

Open bronnen: betrouwbare informatie?

Arno Reuser is werkzaam bij de Militaire Inlichtingen- en Veiligheidsdienst (MIVD) waar hij de Open Source Intelligence Unit heeft opgericht. DB sprak met hem over open bronnen en informatiebetrouwbaarheid. Reuser spreekt in dit artikel op persoonlijke titel, dus niet als vertegenwoordiger van de MIVD.

Erik Bouwer



Je zou verwachten dat open source intelligence gebaat is bij de snel groeiende openbare databrij. Want hoe meer informatie er is, hoe groter de kans dat er bruikbare delen bij zitten. Toch gaat dit niet helemaal op. Ook het inlichtingenwerk heeft te kampen met beperkingen; succesvol vinden wordt vooral bepaald door goede zoekstrategieën. Bij open source intelligence (kortweg OSINT) staat open bron informatie centraal, met andere woorden informatie die vrij verkrijgbaar is. Dat kan variëren van luchtfoto's tot online sociale netwerken en van onderzoeksdata tot publieke overheidsdocumenten.

Na de oprichting van OSINT gaf Reuser jarenlang leiding aan OSINT. Nu is hij Senior Policy Advisor voor OSINT en Cyber bij de MIVD. Daarnaast heeft hij een eigen bedrijf, Reuser Information Services. Ook is hij de beheerder van een bekende discussielijst voor informatieprofessionals, Nedbib-L. Reuser spreekt op de komende editie van het congres Kennis in Praktijk (29 september, Utrecht).

Wat is 'open'?

Reuser: "Een mooie definitie voor OSINT zou kunnen zijn: het afstemmen van de interne vraag naar informatie op het wereldwijde aanbod van informatie door selectie, monitoring, beschrijven en toegankelijk maken en verspreiden van informatie uit openbare bronnen. Het gaat daarbij om informatie die legaal en op ethische wijze kan worden verkregen."

Maar open bronnen vormen een rekbaar begrip, zo licht Reuser toe aan de hand van een voorbeeld: "Als een kind besluit via een peer-to-peer-netwerk

muziekbestanden te uploaden via de laptop van zijn vader en daarbij per ongeluk de volledige administratie van zijn vader meeneemt, is dat vrij beschikbare informatie. Maar het is niet bedoeld om op internet terecht te komen en valt dus niet onder het begrip 'OSINT' – als die informatie toch nodig zou zijn voor het inlichtingenwerk dan zou er in dat geval gebruik moeten worden gemaakt van bestaande wetgeving rondom bijzondere verwervingsmiddelen."

Vrijplaats internet

De Wet op de Inlichtingen- en Veiligheidsdiensten geldt voor beide inlichtingendiensten die ons land kent: de AIVD en de MIVD. "Die wet beschrijft de taken van de diensten en de mogelijkheden die zij hebben voor het verwerven van informatie, alsmede de voorwaarden waaronder dat mag, of niet. Ook controle instanties zijn er door benoemd, zoals de Commissie van Toezicht op de Inlichtingen- en veiligheidsdiensten (CTIVD). De Wet is restrictief en is ook bedoeld om de rechtspositie van burgers te beschermen. Zo geeft de wet onder andere aan dat er niet zo maar brieven mogen worden geopend of in databases mag worden ingebroken."

Ook *commerciële* informatie kan worden gebruikt voor OSINT-activiteiten. "Als de intuïtie – maar vooral ethiek – zegt dat er betaald zou moeten worden voor bepaalde informatie, dan wordt die informatie ingekocht, ook al staat deze illegaal en 'gratis' op een site. Daarnaast geldt dat doel en middelen bij het inlichtingenwerk met elkaar in balans moeten zijn", benadrukt Reuser.

Het internet heeft het informatielandschap aanzienlijk gewijzigd, legt Reuser uit: "Sinds de 16e eeuw is er veel veranderd. In de begintijd was er het gesproken woord en daar kwam het boek bij.

Toen de overheid ontdekte wat de kracht van gedrukte tekst was, begon de censuur: er werden bibliografieën opgesteld waarmee inzichtelijk werd gemaakt wat er allemaal beschikbaar was.” (In oude bronnen is zichtbaar hoe de kerk daarnaast actief censuur toepaste door in gedrukte boeken passages te schrappen of onleesbaar te maken¹, red.).

Herkomst onbekend

Volgens Reuser is het sinds het ontstaan van de censuur steeds moeilijk gebleven om zonder obstakels de vrije mening te verkondigen op een andere manier dan via het gesproken woord: “Voor radio en televisie moet je een zendmachtiging hebben, voor een krant of tijdschrift moet je kapitaal hebben.” Dat verklaart in zijn ogen mede de opkomst van rapporten, bronnen waarmee specifieke groepen onderling informatie konden uitwisselen zonder dat daarvoor de regulering van toezichthouders of uitgevers nodig was. Alleen met het internet lijken we nu voor het eerst aan iedere vorm van toezicht of registratie te kunnen ontsnappen: “Daar kan iedereen publiceren. Daar tegenover staat dat het ontbreekt aan mechanismen voor validatie van informatie – is de informatie juist en betrouwbaar? Ook de authenticatie, de herkomst van informatie, is een probleem: *on the Internet, nobody knows you are a dog*. Je weet nooit echt goed wie er achter bepaalde informatie zit. In veel gevallen zijn internetgebruikers zich daar weinig van bewust. Maar je mening geven kost op het web geen geld of moeite. Overigens denken we wel vaak – en geheel ten onrechte – dat de hele wereld op het internet is aangesloten en we dus iedereen kunnen bereiken. In de praktijk is dat maar een klein gedeelte, zo’n 30 procent van de wereldbevolking.”



FOTO: ERIK BOUWIER

Internet = relatief

Dat expanderende web heeft in de zoektocht naar relevante bronnen niet geleid tot compleet nieuwe zoektechnologieën. Evenmin, zo benadrukt Reuser, is er sprake van het ‘afschrapen’ van het internet: het analyseren van alle content. “Daarvoor is er eenvoudigweg veel te veel informatie aanwezig. In 2002 werd de totale hoeveelheid informatie geschat op 5 Exabyte, zeven jaar later was dat al meer dan 281 Exabyte en in 2010 liep de schatting op tot 1.200 Exabyte. Naast de groeiende omvang van het web speelt mee dat er altijd beperkte middelen zijn.”

De grootste uitdaging voor het inlichtingenwerk en met name OSINT is het krijgen en houden van zicht op het veranderende informatielandschap. Dat vereist diepgaande kennis van zoektechnieken en -strategieën en het goed kunnen definiëren van informatiebehoefte, het selecteren van betrouwbare bronnen en het feitelijk filteren van informatie uit die bronnen. Informatie-uitwisseling tussen inlichtingenmedewerkers speelt hierbij een belangrijke rol – bijvoorbeeld elkaar attenderen op relevante bronnen. In die tweede slag, het filteren van informatie uit die bronnen, is de technologie wel behulpzaam. “Taaltechnologie zoals vertaalsoftware en semantische technieken worden steeds belangrijker vanwege de talen uit opkomende (Aziatische) economieën. Het ICANN heeft bovendien het aantal domeintalen uitgebreid waardoor het aanbod aan internetinformatie uit bijvoorbeeld Azië toeneemt en Engels niet langer de dominante taal is.”

Westerse betrouwbaarheid

Dat het informatielandschap verandert, is uiteraard relevant voor informatieprofessionals. Voor Reuser is niet zo zeer de toenemende *omvang*, als wel de betrouwbaarheid van aangeboden informatie een heikel punt. “Google, de belangrijkste ingang voor consumenten, levert bij een zoekterm als ‘nuclear weapons Iran’ via Google.com heel andere resultaten op dan via Google.de: anders in aantallen en anders in verwijzingen. De zoekresultaten komen bovendien niet transparant tot stand en de gepresenteerde informatie is qua herkomst en inhoud niet gevalideerd. Datzelfde nadeel speelt bij nieuwe publicatiemodellen waarbij auteurs rechtstreeks op het web publiceren.”

“Daarnaast wordt zoekgedrag hoofdzakelijk ge-

Zoekterm Google: *nuclear weapons iran* (query uitgevoerd op 10 juli 2011, resultaten wijzigen in de loop der tijd)

Google.de:	17400.000 resultaten
Googl.nl:	12.500.000 resultaten
Google.com:	9.530.000 resultaten

baseerd op westerse hardware en software. Nokia, Apple, Google, LexisNexis: de apparaten, de programma’s en de zoeksystemen, maar ook de bronnen zijn vrijwel allemaal Amerikaans of soms Europees. Er bestaat geen vergelijkbare concurrent van Google in het Midden-Oosten of Azië die wij gemakkelijk kunnen gebruiken,” aldus Reuser.

Plan of attack

Reuser geeft ook les in zoektechnieken: aan militairen, overheidsorganisaties, think tanks, financieel analisten en aan studenten. “Hoewel er evenveel zoekstrategieën als informatieprofessionals zijn, presteren IP’ers in het algemeen beter dan ‘leken’ – maar er wordt lang niet altijd met een *plan of attack* gewerkt.” Cruciaal voor het opstellen van een zoekstrategie is bijvoorbeeld het vooraf bepalen wanneer je stopt met zoeken – anders verdwaal je. “Resultaten beoordelen en je plan eventueel aanpassen.”

Bij de Universiteit Twente zijn laboratoriumtests gedaan in zoekgedrag met opvallende resultaten waaruit in het algemeen blijkt dat de internetvaardigheden van burgers ernstig worden overschat. Reuser legt uit hoe informatie op het internet klakkeloos voor waar werd aangenomen – ook het definiëren van zoektermen en selecteren van relevante informatie leverde veel problemen op. Het oplossen van een eenvoudig vraagstuk

(een voorkeur uitspreken voor een politieke partij op basis van bepaalde standpunten) lukte slechts in 25 procent van de gevallen. 37 procent van de proefpersonen had moeite een bestand te downloaden. 49 procent lukte het niet een PDF bestand te openen en in een folder op de desktop op te slaan. 25 procent daarvan had absoluut geen idee hoe dat moest.

All cops are bastards?

De rol van een redactie en uitgever is belangrijk, maar toch speelt het ontbreken daarvan voor veel informatiegebruikers geen enkele rol van betekenis. Reuser verwijst hierbij naar een voorbeeld waarbij zelfs Google hits als informatiebron werden ingezet: de rechtszaak over de letters ACAB op een bomberjack. De drager van dat jack kreeg een boete, omdat de afkorting zou staan voor 'All Cops Are Bastards' – goed voor een belediging van een agent. Het gerechtshof had zich bij de uitspraak mede gebaseerd op een zoekactie op basis van de combinatie van de zoektermen 'ACAB' en 'Cops', wat 190.000 treffers opleverde. Op basis daarvan nam de rechter aan dat ACAB over het algemeen bedoeld was als negatief voor politieagenten. Daarmee werd het gebruik van de afkorting op kleding bestraft. De Hoge Raad floot de rechter terug: het aantal treffers bij het zoeken in alle, ook andersstalige, internetsites 'is niet zonder meer redenevend'. Hetzelfde geldt voor het aantal treffers bij de gebruikte zoekmachine waarop het Hof zich heeft beroepen. Het hof had verzuimd na te gaan op wat voor websites de treffers betrekking hadden.

Informatievaardigheden

Reuser deelt de zorgen over de vaardigheid en de kennis op het vlak van internet: "Op YouTube staat een filmpje – een straatinterview met toevallige voorbijgangers op Times Square in New York – waarin naar voren komt dat mensen niet weten wat het verschil is tussen een website, Google en een browser. Dat blijkt zo te zijn in New York, maar ook in Nederland en op vele andere plekken." "Ondertussen veranderen wijzelf als individu allemaal in 'inlichtingendiensten' en geven we ook steeds gemakkelijker informatie over derden door aan overheden of over onszelf op sociale netwerken. Voor opsporings- en inlichtingendiensten is het niet moeilijk om na te gaan of social media accounts fake zijn of niet, maar voor de gewone burger is dat nauwelijks te achterhalen. Terwijl iedereen een account kan aanmaken onder andermans naam. Vervolgens wordt de informatie die we van Facebook terugkrijgen over onze vrienden ook nog eens gemanipuleerd."

Reuser vindt dan ook dat informatievaardigheden meer dan ooit van belang zijn en bij uitstek de expertise van de bibliotheek zijn. Leesbevordering – dat nog altijd van groot belang wordt gevonden binnen in de openbare bibliotheek – daarentegen, is binnen de bibliotheek "een verkwisting van talent en hoort thuis in het basis- en voortgezet onderwijs – waar hebben we anders scholen voor?" **db**

Noot:

1 Zie [Digitale Bibliotheek 5-2011/Digi-review: recensie 'Boeketje Boekwetenschap'](#).

Andere literatuursuggestie:

[Overheid overschat internetvaardigheden burgers](#)