personen (9% van de NL bevolking). Omdat ze eerder van gsm-data gebruik maakten, konden ze op basis van hun ervaringen een vergelijking tussen beide typen data maken, die opgenomen is in dit [document](#).

3. Google popular times

Toelichting

Google is een interessante bron van informatie. In het kader van bezoekersmanagement is o.a. de druktemeter die op de site staat relevant: Google popular times. Google kan populaire tijden, live bezoekinformatie, wachttijden en de typische bezoekduur weergeven van locaties op Google Maps en in Zoeken. Google maakt hiervoor gebruik van cumulatieve en geanonimiseerde gegevens van gebruikers die zich hebben aangemeld voor locatiegeschiedenis.

Bruikbaarheid

Deze google data zijn beschikbaar op niveau van een point of interest (POI), zoals een restaurant of attractie. Het geeft je dus real time informatie over de drukte op een bepaalde locatie. Het is echter niet beschikbaar op niveau van een stad of regio. Daarentegen is bijvoorbeeld het Nationale Park De Hoge Veluwe wel opgenomen als POI en biedt het dus ook oplossingen voor afgebakende natuurgebieden.

Toegankelijkheid

Deze google data is momenteel niet beschikbaar via een API. Wanneer je deze data zou willen gebruiken, zou je met *web scraping* technieken aan de slag moeten gaan. Dit is echter tegen de richtlijnen van Google, dus lijkt geen reële optie. Dat maakt de data per locatie nog steeds interessant, maar is het lastig te gebruiken in druktemonitors of data analyses.

Status

Door verschillende partijen is de afgelopen periode contact gezocht met Google met de vraag of ze data beschikbaar kunnen stellen ten behoeve van analyses over drukte. Ook NBTC heeft contact gezocht, waaruit volgt dat ze enkel de mobility reports (volgende slide) kunnen delen en ze geen uitzonderingen kunnen maken voor één specifieke sector.

Voordelen:

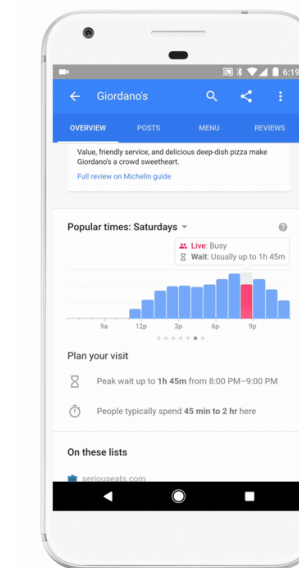
- Data op basis van een grote groep gebruikers, hoge betrouwbaarheid
- Realtime data, inzicht op uurniveau

Nadelen:

- Data niet via API beschikbaar

Input voor:

- Voorspelmodellen van drukte
- Druktemonitoring, inzicht in realtime drukte



4. Google mobility reports

Toelichting

Google geeft de data vanuit Google Maps (deels) vrij voor Corona bestrijdingsdoeleinden: <https://www.google.com/covid19/mobility/>. In de rapporten is te zien op provincieniveau in hoeverre bezoek aan parken, supermarkten, retail & recreatie, openbaar vervoer en werk groeit of daalt.

Bruikbaarheid

De mobility reports zijn op een te globaal niveau om inzichten op het specifieke gebied van bezoekersmanagement te genereren. Het geeft enkel zicht op wat algemene trends voor Nederland en naar provincie op een aantal specifieke clusters.

Toegankelijkheid

De Community Mobility Reports zijn hier te vinden: <https://www.google.com/covid19/mobility>. De data worden regelmatig geupdate en is niet alleen in PDF maar ook als open data (csv formaat) beschikbaar. Deze informatie en data zijn op geaggregeerd landelijk en provincie niveau.

Status

De mobility reports lijken een tijdelijk karakter te hebben, in het kader van corona. De vraag is hoe lang ze beschikbaar blijven. Het betreft dus geen zekere bron van data.

Voordelen:

- Data op basis van een grote groep gebruikers, hoge betrouwbaarheid
- Geeft inzicht in bewegingen van Nederlanders

Nadelen:

- Niet real time
- Niet lokaal genoeg
- Niet duidelijk in hoeverre dit een continue datastroom oplevert (waarschijnlijk tijdelijk)

Input voor:

- Voorspelmodellen van drukte, maar op globaal niveau
- Monitoren van effecten van maatregelen (bijv. effect van coronarestricties) op een hoger schaalniveau (provincies).

5. Google trend data

Toelichting

Google trend data geeft trends weer in de gebruikte zoekwoorden. Wanneer er meer gezocht wordt op bepaalde locaties zijn die populairder dan andere locaties. Door ontwikkelingen in het zoekgedrag naar plekken in de tijd te analyseren, en deze naast daadwerkelijk bezoek aan die plekken te leggen, zou je kunnen afleiden in hoeverre er een correlatie tussen beide is. En dus in hoeverre de zoekdata voorspellend kan zijn voor toekomstig bezoek.

Bruikbaarheid

De bruikbaarheid van deze databron moet verder onderzocht worden. Dat doen we o.a. binnen het DDL project 'Voorspellen van toerismedruk'. We onderzoeken in hoeverre 'Google Trends' (zoekopdrachten) houvast kan bieden bij het voorspellen van het aantal internationale aankomsten in een toeristische regio in Nederland.

Toegankelijkheid

De google trends data is te achterhalen via: <https://trends.google.nl/trends/?geo=NL>. Ook is het mogelijk de data via een dataleverancier als Trendata aan te kopen, waarbij er meer diepgaande analyses mogelijk zijn dan op basis van de openbare google data.

Status

Het DDL project loopt momenteel, de resultaten ervan worden in het eerste kwartaal van 2021 verwacht.

Voordelen:

- Data op basis van een grote groep gebruikers, hoge betrouwbaarheid
- Zoekdata zijn vrij beschikbaar via een online platform

Nadelen:

- Onduidelijk wat de voorspellende waarde van de data is
- Het levert geen directe informatie over drukte, het levert 'slechts' een mogelijke voorspellende indicator
- De trenddata van zoektermen is relatief: als je als bestemming stijgt in de lijst met zoekwoorden, komt dat dan door meer zoekverkeer naar jouw bestemming of hebben andere bestemmingen minder zoekverkeer gezien? Dit is belangrijk om je te realiseren als je met de data gaat werken.

Input voor:

- Voorspelmodellen van drukte

6. Apple data over mobiliteit

Toelichting

Ook Apple geeft soortgelijke inzichten als Google vrij vanwege Corona, de zogenaamde COVID-19-mobiliteitstrends. Het geeft inzicht in bewegingen te voet, per OV of met de auto. De rapporten worden dagelijks gepubliceerd en geven aanvragen voor een routebeschrijving in Apple Kaarten weer.

Bruikbaarheid

De bruikbaarheid van deze databron is beperkt, er worden data gedeeld over Nederland, de provincies en een beperkt aantal steden in Nederland. Wel geeft het je een algemeen beeld van de bewegingstrends. Zo zie je de invloed van de lockdown er duidelijk in terug, evenals het week vs weekend patroon.

Toegankelijkheid

De Apple trends data zijn te achterhalen via: <https://www.apple.com/covid19/mobility>.

Status

De data zijn momenteel nog beschikbaar. Net als bij de Google mobility reports is onduidelijk hoe lang de data nog vrij gedeeld wordt.

Voordelen:

- Data zijn vrij beschikbaar via een online platform

Nadelen:

- Het betreft enkel Apple gebruikers die een routebeschrijving opvragen via de app, geen standaard GPS tracking
- Voor een beperkt aantal bestemmingen beschikbaar
- Niet duidelijk in hoeverre dit een continue datastroom oplevert (waarschijnlijk tijdelijk)

Input voor:

- Voorspelmodellen van drukte, maar op globaal niveau
- Monitoren van effecten van maatregelen (bijv. effect van coronarestricties)

7. NDW data

Toelichting

NDW beschikt als uitvoeringsorganisatie van Rijkswaterstaat over de wegendata in Nederland. Zoals de intensiteit (aantal auto's op bepaalde plek per uur, teruggerekend naar minuut), snelheidsgegevens, reistijden, wegwerkzaamheden etc.

Bruikbaarheid

De data wordt gebruikt voor verkeersinformatie (ANWB), verkeersbeleid van provincies en door CBS. Maar de data wordt bijv. ook gedeeld worden met marktpartijen als Google. Het kan voor toerisme interessant zijn om te onderzoeken of de wegendata een correlatie heeft met toeristische indicatoren. Om zo te bekijken of je de wegendata in de toekomst als indicator voor toeristische druk kunt gebruiken. Daarvoor zou je analyses kunnen doen op historische data, die makkelijk uit het systeem te halen is.

Toegankelijkheid

De data is beschikbaar via hun open data platform, maar is ook via het CBS beschikbaar. De historische data zou je in excel kunnen downloaden. Voor een account tot bijv. de realtime (1 minuut oud) data komen (semi-)overheidsinstanties ook in aanmerking. Wel moet de NDW daar dan iets voor terugkrijgen, bijv. de inzichten uit de uitgevoerde analyses. Let op dat er voor analyses op realtime data een omvangrijk analyseplatform en data science expertise nodig is, dus dat is niet zomaar gedaan.

Status

De data is beschikbaar, op basis van nog verder uit te werken vraagstuk gaan NBTC en NDW op een later moment nogmaals het gesprek aan met een data analist van NDW om te kijken of we tot een concreet onderzoeksproject kunnen komen.

Voordelen:

- Er wordt standaard veel data verzameld
- Data is beschikbaar via open data platform
- Betreft realtime data (1 minuut oud)

Nadelen:

- Dataverwerking vraagt om een uitgebreid analyseplatform
- Niet alle wegen zijn in de data opgenomen
- Meet alleen verkeersdrukke
- Registreert alle bewegingen, dus niet alleen toeristische bewegingen

Input voor:

- Voorspelmodellen van drukke
- Druktemonitoring, inzicht in realtime drukke

8. Floating Car Data

Toelichting

Op basis van floating car data van Google Maps, TomTom of Here geven diverse partijen (oa ODIQ) inzicht in verkeersdruktes, gemiddelde snelheden en vertragingen.

Bruikbaarheid

In eerste instantie is deze data bedoeld om verkeersknooppunten in kaart te brengen, maar het is ook (deels) bruikbaar voor bezoekersmanagement doelen. Om drukte te meten kunnen bijvoorbeeld de laatste delen van aanrijdroutes gemonitord worden. Wanneer het afslagen betreft, waarvan je weet dat die leiden naar een bepaalde attractie of recreatieplek. Ook is te bepalen via welke wegen de auto's van een bepaalde parkeerplaats gekomen zijn.

Toegankelijkheid

De data wordt beschikbaar gesteld door commerciële partijen. De investering is veelal mede afhankelijk van het aantal kilometers dat je wilt meten.

Status

De data wordt momenteel veel gebruikt bij vraagstukken over verkeersstromen en betere bereikbaarheid van locaties. Er zijn bij ons geen voorbeelden bekend van gebruik van de data in relatie tot druktemonitors.

Voordelen:

- Realtime inzichten voor ca. 60% van alle weggebruikers
- Je kunt zelf kiezen welke wegen je wilt monitoren

Nadelen:

- Geen inzicht in absolute aantallen
- Betreft enkel autodata

Input voor:

- Druktemonitoring, inzicht in realtime drukte

9. Satelliet data

Toelichting

Continu monitoren satellieten in de ruimte wat er op de aarde gebeurt. Satellietdata is deels openbaar beschikbaar (open source) en deels tegen betaling. De resolutie van de beelden verschilt echter: de gratis versies kennen een minder goede resolutie en de beelden zijn minder frequent beschikbaar en tegen betaling kan de resolutie sterk vergroot worden en zijn de beelden frequenter beschikbaar.

Bruikbaarheid

Het voordeel van satellietdata is dat er tijdreeksen beschikbaar zijn, waarmee in potentie longitudinale analyses gedraaid kunnen worden die tot aardig ver in het verleden terug kunnen gaan. Bovendien gaat het om beelden (en dus data) die de feitelijke situatie weergeven en niet gebruikersafhankelijk zijn of anderszins door belangen beïnvloedbaar zijn. Om beelden op grote schaal en geautomatiseerd te kunnen analyseren is de inzet van de juiste algoritmes / deep learning methodieken cruciaal. Om te bepalen hoe groot de bruikbaarheid van deze data is voor toeristische vraagstukken, heeft het DDL hier onderzoek naar gedaan. Waarbij gestart is met een test op open source data (Sentinel data). Die test toont aan dat er aanknopingspunten zijn voor toeristische vraagstukken, maar dat de drempel om ermee aan de slag te gaan hoog is.

Toegankelijkheid

Die hoge drempel is ingegeven door kosten (als je betere kwaliteit data wil, brengt dat aanzienlijke kosten met zich mee) en de aanwezige kennis en analyseprogrammatuur. Het werken met deze data is een specialisme en daarmee niet voor iedereen een mogelijkheid.

Status

De resultaten van het DDL project naar de mogelijkheden van satellietdata voor toeristische vraagstukken rapporteren we op korte termijn.

Voordelen:

- Data is reeds beschikbaar
- Afhankelijk van de bron, biedt het zeer accurate afbeeldingen

Nadelen:

- Verstoringen in het landschap kunnen de analysemogelijkheden beperken (denk aan bomen die het zicht op een parkeerplaats wegnemen)
- Vraagt veel kennis en investeringen om hiermee aan de slag te gaan

Input voor:

- Voorspelmodellen van drukte
- Bepalen van druktebeelden op herkenbare en zichtbare plekken

Colofon

Voor de totstandkoming van dit overzicht hebben we input gekregen van verschillende leden van het LDA-netwerk en gebruik gemaakt van openbare bronnen, zoals de websites van data-aanbieders.

Met speciale dank aan Pim Nouwens van Toerisme VAN voor zijn betrokkenheid bij de totstandkoming van dit document.

NBTC Holland Marketing heeft de inhoud van deze publicatie namens de Landelijke Data Alliantie met de grootst mogelijke zorg samengesteld, maar aanvaardt geen aansprakelijkheid ten aanzien van mogelijke onjuistheden.

Wel staan we open voor uw feedback. Neem dan contact met ons op via: info@landelijkedataalliantie.nl.

Alle teksten in de publicatie zijn eigendom van NBTC. Deze mogen NIET worden verveelvoudigd, gekopieerd, gepubliceerd, opgeslagen, aangepast of gebruikt in welke vorm dan ook, online of offline, zonder voorafgaande schriftelijke toestemming van NBTC.

