

Een exploratief  
onderzoek naar  
indicatoren van  
tweets die  
verstuurd worden  
door omstanders  
van een  
natuurramp

juli 19

20

16

Document:	Afstudeerscriptie
1 <sup>e</sup> beoordelaar:	Iris Hendrickx
2 <sup>e</sup> beoordelaar:	Suzan Verberne
Datum:	18 juli 2016
Student:	Timo Calamé (4384911)
Opleiding:	M Communicatie & Beïnvloeding
Cursus:	Afstudeerscriptie
Tel:	06 2311 5026
E-mail:	timo.calame@student.ru.nl

## Inleiding

Het nieuws over de dood van Osama Bin Laden werd eerder verspreid op Twitter dan door nieuwszenders of het Witte Huis (Hu et al., 2012). Sociale media kunnen dus ongelooflijk snel en real-time informatie verschaffen als officiële instanties daar nog niet toe in staat zijn. Sociale media zijn een waardevolle bron van bijna elk soort informatie (Truong et al., 2014). Tweets worden om die reden ook vaak gebruikt als onderzoeksmateriaal. Een andere reden dat tweets onderzocht worden is dat deze meestal openbaar en makkelijk te verzamelen zijn. Tweets zijn daarom aantrekkelijker om te onderzoeken dan bijvoorbeeld Facebook- of Instagramposts.

In het huidige onderzoek worden tweets rondom overstromingen geanalyseerd. Tweets over overstromingen kunnen relevante informatie bevatten voor bijvoorbeeld journalisten, burgers en hulpdiensten. Het doel van dit onderzoek is om indicatoren van die relevante tweets te beschrijven. Door het beschrijven van deze indicatoren kunnen automatische systemen die relevante tweets herkennen verbeterd worden. Relevante tweets worden in het kader van dit onderzoek "*directe tweets*" genoemd. Directe tweets worden verstuurd door omstanders van een gebeurtenis.

Aan de hand van een literatuurstudie worden mogelijke indicatoren van directe tweets beschreven in de vorm van hypothesen. De directheid van tweets en de mogelijke indicatoren van tweets worden vervolgens in een corpus van tweets rondom een natuurramp geannoteerd. Het annoteren zal uitgevoerd worden in drie fasen. In de eerste fase wordt ruis uit het corpus van tweets verwijderd. In de tweede fase wordt de directheid van tweets geannoteerd en wordt getoetst of de indicatoren die handmatig geannoteerd moeten worden significante resultaten opleveren. Indicatoren van tweets die automatisch geannoteerd kunnen worden, zullen automatisch geannoteerd worden in de derde fase. Nadat het corpus geannoteerd is zullen er statistische toetsen uitgevoerd worden om de hypothesen die opgesteld zijn te kunnen evalueren.

## Verwante literatuur

Een voorbeeld van een toepassing van Twitter is het analyseren van de inhoud of het meten van het sentiment tijdens een epidemie (Chew & Eysenbach, 2010; Signorine, Segre & Polgreen, 2011). Tweets tijdens de H1N1-epidemie (Mexicaanse griep) konden gebruikt worden om het aantal H1N1 geïnficeerden sneller vast te stellen dan dat traditionele methoden dat konden (Chew & Eysenbach, 2010). Dit laat zien dat tweets relevante informatie bevatten. In ander onderzoek werden tweets tijdens rampen geanalyseerd om een beeld te krijgen over wat mensen tweeten over een ramp en om te onderzoeken of er relevante informatie in tweets staan (Takahashi, Tandoc jr. & Carmichael, 2015; Truong et al. 2014). Het is belangrijk om dit soort onderzoek te doen omdat er tijdens rampen over het algemeen weinig betrouwbare informatie voor handen is. Truong et al. (2014) hebben

getracht om tweets automatisch te annoteren als informatie- of conversatietweets. Het automatisch annoteren van conversatietweets was succesvol (precision = .92). De annotatie van informatietweets was daarentegen minder betrouwbaar (precision = .41) terwijl relevante informatie waarschijnlijk hetgeen is waar mensen naar op zoek zijn tijdens rampen. Verbeteringen van dergelijke systemen zijn daarom wenselijk en noodzakelijk.

Een ander probleem waar dit soort onderzoek vaak mee van doen heeft, is dat tweets foutieve informatie bevatten. Tweets die viral gingen na de aanslag in Boston op 15 april 2013 bevatten bijvoorbeeld in 29% van de gevallen onjuiste informatie (Gupta, Lamba & Kumaraguru, 2013). Gupta et al. (2013) concluderen dat tweets vaak foutieve informatie bevatten omdat mensen die geraakt worden door een ramp de behoefte hebben om informatie te delen ondanks het feit dat het van een onbekende bron afkomstig is. Het gebruik van tweets als bron van informatie over rampen is dus een risico. Een andere uitdaging is dat Twitter een rumoerig ("noisy") medium is (Truong et al, 2014). Er worden veel tweets gestuurd die niet de relevante informatie bevatten voor mensen die in een rampsituatie verkeren (Helsloot & Groenendaal, 2013)

Onderzoek naar tweets tijdens rampen is belangrijk, want tijdens een ramp kan er veel onduidelijkheid zijn onder burgers, journalisten, hulpverleners en crisismanagers (hierna: "crisisstakeholders"). Tweets kunnen, door het real-time karakter, snel informatie geven over de ontwikkeling van een ramp. De uitdaging van onderzoek naar dit onderwerp is om tweets te selecteren die waardevolle, nieuwe en betrouwbare informatie bevatten. Door het selecteren van relevante tweets kunnen crisisstakeholders tijdig geïnformeerd worden over een ramp en zijn daardoor in staat maatregelen nemen. Tevens hebben journalisten en crisismanagers in dat geval toegang tot waardevolle informatie om nieuws te brengen of om beslissingen te nemen (Hirschberg & Manning, 2015). Het doel van dit onderzoek is daarom het beschrijven van indicatoren van relevante tweets tijdens rampen. Dit draagt bij aan een automatische inhoudsanalyse om in een korte tijd de meest relevante informatie op twitter te kunnen detecteren als een ramp zich voordoet.

### **Literatuursectie**

Om relevante tweets te selecteren tijdens een ramp zijn er grofweg twee manieren. De eerste mogelijkheid is het handmatig annoteren of lezen van tweets. Deze methode is echter geen interessante optie aangezien het te veel tijd kost om alle tweets over een ramp door te nemen. Over de brand in Moerdijk waarbij een chemische opslag in brand stond werden bijvoorbeeld binnen 24 uur 52,806 tweets verstuurd (Helsloot & Groenendaal, 2013). Het is door deze grote hoeveelheid tweets onmogelijk deze met de hand te annoteren. Daarnaast valt met het handmatig annoteren van tweets tijdens een ramp het real-time karakter van de informatie weg.

De tweede mogelijkheid is om relevante tweets door middel van een automatische inhoudsanalyse te selecteren. Automatische inhoudsanalyse van tweets is sneller dan het handmatig annoteren van tweets. Dit maakt automatische inhoudsanalyse een geschiktere methode dan handmatige analyse.

Er zijn verschillende onderzoeken gedaan naar het automatisch selecteren van relevante tweets tijdens rampen. Truong et al. (2014) hebben bijvoorbeeld door middel van automatische inhoudsanalyse, informatieve tweets geprobeerd te onderscheiden van conversatietweets. Een informatietweet bevat volgens dit onderzoek waardevolle en concrete informatie over de ramp voor iedereen die de tweet zou lezen. Informatieve tweets zouden volgens Truong et al. (2014) relevant kunnen zijn voor crisisstakeholders. Conversatietweets bevatten namelijk geen waardevolle en concrete informatie en informatieve tweets wel. Een nadeel van deze operationalisaties is dat informatieve tweets ook onbetrouwbare informatie kunnen bevatten.

Andere onderzoekers hebben getracht automatisch tweets te selecteren die individuen kunnen gebruiken om beslissingen te nemen (Nguyen, Kitamoto & Nguyen, 2015; Verma et al., 2011). Dit werden respectievelijk *actionable*- en *situational awareness tweets* genoemd. Het lezen van deze tweets kan voor individuen, die hinder ondervinden van een ramp, van belang zijn omdat er meestal weinig actuele informatie voor handen is. Vraagtekens kunnen echter gezet worden bij de betrouwbaarheid van deze informatie. Dat een tweet relevante informatie bevat (dat gebruikt kan worden om beslissingen te nemen) betekent niet noodzakelijk dat deze informatie ook betrouwbaar is. Bovendien kan opgemerkt worden dat deze informatie meestal niet nieuw is. Een “actionable tweet” kan namelijk ook refereren aan het nieuws over een ramp.

Omdat de betrouwbaarheid van tweets niet altijd te garanderen is kan het nuttig zijn om de betrouwbaarheid van Twittergebruikers te analyseren (Abbasi & Liu, 2013). Het blijkt dat de profielinformatie van een gebruiker relatief goed kan voorspellen of een tweet betrouwbaar is (Bodnar et al., 2014). Andere onderzoeken focussen zich niet zozeer op de profielinformatie van een tweet maar ook op de indicatoren van tweets zelf, zoals het aantal woorden en of de tweet een URL bevat (Castillo, Mendoza & Poblete, 2012; Ross, 2015). Deze systemen lijken de betrouwbaarheidstekortkomingen van onder andere Truong et al. (2014) en Nguyen et al. (2015) op te lossen. Als automatische inhoudsanalyses er in slagen om betrouwbare tweets met relevante informatie te detecteren zouden crisisstakeholders die kunnen gebruiken om maatregelen te nemen.

Een kanttekening is dat betrouwbare informatie uit tweets niet per definitie nieuwe informatie is. Voor bepaalde typen crisisstakeholders, zoals journalisten en crisismanagers, kan het van belang zijn om niet alleen betrouwbare informatie te detecteren maar ook nieuwe informatie. Een tweet die refereert aan een nieuwsbron is bijvoorbeeld wel betrouwbaar maar bevat hoogstwaarschijnlijk geen nieuwe informatie. In het huidige onderzoek wordt gesteld dat tweets van omstanders

mogelijkerwijs nieuwe én betrouwbare informatie bevatten. Omstanders van een ramp wonen de gebeurtenis fysiek bij waardoor ze mogelijk goed in staat zijn nieuwe en betrouwbare informatie te verschaffen over een ramp dan mensen die een ramp niet fysiek meemaken. Bovendien is deze informatie mogelijk nog niet op een andere plek in het informatie-universum beschikbaar (Starbird et al., 2010). Een tweet wordt in het huidige onderzoek als relevant gezien als deze ogenschijnlijk door een omstander van de ramp is verstuurd.

Tweets die door omstanders zijn verstuurd, worden volgens de ontologie van Imran et al. (2013) “directe” tweets genoemd. Tweets die bijvoorbeeld refereren aan het nieuws, de krant of de televisie worden “indirecte” tweets genoemd. In het huidige onderzoek zal de terminologie van Imran et al. (2013) gehandhaafd worden. Om een bijdrage te leveren aan automatische inhoudsanalyse voor het selecteren van directe tweets is het van belang om indicatoren van directe tweets te beschrijven. De onderzoeksvraag van het huidige onderzoek is daarom als volgt: *“Welke indicatoren hebben directe tweets?”*.

### **Indicatoren van directe tweets**

Een belangrijke indicator van een directe tweet zou het aantal “retweets” kunnen zijn (Starbird & Palen, 2012). Retweets werken volgens Starbird en Palen (2012) net zoals een aanbeveling-systeem waarbij een retweet als het ware een stem is op de originele tweet (recommendation). Starbird en Palen (2012) hebben handmatig ruim 250 tweets geannoteerd die relatief veel retweets kregen tijdens de Arabische lente. Het aantal retweets bleek een goede voorspeller te zijn voor een directe tweet. Dit effect werd alleen gevonden als het aantal volgers (followers) van de gebruikers relatief laag was. Op basis van dit resultaat is de volgende hypothese opgesteld:

H<sub>1</sub>: Directe tweets krijgen significant meer retweets dan indirecte tweets. Dit effect treedt alleen op als de gebruiker die de tweet heeft verstuurd minder dan 500 volgers heeft.

De eigenschappen van het medium Twitter maken het voor gebruikers mogelijk om een bericht van maximaal 140 tekens te versturen. Omdat 140 tekens soms onvoldoende is, kunnen Twitter-gebruikers een URL in het bericht plaatsen. Deze URL verwijst door naar een andere webpagina. Ongeveer 22% van de tweets bevat een URL (Boyd, Golder & Lotan, 2010). Tweets die een URL bevatten worden tevens vaker ge-retweet dan tweets die geen URL bevatten (Boyd et al., 2013; Tanaka, Sakamoto & Honda, 2014). De resultaten van een experiment van Tanaka et al. (2014) suggereren dat tweets die een URL bevatten vaker worden ge-retweet ondanks dat de URL niet daadwerkelijk doorverwijst naar een andere webpagina.

De aanwezigheid van een URL voorspelt mogelijk of een tweet direct of indirect is. Omstanders van een ramp hebben namelijk eerstehands informatie. Een directe tweet mét URL naar een andere webpagina is daarom minder waarschijnlijk. Bovendien is een URL naar een webpagina voor dit onderzoek minder interessant omdat deze verwijst naar informatie die al bekend is. Daarom wordt in het huidige onderzoek verwacht dat directe tweets geen URL bevatten.

H<sub>2</sub>: Indirecte tweets bevatten significant vaker een URL dan directe tweets.

Woordkeuze van mensen kan iets zeggen over henzelf, hun publiek en de situatie waarin ze zich begeven (Pennebaker et al., 2003). De Construal Level Theory (CLT) zou een aanwijzing kunnen geven voor het woordgebruik in directe tweets. Deze theorie gaat er vanuit dat de ervaring van een gebeurtenis afhangt van de mate waarin iemand de gebeurtenis als dichtbij of ver weg ervaart (Trope, Liberman & Wakslak, 2011). Dit wordt ook wel psychologische afstand genoemd. Psychologische afstand is de subjectieve perceptie dat iets ver weg of dichtbij is (Trope & Liberman, 2010). Er zijn vier vormen van psychologische afstand: temporele afstand (nu of in de toekomst), ruimtelijke afstand (hier of daar), sociale afstand (jezelf of anderen) en hypothetische afstand (echt of ingebeeld). Directe tweets refereren aan mensen die een ramp hebben meegemaakt. De ruimtelijke en temporele afstand voor mensen die een ramp hebben meegemaakt is daarom waarschijnlijk klein. Volgens de CLT is de mentale representatie van een gebeurtenis concreter bij een kleine psychologische afstand dan bij een grote psychologische afstand (Trope et al., 2011).

Uit een experiment van Fujita et al. (2006) bleek dat proefpersonen die een gebeurtenis moesten beschrijven die dichtbij was (kleine ruimtelijke afstand) concretere werkwoorden gebruikten dan proefpersonen die een gebeurtenis moesten beschrijven die ver weg was (grote ruimtelijke afstand). Proefpersonen die een gebeurtenis moesten beschrijven die ver weg was gebruikten meer abstracte werkwoorden dan concrete werkwoorden. Een concreet werkwoord is bijvoorbeeld "lopen", een abstract werkwoord is bijvoorbeeld "haten". Deze resultaten suggereren een verband tussen taalgebruik en de ruimtelijke afstand die mensen ervaren ten aanzien van een gebeurtenis. Aangezien omstanders van een ramp de gebeurtenis waarschijnlijk als dichtbij ervaren wordt in het huidige onderzoek verwacht dat directe tweets concretere werkwoorden bevatten dan indirecte tweets.

H<sub>3</sub>: Werkwoorden in directe tweets zijn significant concreter dan werkwoorden in indirecte tweets.

Een andere indicatie voor directe tweets zou het gebruik van werkwoordtijden kunnen zijn. Werkwoorden kunnen namelijk refereren naar de toekomst, het heden of het verleden (Bos et al., 2013). “Ik zie een dijk instorten” refereert bijvoorbeeld naar het heden terwijl “ik zag een dijk instorten” refereert naar het verleden. Dit gegeven kan een interessante aanwijzing zijn voor het voorkomen van directe tweets. Werkwoordtijd geeft namelijk informatie over wanneer een gebeurtenis plaatsvindt/heeft plaatsgevonden (Bos et al., 2013). Verwacht wordt daarom dat de werkwoordtijd die gebruikt wordt in tweets een aanwijzing kan zijn voor het voorkomen van directe tweets. Omdat een directe tweet refereert aan een gebeurtenis die iemand zelf heeft meegemaakt wordt verwacht dat deze geschreven zijn in de verleden tijd.

H<sub>4</sub>: Directe tweets zijn significant vaker geschreven in de verleden tijd dan indirecte tweets.

### **Overstromingen in het Verenigd Koninkrijk**

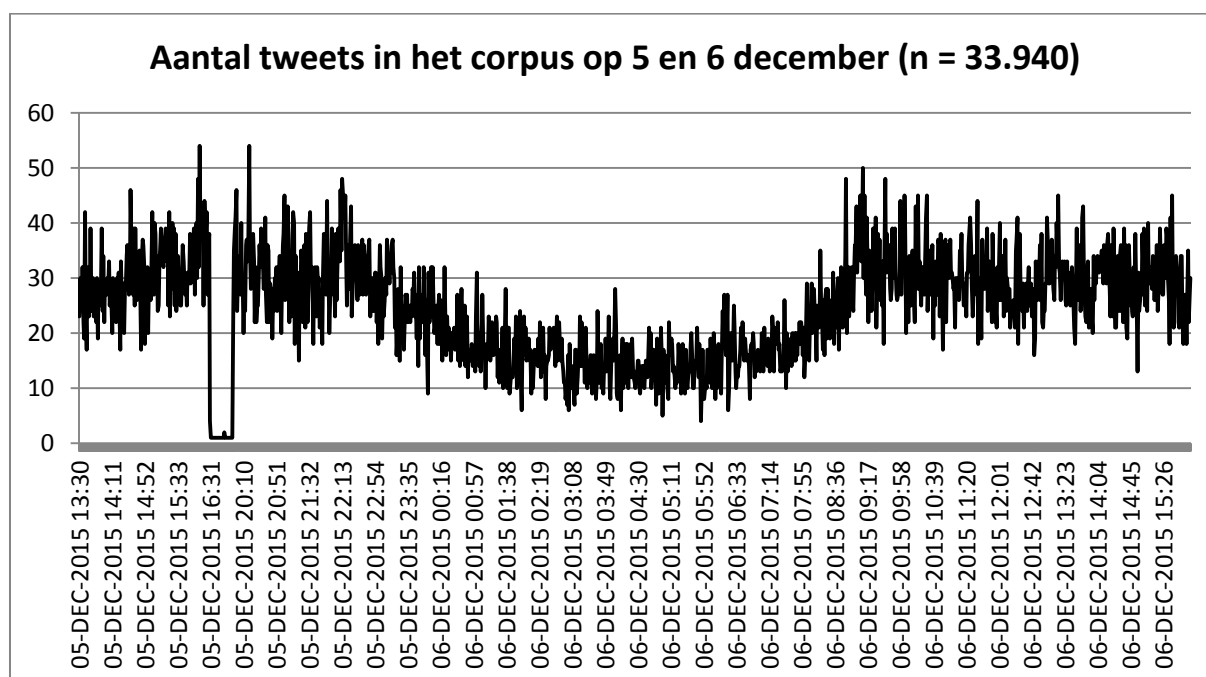
In het huidige onderzoek worden tweets geanalyseerd tijdens overstromingen in het Verenigd Koninkrijk. Op 5 en 6 december 2015 trok er een hevige storm over het Verenigd Koninkrijk (VK) en Ierland, genaamd “*Desmond*”. Deze storm zorgde voor ongewoon veel regenval. Binnen 24 uur viel er bijna 35 cm regen. De regenval op 5 en 6 december was de oorzaak van hevige overstromingen. Duizenden mensen moesten worden geëvacueerd, trein en autoverkeer was in delen van het land onmogelijk en sportevenementen werden afgelast. In de maand december hebben twee andere stormen het Verenigd Koninkrijk geteisterd (Frank en Eva). De schade van deze stormen was kleiner dan storm Desmond. In het huidige onderzoek is er daarom voor gekozen tweets rondom storm Desmond te analyseren.

## Methode

In samenwerking met FloodTags werd er een corpus van tweets verzameld ( $n = 33.940$ ) die verstuurd zijn tijdens de overstromingen in het Verenigd Koninkrijk. Dit corpus is gebruikt om indicatoren van directe tweets te kunnen beschrijven. FloodTags is een organisatie die tweets rondom overstromingen verzamelt en analyseert ([www.floodtags.com](http://www.floodtags.com)). De tweets zijn in drie fasen geanalyseerd. In de eerste fase is de ruis in de tweets verwijderd. Nadat de ruis is verwijderd, zijn er in de twee opvolgende fasen de tweets zowel handmatig als automatisch geannoteerd. De abstractie van werkwoorden kon volgens de handleiding die is gebruikt alleen handmatig uitgevoerd worden. Er is bij een kleine subset van tweets getoetst of de abstractie van tweets een indicator is van directe tweets. Vervolgens is over de hele set tweets een automatische analyse gedaan.

### Materiaal

Het corpus bestond uit Engelstalige tweets die minstens één woord bevatten over overstromingen en de gevolgen daarvan. Hierbij kan gedacht worden aan woorden zoals: “*floods*”, “*inundations*”, “*embankment burst*” en “*dyke fall*” (voor de complete woordenlijst zie bijlage 1). Er zijn 33.940 tweets verzameld tijdens storm Desmond van 5 december 13.30u (lokale tijd) tot en met 6 december 16.00u. Er is gekozen voor dit tijdsinterval aangezien er in die periode overstromingen en andere gerelateerde ongevallen zijn gerapporteerd door de media (Nort-West Evening Mail, 2015).



Figuur 1: Het aantal tweets per minuut uit het corpus tussen 5 december 13.30u en 6 december 16.00u.



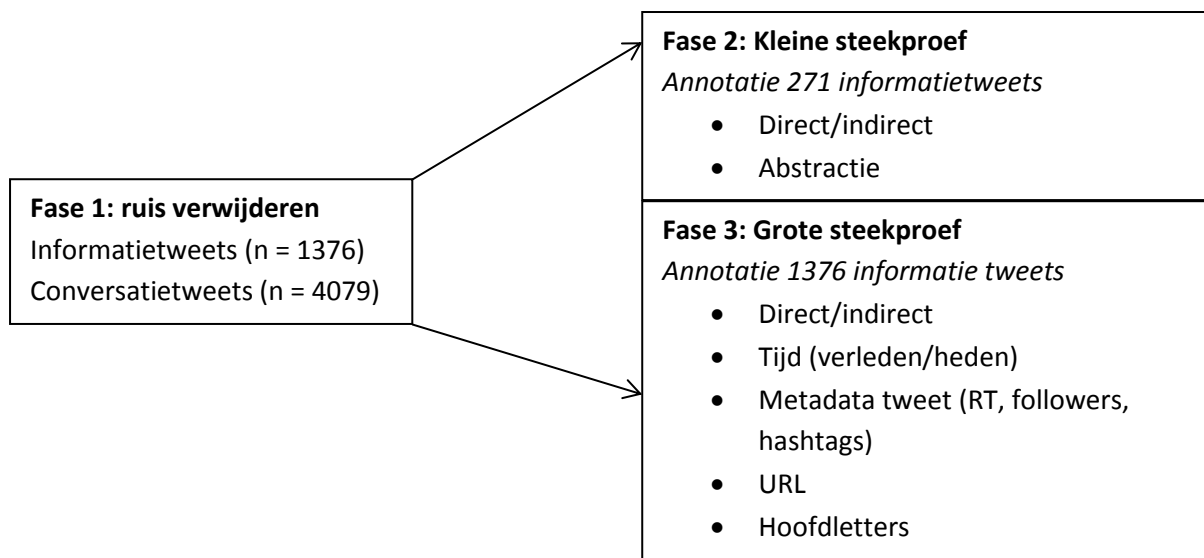
## Procedure

Het annoteren van tweets is in drie fasen uitgevoerd. In de eerste fase is het corpus (n=33.940) geannoteerd op een binaire schaal van "*informatie-*" en "*conversatietweets*" op eenzelfde manier als Truong et al. (2014). Dit onderscheid is gemaakt om de niet relevante (conversatie) tweets op voorhand uit het corpus te halen omdat dit per definitie geen directe tweets zijn. Tevens zijn automatische filtersystemen succesvol om onderscheid te maken tussen informatie- en conversatietweets (Truong et al., 2014).

Het corpus is in random volgorde geannoteerd totdat er 1376 tweets geannoteerd werden als informatietweets en 4079 tweets werden geannoteerd als conversatietweets. Tweets werden in het huidige onderzoek geannoteerd als informatietweets als deze waardevolle en concrete informatie bevatten voor crisisstakeholders zoals: aantal doden en gewonden, nieuws over de ramp, persoonlijke ervaringen met de overstromingen en andere relevante informatie. Alle andere tweets werden geannoteerd als conversatie (ruis). Tweets die geschreven waren in een andere taal dan Engels en tweets waarvan de inhoud niet overstroming-gerelateerd waren zijn geannoteerd als conversatietweets. Bij twijfelgevallen stelden codeurs zich de vraag of de tweet mogelijk relevante informatie zou kunnen bevatten voor crisisstakeholders. 200 tweets werden dubbel geannoteerd. De interbeoordelaarsbetrouwbaarheid van deze schaal was goed (Cohen's  $\kappa = .78$ ). Deze annotatiefase leverde een corpus op van 1376 informatietweets.

In de tweede fase zijn 271 informatietweets uit de eerste fase opnieuw geannoteerd. Tweets werden geannoteerd als "*direct*" of "*indirect*". De enige indicator uit de literatuursectie dat niet automatisch geannoteerd kon worden (abstractie van werkwoorden) werd ook geannoteerd. Nadat deze 271 tweets geannoteerd waren, is er een t-toets uitgevoerd om te bepalen of werkwoordabstractie een indicator kon zijn van directe tweets. Dit bleek niet het geval te zijn.

Op basis van het annoteren van de kleine steekproef kwamen nog andere mogelijke indicatoren van directe tweets naar boven. Deze indicatoren zijn vervolgens in de grote steekproef geannoteerd. Indirecte tweets leken minder hoofdletters en hashtags leken te bevatten dan directe tweets. Deze indicatoren zijn daarom geannoteerd in de grote steekproef. Vervolgens zijn in fase 3 alle informatietweets uit het corpus geannoteerd (n = 1376). Deze tweets bevatten ook tweets uit de kleine steekproef omdat de werkwoordabstractie geen indicator bleek te zijn van directe tweets.



*Figuur 2: Overzicht van de procedure.*

Een aantal indicatoren van de tweets uit de grote steekproef zijn middels een PHP script automatisch geannoteerd. Dit waren de aanwezigheid van een URL en aantal hoofdletters. Het aantal retweets, het aantal volgers van de gebruiker op het moment van versturen van de tweet, het aantal mentions, het aantal hashtags, het aantal likes en of de tweets een foto bevatten is geannoteerd door gebruik te maken van de API die twitter aanbiedt. De werkwoordtijden zijn automatisch gecodeerd door gebruik te maken van een part of speech tagger. De concreetheid van de werkwoorden en of de tweet direct of indirect was, is handmatig geannoteerd in de kleine en de grote steekproef.

Tweets werden geannoteerd als “*direct*” of “*indirect*” op dezelfde manier als Starbird et al. (2012) dat hebben gedaan. In het corpus zijn 191 tweets geannoteerd als direct en 1185 als indirect. Een tweet werd geannoteerd als direct wanneer iemand informatie twitterde die van de rampplek zelf kwam en als die persoon ook zelf op de rampplek aanwezig was. Een tweet werd geannoteerd als indirect als iemand niet op de rampplek aanwezig was of als iemand geen informatie twitterde over de ramp. Bij twijfelgevallen beoordeelde de codeurs of de persoon die de tweet had gestuurd zelf iets van de ramp kon waarnemen met de zintuigen van die persoon. Als de tweet refereerden aan een televisieverslag dan werd deze tweet geannoteerd als indirect omdat de persoon de ramp niet zelf heeft waargenomen. Als een persoon bijvoorbeeld een situatie rondom zijn eigen huis beschrijft die ogenschijnlijk door de twitteraar zelf is waargenomen dan is die tweet geannoteerd als direct. De interbeoordelaarsbetrouwbaarheid van deze schaal was goed (Cohen’s  $\kappa = .71$ ).

De abstractie van (werk)woorden in tweets werd geannoteerd volgens het “*Linguistic Category Manual*” (LCM). Het LCM is een instrument dat de abstractie van taalgebruik kan kwantificeren (Coenen, Hedebouw & Semin, 2006). Het handboek categoriseert vier verschillende

maten van abstractie bij woorden. Bijvoeglijke naamwoorden, zoals “grappig” zijn volgens het LCM de meest abstracte woorden. Woorden die refereren aan een actie met een duidelijk begin en einde, zoals “vallen”, zijn volgens het LCM de meest concrete woorden. De vier verschillende categorieën van abstractieniveau worden in de volgende vier alinea’s toegelicht. Bij twijfelgevallen werd gebruik gemaakt van het LCM handboek (Coenen, Hedeboom & Semin, 2006). De meest abstracte woorden zullen als eerst worden beschreven alvorens er toegewerkt wordt naar meer concrete woorden. De interbeoordelaarsbetrouwbaarheid van deze schaal was voldoende (Cohen’s  $\kappa = .66$ ).

Het verschil tussen bijvoeglijke naamwoorden (BN) en de andere drie categorieën is dat het géén werkwoorden zijn (Coenen et al., 2006). Een BN refereert namelijk naar de kwaliteiten of indicatoren van een persoon of object (sociaal, mooi) en een werkwoord refereert naar een actie (lopen, praten) of psychologische staat (haten, liefhebben). Een BN verwijst dus niet naar wat een persoon doet maar naar de karakteristieken van een persoon of object. Een BN kan daarnaast ook iets zeggen over een gebeurtenis of object. Bijvoeglijke naamwoorden zijn geannoteerd door gebruik te maken van de Stanford Part Of Speech tagger (SNLP, n.d.). “Adjective”, “Adjective comparative” en “Adjective Superlative” uit de tagger zijn in het huidige onderzoek samengevoegd tot één variabele namelijk: BN (bijvoeglijk naamwoord).

Van de werkwoorden in het LCM worden “*state-verbs*” (SV) als het meest abstract gezien. SV’s refereren naar de psychologische staat van een persoon ten opzichte van een andere persoon of de verandering daarin (Coenen et al., 2006). Deze werkwoorden hebben geen duidelijk begin of eind. “*Bewonderen*”, “*houden van*” en “*haten*” zijn voorbeelden van SV’s. Aangezien er niet veel verschillende SV’s zijn worden er geen problemen verwacht met het annoteren ervan. Er is bijvoorbeeld een corpus van SV’s die geraadpleegd kan worden bij twijfelgevallen.

Ten opzichte van state verbs zijn “*state action verbs*” (SAV) en “*interpretative action verbs*” (IAV) concretere werkwoorden (Coenen et al., 2006). Deze twee typen werkwoorden zijn even abstract en lijken dermate veel op elkaar dat ze in het LCM samengevoegd zijn onder één categorie. SAV en IAV worden in het huidige onderzoek daarom ook samengevoegd en “*interpretative action verbs*” (IAV) genoemd. IAV’s refereren naar een cluster van acties zonder duidelijk begin en eind (Coenen et al., 2006). Een IAV is een verzamelwerkwoord voor verschillende soorten acties. “*Ik heb mijn oma vandaag geholpen*”. Helpen is in dit geval een cluster van verschillende acties die iemand kan uitvoeren om zijn oma te helpen. “*Ik heb vandaag het bed van mijn oma opgemaakt*”. Opmaken is in deze context een actie die onder het cluster “helpen” zou kunnen vallen. Werkwoorden in het corpus die refereerden aan een cluster van verschillende acties werden geannoteerd als IAV. Werkwoorden die refereerden naar specifieke acties, zoals “*opmaken*”, werden geannoteerd als “*descriptive action verbs*”. DAV’s refereren aan een specifieke actie met een duidelijk begin en einde. DAV’s refereren bovendien aan acties die visueel direct voorstelbaar zijn.

Om te bepalen hoe concreet een tweet in het corpus was, zijn het aantal woorden dat geannoteerd was geteld. Dit is op eenzelfde manier gedaan zoals in het LCM wordt beschreven. Elk type woord kreeg een waarde: BN (4), SAV (3), IAV (2), DAV (1). Deze waarden werden per tweet bij elkaar opgeteld en gedeeld door het aantal woorden dat per tweet geannoteerd is. Een tweet waarin 2 BN en 1 DAV zat kreeg bijvoorbeeld de abstractiewaarde  $\underline{3}$   $((2 \times 4 + 1 \times 1)/3 = 3)$ . Hoe hoger de score van een tweet hoe hoger het abstractieniveau.

Werkwoordtijden zijn, net zoals bijvoegelijke naamwoorden, geannoteerd door gebruik te maken van de Stanford Part Of Speech tagger (SNLP, n.d.). Omdat tweets vaak taalfouten bevatten en omdat de tagger hoe dan ook fouten maakt, moet hier rekening mee gehouden worden. In het huidige onderzoek zijn de fouten die de tagger maakt meegenomen in het onderzoek.

### Statistische toetsing

Om te bepalen op welke indicatoren directe en indirecte tweets van elkaar verschilden zijn er twee verschillende statistische toetsen gebruikt. Namelijk t-toetsen en een Chi-square toets. De Cohen's Kappa is gebruikt om de interbeoordelaarsbetrouwbaarheid van de concreetheidschaal en de directheid van tweets te meten.

## Resultaten

### Resultaten kleine steekproef

De kleine steekproef ( $n = 271$ ) is geannoteerd om te onderzoeken of indirecte tweets abstracter zijn dan directe tweets. De abstractiescore verschilde niet tussen de twee groepen (indirect en direct). Uit een t-toets voor Abstractie per tweet met als factor Directheid van de tweet bleek er geen significant verschil te zijn tussen directe en indirecte tweets ( $t(269) = .03, p = .973$ ). Abstractie van tweets is om die reden niet geannoteerd in de grote steekproef.

### Resultaten grote steekproef

#### Retweets

Er is geen significant verschil gevonden tussen het Aantal retweets bij Indirecte en Directe tweets als gebruikers met meer dan 200, 500 of 1000 followers niet werden meegenomen. Als alle tweets werden meegenomen werd er wel een significant verband gevonden. Uit een t-toets voor het Aantal retweets met als factor directheid van de tweet bleek er een significant verschil te zijn tussen directe en indirecte tweets ( $t(1374) = -2.03, p = .001$ ). Indirecte tweets ( $M = 3.83, SD = 15.21$ ) bleken meer retweets te krijgen dan directe tweets ( $M = 1.56, SD = 6.67$ ).

#### URL's

Uit de  $\chi^2$ -toets tussen directheid van de tweet en de aanwezigheid van een URL bleek er een significant verband te bestaan ( $\chi^2(1) = 70.28, p < .001$ ). De gestandaardiseerde residuen in Tabel 1 suggereren dat URL's significant minder vaak voorkwamen dan verwacht bij directe tweets. Daarnaast kwamen URL's significant vaker voor bij indirecte tweets dan op basis van toeval verwacht kon worden. Bovendien kwamen directe tweets zonder URL vaker voor dan verwacht.

Tabel 1. Directheid en de aanwezigheid van URL's: kruistabel voor percentage tweets met URL's en directe/indirecte tweets, met gestandaardiseerde residuen tussen haakjes.

Direct/indirect	URL in tweet (n = 422)	Geen URL in tweet (n = 954)	Totaal
Direct	5% (-6.48)*	95% (4.31)*	100%
Indirect	35% (2.60)*	65% (-1.73)	100%

\*  $residu > 1.96$  of  $< -1.96$

### Werkwoordtijd

Uit een t-toets voor Aantal werkwoorden in de verleden tijd per tweet met als factor Directheid van de tweet bleek er een significant verschil te zijn tussen het aantal werkwoorden in de verleden tijd voor directe en indirecte tweets ( $t(1374) = 5.15, p < .001$ ). Directe tweets ( $M = 1.20, SD = 1.01$ ) bleken meer werkwoorden in de verleden tijd te bevatten dan Indirecte tweets ( $M = .83, SD = .92$ ).

Uit een t-toets voor Aantal werkwoorden in het heden per tweet met als factor Directheid van de tweet bleek er een significant verschil te zijn tussen het aantal werkwoorden in het heden voor directe en indirecte tweets ( $t(1374) = 3.20, p = .001$ ). Directe tweets ( $M = 2.12, SD = 1.62$ ) bleken meer werkwoorden in het heden te bevatten dan Indirecte tweets ( $M = 1.73, SD = 1.53$ ).

Uit een t-toets voor Aantal werkwoorden per tweet met als factor Directheid van de tweet bleek er een significant verschil te zijn tussen het aantal werkwoorden voor directe en indirecte tweets ( $t(1374) = 5.09, p < .001$ ). Directe tweets ( $M = 3.32, SD = 1.80$ ) bleken meer werkwoorden te bevatten dan Indirecte tweets ( $M = 2.56, SD = 1.93$ ).

*Tabel 2.* T-toetsen met gemiddelden en standaarddeviaties van het aantal werkwoorden in de verleden tijd, in het heden en het totaal aantal werkwoorden per tweet.

	Directe tweets		Indirecte tweets	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Verleden*	1.20	1.01	.83	.92
Heden**	2.12	1.62	1.73	1.53
Alle werkwoorden* (=heden + verleden)	3.32	1.80	2.56	1.93

\*\*  $p < .001$ , \*  $p = .001$

### Hashtags

Uit een t-toets voor het aantal hashtags per tweet met als factor directheid van tweets bleek er een significant verschil te zijn tussen directe en indirecte tweets ( $t(380.62) = -6.17, p < .001$ ). Indirecte tweets ( $M = .62, SD = 1.11$ ) bleken meer hashtags te bevatten dan directe tweets ( $M = .26, SD = .67$ ).

### Hoofdletters

Uit een t-toets voor het aantal hoofdletters per tweet met als factor directheid van tweets bleek er een significant verschil te zijn tussen directe en indirecte tweets ( $t(462.66) = -6.47, p < .001$ ). Directe tweets ( $M = 2.36, SD = 3.54$ ) bleken minder hoofdletters te bevatten dan Indirecte tweets ( $M = 4.45, SD = 6.86$ ).

## Conclusie en discussie

Dit onderzoek had als doel het beschrijven van de indicatoren van directe tweets. Er is een corpusonderzoek uitgevoerd om dit doel te bereiken. In een corpus van tweets rondom overstromingen zijn er mogelijke indicatoren automatisch en handmatig geannoteerd. Op basis van statistische toetsing kon bepaald worden wat indicatoren van directe tweets zijn.

Verwacht werd dat directe tweets minder vaak een URL bevatten dan indirecte tweets (hypothese 2). Hypothese 2 kon op grond van de resultaten bevestigd worden. Directe tweets bevatten namelijk significant minder URL's dan indirecte tweets. Daarnaast werd verwacht dat directe tweets vaker geschreven zijn in de verleden tijd dan indirecte tweets. Op basis van de resultaten kan deze verwachting (hypothese 4) deels bevestigd worden. Directe tweets bevatten namelijk meer werkwoorden in de verleden tijd dan indirecte tweets. Daarnaast bevatten directe tweets significant meer werkwoorden die refereren aan het heden. Bovendien suggereren de resultaten dat indirecte tweets in totaal meer werkwoorden bevatten dan indirecte tweets. Dit resultaat lijkt logisch aangezien directe tweets meer werkwoorden bevatten die refereren naar het heden en naar het verleden.

Verwacht werd dat directe tweets meer retweets zouden krijgen dan indirecte tweets (hypothese 1). Deze hypothese moest op grond van de resultaten worden ontkracht. Er is namelijk geen verband gevonden tussen het aantal retweets en de directheid van tweets. Er trad ook geen effect op als gebruikers met meer dan 500 volgers niet werden meegenomen in de toetsing. De verwachting dat directe tweets concretere woorden bevatten kon ook niet bevestigd worden (hypothese 3). De resultaten suggereren namelijk geen verband tussen de concreetheid van woorden en de directheid van tweets.

Naast de hypothesen, die geformuleerd zijn op basis van literatuur, zijn er ook indicatoren van directe tweets gevonden tijdens het annoteren. Directe tweets leken bijvoorbeeld minder hashtags te bevatten dan indirecte tweets. De resultaten suggereren dan ook dat directe tweets minder hastags bevatten dan indirecte tweets. Ook bleken directe tweets significant minder hoofdletters te bevatten dan indirecte tweets.

De hypothese dat retweets een indicator zouden kunnen zijn voor directe tweets kon niet bevestigd worden. Dit resultaat kan als verrassend gezien worden aangezien Starbird en Palen (2012) wel een verband vonden tussen retweets en de directheid van tweets. Een van de verklaringen voor het verschil in resultaat tussen het huidige onderzoek en het onderzoek van Starbird en Palen (2012) is dat het onderwerp van de tweets (politiek en natuurramp) en de tijdsperiode (2011 en 2015) van de gebeurtenis elkaar verschilden. Het tweetgedrag van mensen zou bij politieke gebeurtenissen mogelijk anders zijn dan bij natuurrampen. Daarnaast kan twittergedrag na verloop van tijd

veranderen. Retweets zouden, volgens de gedachtegang van Starbird en Palen (2012), een aanbevelingssysteem zijn waarin mensen als het ware een stem uitbrengen op een tweet door te retweeten. De manier waarop mensen Twitter gebruiken zou in 2015 mogelijk anders kunnen zijn waardoor er in het huidige onderzoek andere resultaten worden gevonden dan in het onderzoek van Starbird en Palen (2012). Bovendien was de steekproef bij Starbird en Palen (2012) anders geselecteerd dan in het huidige onderzoek. Starbird en Palen (2012) hebben namelijk de 1000 meest getweete tweets als steekproef gebruikt in tegenstelling tot het huidige onderzoek waarin een random steekproef is getrokken uit een grote verzameling van tweets.

De hypothese dat directe tweets concretere woorden bevatten dan indirecte tweets kon niet bevestigd worden. Uit een experiment van Fajita et al. (2006) bleek dat proefpersonen met een kleine ruimtelijke afstand tot een gebeurtenis concretere woorden gebruikten dan mensen met een grotere ruimtelijke afstand. Om die werd in het huidige onderzoek verwacht dat directe tweets concretere woorden zou bevatten dan indirecte tweets. Een verklaring zou kunnen zijn dat mensen waarschijnlijk via beelden (video/foto) op de hoogte gebracht van de overstromingen, vooral als mensen de overstroming niet zelf meemaken. Iemand die op die manier het nieuws over de overstromingen tot zich neemt heeft een concreet beeld van de overstromingen maar schrijft geen directe tweets omdat die persoon waarschijnlijk geen persoonlijke ervaring heeft met de overstromingen en dus geen directe tweet kan schrijven. De mentale representatie kan door beelden in de vorm van foto's of video's dus gelijk zijn tussen mensen die de overstromingen wel en niet hebben meegemaakt. Dit zou een verklaring kunnen zijn dat de concreetheid van indirecte tweets niet significant verschilt van directe tweets. Een andere verklaring voor dit resultaat kan zijn de manier waarop concreetheid van woorden is geoperationaliseerd. Er zijn in het huidige onderzoek verschillende woordklassen geannoteerd (SAV, DAV, IAV en BN) terwijl er ook andere methoden zijn om concreetheid van woorden te toetsen.

De verwachting dat URL's minder vaak voor zouden komen bij directe tweets kon bevestigd worden. 5% van de directe tweets bevatten namelijk een URL en 35% van de indirecte tweets bevatten een URL. Mensen die een persoonlijke ervaring delen over de overstroming gebruiken dus minder vaak een URL die meestal doorverwijst naar tweedehands informatie. Dit lijkt ook logisch aangezien directe tweets een eigen ervaring beschrijven rondom de overstromingen, een URL kan in een directe tweet overbodig zijn.

De hypothese dat directe tweets meer werkwoorden bevatten die refereren aan de verleden tijd dan aan het heden kon deels worden bevestigd. Indirecte tweets bevatten inderdaad vaker werkwoorden die refereerden aan de verleden tijd maar dat gold ook voor werkwoorden die refereren aan het heden. Concluderend bevatten directe tweets meer werkwoorden die refereren aan het heden en het verleden dan indirecte tweets. Een mogelijke verklaring voor dit resultaat zou



kunnen zijn dat de tijd van een werkwoord refereert aan een bepaalde tijd en directe tweets refereren aan een bepaalde gebeurtenis. Een gebeurtenis vindt altijd plaats op een bepaald moment in de tijd waardoor een directe tweet meer werkwoorden zou kunnen bevatten.

Concluderend zou op basis van de resultaten gezegd kunnen worden dat directe tweets verschillende indicatoren hebben. Directe tweets bevatten namelijk: minder URL's, meer werkwoorden, minder hashtags en minder hoofdletters te bevatten dan indirecte tweets. Aangezien deze resultaten zo algemeen zijn voor tweets zou het mogelijk zijn dat deze indicatoren ook van toepassing zijn op andere rampen dan overstromingen.

In het huidige onderzoek is gebruik gemaakt van een part of speech tagger die automatisch woorden classificeert. Werkwoorden en bijvoeglijke naamwoorden zijn met deze tagger automatisch geclassificeerd. Bij het interpreteren van de resultaten moet rekening gehouden worden met het feit dat taggers fouten maken. Een tagger maakt vanzelfsprekend meer fouten bij tekst met spelfouten, zoals tweets. Bovendien is het aantal directe tweets ( $n = 191$ ) ten opzichte van het aantal indirecte tweets ( $n = 1185$ ) erg klein. Als een woord of indicator per toeval vaker voorkomt bij indirecte tweets is de kans dat er een significant verband gevonden wordt groter. Bij het interpreteren van de resultaten moet daarom rekening gehouden worden met de scheve klassenverdeling tussen directe en indirecte tweets.

Ondanks de kleinschalige casestudy zijn de resultaten nuttig te noemen. In dit paper worden namelijk verschillende aanwijzingen gegeven voor mogelijke indicatoren van directe tweets. Met deze aanwijzingen kunnen mogelijk automatische systemen ontwikkeld worden om directe tweets tijdens een natuurramp te kunnen herkennen. Dit is van groot belang omdat natuurrampen slachtoffers en financiële schade kunnen opleveren. De detectie van directe tweets kunnen deze schade mogelijk tegengaan omdat directe tweets real-time en betrouwbare informatie kunnen verschaffen aan crisisstakeholders zoals journalisten, hulpdiensten en burgers.

Het huidige onderzoek heeft niet alleen praktische relevantie voor crisisstakeholders maar ook voor onderzoek naar automatische inhoudsanalyse naar tweets rondom natuurrampen. Er worden in dit paper namelijk verschillende indicatoren beschreven van indirecte tweets. Hierdoor wordt er binnen onderzoek naar indirecte tweets specifiek beschreven wat de indicatoren van directe tweets precies zijn. Dit is bij de automatische classificatie van tweets niet het geval.

Het zou nuttig zijn om in een vervolgonderzoek de indicatoren die naar voren kwamen op basis van het huidige onderzoek te testen op een andere natuurramp. Hierdoor kan getoetst worden of de indicatoren die zijn gevonden in het huidige onderzoek ook van toepassing zijn op andere rampen of andere tijdsperiodes. Daarnaast zou een onderzoek uitgevoerd kunnen worden om te toetsen of de indicatoren die gevonden zijn in het huidige onderzoek gegeneraliseerd kunnen worden naar anderen typen rampen zoals een terroristische aanslag.

## Literatuur

- Bos, L.S., Dragoy, O., Stowe, L.A., & Bastiaanse, R. (2013). Time reference teased apart from tense: Thinking beyond the present. *Journal of Neurolinguistics*, 26, 283-297.
- Boyd, D., Golder, S., & Lotan, G. (2010). *Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter*. In Proceedings of the 43rd Hawaii International Conference on System Sciences.
- Castillo, C., Mendoza, M., & Poblete, B. (2012). Predicting information Credibility in time-sensitive social media. *Internet Research*, 5(23), 560-588.
- Coenen, L. H. M., Hedeboom, L., & Semin, G. R. (2006). *Measuring language abstraction: The linguistic category manual (LCM)*. Free University: Amsterdam.
- Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one*, 5(11).
- Fujita, K., Henderson, M. D., Eng, J., Trope, Y., & Liberman, N. (2006). Spatial distance and mental construal of social events. *Psychological Science*, 17(4), 278-282.
- Gupta, A., Lamba, H., & Kumaraguru, P. (2013). *\$1.00 per rt# bostonmarathon# prayforboston: Analyzing fake content on Twitter*. In eCrime Researchers Summit (eCRS), 1-12.
- Helsloot, I., & Groenendaal, J. (2013). Twitter: An Underutilized Potential during Sudden Crises? *Journal of Contingencies and Crisis Management*, 3(21), 178-183.
- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261–266.
- Hu, M., Liu, S., Wei, F., Wu, Y., Stasko, J., & Ma, K. (2012). Breaking News on Twitter. In: CHI, 2751-2754.
- Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., & Meier, Patrick. (2013). *Extracting Information Nuggets from Disaster-Related Messages in Social Media*. In Proceedings of the 10th International ISCRAM Conference.
- Mendoza, M., Poblete, B., & Castillo, C. (2010). *Twitter under crisis: Can we trust what we RT?* Proceedings of the first workshop on social media analytics, 71-79.
- Nguyen, M., Kitamoto, A., & Nguyen, T. (2015). *TSum4act: A Framework for Retrieving and Summarizing Actionable Tweets during a Disaster for Reaction*. In: Advances in Knowledge Discovery and Data Mining, 19<sup>th</sup> Pacific-Asia Conference, 64-75.
- North-West Evening Mail (2015). *TIMELINE: How Storm Desmond disaster unfolded*. Geraadpleegd op 2 maart 2016 van <http://www.nwemail.co.uk/TIMELINE-How-Storm-Desmond-disaster-unfolded-34017ac6-af2e-4248-85b0-b01fcac74579-ds>

- Pennebaker, J., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54, 47–77.
- Ross, J.W. (2013). *Features for Ranking Tweets Based on Credibility and Newsworthiness*, MA Thesis, Computer Science and Engineering, Wright State University, Dayton.
- Signorini, A., Segre A.M., & Polgreen, P.M. (2011). The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza AH1N1 Pandemic. *PLoS ONE* 6(5).
- SNLP (n.d.). Stanford Log-linear Part-Of-Speech Tagger. Opgevraagd op 23 april 2016, van: <http://nlp.stanford.edu/software/tagger.html>.
- Starbird, K., & Palen, L. (2012). *(How) Will the Revolution be Retweeted?: Information Propagation in the 2011 Egyptian Uprising*. In Proceedings of the 2012 Conference on Computer Supported Cooperative Work (CSCW), TBA. Bellevue, WA: ACM.
- Starbird, K., Muzny, G., & Palen, L. (2012). *Learning from the Crowd: Collaborative Filtering Techniques for Identifying On-the-Ground Twitterers during Mass Disruptions*. In Proceedings of the 9th International ISCRAM Conference.
- Starbird, K., Palen, L., Hughes, A. & Vieweg, S. (2010). Chatter on The Red: What hazards threat reveals about the social life of microblogged information. Proc of CSCW 2010. ACM, 241-250.
- Takahashi, B., Tandroc Jr., E.C., & Carmichael, C. (2015). Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines. *Computers in Human Behavior*, 50, 392-398.
- Tanaka, Y., Sakamoto., Y. & Honda., H. (2014). The Impact of Posting URLs in Disaster . related Tweets on Rumor Spreading Behavior. *Conference on System Sciences*, 520-529.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.
- Trope, Y., Liberman, N., & Wakslak, C. (2007). Construal Levels and Psychological Distance: Effects on Representation, Prediction, Evaluation, and Behavior. *Journal of Consumer Psychology*, 17(2), 83-95.
- Truong, B., Caragea, C., Squicciarini, A., & Tapia, A.H. (2014). *Identifying Valuable Information from Twitter During Natural Disasters*. 77<sup>th</sup> ASIS&T Annual Meeting.
- Verma, S., Vieweg, S., Corvey, W.J., Palen, L., Martin, J.H., Palmer, M., Schram, A. & Anderson, K.M. (2011). *Natural Language Processing to the Rescue?: Extracting "Situational Awareness" Tweets During Mass Emergency*. In: Fifth International AAAI Conference on Weblogs and Social Media.

## Bijlage 1: Woordenlijst tweets

Flood  
Floods  
Flooding  
Flooded  
Inundation  
Inundations  
Inundated  
Landslide  
Dam break  
Dam burst  
Dam bursting  
Dam breached  
Dam fail  
Dam failed  
Dam failing  
Dam failure  
Dam broken  
Dam collapse  
Dyke break  
Dyke burst  
Dyke bursting  
Dyke breached  
Dyke fall  
Dyke failing  
Dyke falling  
Dyke failure  
Dyke broken  
Dyke collapse  
Embankment break  
Embankment burst  
Embankment bursting  
Embankment breached  
Embankment fail  
Embankment failing  
Embankment failure  
Embankment broken  
Embankment collapse