Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie
ISSN: (Online) 2222-4173, (Print) 0254-3486

Oopbronintelligensie (OSINT) vir veiligheids-doeleindes: Die ontwikkeling van ’n data-ontledingspyplyn om relevante WhatsApp-boodskappe te ontleed

Outeurs:
Burgert A Senekal,
Eduan Kotzé

Affiliasie:
Universiteit van die Vrystaat,
Posbus 339, Bloemfontein,
9300

Korresponderende outeur:
Burgert Senekal
E-pos:
burgertsenekal@yahoo.co.uk

Datums:
Ontvang: 31/01/19
Aanvaar: 22/05/19
Gepubliseer: 12/06/19

Hoe om hierdie artikel aan te haal:
Burgert A Senekal, Eduan Kotzé, Oopbron-intelligensie (OSINT) vir veiligheidsdoeleindes: Die ontwikkeling van ’n data-ontledingspyplyn om relevante WhatsApp-boodskappe te ontleed, *Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie* 38(1) (2019). https://doi.org/10.36303/SATNT.2019.38.17.717

Oopbronintelligensie het in die laaste paar dekades die dominante intelligensiedissipline geword. Die inligtingsontploffing het egter daartoe gelei dat geoutomatiseerde metodes aangewend moet word om inligting uit groot hoeveelhede data, wat gereeld in ’n ongestruktureerde formaat bestaan en deurlopend gegeneer word, te onttrek. Boonop het nuwe inligtingskanale soos sosiale media ontstaan wat beteken dat nie slegs die hoofstroommedia gemonitor kan word om gebeurtenisse te identifiseer nie. Hierdie artikel bespreek hoe ’n dataversamelings en -ontledingspyplyn saamgestel kan word om inligting uit WhatsApp boodskappe, wat op groepe met hul eie sienings van veiligheidsbegrippe of probleme voorkom, en onder meer verwys na gewelddadige betogings, plaasaanvalle, grondgrype en ander misdaad in Suid-Afrika, te onttrek. Daar word bespreek hoe teks skoongemaak word en gebeurtenisse en plekname onttrek word en dan outomaties na ’n visuele en interaktiewe gebruikerskoppelvlak uitgevoer word om die voorkoms van hierdie gebeurtenisse in kleiner areas of sektore en later landswyd te kan monitor. Uitdagings en probleme word bespreek, sowel as verdere navorsingsgeleenthede, wat insluit om so ’n stelsel met veiligheidheidsreaksiemagte soos die Polisiediens se databasisse te integreer ten einde waardevolle reaksie vir die gemeenskap ten opsigte van veiligheid te weeg te bring.

Kernwoorde: oopbronintelligensie, OSINT, WhatsApp, grootdata, datapyplyn, protes, Suid-Afrika

Open Source Intelligence (OSINT) for security purposes: Developing a data analysis pipeline to analyse relevant WhatsApp messages: Open source intelligence has become the dominant intelligence discipline over the past few decades. However, the information explosion has led to a need for automated methods to be used to extract information from large amounts of data, which is often in an unstructured format and continuously generated. In addition, new information channels such as social media have emerged which mean that not only the mainstream media can be monitored to identify events. This article discusses how a data collection and analysis pipeline can be compiled for extracting information from WhatsApp messages, which appear on groups that focus on security concepts or problems, and refer, inter alia, to violent protests, farm attacks, land grabs and other crime in South Africa. It discusses how to clean text and extract events and place names, and then automatically exports the results to a visual and interactive user interface to monitor the occurrence of these events in smaller areas or sectors and later nationwide. Challenges and problems are discussed, as well as further research opportunities, which include integrating such a system with security response forces such as the Police Service’s databases in order to provide valuable response for the community in terms of safety.

Key words: open source intelligence, OSINT, WhatsApp, big data, data pipeline, protest, South Africa

Inleiding

Oopbronintelligensie ("Open Source Intelligence" of OSINT) verwys na die versameling en ontleding van bronne wat in die openbare domein is, byvoorbeeld nuusberigte, webblaaie en navorsingspublikasies, en die aanwending daarvan vir intelligensiedoeleindes. Meer formeel word dit gedefinieer as "publicly available information that has been discovered, determined to be of intelligence value, and disseminated by a member of the IC [Intelligence Community]" (Williams & Blum, 2018:8). Dié dissipline het in die laaste paar dekades die dominante intelligensiedissipline geword (Arslan & Yanık, 2015) (Gibson, 2014) (Hobbs, et al., 2014) (Brandt, et al., 2011), tot 'n groot mate as gevolg van die inligtingsontploffing wat met die Digitale Era gepaard gaan. Die VSA se Central Intelligence Agency (CIA) (Central Intelligence Agency, 2010) voer aan: "OSINT has always been an important part of all-source analysis, but continuing advances in information technology have given a voice to even larger numbers of people and made it possible to address new intelligence questions."

Die inligtingsontploffing sedert die negentigerjare het daartoe gelei dat geoutomatiseerde metodes aangewend moet word om inligting te versamel, te ontleed en te versprei, aangesien die hoeveelheid inligting wat beskikbaar is, sulke groot volumes opneem dat dit nie deur die mens ontleed kan word nie (Williams & Blum, 2018). Daar word byvoorbeeld daagliks omtrent 500 miljoen twiets op Twitter geplaas, 50 miljoen foto's op Instagram, 3 miljard video's word op YouTube gekyk en 3 miljoen blog artikels word geskryf (Internet Live Stats, 2019). Ten einde inligting uit sulke groot volumes te kan onttrek, moet daar gebruik gemaak word van Natuurlike Taalverwerkingstegnieke (Natural Language Processing of NLP), masjienleer, neurale netwerke en dergelike. Die gebruik van hierdie tegnieke plaas huidige oopbronintelligensie duidelik binne die grootdataparadigma (Dencik, et al., 2017) (Williams & Blum, 2018).

Daar is reeds verskeie inisiatiewe oorsee geloods om inligting rekenaarmatig uit nuusbronne te onttrek en vir intelligensiedoeleindes aan te wend. Voorbeelde sluit in Bueno de Mesquitase se Policon model (Bueno de Mesquita, 1981) (Bueno de Mesquita, et al., 1985) (Bueno de Mesquita & Stockman, 1994), wat deur die CIA gebruik is (O'Brien, 2010) en Senturion (Abdollahian, et al., 2006), wat deur die VSA se Departement van Verdediging gebruik is (O'Brien, 2010). Lockheed Martin se Integrated Crisis Early Warning System (ICEWS) onttrek gestruktureerde data uit nuusberigte en het ten doel: "to develop a comprehensive, integrated, automated, generalizable, and validated system to monitor, assess, and forecast national, sub-national, and international crises in a way that supports decisions on how to allocate resources to mitigate them" (O'Brien, 2010:88). Ander soortgelyke projekte sluit in die CIA se Political Instability Task Force (PITF) (Esty, et al., 1995) (Esty, et al., 1998) (Bates, et al., 2003) (Goldstone, et al., 2010), die Armed Conflict Location and Event Data (ACLED) projek (Raleigh, et al., 2010), die Uppsala Conflict Data Program (UCDP) (Sundberg & Melander, 2013) (Croicu & Sundberg, 2017), en die Social Conflict Analysis Database (SCAD) (Salehyan, et al., 2012) (sien Schrodt, 2012 vir 'n bespreking van soortgelyke projekte). Privaat Militêre Maatskappye soos CACI Inc. en Third Point Systems het ook reeds data gebruik wat deur stelsels soos ICEWS gegeneer is (Gerner, et al., 1994).

Toe die dissipline van oopbronintelligensie tydens die Tweede Wêreldoorlog ontwikkel is, was die aanvanklike bronne van inligting hoofsaaklik radio uitsendings, maar teen die sestigerjare het die fokus na die gedrukte media verskuif. Sedert die koms van die wêreldwye web in die negentigerjare het die klem na die web verskuif, maar sosiale media kanale het die inligtingslandskap verder verander (Williams & Blum, 2018). Mense maak nie meer uitsluitlik staat op die hoofstroom media om inligting te bekom nie, en platforms soos Facebook en Twitter het belangriker bronne van nuus geword. Onlangs het WhatsApp ook in belangrikheid toegeneem wanneer dit by nuusbronne kom en 'n studie van die Reuters Institute het bevind dat mense toenemend hulle nuus eerder van WhatsApp as van Facebook kry (Anoniem, 2018). In baie lande het WhatsApp se gebruik vir nuus amper verdriedubbel (Anoniem, 2018). Ook in Suid-Afrika word WhatsApp gereeld gebruik om inligting rakende noodgevalle te deel (Anoniem, 2015). Wanneer 'n mens 'n omvattende beeld wil opbou oor wat in Suid-Afrika gebeur, is dit met ander woorde belangrik om so veel as moontlik inligtingskanale te betrek, wat WhatsApp en ander sosiale media kanale moet insluit.

Die huidige ondersoek ontleed die boodskappe wat op verskillende WhatsApp groepe geplaas word wat in veiligheidskwessies in Suid-Afrika belangstel. Dié groepe deel hoofsaaklik inligting rakende protesoptrede, plaasaanvalle, grondgrype en ander ernstige misdaad. Die doel van die huidige projek is om 'n dataversamelings en -ontledingspyplyn te ontwikkel wat inligting uit ongestruktureerde WhatsApp boodskappe onttrek en gebeurtenisse intyds ontleed en visualiseer ten einde onveilige insidente in Suid-Afrika in 'n gebruikervriendelike formaat te kan monitor.

WhatsApp as inligtingskanaal

Die projekte wat in die inleiding genoem is, byvoorbeeld ICEWS en ACLED, maak gebruik van stelsels wat inligting van hoofstroom nuusagentskappe versamel. Alhoewel hierdie stelsels op 'n globale skaal met vrug aangewend kan word om konflik te monitor, is dit onses insiens nie die mees geskikte platform vir die monitering van die geweldsituasie in Suid-Afrika nie, maar eerder WhatsApp omdat dit 'n aantal beduidende voordele bo die hoofstroom media het.

Eerstens word alle gebeurtenisse nie deur die hoofnuusagentskappe gedeel nie, aangesien ruimte en hulpbronne hulle verplig om op groter gebeure te fokus (Gerner, et al., 1994). Navorsing deur Hendrix en Salehyan (2015) het

byvoorbeeld aangetoon dat die hoofstroom media meer geneig is om oor 'n gebeurtenis te rapporteer indien meer mense en meer sterftes betrokke is. Daar kan nie van nuusagentskappe soos Netwerk24 verwag word om elke protesaksie in Suid-Afrika te dek nie, maar dit is juis wat nodig is vir 'n veiligheidstelsel. WhatsApp laat gebruikers toe om self inligting te deel en ook berigte van plaaslike nuusagentskappe te deel, wat beteken dat ook kleiner protesaksies ingesluit word.

WhatsApp het verder die voordeel dat inligting onmiddellik gedeel kan word. Anders as hoofstroom nuuskanale hoef gebruikers nie te wag vir kopieskrywers, redakteurs en teksversorgers voor inligting gedeel word nie. Dit bring mee dat inligting vinniger op WhatsApp beskikbaar is as wat die geval is met hoofstroom mediastelsels. Hierdie onmiddellike beskikbaarstelling van inligting het van WhatsApp 'n belangrike inligtingskanaal gemaak, nie alleen vir persoonlike kommunikasie of binne 'n werksopset nie, maar ook vir noodgevalle (Malik, 2016) (Debnath, et al., 2016) (Mazzarella, 2016) (Moreno, et al., 2017). Gemeenskapspolisiëringsforums landswyd gebruik WhatsApp, byvoorbeeld in Johannesburg (Anoniem, 2015), Stellenbosch (Anoniem, 2017), Bloemfontein (Damons, 2017), Kroonstad (Hurter, 2018) en Port Alfred (Knowles, 2018), terwyl WhatsApp ook in landelike gebiede gebruik word om plaasaanvalle te bekamp (Smith, 2014) (Bruwer, 2015) (Cronjé, 2016) (Damons, 2017), veediefstal teë te werk (Anoniem, 2016b) en ook vir noodbystand op paaie (Waterworth, 2017). Boonop gebruik instansies soos die Nasionale Seereddingsinstituut (NSRI) ook WhatsApp om inligting rakende noodgevalle te versprei (Rodgers, 2018). Nieregeringsinstansies gebruik ook WhatsApp om nuus te deel, byvoorbeeld AfriForum (AfriForum, 2017) (Damons, 2017) (Anoniem, 2016a) en die Suidlanders (Suidlanders, 2016). Privaat veiligheidsmaatskappye soos Willshir & Associates en N-ERT gebruik ook WhatsApp om inligting te kommunikeer, asook mediese maatskappye soos Crisis Medical. Verskeie Suid-Afrikaanse nuusagentskappe, soos Eye Witness News, North Coast Courier en Maroela Media, gebruik ook WhatsApp om nuusberigte te versprei, terwyl die meeste nuusagentskappe WhatsApp gebruik om inligting te versamel.

WhatsApp het verder die voordeel dat dié platform gebruikers toelaat om nuusberigte te deel, van nuusagentskappe soos Netwerk24 of Maroela Media, sosiale media platforms soos Twitter en Facebook, asook van ander WhatsApp groepe. Boonop stel WhatsApp mense in staat om multimedia te deel: benewens boodskappe en skakels is daar ook 'n groot hoeveelheid foto's, video's en stemboodskappe wat op hierdie platform versprei word. Dít bring mee dat WhatsApp as 'n versamelplatform gebruik kan word wat inligting uit verskeie oorde versamel.

Die feit dat WhatsApp ook die boodskaptoepassing is wat die meeste in Suid-Afrika gebruik word, beteken dat dié platform gebruik kan word om netwerke van bydraers met mekaar te verbind en die deel van inligting 'n samewerkingspoging te maak. WhatsApp is die populêrste boodskaptoepassing in Afrika (Dahir, 2018) en die meerderheid data wat in Afrika deur die Internet vloei, vloei deur WhatsApp (Thompson, 2018). In Suid-Afrika is WhatsApp een van die gewildste platforms om inligting te deel en gewoonlik een van die eerste toepassings wat mense op hul slimfone installeer (Shapshak, 2015). Volgens Arthur Goldstuck, die besturende direkteur van World Wide Worx, domineer WhatsApp die mark in Suid-Afrika met byna 20 miljoen gebruikers, ten spyte van kompetisie van veral Telegram: "WhatsApp has almost become the default messaging app for most South Africans" (Rawlins, 2018).

Daar bestaan egter 'n aantal uitdagings daarmee om inligting uit WhatsApp boodskappe te onttrek. Eerstens is WhatsApp, anders as byvoorbeeld Twitter, 'n geslote platform waartoe 'n mens eers toegang moet kry. Dít beteken dat daar nie so 'n groot verskeidenheid bronne gemonitor kan word soos wat byvoorbeeld die geval met Twitter is nie. Vir die stelsel wat hier bespreek word, word daar slegs 16 groepe se boodskappe gebruik, waarop die outeur(s) eers geregistreer moes word.

Tweedens is daar etiese kwessies by betrokke wanneer persoonlike boodskappe ter sprake is. Om etiese redes kan die groepe se name, lede en stigters nie hier vermeld word nie.

Derdens is daar altyd vrae oor geldigheid wanneer gebruikers self inligting kan deel. Die groepe wat hier ingesluit word, is egter oor 'n tydperk van ten minste drie maande gemonitor en daar is nog geen valse nuus per groepsdefinisies hierdeur versprei nie.

Vierdens word boodskappe nie met 'n bepaalde struktuur geplaas nie: sommige boodskappe verwys na meer as een plek of gebeurtenis, gebruik nie leestekens of spellings konsekwent nie, en gebruik verskillende terme om na 'n gebeurtenis te verwys. WhatsApp boodskappe is met ander woord "wild" in die sin daarvan dat dit taal is wat natuurlik voorkom – taal soos mense dit gebruik. Dit beteken dat boodskappe eers skoongemaak moet word voor ontleding. Meer hieroor in die afdeling data voor- en naverwerking.

Gebeurtenisdata

Projekte soos die ICEWS genereer sogenaamde gebeurtenisdata ("event data"), wat in eenvoudige terme beteken wie doen wat aan wie, waar, wanneer en hoekom. Azar en Ben-Dak (1975:1) definieer 'n gebeurtenis as:

> ... some activity undertaken by an international actor (a nation-state, a major subunit of a nation-state, or international organization) ... at a specific time and which is directed toward another actor for the purposes of conveying interest (even non-interest) in some issue. Thus, an event involves (1) an actor, (2) a target, (3) a time period, (4) an activity, and (5) an issue about which the activity revolves.

Hierdie definisie is egter tydens die Koueoorlog geformuleer toe die politieke landskap heelwat anders daar

uitgesien het. Tydens die Koueoorlog is konflik gekenmerk deur die betrokkenheid van internasionale moondhede by oorloë in ontwikkelende lande, byvoorbeeld die vele konflikte in Suider-Afrika. Ná die Koueoorlog het konflikte egter versplinter in wat Van Creveld (1991) (2008) *lae-intensiteit konflik* noem, Münkler (2005) en Kaldor (2006) *nuwe oorloë*, of Lind, Schmitt en Wilson (Lind, et al., 1989) *vierde generasie oorlogvoering* noem. Dié teoretici se onderskeie definisies verskil, maar in breë trekke word *lae-intensiteit konflik*, *nuwe oorloë* en *vierde generasie oorlogvoering* gekenmerk deur 'n vervaging van verskeie grense: tussen oorlog en vrede, die burgerlike en die militêre, die interne en die eksterne, en tussen misdaad en oorlogvoering. Verder speel identiteitspolitiek 'n groter rol wanneer die teiken 'n kultuur self kan wees.

Azar en Ben-Dak se bostaande definisie ondervang met ander woorde nie meer belangrike gebeurtenisse in die hedendaagse konflikmilieu nie, omdat substaatrolspelers en veral burgerlikes meer belangrik geword het. Verder is die akteur nie altyd identifiseerbaar nie, byvoorbeeld in gewelddadige betogings, en soms is die teiken en die kwessies ter sprake ook onduidelik.

Omdat die konsep van gebeurtenisdata verouderd is in die Suid-Afrikaanse konteks waar byvoorbeeld misdaad en protesaksies 'n veel groter veiligheidsrisiko inhou as interstaatoorlogvoering, pas ons die konsep van 'n gebeurtenis aan om te beteken: iets gebeur iewers op 'n gegewe dag. Hierdie eenvoudige omskrywing beteken dat 'n handeling onttrek word (protes, misdaad, plaasaanval of onwettige grondonteiening), 'n plek en 'n datum. Die rolspelers betrokke is gereeld onduidelik of onbekend, byvoorbeeld in onwettige protesaksies of ander misdaad, en daarom word rolspelers nie geïdentifiseer nie.

Die dataversameling en -ontledingspyplyn

Dataverwerking ("data processing") bestaan gewoonlik uit 'n reeks bedrywighede wat op data uitgevoer word, meestal deur 'n rekenaar, om data te onttrek en te ontleed. Dataverwerking word al vir 'n geruime tyd in die veld van datapakhuise en besigheidsintelligensie gebruik en is deel van die spreekwoordelike ETL-proses ("Extract Transform and Load") om data te onttrek, te verwerk en in 'n datapakhuis te laai (Dayal, 2009). Met die verskuiwing van tradisionele relasionele databronne na grootdatabronne wat meestal semi- of ongestruktureerd is, word die term *dataverwerking* hedendaags met bondel- en intydse verwerking geassosieer (Hashem, 2015). Hierdie verwerkingsmetodes sluit gewoonlik 'n dataversameling- en ontledingspyplyn in. 'n Tipiese dataverwerkingspyplyn behels die volgende stappe: dataskepping, dataverkryging (wat insluit data-voorverwerking en dataversameling), databerging en data-ontleding (Taleb & Serhani, 2015). Data voorverwerking ("data preprocessing") is 'n kritieke komponent en kan nie afgeskeep word nie, aangesien dié stap datakwaliteit moet verseker. Vir hierdie studie word na hierdie stap verwys as teks-voorverwerking, aangesien ons databronne uit natuurlike taal (teks) bestaan en nie 'n tradisionele databasis nie.

Die res van hierdie afdeling bespreek hoe die datapyplyn saamgestel is.

Dataskepping

Gebruikers plaas boodskappe, stuur boodskappe van ander groepe aan en deel nuusberigte van 'n verskeidenheid platforms soos Facebook en Twitter op WhatsApp groepe, wat dan uitgevoer word as 'n plein tekslêer. Dié lêer word vanaf 'n iPhone na iCloud gestoor, wat sinchroniseer met 'n MacBook. 'n Mens hoef met ander woorde nie fisies in 'n kantoor te wees om boodskappe tot die stelsel by te dra nie; daar is vantevore al toevoegings vanuit die buiteland gemaak. Daar is ook besluit om eerder van 'n MacBook as 'n rekenaar of iMac gebruik te maak omdat dit beskerming teen kragonderbrekings bied. Sodra die lêer gesinchroniseer het, word dit outomaties na 'n vouer in Google Drive gekopieer. Die rede hiervoor is dat Google Drive gebruikers toelaat om vouers te deel, wat iCloud nie toelaat nie, maar iCloud se integrasie met 'n iPhone is beter as wat die Google Drive toepassing se integrasie met die iPhone is.

Dataverkryging

Teks-voorverwerking

'n Python 3.7 toepassing is ontwikkel om die teks-voorverwerking te hanteer. Volgens Dey en Haque (2009) en Palmer (2012) is teks-voorverwerking noodsaaklik op enige sosiale media databron aangesien dit baie ruis bevat. Om datakwaliteit te handhaaf, is die toepassing opgestel op 'n bediener wat gekoppel is aan ononderbroke kragtoevoer en 'n kragopweker. Hierdie dienste word verskaf deur ons Universiteit se IKT-datasentrum.

Volgens Jianqiang en Xiaolin (2017) bestaan 'n effektiewe teks-voorverwerking gewoonlik uit teksreiniging, tekseenheididentifisering ("tokenisation"), sowel as sintaktiese ontleding ("syntatic parsing"). Die volgende paragrawe bespreek die stappe wat ons as deel van teks-voorverwerking gevolg het.

1. WhatsApp data word ingelees en boodskappe word geïdentifiseer deur middel van gereelde uitdrukkings. Die formaat van WhatsApp boodskappe is soos volg: [datum, tyd] +[Van Wie]: [Boodskap]. Ten einde karakterstelomskakelingsfoute te vermy, is UTF-8 lêer enkodering gebruik om die datalêers in te lees, en sodra al die data ingelees is, die teks om te skakel na 'n Unicode karakterstel.
2. Boodskappe verdeling: Die ingeleesde boodskappe het heelwat bondel-boodskappe bevat. 'n Bondelboodskap is 'n WhatsApp boodskap wat linguisties uit verskillende sinne bestaan, en waar elke sin 'n enkele WhatsApp boodskap verteenwoordig. Dit is belangrik om hierdie boodskappe te verdeel as enkel (of verskillende) boodskappe voordat dit verder gebruik kan word, omdat sulke boodskappe na verskillende gebeurtenisse en plekke verwys

het. Python-gereelde uitdrukkings tesame met lusse (FOR-LOOP) is gebruik om 'n bondel-boodskap te identifiseer en te verdeel.

3. Teksreiniging ("text cleaning"): teks is skoongemaak van alle spesiale karakters, punte en kommas. Dit sluit emotikons ("emoticons") in aangesien hierdie voorstelling van gesigsuitdrukkings nie 'n rol gaan speel in die klassifikasie nie en ook nie nuttig is as deel van die woordeskat nie. Verder is alle "http", "https" en "www" van webadresse verwyder maar nie die hiperskakel self nie aangesien dit nuttig kan wees om 'n gebeurtenis te identifiseer (byvoorbeeld: https://citizen.co.za/uncategorized/2037294/mpumalanga-businessman-narrowly-escapes-death-after-being-shot). Kontraksies is ook geïdentifiseer deur middel van gereelde uitdrukkings en vervang met die korrekte tekseenheid uit die woordeboek (byvoorbeeld "ek't" is omgeskakel na "ek het" en "isn't" na "is not"). Koppeltekenwoorde is ook korrek bymekaar gevoeg, aangesien koppeltekens ("-") wat 'n spesiale karakter is, verwyder was tydens teksreiniging. Weereens is daar van 'n woordeboek gebruik gemaak om die koppeltekenwoorde saam te stel (byvoorbeeld "no go" is omgeskakel na "no-go" en "hi jacking" na "hi-jacking"). Laastens is informele taalgebruik en afkortings reggestel met die gepaste woord of woorde (byvoorbeeld "kzn" word "KwaZulu-Natal" en "gonna" word "going to").
4. Tekseenheididentifisering ("tokenisation"): Nadat die teks skoongemaak is, is sin en woord tekseenheididentifisering toegepas. Tekseenheididentifisering is 'n linguistiese verwerkingsproses waartydens woorde en leestekens van mekaar verdeel word om 'n korpus van tekseenhede te vorm. Aangesien leestekens reeds in 'n vorige stap verwyder is, was dit net nodig om woorde te verdeel en ekstra spasies te verwyder. Ten einde te verseker dat net woorde oorbly, is daar getoets of die lengte van 'n woord groter as 1 karakter is.
5. Plekname: Aangesien dit belangrik is om elke gebeurtenis te koppel aan 'n plek waar dit voorkom, is Raper, Möller en Du Plessis se *Dictionary of Southern African Place Names* (2014) gebruik om geldige plekname te identifiseer. Indien meer as een pleknaam in 'n WhatsApp-boodskap voorkom, is die pleknaam met die hoogste woordfrekwensie gekies as die plek waar 'n gebeurtenis plaasgevind het. Verder, indien daar meer as een pleknaam (byvoorbeeld "Westdene" in Johannesburg en "Westdene" in Bloemfontein) in die plekname woordeboek verskyn, is die eerste pleknaam gekies vir verdere verwerking.

Teksklassifikasie

Die doel van teksklassifikasie is om outomaties 'n etiket of kategorie aan 'n teksgedeelte (sin, paragraaf of dokument) volgens die teks se inhoud toe te ken (Manning & Schütze, 1999). Teksklassifikasie maak meestal gebruik van reëlgebaseerde en masjienleer stelsels om teks te klassifiseer (Medhat, et al., 2014). Masjienleer, 'n subveld van kunsmatige intelligensie, benodig 'n handmatig-geannoteerde korpus as opleidingsdata om 'n masjienleeralgoritme te leer wat die verskillende assosiasies tussen die teksgedeeltes en 'n uitset (byvoorbeeld etiket) is. Reëlgebaseerde stelsels maak staat op handgemaakte reëls vir teksklassifikasie en het nie soos masjienleeralgoritmes, 'n handmatige geannoteerde korpus nodig nie. Geen leer vind plaas in 'n reëlgebaseerde stelsel nie.

'n Reëlgebaseerde benadering is gevolg om die WhatsApp-boodskappe te klassifiseer volgens "veilig" en "onveilig" aangesien daar nie geannoteerde opleidingsdata beskikbaar was nie. Die reëlgebaseerde benadering het staatgemaak op vooraf opgestelde woordeboeke. Woordeboeke is saamgestel volgens die N-gram benadering wat opeenvolgende woorde in 'n lys kombineer volgens grootte *N*. Beide enkelgram ("unigrams") (1-gram) en tweegram ("bigrams") (2-gram) woordelyste is opgestel en elke woord is met 'n kategorie geassosieer. Vier "onveilige" kategorieë is gebruik vir beide enkelgram en tweegram woordelyste: protes, plaasaanval, misdaad en grondonteiening. Let daarop dat hierdie kategorieë aangepas sal moet word indien 'n mens dit met die Suid-Afrikaanse Polisiediens se databasisse wil integreer. Gewigte is toegeken aan elke kategorie en gebruik tydens teksklassifikasie. Die gewigte is gekies op grond van die erns en potensiële impak van die gebeure wat met elke kategorie geassosieer is. Sien Tabel 1 vir die kategorieë en die gewigte van "onveilige" boodskappe.

TABEL 1: Kategorieë en die gewigte van "onveilige" boodskappe

Onveilig	Gewig	Voorbeeld van enkelgram	Voorbeeld van tweegram
Misdaad	+0.75	aanval	armed,robbery
		geskiet	gun,fire
		robbed	
Protes	+0.75	betoog	burning,tyres
		betogings	picket,action
		picketing	no,go
		protests	toi,toi
Plaasaanval	+1.25	plaasaanval	farm,attack
		plaasmoord	plaas,aanval
Grondonteiening	+0.5	besetting	land,grab
		invade	land,invasion
		landgrab	takes,farm

Hierna is 'n puntetoekenning op die WhatsApp-boodskappe uitgevoer. Gegewe boodskap *m*, is woorde eers geïdentifiseer en met 'n kategorie gekoppel deur 'n soektog in die enkelgram- en tweegramwoordeboeke. Dit het 'n woordfrekwensielys in die formaat [*t, w, i*] geproduseer, waar *t* die woordeboeksoort van die woord *w* was en *i* die woordtelling (frekwensie). Ons het dan die klassifikasietelling van elke boodskap *m* bereken deur die produk van woord *w* met die woordeboekgewig *t* te sommeer waaraan 'n woord behoort het. Die kategorie met die hoogste telling is dan aan elke WhatsApp boodskap toegeken. Hierdie benadering verseker dat net een tipe kategorie (gebeurtenis) aan 'n boodskap toegeken word aangesien ons wou fokus op enkel-etiket klassifikasie.

Met die teks verwerk en geklassifiseer as "veilig" of "misdaad, protes, plaasaanval, grondonteiening", word die WhatsApp-boodskappe in 'n datastruktuur verander. Pandas dataframe[1] is hiervoor gebruik aangesien dit 'n oopbron biblioteek voorsien wat hoë werkverrigting vir data-ontleding en manipulasie verseker. Die Pandas-datastel bestaan uit die volgende attribute: "datetime", "times", "classification", "placename1", "placename2", "original_message" en "cleaned_message".

Databerging

Die volgende fase maak die skoongemaakte data in die Pandas-datastel beskikbaar vir verdere data-ontleding. Ten einde dit te bewerkstelling, moes 'n paar stappe gevolg word. Eerstens moes die Python-toepassing OAuth2 toestemming vanaf die Google Ontwikkelaarsportaal verkry.[2] Dit verleen toegang vir die Python-toepassing tot Google Sheets en Google Drive sonder handmatige verifikasie. Die toestemming is in 'n JSON lêer (client_secret.json) gestoor en het toegang verleen aan die toepassingprogrammering-koppelvlak (Application Programming Interface) van Google Drive[3] en Google Sheets[4]. Sodra ontwikkelaar-toestemming vanaf Google verkry is, is die APIs geaktiveer en toeganklik gemaak vir Pygsheets. Pygsheets is 'n oopbron Python-biblioteek vir Google Sheets en laat 'n ontwikkelaar toe om sigblaaie te skep, en toegang tot sigblaaie te beheer, asook werkkaarte en selle deur Google Sheets API v4. Deur 'n kombinasie van Pygsheets en Pandas-funksionaliteit is die data geskryf na die Google Sheets op Google Drive en gesinkroniseer met die MacBook waar die skoongemaakte roudata analities verwerk kon word.

Data-ontleding

Die uitvoer wat in die vorige afdeling bespreek is, word na 'n Google Sheets blad geskryf, waar:

1. Die bron van die boodskap met behulp van 'n gereelde uitdrukking geïdentifiseer word, byvoorbeeld Maroela Media of News24, en indien geen bron in die boodskap aangedui is nie, word "WhatsApp" as verstek aangedui. Sodoende word persoonlike plasings en delings van ander groepe met die bron as WhatsApp aangedui; indien 'n ander bron gebruik is, word dit altyd in die boodskap vermeld.
2. Verwagte betogings uit die datastel onttrek word omdat dit nie noodwendig realiseer nie, dus word slegs bestaande betogings verder aangedui.
3. Gebeurtenisse sonder plekname uit die datastel verwyder word. Dit baat 'n mens min om te weet dat daar betogings is as 'n mens nie weet waar die betogings is nie.
4. Die pleknaam outomaties in Afrikaans vertaal word (byvoorbeeld Cape Town word omgeskakel na Kaapstad).
5. Die tyd wat in die boodskap genoem is, gestandaardiseer word na die formaat hh:mm (indien 'n tydsaanduiding in die boodskap voorgekom het), en indien daar geen tyd in die boodskap genoem is nie, word die tyd wat die boodskap geplaas is as tyd aangedui. Die tyd wat met die boodskap geassosieer word, dui met ander woorde op die nuutste inligting.
6. Die tyd wat in 5 toegeken is, in 'n uur omgeskakel word, sowel as in die tyd van die dag (<=4, "nag", <=8,"vroegoggend", <=12,"oggend", <=17, "middag", <=20, "vroeg aand", <=23, "aand", "geen tyd"). Dit maak dit moontlik om byvoorbeeld te bepaal teen watter tyd van die dag protesaksies plaasvind.
7. Die teenwoordigheid van wapens onttrek word, byvoorbeeld vuurwapens, messe, pangas en klippe.
8. Beserings onttrek word.
9. 'n Eenvoudige sin gegenereer word, byvoorbeeld: "Protes in Kaapstad op 2018/12/01." Hierdie sin maak dit moontlik om nie alleen die getal boodskappe te tel nie, maar ook die getal gebeurtenisse. Wanneer daar byvoorbeeld tien boodskappe voorkom wat verwys na protes in Kaapstad op 2018/12/01, sal slegs een gebeurtenis aangedui word.
10. Hierna word data in Google Data Studio ingetrek en op 'n kaart gevisualiseer. Daar word ook 'n opsomming van insidente verskaf, sowel as 'n lys van die mees onlangse insidente. Google Data Studio is gekies omdat dit op enige slimfoon, rekenaar of tablet toeganklik is en maklik deelbaar is sonder dat verdere sagteware geïnstalleer hoef te word. Die paneelbord ("dashboard") word hieronder bespreek.

Dieselfde boodskappe word gereeld tussen groepe gedeel, maar in ander gevalle kom 'n boodskap net op een groep voor. Alle groepe en boodskappe word in hierdie bostaande

1. https://*pandas.pydata.org*/
2. https://console.developers.google.com/apis/
3. http://drive.googleapis.com
4. http://sheets.googleapis.com

stelsel geïntegreer om mekaar aan te vul, maar die stelsel verwyder ook duplikate wanneer boodskappe op meer as een groep gedeel is. Hierdie pyplyn stel 'n mens in staat om 'n byna intydse ontleding van die geweldsituasie in Suid-Afrika te verskaf, terwyl die verdere toevoeging van inligting soos die teenwoordigheid van wapens, tyd van die dag en ander 'n mens in staat stel om later ondersoek in te stel na fyner nuanses.

Die paneelbord kan in Figuur 1 gesien word.

Links bo word die getal gebeurtenisse aangedui, wat 2 446 is op 25 Januarie 2019. Daaronder word die getal plekke aangedui (592), sowel as die getal rekords wat na daardie plekke verwys (3 962). Onthou dat hierdie getal dui op die getal relevante rekords; boodskappe wat geen pleknaam aandui nie, slegs verwagte protes aandui of nie andersins geklassifiseer is nie, is reeds verwyder, soos hierbo bespreek. Onder die getalle word die plekke aangedui waarvoor die meeste gebeurtenisse aangeteken is, en daaronder die getal bronne (36) en getal WhatsApp groepe (16) wat gemonitor word. Die tydlyn dui die getal gebeurtenisse per maand aan, en die sektordiagram dui die getal gebeurtenisse volgens klassifikasie (protes, plaasaanval, grondbesetting of misdaad) aan. Onder die kaart word die oorspronklike boodskap aangedui om verdere konteks te verskaf en om ons in staat te stel om te kontroleer of die boodskap korrek geklassifiseer is. Regs van die kaart kan gefiltreer word volgens gebeurtenis, datum of plek.

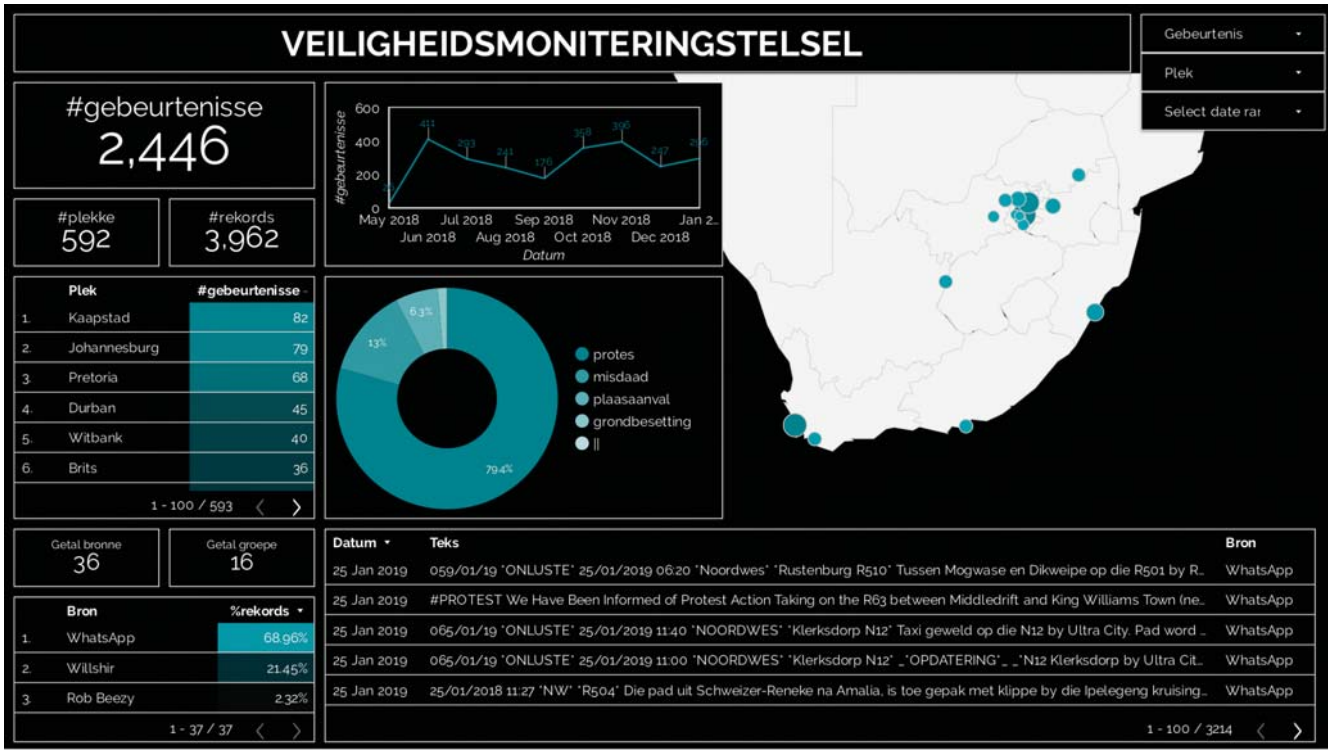
Ter opsomming kan bogenoemde pyplyn diagrammaties voorgestel word soos in Figuur 2.

Hierdie pyplyn maak dit vir ons moontlik om byna intyds 'n beeld te bied oor wat ten opsigte van hierdie geweldsituasies in die land gebeur. Indien 'n mens egter hierdie verwerkte inligting sou wou beskikbaar stel aan die SAPD, privaat maatskappye of nie-regeringsorganisasies, sal dit aangepas moet word by hulle behoeftes.

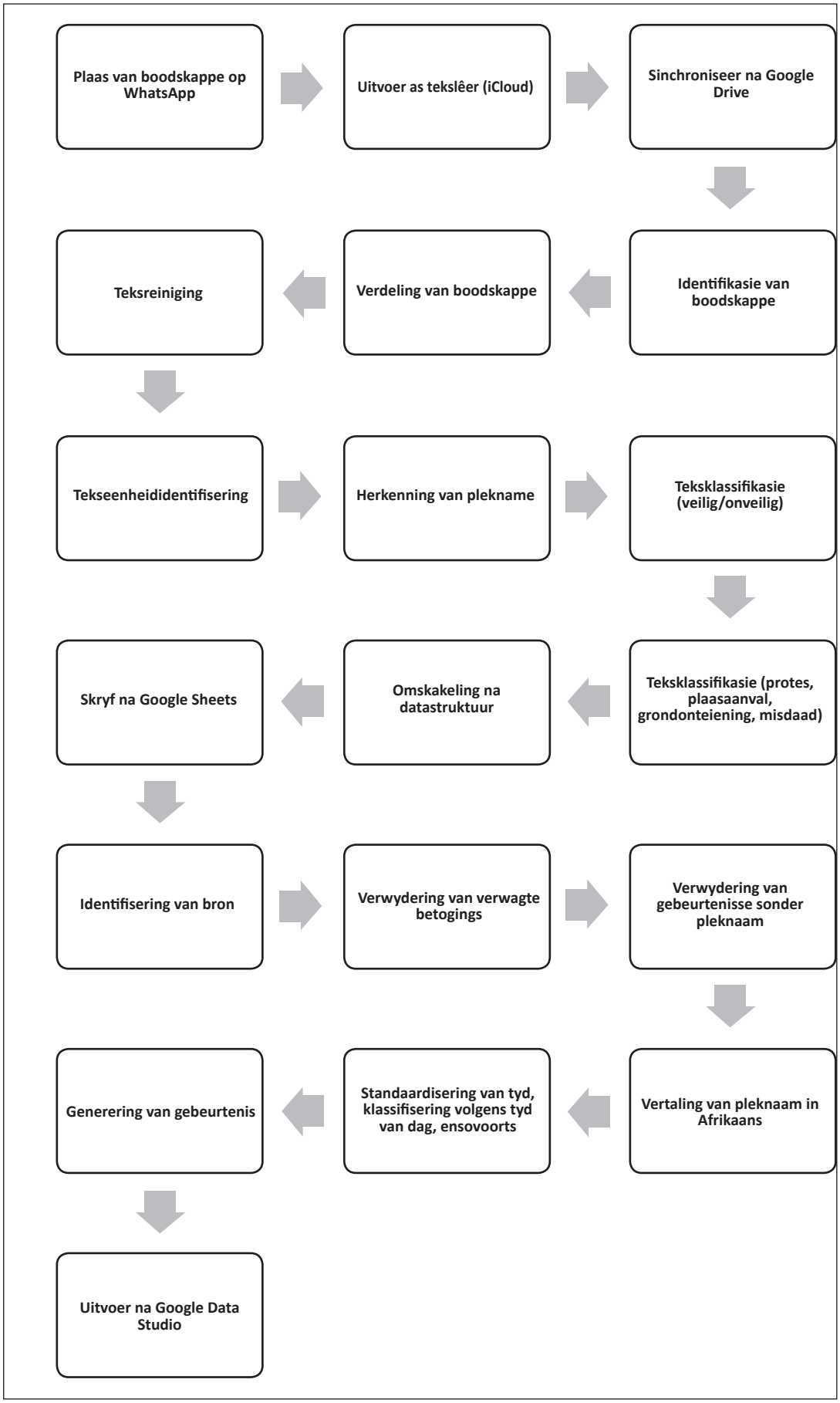
Verdere navorsing

'n Aantal vals positiewe het voorgekom as gevolg van die reëlgebaseerde benadering wat gevolg is. In Januarie 2019 het grootskaalse brande byvoorbeeld in die Wes-Kaap voorgekom, en aangesien "brand" een van die kernwoorde was wat gebruik is om protes aan te dui, is hierdie brande verkeerdelik as protes bestempel. Masjienleer kan gebruik word om sulke vals positiewe te verwyder, wat die onderwerp van 'n verdere navorsingsprojek is. Ons beplan om gebruik te maak van steunvektorklassifiseerders (SVM), Naïewe Bayes en Logistieke Regressie-klassifiseerders, sowel as meer gespesialiseerde neurale netwerkklassifiseerders soos Konvolusie Neurale Netwerke (CNN) en Herhalende Neurale Netwerke (RNN) vir outomatiese teksklassifikasie.

Die huidige studie het slegs inligting uit teks onttrek, maar daar is ook 'n groot aantal video's, foto's en stemboodskappe



FIGUUR 1: Die paneelbord



FIGUUR 2: Die dataversamelings-en ontledingspyplyn

op hierdie groepe gedeel. Verdere tegnieke soos beeldherkenning en rekenaarmatige transkripsies sal toegepas moet word om ook inligting uit sulke boodskappe te onttrek.

Alhoewel WhatsApp reeds 'n bruikbare platform is, kan platforms soos e-pos ook met vrug aangewend word om geweldsituasies of ander misdaad te monitor. Eksperimente is ook reeds onderneem om RSS-voere en Twitter by hierdie stelsel te integreer om 'n meer omvangryke beeld te bied.

Indien 'n mens 'n soortgelyke stelsel met die SAPD se databasisse wil integreer, sal terme en klassifikasies aangepas moet word. Bostaande prosesse is ontwikkel vir privaat toepassings, maar sou aangepas kon word om saam met die polisie te kan werk.

Slot

Die veranderende inligtingslandskap van die hedendaagse wêreld bied nuwe geleenthede en uitdagings aan diegene wat inligting wil versamel. Hierdie studie het aangetoon hoe bruikbare inligting uit groot hoeveelhede ongestruktureerde teks, wat deurlopend gegenereer word, onttrek kan word. Een van die grootste voordele van so 'n stelsel is dat inligting onmiddellik gedeel kan word, soos gebeure plaasvind. 'n Aantal uitdagings en verdere navorsingsgeleenthede is ook bespreek.

Verwysings

Abdollahian M, Baranick M, Efird B, Kugler J. 2006. Senturion: a predictive political simulation model. Center for Technology and National Security Policy National Defense University.

AfriForum. 2017. AfriForum neighbourhood watch in Fichardt Park catches criminals. Beskikbaar by: https://www.afriforum.co.za/free-state/ [Toegang tot 30 November 2018].

Anoniem. 2015. South Africans turn to technology to fight crime. Beskikbaar by: http://visiontactical.co.za/2015/07/08/south-africans-turn-to-technology-to-fight-crime/ [Toegang tot 30 November 2018].

Anoniem. 2016a. Fichardtpark-tak spoor Alzheimer-lyer op. Beskikbaar by: https://www.bloemfonteincourant.co.za/fichardtpark-tak-spoor-alzheimer-lyer-op/ [Toegang tot 30 November 2018].

Anoniem. 2016b. Whatsapp groepe fnuik skaapdiewe. Beskikbaar by: https://parysgazette.co.za/8543/whatsapp-groepe-fnuik-skaapdiewe/ [Toegang tot 30 November 2018].

Anoniem. 2017. Stellenbosch women students start WhatsApp safety groups. Beskikbaar by: https://www.news24.com/SouthAfrica/News/stellenboschwomen-students-start-whatsapp-safety-groups-20170531 [Toegang tot 30 November 2018].

Anoniem. 2018. Study finds Facebook news use declining, WhatsApp growing. Beskikbaar by: http://www.sabcnews.com/sabcnews/study-finds-facebooknews-use-declining-whatsapp-growing/ [Toegang tot October 2018].

Arslan C, Yanık M. 2015. How to make social media more effective as an exploitation area? *Journal of Military and Information Science*, 3(3), pp. 79-87.

Azar EE, Ben-Dak J. 1975. Theory and practice of events research. New York: Gordon and Breach.

Bates RH, et al. 2003. Political instability task force report: phase IV findings. McLean, VA: Science Applications International Corporation.

Brandt PT, Freeman JR, Schrodt PA2011. Real time, time series forecasting of inter- and intra-state political conflict. *Conflict Management and Peace Science*, 28(1), p. 41-64.

Bruwer P. 2015. WhatsApp-groep help boere om boewe aan te keer. Beskikbaar by: https://maroelamedia.co.za/nuus/sa-nuus/fotos-whatsapp-groep-helpboere-om-groep-boewe-aan-te-keer/ [Toegang tot 30 November 2018].

Bueno de Mesquita B, Stockman FN. 1994. European community decision making. New Haven: Yale University Press. Page 9 of 10 Oorspronklike Navorsing.

Bueno de Mesquita B. 1981. The war trap. New Haven: Yale University Press.

Bueno de Mesquita B, Newman D, Rabushka A. 1985. Forecasting political events: the future of Hong Kong. New Haven: Yale University Press.

Central Intelligence Agency, 2010. INTellingence: Open Source Intelligence. Beskikbaar by: https://www.cia.gov/news-information/featured-storyarchive/2010-featured-story-archive/open-source-intelligence.html [Toegang tot 16 Januarie 2019].

Croicu M, Sundberg R. 2017. UCDP GED Codebook version 17.1. Uppsala: Department of Peace and Conflict Research, Uppsala University.

Cronjé SM. 2016. GPF waarsku plaasinwoners om veiligheid op te skerp. Beskikbaar by: https://ridgetimes.co.za/77903/gpf-waarsku-plaasinwonersom-veiligheid-op-te-skerp/ [Toegang tot 30 November 2018].

Dahir AL. 2018. WhatsApp is the most popular messaging app in Africa. Beskikbaar by: https://qz.com/africa/1206935/whatsapp-is-the-mostpopular-

messaging-app-in-africa/ [Toegang tot 30 November 2018].

Damons A. 2017. Groepe op WhatsApp kan help – as dit reg gebruik word. Beskikbaar by: https://www.pressreader.com/south-africa/volksblad/20171004/281685435050506 [Toegang tot 30 November 2018].

Dayal UCMSA, WK. 2009. Data integration flows for business intelligence. s.l., s.n., pp. 1-11.

Debnath P, Haque S, Bandyopadhyay S, Roy S. 2016. Post-disaster situational analysis from WhatsApp group chats of emergency response providers. Rio de Janeiro, s.n.

Dencik L. Hintz A, Carey Z. 2017. Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom. *New Media & Society*, pp. 1-18.

Dey L, Haque SM. 2009. Opinion mining from noisy text data. *International Journal on Document Analysis and Recognition (IJDAR)*, 12(3), p. 205-226.

Esty DC, et al. 1995. State failure task force report. McLean, VA: Science Applications International Corporation.

Esty DC, et al.1998. State failure task force report: Phase II findings. McLean, VA: Science Applications International Corporation.

Gerner DJ, Schrodt PA, Francisco RA, Weddle JL. 1994. The machine coding of events from regional and international sources. *International Studies Quarterly*, Volume 38, pp. 91-119.

Gibson SD. 2014. Exploring the role and value of open source intelligence. In: C. Hobbs, M. Moran & D. Salisbury, reds. Open Source Intelligence in the Twenty-First Century. New Approaches and Opportunities. New York: Palgrave Macmillan, pp. 9-23.

Goldstone JA, et al. 2010. A global model for forecasting political instability. *American Journal of Political Science*, 54(1), p. 190-208.

Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU. 2015. The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, Volume 47, p. 98-115.

Hendrix CS, Salehyan I. 2015. No news is good news: mark and recapture for event data when reporting probabilities are less than one. *International Interactions*, 41(2), pp. 392-406.

Hobbs C, Moran M, Salisbury D. 2014. Introduction. In: C. Hobbs, M. Moran & D. Salisbury, reds. Open Source Intelligence in the Twenty-First Century. New Approaches and Opportunities. New York: Palgrave Macmillan, pp. 1-8.

Hurter T. 2018. Wakende oog puik. *Kroonnuus*, 23 Oktober, p. 1. Internet Live Stats, 2019. Beskikbaar by: hk p://www.internetlivestats.com [Toegang tot 19 Januarie 2019].

Jianqiang Z, Xiaolin G. 2017. Comparison research on text pre-processing methods on Twitter sentiment analysis. *IEEE Access*, Volume 5, p. 2870-2879.

Kaldor M. 2006. New & old wars. Organized violence in a global age. Cambridge: Polity.

Knowles L. 2018. WhatsApp group raise crime awareness, CPF says. Beskikbaar by: https://www.pressreader.com/south-africa/talk-of-thetown/20180517/281517931772790 [Toegang tot 30 November 2018].

Lind, WS, Schmitt JF, Wilson GI. 1989. The changing face of war: into the fourth generation. *Marine Corps Gazette,* Oktober, pp. 22-26.

Malik D. 2016. Emergency response in the Whatsapp era. Beskikbaar by: http://blogs.worldbank.org/endpovertyinsouthasia/emergency-responsewhatsapp-era [Toegang tot 30 November 2018].

Manning CD, Schütze H. 1999. Foundations of statistical natural language processing. s.l.:MIT Press.

Mazzarella J. 2016. What's up with WhatsApp for emergency communications? Beskikbaar by: https://urgentcomm.com/collections/whats-up-with-whatsappfor-emergency-communications/ [Toegang tot 30 November 2018].

Medhat W, Hassan A, Korashy H. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), p. 1093-1113

Moreno A, Garrison P, Bhat K. 2017. WhatsApp for monitoring and response during critical events: Aggie in the Ghana 2016 Election. Albi, France, s.n.

Münkler H. 2005. The new wars. Cambridge: Polity.

Newman N, et al. 2018. Digital News Report 2018, Oxford: Reuters Institute for the Study of Journalism.

O'Brien SP. 2010. Crisis early warning and decision support: contemporary approaches and thoughts on future research. *International Studies Review*, Volume 12, p. 87-104.

Palmer D. 2012. Text preprocessing. In: N. Damerau & F. J. Indurkhya, reds. Handbook of Natural Language Processing. s.l.:CRC Press, p. 9-30.

Raleigh C, Linke A, Hegre H, Karlsen J. 2010. Introducing ACLED – Armed Conflict Location and Event Data. *Journal of Peace Research*, 47(5), pp. 651-660.

Raper P, Möller L, Plessis LD. 2014. Dictionary of Southern African Place Names. Johannesburg: Jonathan Ball.

Rawlins LK. 2018. SA radio station says WhatsApp out, Telegram in. Beskikbaar by: https://www.itweb.co.za/content/o1Jr5qxEX2DvKdWL [Toegang tot 30 November 2018].

Rodgers M. 2018. NSRI Plett assist a crewman who had fallen overboard and a man who had fallen off from his mountain bike. Beskikbaar by: https://www.nsri.org.za/2018/11/nsri-plett-assist-a-crewman-who-had-fallen-over-boardand-was-hypothermic-and-a-man-who-had-fallen-off-from-his-mountainbike-and-injured-his-shoulder/ [Toegang tot 30 November 2018].

Salehyan I, et al. 2012. Social conflict in Africa: A new database. *International Interactions*, 38(4), pp. 503-511.

Schrodt PA. 2012. Precedents, progress, and prospects in political event data. *International Interactions*, 38(4), pp. 546-569.

Shapshak T. 2015. Why WhatsApp is South Africa's favourite App. Beskikbaar by: https://www.forbes.com/sites/tobyshapshak/2015/09/04/why-whatsapp-issouth-africas-favourite-app/ [Toegang tot 30 November 2018].

Smith M. 2014. Fnuik plaasaanvallers met WhatsApp. Beskikbaar by: https://www.netwerk24.com/landbou/Nuus/fnuik-plaasaanvallers-metwhatsapp-20170914 [Toegang tot 30 November 2018].

Suidlanders. 016. Nuwe Whatsapp waarskuwingsboodskappe. Beskikbaar by: https://www.suidlanders.co.za/nuwe-whatsapp-waarskuwingsboodskappe/ [Toegang tot 30 November 2018].

Sundberg R, Melander E. 2013. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50(4), pp. 523-532.

Taleb RD, Serhani MA. 2015. Big data pre-processing: a quality framework. New York, s.n., pp. 191-198.

Thompson A. 2018. Why most of Africa's data is used on WhatsApp. Beskikbaar by: https://theculturetrip.com/africa/articles/why-most-of-africas-data-isused-on-whatsapp/ [Toegang tot 30 November 2018].

Van Creveld M. 1991. The transformation of war. London: The Free Press.

Van Creveld M. 2008. The changing face of war. Combat from the Marne to Iraq. New York: Ballantine.

Waterworth T. 2017. Durban's highways of terror. Beskikbaar by: https://www.iol.co.za/news/south-africa/kwazulu-natal/durbans-highways-of-terror-10191779 [Toegang tot 30 November 2018].

Williams HJ, Blum I. 2018. Defining second generation open source intelligence (OSINT) for the Defense Enterprise. Santa Monica: RAND Corporation.