



Universiteit  
Leiden

The Netherlands

## **Towards advanced social media metrics: understanding the diversity and characteristics of Twitter interactions around science**

Fang, Z.

### **Citation**

Fang, Z. (2021, December 21). *Towards advanced social media metrics: understanding the diversity and characteristics of Twitter interactions around science*. Retrieved from <https://hdl.handle.net/1887/3247587>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3247587>

**Note:** To cite this publication please use the final published version (if applicable).

# Nederlandse Samenvatting

Met de toenemende populariteit van het wetenschappelijke gebruik van sociale media, worden er dagelijks talloze digitale sporen van sociale media-interacties rond wetenschap achtergelaten in de online omgevingen. De analyse van deze sporen heeft geleid tot het ontstaan van de *sociale mediametriek van de wetenschap*. Twitter is aantoonbaar het populairste sociale media platform voor de communicatie van wetenschappelijke informatie. De analyse van Twitter-engagement rond wetenschap opent een waaier van mogelijkheden om de aandacht voor wetenschappelijke ontwikkelingen buiten de academische wereld te vatten en te karakteriseren. Er bestaan verschillende vormen van Twitter-interacties, mogelijk gemaakt door de vele interactieve functies die het platform biedt, die de verhalen vastleggen over hoe wetenschappelijke informatie wordt gedeeld, verspreid, besproken en gebruikt door Twitter-gebruikers.

Voortbouwend op een voorgesteld conceptueel kader van wetenschap-sociale media interacties, is het hoofddoel van deze doctoraatsverhandeling het karakteriseren van de verschillende vormen van Twitter interacties rond wetenschap, en het benaderen van meer geavanceerde Twitter-gebaseerde metrieken door systematisch rekening te houden met de diversiteit en karakteristieken van Twitter interacties. Dit proefschrift bestaat uit zeven hoofdstukken:

**Hoofdstuk 1** geeft een algemene inleiding tot de interacties tussen wetenschap en sociale media. Dit hoofdstuk begint met de definities van sociale media, wetenschappelijk gebruik van sociale media, en sociale mediametriek van de wetenschap. Het hoofdstuk stelt een conceptueel kader voor van interacties tussen wetenschap en sociale media, dat de diverse vormen van interacties binnen en tussen de wetenschappelijke en sociale media-omgevingen conceptueel schetst en categoriseert, evenals de informatiestromen die door de overeenkomstige interacties worden opgewekt. Op basis van het voorgestelde conceptuele kader wordt in dit hoofdstuk een systematisch overzicht gegeven van de relevante literatuur over de wisselwerking tussen wetenschap en Twitter. Ten slotte worden in dit hoofdstuk, rekening houdend met de mogelijkheden en uitdagingen voor wetenschappelijke Twitter-metrieken, de hoofddoelstelling en de te behandelen onderzoeksvragen naar voren gebracht.

**Hoofdstuk 2** presenteert een state-of-the-art analyse van de aanwezigheid van Twitter-vermeldingsgegevens bij wetenschappelijke artikelen, in vergelijking met andere elf soorten altmetrische gegevensbronnen die door Altmetric.com worden gevolgd. Verschillende altmetrische gegevensbronnen vertonen aanzienlijke verschillen in de opname van wetenschappelijke papers, wat de heterogeniteit van altmetrische gegevens bevestigt en het belang om ze zowel in onderzoek als in de praktijk gescheiden te houden. Over het geheel genomen bestrijken Twitter-vermeldingen meer dan een derde van de recente

wetenschappelijke papers, en vormen ze de meest globale bron van bewijsmateriaal voor sociale media-interacties rond wetenschap, slechts tweede na Mendeley. In dit hoofdstuk wordt ook gewezen op de recente vertekening van Twitter ten aanzien van wetenschappelijke artikelen die in de afgelopen jaren zijn gepubliceerd, en op de vertekening in de verschillende disciplines ten aanzien van artikelen op het gebied van *sociale en geesteswetenschappen*, *biomedische en gezondheidswetenschappen*, en *biowetenschappen en aardwetenschappen*. Ten slotte vergelijkt dit hoofdstuk de Twitter-acceptatie van wetenschappelijke papers op het niveau van het onderzoeksthema, wat licht werpt op een nieuwe manier om populaire onderzoeksonderwerpen in de ogen van Twitter-gebruikers te identificeren.

**Hoofdstuk 3** bestudeert de Twitter-opname van wetenschappelijke papers vanuit het oogpunt van snelheid, namelijk hoe snel wetenschappelijke papers op Twitter worden vermeld nadat ze zijn gepubliceerd. Hoofdstuk 3 presenteert ook een vergelijkende analyse van de gegevens accumulatie snelheid tussen twaalf altmetrische gegevensbronnen die worden bijgehouden door Altmetric.com. Verschillende bronnen laten significante discrepantie zien in de opnamesnelheid van wetenschappelijke artikelen. In het algemeen bestaan er *snelle bronnen* die een relatief hogere snelheid vertonen bij het vermelden van wetenschappelijke papers na hun publicatie, zoals Twitter, Reddit, Facebook en nieuwsmedia, en *trage bronnen* die een relatief lagere snelheid vertonen, zoals beleidsdocumenten, Wikipedia en platforms voor collegiale toetsing. Deze discrepantie versterkt het idee dat het gescheiden houden van altmetrische gegevens een belangrijke aanbeveling is, en benadrukt de noodzaak van het selecteren van geschikte tijdvensters voor verschillende bronnen van altmetrische gegevens in altmetrisch onderzoek. In hoofdstuk 3 wordt ook opgemerkt dat Twitter, als een van de snelste bronnen, de neiging heeft zich sneller te accumuleren voor documenttypes als redactioneel materiaal en brieven, en wetenschappelijke papers op het gebied van *natuur- en ingenieurswetenschappen* en *biowetenschappen en aardwetenschappen*.

**Hoofdstuk 4** richt zich op het gedrag van gebruikers rond originele wetenschappelijke tweets waarin wetenschappelijke artikelen worden genoemd, om te onderzoeken hoe originele wetenschappelijke tweets door Twittergebruikers verder worden gebruikt door ze te *liken*, te *retweeten*, te *citeren* en te *beantwoorden*. Het blijkt dat slechts de helft van de originele wetenschappelijke tweets minstens één van de vier types van gebruikersbetrokkenheid uitlokten, wat impliceert dat originele wetenschappelijke tweets anders presteren in het trekken van bredere aandacht in de Twittersphere. Originele wetenschappelijke tweets over *sociale en geesteswetenschappen*, *biomedische en gezondheidswetenschappen*, en *biowetenschappen en aardwetenschappen* zullen waarschijnlijk meer gebruikersbetrokkenheid op Twitter aantrekken. Gebaseerd op correlatie- en regressieanalyses tonen de resultaten aan dat metriek van gebruikersbetrokkenheid meer correleert met andere Twitter-gebaseerde factoren (bv. volgers van Twitter-gebruikers of vermelde gebruikers in wetenschappelijke tweets) dan met wetenschapsgebaseerde factoren (bv. citaties of Mendeley-lezers van wetenschappelijke papers), wat de intrinsieke relaties

tussen Twitter-elementen en -activiteiten suggereert, alsook de differentiële neigingen van interacties die plaatsvinden in de wetenschap en op Twitter voor wetenschappelijke outputs.

**Hoofdstuk 5** onderzoekt de mogelijke onbeschikbaarheid van wetenschappelijke tweets en de invloed daarvan op de stabiliteit van op Twitter gebaseerde metrieken. Door het opnieuw controleren van de (on)beschikbaarheid van een set wetenschappelijke tweets van hooggetwitterde wetenschappelijke papers die tot oktober 2017 door Altmetric.com zijn geregistreerd, meldt dit hoofdstuk dat er in april 2019 14,3% van de tweets onbeschikbaar was geworden voor het publiek, voornamelijk als gevolg van het verwijderen van tweets, schorsing van gebruikersaccounts en bescherming van gebruikersaccounts. Het niet beschikbaar zijn van wetenschappelijke tweets kan de stabiliteit van op Twitter gebaseerde metrieken ernstig beïnvloeden. Door een onderscheid te maken tussen originele tweets en retweets, en daartussen vervolgens te netwerken op basis van de retweeting-relaties, toont hoofdstuk 5 dat wetenschappelijke artikelen met minder originele tweets en veel retweets kwetsbaarder zijn in hun verspreidingsstructuur op Twitter, en dus een groter risico lopen om onstabiele, op Twitter gebaseerde statistieken te creëren. Deze bevindingen weerspiegelen de potentiële instabiliteit van op Twitter gebaseerde metingen als gevolg van de veranderlijke aard van Twitter-gegevens, en benadrukken de noodzaak om niet alleen een onderscheid te maken tussen originele tweets en retweets, maar ook om de Twitter-verspreidingsstructuur van papers te analyseren met een netwerkbenadering.

**Hoofdstuk 6** onderzoekt het klikgedrag rond wetenschappelijke URL's die verwijzen naar wetenschappelijke artikelen die zijn ingebed in wetenschappelijke tweets, waardoor Twitter-gebruikers toegang krijgen tot de originele webpagina's van de getweete wetenschappelijke artikelen. Gebaseerd op de klikstatistieken van Bitly voor zijn gegenereerde korte links, stelt hoofdstuk 6 vast dat slechts ongeveer de helft van de wetenschappelijke URL's met succes op Twitter werden aangeklikt en Twitter-gebruikers dus naar de oorspronkelijke wetenschappelijke inhoud leidden. De meeste Twitter-kliks op de getweete wetenschappelijke URL's concentreren zich in de eerste paar dagen nadat de wetenschappelijke URL's waren getweet. Net als bij andere Twitter-interacties lijken Twitter-kliks vaker voor te komen bij wetenschappelijke URL's met betrekking tot *sociale en geesteswetenschappen, biomedische en gezondheidswetenschappen, en biowetenschappen en aardwetenschappen*. Het verschil in het aantal Twitter-kliks tussen wetenschappelijke URL's wijst ook op de verschillende prestaties van wetenschappelijke tweets bij het op gang brengen van een bredere aandacht, wat de mogelijkheid biedt om de doeltreffendheid te meten van wetenschappelijke tweets bij de verspreiding van wetenschappelijke informatie en de feedback van Twitter-verspreiding op de wetenschap zelf (bv. door meer bezoeken aan wetenschappelijke papers).

**Hoofdstuk 7** vat de belangrijkste bevindingen van de hoofdstukken 2 tot en met 6 samen en bespreekt de implicaties en enkele toekomstperspectieven op basis van de belangrijkste

bevindingen. Onderzoek naar de opname van wetenschappelijke papers via Twitter (hoofdstukken 2 en 3) toont aan dat Twitter een wereldwijde en onmiddellijke bron van bewijs is van sociale media-interacties rond wetenschap. Onderzoek naar de verschillende soorten Twitter-engagement metrieken voor wetenschappelijke tweets (hoofdstukken 4 en 6) onthult de mogelijkheid om een meer fijnmazig indicatorsysteem van Twitter-gebaseerde metrieken op te zetten, dat ook de metingen van meer diepgewortelde Twitter-ontvangst van wetenschappelijke informatie mogelijk maakt. Het onderzoek naar de mogelijke onbeschikbaarheid van wetenschappelijke tweets (hoofdstuk 5) weerspiegelt de invloed die de vluchtige aard van tweets kan hebben op de stabiliteit van op Twitter gebaseerde metriek, die wordt aanbevolen om in Twitter-gerelateerd onderzoek en evaluatie met de nodige voorzichtigheid te behandelen. Om de aard van de diverse Twitter-interacties beter te begrijpen, worden in hoofdstuk 7 ook enkele richtingen voor toekomstig onderzoek voorgesteld, met name de contextualisering van Twitter-interacties door rekening te houden met de informatie van de betrokken Twitter-gebruikers en de gedetailleerde tweet-inhoud. Bovendien zou het een belangrijke toekomstige stap zijn om het conceptuele raamwerk van wetenschap-sociale media-interacties (in hoofdstuk 1) en het ontwerp van een fijnmaziger indicatorsysteem van sociale media-metrieken (in hoofdstuk 7) te veralgemenen naar een breder scala van sociale media bronnen.