



De waarde van online sentiment data – case bestemming NL

Project van Data & Development Lab (DDL)

Auteurs: Diana Korteweg Maris (HZ Kenniscentrum Kusttoerisme), Evelien Visser (NBTC Holland Marketing)



Aanleiding

Op sociale media wordt veel gesproken over toeristische bestemmingen in zowel positieve als negatieve zin. Toeristen delen vaak hun ervaringen via sociale media met vrienden/familie en kennissen. Er zijn diverse partijen die deze data verzamelen, analyseren en vervolgens aanbieden aan bestemmingen. Bestemmingen kunnen deze data gebruiken om het sentiment over de bestemming te meten en te volgen. Deze data kunnen gebruikt worden als input voor het meten van het imago en de tevredenheid over een bestemming. Daarnaast biedt het inzicht in de verschillen in het sentiment over tijd (real time monitoring). Hierdoor kunnen bestemmingen sneller reageren op zowel positieve als negatieve sentimenten.

Voor veel bestemmingen zijn deze data vrij nieuw. Vanuit het Data & Development Lab (DDL) is een project gestart om de data te verkennen en inzicht te krijgen in het online sentiment van (potentiële) bezoekers van bestemming Nederland. NBTC heeft, middels een dashboard, toegang gekregen tot de online sentiment data van één van de aanbieders van online sentiment data. Door een aantal verdiepende analyses is inzicht verkregen in de praktische bruikbaarheid van deze data: Hoe kunnen bestemmingen in Nederland deze data gebruiken? Wat kunnen bestemmingen doen met de negatieve en positieve sentimenten? Kan hierop geacteerd worden? En zo ja, hoe? Ook is een beeld gevormd van de methodologische mogelijkheden en beperkingen.

Binnen het DDL zijn deze analyses uitgevoerd in samenwerking tussen NBTC Holland Marketing, afdeling Intell & Insights en HZ University of Applied Sciences, lectoraat Data Sciences en Kenniscentrum Kusttoerisme. In dit document worden de belangrijkste bevindingen van de analyses samengevat.



Praktische bruikbaarheid (1/2)

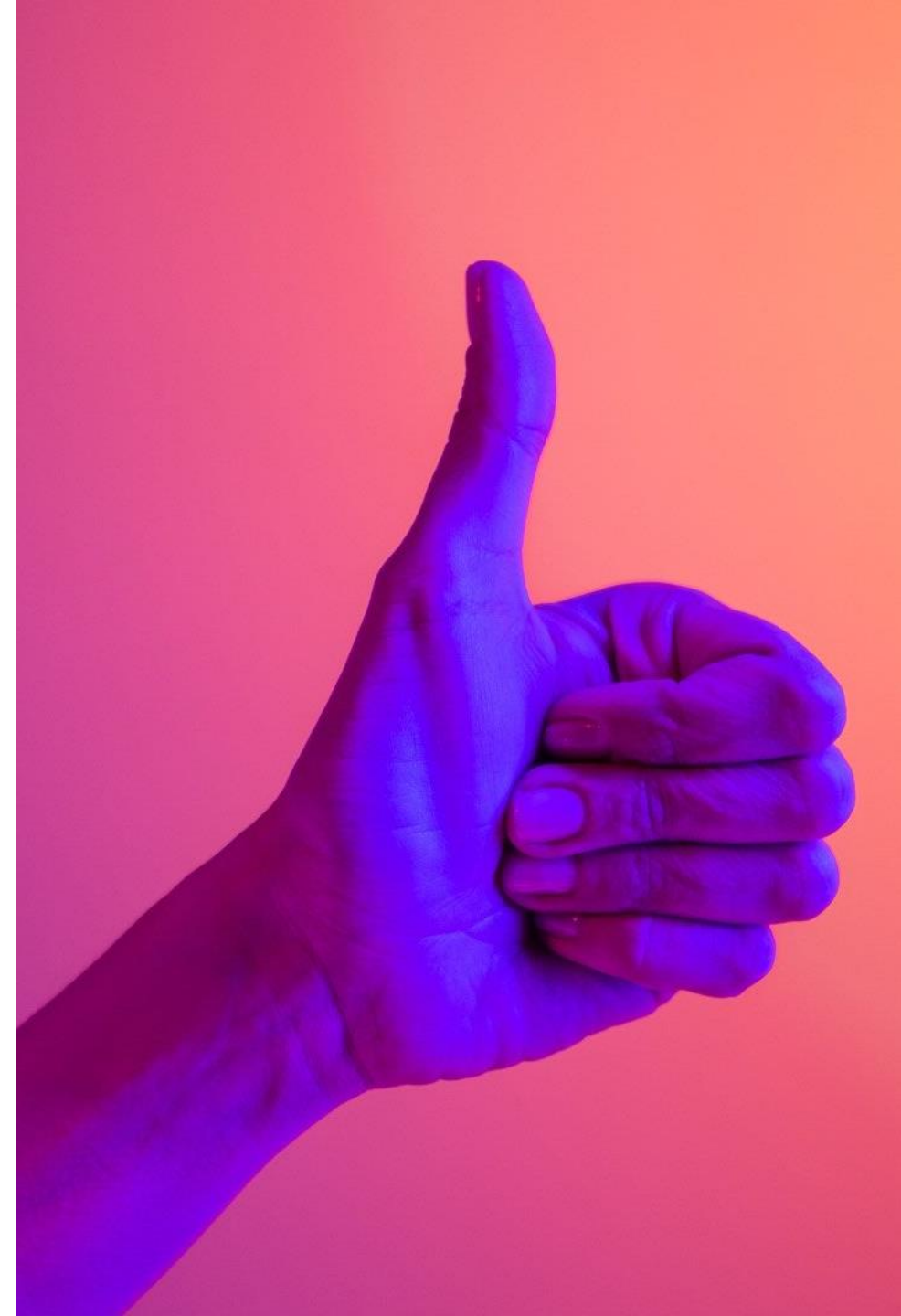
Waar kunnen bestemmingen online sentiment data voor gebruiken?

- Om inzicht te krijgen in de ontwikkeling van het online sentiment van een bestemming. Hoe wordt er over een bestemming online gesproken? En is dit positief of negatief? Hierbij moet in gedachten worden gehouden **dat online sentiment niet kan worden gezien als tevredenheidsmeting; de gebruikers van sociale media zijn niet representatief voor bezoekers van een bestemming en hebben de bestemming soms niet eens bezocht.**
- **Online sentiment data zijn vooral geschikt voor specifieke bestemmingen, zoals een stad of specifieke locatie (attractie, museum).** In deze gevallen kunnen de benodigde tags (bijvoorbeeld #Amsterdam of #Keukenhof) goed worden gekozen. De data zijn dan meer specifiek en het is beter mogelijk het verloop van de index te duiden. Wanneer de data voor grotere gebieden of een heel land worden gebruikt, dan is het lastig om volledig te zijn in de keuze van tags, waardoor een incompleet beeld ontstaat. Bovendien is het dan moeilijker te doorzien welke zaken het verloop van de index beïnvloeden.
- **Bij de inzet van influencer marketing kun je via deze data de ontwikkeling van het online sentiment, en daarmee het effect van de marketing-inspanningen, in beeld brengen.** De data geven op dagniveau inzicht in de ontwikkeling van het sentiment, zodat bijsturen eventueel mogelijk is.



Praktische bruikbaarheid (2/2)

- **In combinatie met andere databronnen kunnen online sentiment data 'kleuring' geven aan feitelijke ontwikkelingen.** De toename of afname van bijvoorbeeld het aantal hotelboekingen of aantal vluchtboekingen kan worden geduid aan de hand van de ontwikkeling van het online sentiment.
- Online sentiment data kunnen **helpen bij het bepalen van de producten en attracties** (per herkomstmarkt) **die belangrijk zijn voor het online sentiment van een bestemming.** Hierbij is het wel extra belangrijk om te beseffen dat gebruikers van sociale media niet representatief zijn voor een gehele herkomstmarkt, het betreft een zeer specifieke groep consumenten.



Wat kan wel?

- ✓ Online sentiment data gebruiken als duiding voor feitelijke ontwikkelingen
- ✓ Meten van effect van de inzet van influencer marketing
- ✓ Analyse van online sentiment voor één stad, één evenement, één bedrijf

Wat kan niet?

- X Online sentiment data gebruiken als tevredenheidsmeting, de personen die sociale media-berichten geplaatst hebben zijn niet representatief voor alle bezoekers / alle consumenten
- X Analyse van online sentiment voor een grote regio of een heel land, het is te ingewikkeld om door middel van hashtags een volledig beeld te krijgen



Methodologische ervaringen (1/3)

Toegang tot data

- De meeste datasets van aanbieders van online sentimentdata zijn niet ingericht op het maken van eigen analyses; het is zoveel mogelijk gericht op het gebruik van gestandaardiseerde dashboards. Toegang tot de data is daardoor complex. De onbewerkte data zijn zelden toegankelijk, de data zijn vrijwel altijd voorbereid door middel van het eigen algoritme van de aanbieder.

Inhoud dataset

- De dataset wordt vaak opgebouwd door een keuze van hashtags door de opdrachtgever. Dit leidt tot incomplete data, omdat niet altijd bekend is welke hashtags gekozen worden door bezoekers. Hierdoor kunnen bepaalde onderdelen van het toeristisch aanbod in een regio ondervertegenwoordigd worden in de dataset en blijven de daaraan gerelateerde doelgroepen / herkomstlanden onderbelicht. Het is van belang dat men zich hiervan bewust is bij het interpreteren van de data.
- De inhoud van de dataset wordt bepaald door de selectie van #hashtags. Het is lastig om bij de selectie van hashtags goed aan te sluiten op de woordkeuze / formulering in de geselecteerde herkomstlanden. De indruk is dat hierdoor bepaalde herkomstlanden ondervertegenwoordigd zijn in de dataset.
- De relatie tussen de inhoud van social media-berichten en de typering van toeristisch aanbod die de data-aanbieder eraan geeft is niet altijd helder. Dat maakt het lastig om te bepalen welke onderdelen van het toeristisch product invloed hebben op de ontwikkeling van het online sentiment.



Methodologische ervaringen (2/3)

- Data-aanbieders die bewerkingen maken van sociale media content geven zelden toegang tot de achterliggende data en veelal wordt gesteld dat het juist gaat om de trend die wordt weergegeven. Het is daarom des te relevanter om te begrijpen hoe zo'n trend tot stand komt. Wat wordt precies weergegeven? In de dataset die in deze verkenning werd gebruikt was sprake van een index die daalde van 23 naar 20. Het was a) niet geheel duidelijk wat deze waarde betekent - is het hoog of laag - en b) waardoor de dalende trend wordt veroorzaakt. Daardoor verliest zo'n index aan betekenis.
- Data-aanbieders creëren vaak indexen voor bepaalde deelonderwerpen, zoals een index voor de perceptie van het toeristisch product, het weer of veiligheid. Wanneer het exacte verloop van zo'n index op dagniveau wordt geanalyseerd, blijkt dat het vooral gaat om het in weergeven van gespreksonderwerpen: een actuele gebeurtenis die (internationaal) het nieuws haalt, actuele weersomstandigheden etc. Bij bijzondere gebeurtenissen zijn scherpe pieken of dalen in zo'n index zichtbaar, maar de dag erna is de index weer op het oude niveau. Is de perceptie van het betreffende onderwerp daadwerkelijk zo snel hersteld? Of wordt dit onderwerp niet meer benoemd in sociale media posts? **Dat roept vragen op over de validiteit; meten zulke indexen wel hetgeen gesuggereerd wordt?** Een gespreksonderwerp is immers niet persé hetzelfde als perceptie.



Methodologische ervaringen (3/3)

Technische beperkingen

- **Niet alle achtergrondkenmerken (land van herkomst, leeftijd) zijn van alle gebruikers bekend.** In de gebruikte dataset kon zo'n 20% van alle berichten niet worden toegewezen aan een land of leeftijdsgroep. De analyse op achtergrondkenmerken is daardoor onvolledig.
- De praktijk leert dat **hashtags over Nederland ook gebruikt worden door mensen die niet in Nederland zijn geweest.** Idealiter zou dit kunnen worden ondervangen door alleen berichten te gebruiken die geplaatst zijn tijdens het verblijf in Nederland (op basis van geolocatie). Helaas voegen niet alle gebruikers een (nauwkeurige) locatie toe.

Algemene overweging

- Het is van belang ons te realiseren dat het gebruik van sociale media door consumenten continu verandert. En ook de gebruiksmogelijkheden van sociale media platforms wijzigen continu, onder invloed van allerlei maatschappelijke ontwikkelingen en wijzigingen in wetgeving (denk aan privacywetgeving, het tegengaan van fake news, etc). Deze **wijzigingen in gedrag en in aanbod zijn van grote invloed op de inhoud van de sociale media berichten en daarmee de inhoud van de data. Dit maakt het voor alle aanbieders van analyses van sociale media content lastig om constante indexen te creëren** en belemmert vergelijkbaarheid door de tijd. Desondanks wordt het volgen van de trend vaak genoemd als de reden om inzicht te hebben in het online sentiment. Dit staat dus op gespannen voet met elkaar.

