Avointen lähteiden tiedustelun vaikutus yksityisyyden suojaan

Tietojenkäsittelytiede
Tietotekniikan laitos, Teknillinen tiedekunta
LuK-tutkielma

Laatija:
Ville Veikko Vartia

Ohjaajat:
Sampsa Rauti

Toukokuu 2024

LuK-tutkielma
Tietotekniikan laitos, Teknillinen tiedekunta
Turun yliopisto

Oppiaine: Tietojenkäsittelytiede
Tutkinto-ohjelma: Tietojenkäsittelytieteiden tutkinto-ohjelma
Tekijä: Ville Veikko Vartia
Otsikko: Avointen lähteiden tiedustelun vaikutus yksityisyyden suojaan
Sivumäärä: 22 sivua
Päivämäärä: Huhtikuu 2024

Avoimen datan määrä lisääntyy internetissä jatkuvasti ja sitä mukaan myös tiedonhankintamenetelmät niin siviili- kuin myös sotilastiedustelussa muuttuvat. Tänä päivänä ei ole enää ongelmaa tiedon puutteesta, vaan enemmänkin sen järkevästä ja tehokkaasta louhinnasta sekä käsittelystä. Tähän ongelmaan avointen lähteiden tiedustelu tarjoaa ratkaisun. Tämän kyseisen tiedustelumenetelmän suosio on ollut jo pitkään hyvin jyrkässä kasvussa ja sen avulla onkin saatu lupaavia tuloksia muun muassa rikollisuuden ja terrorismin torjunnassa. Sen lisäksi sitä on käytetty paljastamaan valtion salaisuuksia. Avointen lähteiden tiedustelun metodit herättävät myös huolta tavallisen internet-käyttäjän yksityisyyden suojasta, sillä tätä menetelmää voi yhtä hyvin soveltaa rikollisten lisäksi jokaiseen internettiä käyttävään lainkuuliaiseen kansalaiseen. Tästä onkin jo löytynyt todisteita ja näitä esimerkkitapauksia niiden seurauksineen käydään läpi tässä tutkielmassa. Tämän kandidaatintutkielman tavoitteena on selvittää, mikä selittää avointen lähteiden tiedustelun suosion kasvun ja minkälaisia vaikutuksia kyseisellä menetelmällä voisi olla tavalliselle internetin käyttäjälle. Aiheeseen tutustutaan teorian lisäksi käytännönläheisestä näkökulmasta ottaen myös huomioon menetelmän juridiset sekä eettiset pulmat. Tutkielma on toteutettu kirjallisuuskatsauksena. Tutkielman tulokset antavat ymmärtää, että avointen lähteiden tiedustelun suosiota selittää muun muassa menetelmän tehokkuus, työkalujen helppo saatavuus, avoimen datan valtava määrä sekä menetelmän kustannustehokkuus. Sen lisäksi kävi ilmi, että kyseessä olevalla menetelmällä on hyvin huolestuttavia vaikutuksia tavallisen internet-käyttäjän yksityisyyden suojaan, etenkin sellaisissa maissa, joissa ihmisoikeuksia ei pidetä kovin suuressa arvossa.

Asiasanat: Avointen lähteiden tiedustelu, OSINT, yksityisyyden suoja, tietosuoja, GDPR

Sisällysluettelo

1 Johdanto

Internetin ja nimenomaan WWW:n käytön vakiintumisen myötä ikivanhat tiedustelumenetelmät ovat saaneet uuden alustan ja käyttötarkoituksen. Sitä mukaa kun digitalisaatio jatkaa kehittymistään ja valloittaa aina vain uusia osa-alueita elämästämme, internetissä oleva datan määrä vain lisääntyy. Osa tästä datasta on luonteeltaan avointa, eli se on teoriassa kaikkien saatavilla, ja siksi sitä kutsutaankin avoimeksi dataksi [3]. Avoin data on keskiössä erään uudelleen päätä nostavan ja voisi sanoa jopa tänä päivänä melko kriittisen tiedustelumenetelmän roolissa. Kyseessä on *avointen lähteiden tiedustelu* (engl. OSINT, Open Source Intelligence) joka lyhykäisyydessään tarkoittaa yksinkertaisesti sellaista tiedustelumenetelmää, jossa tiedustelutietoa kerätään kaikille avoimista lähteistä. Tätä menetelmää hyödyntävät niin virkavalta kuin myös yksityiset toimijat, kuten yritykset ja tutkivat journalistit [6][15].

Tässä tutkielmassa tutustutaan avointen lähteiden tiedusteluun ensiksi yleisellä tasolla luvussa kaksi, jonka jälkeen syvennytään analysoimaan, miten tätä menetelmää käytännössä harjoitetaan, muun muassa selvittämällä minkälaisia työkaluja siihen on tarjolla ja miten näitä työkaluja voi hyödyntää. Sen jälkeen etsitään vastausta ensimmäiseen tutkimuskysymykseen: ”**Mikä selittää avointen lähteiden tiedustelumenetelmän suosion tänä päivänä?**”. Myöhemmin luvussa kolme perehdytään tietosuoja-asetuksiin lähinnä GDPR:n kautta, tavoitteena tutkia tämän tiedustelumenetelmän laillisuutta. Lopuksi luvussa neljä pohditaan eettisiä sekä juridisia pulmia ja näitä teemoja yhdistelemällä vastausta toiseen tutkimuskysymykseen: ”**Miten avointen lähteiden tiedustelu vaikuttaa tavallisen internet-käyttäjän yksityisyyteen?**”. Luvussa viisi on pohdintaa aiheesta sekä lyhyt yhteenveto tutkielmasta.

Tässä tutkimuksessa on käytetty akateemisten lähteiden lisäksi esimerkiksi uutistoimisto Vice Newsin palkittua dokumenttielokuvaa ”Selfie soldiers: Russia checks into Ukraine”, jonka avulla aiheeseen voi perehtyä hieman enemmän käytännönläheisemmän näkökulman kautta. Tieteellisiä artikkeleja ja muuta kirjallisuutta löytyi hyvin paljon Google Scholarista sekä IEEE:n tietokannasta hakusanoilla ”Open source intelligence AND ”privacy*” sekä ”Open source intelligence” AND GDPR”. Tuloksia tuli sadoittain ja vain oleellisimmat päätyivät käytettäväksi. Hyvin moni valituista lähteistä oli julkaistu viimeisen 5-10-vuoden sisällä, mikä taas toisaalta kertoo aiheen ajankohtaisuudesta.

2 Avointen lähteiden tiedustelu

On arvioitu, että jopa 80–90 % kaikesta tiedustelutiedosta tulee avoimista lähteistä [1][9]. Tämä väite kuvastaa hyvin sitä maailmaa, missä elämme tänä päivänä. Elämme internetin aikakaudella, jossa suurin osa ihmisistä on jollakin tavalla kytköksissä internetiin [19]. Oikeastaan internet on jo niin suuri osa elämäämme, että voisimme kuvitella ihmisen elävän jossakin määrin ikään kuin kahdessa todellisuudessa: oikeassa elämässä sekä internetissä. Tämän seurauksena oikea elämä, kaikkine päivän ajatuksineen, heijastuu myös internetiin esimerkiksi sosiaalisen median palveluiden, pikaviestimien ja keskustelupalstojen julkaisujen kautta. Näin internettiin on alkanut räjähdysmäisesti kertymään valtava määrä dataa, josta osa on käytännössä kaikille saatavilla olevaa, niin kutsuttua avointa dataa.

Internetissä tapahtuva kommunikaatio on suuressa roolissa avointen lähteiden tiedustelussa. Kun aikaisemmin radikaalit ajatukset ja terrorismi levisivät suullisesti tai kirjeiden avulla, nykyään on helpompi löytää samanmielisiä ihmisiä ja aatetovereita esimerkiksi keskustelupalstojen kautta. Esimerkiksi tanskalaiset turvallisuusviranomaiset ovat havainneet internetin olevan tärkein kanava vaarallisen ääri-islamilaisen jihadismin levittämisessä [3]. Myös valtion viranomaiset ovat huomanneet tämän ilmiön ja luonnollisesti myös ryhtyneet toimenpiteisiin. Tästä on seurannut avointen lähteiden tiedustelun käyttöönotto. Avointen lähteiden tiedustelu on sittemmin saanut oivan käyttötarkoituksen: tehokas massavalvonta. Massavalvonnasta onkin jo esimerkkejä, sillä muun muassa eräs suuri turvallisuusalan yritys nimeltä Raytheon on kehittänyt ohjelman nimeltä RIOT (Rapid Information Overlay Technology) joka kerää ja käy läpi valtavaa määrää dataa sosiaalisesta mediasta ennustaakseen ihmisten toimintaa [6]. Muita esimerkkejä käydään läpi myöhemmin esimerkiksi neljännessä luvussa.

Avointen lähteiden tiedustelun mahdollistama massavalvonta herättää ajatuksia sen laillisuudesta ja eettisyydestä. Uskaltaako sosiaalisessa mediassa enää ottaa kantaa esimerkiksi poliittisiin aiheisiin? Profiloidaanko meitä aktiivisesti poliisin, turvallisuuspalveluiden tai yksityisten toimijoiden toimesta sosiaalisen median tiliemme kautta? Mitä tiedoillamme tehdään ja kenelle niitä jaetaan? Näitä kysymyksiä ihmiset joutuvat luultavasti pohtimaan erityisesti sellaisissa maissa, jossa sananvapautta ja muitakin yksilön vapauksia rajoitetaan ja joissa rikkomuksista voi seurata vakavia seuraamuksia. Tämä luonnollisesti herättää kysymyksen siitä, että voiko avointen lähteiden tiedustelulta suojautua, jos kokee sen aiheelliseksi. Luvussa 4 käydään läpi muutamia esimerkkejä tällaisista

skenaarioista ja pohditaan hieman myös suojautumistoimenpiteitä OSINT-tiedustelua vastaan. Sitä ennen perehdytään hieman syvällisemmin siihen, että minkälaisesta menetelmästä on oikein kyse.

2.1 Avointen lähteiden tiedustelu käytännössä

OSINT-menetelmissä kerättävä data on kaikille avointa ja siksi myös verrattain helposti saatavilla. Tämän tyyppiseen dataan viitataan toisinaan termillä OSINF (Open Source Information) [3]. Monille voi tulla sanasta ”tiedustelu” mieleen arveluttavat tiedonkeruumenetelmät, kuten telekuuntelu tai muut vastaavat menetelmät, joilla saadaan ehkä usein salaiseksi luokiteltua tietoa kohteesta. Sen sijaan avointen lähteiden tiedustelussa hyödynnettävä data *ei* ole missään nimessä sellaista dataa, joka on alkuperäiseltään luonteeltaan jollakin tavoin salattua, omistusoikeudellisesti rajoitettua tai kerätty salakähmäisin tai epärehellisin menetelmin [4]. Tässä tiedustelumenetelmässä toimitaan siis monessa suhteessa hieman eri tavalla kuin perinteisissä viranomaisten tiedustelumenetelmissä. Merkittävin ero on tiedustelutiedon alkuperä. Koska informaatio on julkisesti saatavilla, myös suurelle yleisölle eli tavallisille kansalaisille, informaation luonteen voidaan olettaa olevan eri laatuista kuin esimerkiksi paljon raskaammalla menetelmällä kuten televalvonnalla tai -kuuntelulla saatu informaatio. Televalvonta tarkoittaa sitä, että poliisi voi hankkia salassa pidettäviä tunnistamistietoja kuten puhelinnumeron ja laitteen sijaintitiedon, sekä myös sulkea epäillyn liittymän tai laitteen. Televalvontaa voi lain mukaan harjoittaa vain mikäli ”…jotakuta on syytä epäillä rikoksesta, josta säädetty ankarin rangaistus on vähintään neljä vuotta vankeutta” tai jos häntä epäillään eräistä muista vakavista rikoksista [5]. Televalvonta ja -kuuntelu ovat perinteisempiä tiedustelun menetelmiä. Avointen lähteiden tiedusteluun ryhtyminen ei välttämättä vaadi yhtä suurta kynnystä viranomaisilta, kuin mitä perinteiset tiedustelumenetelmät vaatisivat [1]. Toki tähän vaikuttavat aina paikalliset lait ja asetukset.

Sen sijaan avointen lähteiden tiedustelussa liikutaan lain harmaalla alueella, jos tiedustelun harjoittaja on viranomaisten sijaan esimeriksi yksityinen yritys. Selkeästi avointen lähteiden tiedustelumenetelmät voivat rikkoa Euroopan Unionin yleisen tietosuoja-asetuksen GDPR:n määräyksiä, sillä GDPR:ssä nimenomaan kielletään muun muassa henkilötietojen kerääminen ja käsittely ilman käyttäjän suostumusta [2][6][7]. Toki tähän löytyy myös poikkeuksia, kuten GDPR:n ”oikeutettu etu”. Lyhykäisyydessään oikeutettu etu mahdollistaa henkilötietojen keräämisen ja käsittelyn, mikäli se on välttämätöntä, kuten suoramarkkinointia varten. [7]

Turvallisuusviranomaisten lisäksi myös journalistit käyttävät avointen lähteiden tiedustelun menetelmiä tutkimustyössään [6]. Tästä eräs konkreettinen esimerkkitapaus sijoittuu Ukrainan sodan alkutaipaleeseen vuoteen 2015. Silloin Amerikkalainen Vice News yhdessä Brittiläisen Bellingcatin kanssa tutki ja paljasti venäläisten sotilaiden läsnäolon Ukrainassa, siis silloin kun Venäjä vielä kiisti sotilaidensa läsnäolon Ukrainan maaperällä [8]. Vice News uutistoimiston Simon Ostrovsky tutki erään venäläisen sotilaan sosiaalisen median tiliä VKontakte -palvelussa. Kävi ilmi, että hän todellakin oli siihen aikaan Venäjän asevoimien palveluksessa ja että hänen yksikkönsä oli lähetetty Ukrainaan.

Kuva 1.1: Bato Dambaev ja Simon Ostrovsky Ukrainassa täsmälleen samassa paikassa. Kuvakaappaus Vice Newsin dokumenttielokuvasta ”Selfie soldiers: Russia checks into Ukraine” [8]

Kuvassa 1.1 vasemmalla on edellä mainittu Venäjän asevoimien palveluksessa oleva henkilö ja oikealla toimittaja Simon Ostrovsky. Bato Dambaev oli ottanut vasemmalla olevan kuvan itsestään vuoden 2014 helmikuussa Ukrainan Vuhlehirskissä. Sitä ennen hän oli ottanut useita kuvia Venäjän asevoimien uniformu päällänsä itsestään Venäjällä. Näyttää siltä, että hän paljasti vahingossa Venäjän joukkojen läsnäolon Ukrainassa, vaikka siihen aikaan Venäjän presidentti Vladimir Putin vielä tämän tosiasian kiisti, kuten Vice Newsin dokumenttielokuvassa käy ilmi [8].

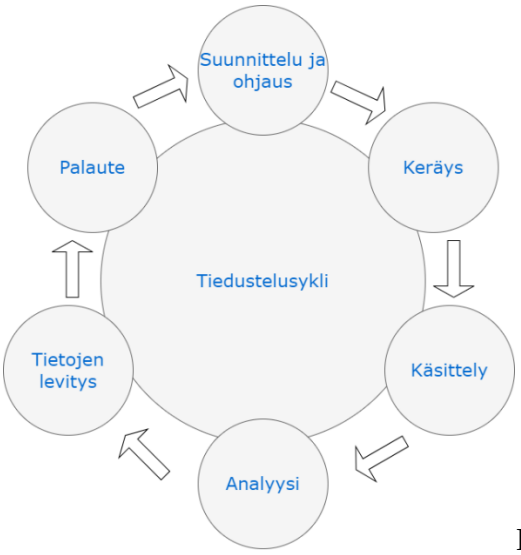
Tästä tutkivan journalismin esimerkistä voi huomata ensinnäkin sen, miten tehokasta avointen lähteiden tiedustelu voi olla hyvinkin arkaluonteisten asioiden selvittämisessä ja toisaalta tämä myös näyttää kuinka monikäyttöinen kyseinen menetelmä voi olla. Avointen lähteiden tiedustelu sopii siis yhtä hyvin tutkivien journalistien työkaluksi, eikä vain turvallisuusviranomaisten käyttöön.

Myös Puolustusvoimissa on reagoitu siihen, että varusmiespalveluksessa olevat kansalaiset ja toisaalta myös kantahenkilökunta saattaa tahattomasti paljastaa arkaluonteisia tietoja koko

maailman nähtäväksi. Suomen turvallisuuden kannalta onneksi tähän on myös osattu reagoida. Uhkakuvaan on reagoitu muun muassa siten, että Puolustusvoimien virallisessa yleisessä palvelusohjesäännössä (YLPALVO) on otettu älypuhelimista koituvat tietoturvariskit huomioon. YLPALVO:ssa mainitaan muun muassa, että matkapuhelima käyttäessä ”…operaatio- ja palvelusturvallisuus ei saa vaarantua.” (YLPALVO 2017, 183). YLPALVO:ssa huomautetaan myös, että mitään tallenteita tai muutakaan mediatiedostoja ei saa tehdä ilman yksikön päällikön antamaa lupaa. Vaikka varusmies voi vapaa-ajalla kuvata kasarmilla, silloinkin on otettava huomioon puolustusvoimien turvallisuus- ja salassapitomääräykset ja myös muut lait, jotka koskevat viranomaisten toiminnan julkisuutta (YLPALVO 2017, 184). Voisi spekuloida, että nämä määräykset ovat asetettu nimenomaan vihamielisen toimijan kohdistamat avointen lähteiden tiedustelun vaarat mielessä. Olisi melko noloa ja pahimmassa tapauksessa jopa vaarallista, jos varusmiehet tai ammattisotilaat vahingossa paljastaisivat jotain arkaluonteista tietoa Puolustusvoimista ja sen protokollista tai kalustosta.

2.2 Tiedustelun vaiheet

Tiedustelun yhteydessä voi usein törmätä termiin ”tiedustelusykli” (engl. intelligence cycle). Tiedustelusykli tarkoittaa prosessia tai mallia, missä tarkoitus on päästä raa’asta datasta informaatioon ja edelleen aina käyttökelpoiseen tiedustelutietoon asti. Sitä voi soveltaa myös OSINT-tiedusteluun. Tiedon avulla asianmukaiset tahot voivat myöhemmin tehdä johtopäätöksiä ja tarvittaessa ryhtyä asianmukaisiin toimiin. Eräs määritelmä tiedostelusyklille menee niin, että se koostuu kuudesta askeleesta: suunnittelu ja ohjaus, keräys, käsittely, analyysi ja tuotanto, tietojen levitys sekä palaute. [6]



Kuva 1.2 Tiedustelusykli

Suunnittelu ja ohjaus -vaiheessa valmistellaan OSINT-operaatiota pohtimalla, mitä operaatiolla on tarkoitus saavuttaa ja samalla määritellään, minkälaista tiedustelutietoa on ylipäätään tarkoitus saada operaation lopputuloksena. Tyypillisesti OSINT-operaatio voi olla asiakkaan antama tehtävänanto [6], mutta esimerkiksi armeijan kontekstissa se tulisi todennäköisesti käskynä johtoportaalta. Tehtävänä voi olla esimerkiksi henkilön tai tapahtuman uhka-arviointi [6]. Henkilön uhka-arviointi voi tarkoittaa esimerkiksi Suojelupoliisin tekemää turvallisuusselvitystä yksityishenkilöstä, joka hakee turvallisuusselvityslaissa määriteltyihin tehtäviin.[1] On hyvin mahdollista, että SuPo käyttää turvallisuusselvityksissä myös avointen lähteiden tiedustelua henkilöitä arvioidessaan. Suunnittelu ja ohjaus -vaiheessa on myös tärkeää löytää tiedustelukohteelle tunnisteita (engl. identifiers), jotka voivat henkilökohteiden tapauksessa olla esimerkiksi käyttäjänimiä, oikeita nimiä tai sähköpostiosoitteita [6]. Periaatteessa mitkä tahansa yksilöivät tiedot voivat käytännössä toimia tunnisteina. Näitä tunnisteita käytetään seuraavassa vaiheessa, kun aloitetaan varsinainen datan keruu [6].

Keräys-vaiheessa operaation fokus siirtyy itse datan keräämiseen. Tarkoituksena on hakea systemaattisesti dataa käyttäen suunnitteluvaiheessa määriteltyjä tunnisteita. Datasta on mahdollista tehdä löydöksiä, joiden perusteella voi myöhemmin tehdä oivalluksia ja johtopäätöksiä rakentaakseen niin ikään kokonaiskuvaa kohteesta tai tilanteesta. OSINT-operaatiossa voidaan käyttää esimerkiksi erilaisia hakukoneita, joilla voi manuaalisesti hakea tiedustelutietoa käyttäen määriteltyjä tunnisteita, tai vaihtoehtoisesti niin sanottuja ”web crawler” -botteja. [6][10][13] Nämä botit selaavat annettujen ohjeiden mukaan internet-sivuja systemaattisesti. Botit ovat hyödyllisiä siinä vaiheessa, kun avointa tiedusteludataa on hyvin paljon saatavilla, mutta sen manuaaliseen läpikäymiseen ei riitä aikaa tai resursseja [10]. Kuvitellaan tilanne, jossa Turun yliopiston verkkosivuja haluttaisiin tutkia OSINT-menetelmillä. Tavoitteena voisi olla kerätä mahdollisimman paljon tietoa Turun yliopistosta ja sen henkilökunnasta. Tässä tunnisteina voisi toimia joko sivun domain, eli utu.fi, tai sitten sivuston IP-osoite. Yksi vaihtoehto olisi manuaalisesti navigoida selaimella tähän osoitteeseen, mutta voisi olla tehokkaampaa automatisoida prosessia käyttämällä web crawleria, joka käsketään keräämään yhteystietoja, osoitteita ja muuta asianmukaista tietoa, joilla olisi käyttöä tässä operaatiossa. Web crawler -botti voi aloittaa yliopiston pääsivulta ja

1 https://supo.fi/miksi-turvallisuusselvitys-tehdaan

edelleen seurata linkkejä myös muille yliopiston sivuille etsiäkseen ja kerätäkseen dataa. Näin operaatiossa voi vähentää merkittävästi resursseja ja aikaa, kun ainakin osa operaatiosta automatisoidaan.

Käsittely-vaiheen tarkoitus tiivistettynä on prosessoida kaikki edellisessä vaiheessa kerätty raaka data *informaatioksi* [6]. Tämä tarkoittaa käytännössä sitä, että data käsitellään ihmisille ymmärrettävään muotoon, jotta sitä voidaan tulkita ja analysoida seuraavassa vaiheessa. Jos data on vieraskielistä tekstiä, se saatetaan joutua kääntämään. Välillä data voi olla jollakin tavoilla kryptistä, esimerkiksi salattua ja silloin se täytyy dekoodata. Tässä vaiheessa voidaan myös verifioida, onko käsiteltävä data edes varteenotettavaa tai oleellista tutkinnan kannalta. [6]

Analyysin aikana edellisessä vaiheessa saadusta informaatiosta pyritään samaan tiedustelutietoa [6]. Tiedustelutiedon saaminen on OSINT-operaatioiden päätavoite [2]. Informaatiosta saadaan tiedustelutietoa yhdistelemällä, analysoimalla ja arvioimalla informaatiota [6]. Kuten luvun 2.1 esimerkkitapauksessa käy ilmi, informaatiota yhdistelemällä, analysoimalla ja arvioimalla voi luoda hyvin tarkan kokonaiskuvan kohteesta. Journalisti onnistui informaation pätkiä yhdistelemällä selvittämään jopa kohteen perheenjäseniä ja asuinpaikan [8]. Vice Newsin toimittaja onnistui selvittämään alueen, missä tutkinnan kohde Bato Dambaev asuu tai asui, mutta aluksi ilman tarkempaa sijaintia. Kuitenkin kyselemällä naapuruston asukkailta, journalistille selvisi pian kohteen osoite ja asunto, mistä löytyi itse kohteen sijaan hänen vaimonsa [8]. Tässä kohtaa on taas hyvä tarkastella avointen lähteiden tiedustelun eettisiä pulmia. Oliko todellakin nämä kaikki henkilötiedot tarpeen paljastaa koko maailmalle dokumenttielokuvan kautta?

Tietojen levitys ja **palaute** ovat OSINT-operaation viimeisiä vaiheita. Tietojen levitys tarkoittaa tässä tapauksessa esimerkiksi sitä, että operaatiosta tehdään raportti, joka välitetään sitten siitä kiinnostuneelle taholle tai tahoille [6]. Esimerkiksi SuPon turvallisuusselvityksen tulokset voidaan välittää eteenpäin sen tilanneelle organisaatiolle. Palaute-vaihe ei ole varsinaisesti mikään virallinen tiedustelusyklin vaihe, ainakaan CIA:n tiedustelusyklissä.[2] Palautevaihe on ikään kuin retrospektiivi suoritetusta operaatiosta [6].

[2] https://www.cia.gov/spy-kids/parents-teachers/docs/Briefing-intelligence-cycle.pdf

2.3 Avointen lähteiden tiedustelun suosion perusta

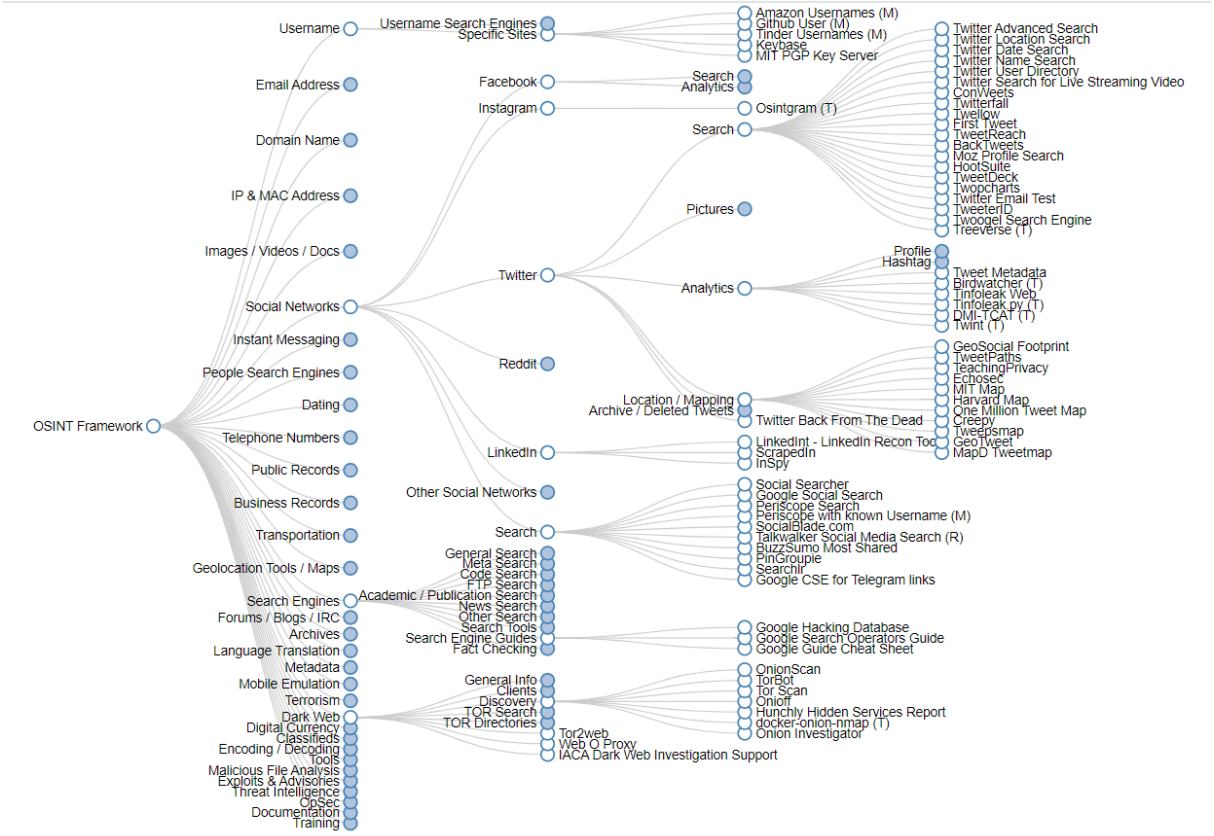
Monet turvallisuusviranomaiset pitävät edelleen avointen lähteiden informaatiota sekundana verrattuna perinteisien tiedustelumetodien tuottamaan salaiseen tietoon. Kuitenkin esimerkiksi Yhdysvaltojen tiedusteluyhteisössä on hiljattain alettu uudelleenarvioimaan avointen lähteiden informaation arvokkuutta verrattuna salaiseen informaation. Uudelleenarvioinnissa on noussut esille muutamia seikkoja, kuten valtavan ja helposti saatavilla olevan informaation vaikutus tiedusteluun sekä avointen lähteiden tiedustelun työkalujen kehitys ja niiden käyttöönotto. [1]

Yhdysvalloissa myös 9/11-iskut vaikuttivat OSINT-menetelmien hyödyn uudelleenarviointiin sikäläisessä tiedusteluyhteisössä. Tavoitteena oli nimenomaan ennustaa ja ennaltaehkäistä terrori-iskuja ja toisaalta myös välttää tiedustelun epäonnistumiset jatkossa. [3] Olisi harkitsematonta jättää tehokkaaksi osoittautuva tiedustelumetodi pois turvallisuuspalveluiden arsenaalista. Eräs tärkeimpiä seikkoja onnistuneessa virkavallan operaatiossa on kerätä ajankohtaista, luotettavaa ja toimintakelpoista tiedustelutietoa, joka liittyy meneillä olevaan tutkintaan tai operaatioon [10]. On myös huomattu, että avointen lähteiden informaation kerääminen ja analysointi on edullisempaa ja vähemmän riskialtista kuin muissa tiedustelun metodeissa [1]. Edullisempi ja riskittömämpi toimintamalli on luultavasti houkutteleva vaihtoehto varsinkin valtion turvallisuuspalveluille kuten turvallisuusviranomaisille ja poliisille, sillä nämä organisaatiot ovat kuitenkin riippuvaisia julkisesta rahoituksesta. On vaikea uskoa, että julkisen rahoituksen nojalla toimivilla viranomaisilla olisi mitään syytä ottaa turhia riskejä tai olla säästämättä kuluissa tai henkilöstöresursseissa.

Vaikka avointen lähteiden tiedustelu syntyi jo ennen internetin vallankumousta, silloin kun tiedustelun kohteena oli esimerkiksi uutiset ja radiolähetykset [6], tämä metodi on saanut uuden käyttötarkoituksen siitä lähtien kun internetin käyttö on alkanut yleistymään kaikkialla maailmassa. Pidän avointen lähteiden tiedustelunmetodien kasvavaa suosiota nimenomaan internetin vallankumouksen tuloksena, sillä internet on täynnä avointa dataa. Koskaan ennen ei ole ollut näin paljon avointa dataa saatavilla kuin nyt internetin aikakaudella. Tiedustelun keskeinen ongelma ei siis enää ole datan puute, vaan sen tehokas louhinta, yhdistely ja jäsentelyn tehokkuus, johon OSINT-menetelmät tarjoavat loistavat työkalut.

Eräs toinen syy OSINT-menetelmien kasvavaan suosioon niin turvallisuusviranomaisten kuin myös yrityksien ja journalistien keskuudessa voi olla tiedustelutyökalujen helppo saatavuus.

Kuvassa 1.2 on kuvakaappaus *osintframework.com* -verkkosivulta, joka toimii käytännössä kokoelmana enemmän tai vähemmän käyttökelpoisille OSINT-työkaluille.



Kuva 1.3 Kuvakaappaus sivustolta osintframework.com

Kuten kuvasta 1.3 voi nähdä, avointen lähteiden tiedusteluun on hyvin paljon erilaisia työkaluja tarjolla. Melkein kaikki sivustolla näkyvät työkalut ovat joko kokonaan tai lähes ilmaisia käyttää. Tämä luonnollisesti madaltaa kynnystä harjoittaa tätä menetelmää, sillä tällä menetelmällä voi selkeästi säästää sekä aikaa, rahaa että muita resursseja, kuten ihmistyövoimaa.

2.4 Esimerkkejä OSINT-työkalujen käytöstä

Näitä työkaluja käyttäen myös yksityishenkilöt voivat ilman kovin suurta teknistä ymmärrystä vakoilla toisia ihmisiä. Sivustolta löytyy muun muassa työkaluja kuvan exif-datan (Exchangable image file format) analysoinnille. Exif-datalla tarkoitetaan käytännössä esimeriksi digitaalisen kuvan *metatietoja*, josta voi selvittää tietoja käytetystä laitteesta ja sen sijainnista kuvan ottohetkellä. Näillä tiedoilla internettiin julkaistun kuvan avulla voi siis jopa jäljittää ihmisiä, jos on vain tarpeeksi dataa saatavilla. Tosin jotkut sosiaalisen median

palvelut kuten Instagram ja Facebook pyyhkivät näitä metatietoja pois julkaisujen kuvista ja videoista.

Make	samsung
Model	SM-A536B
Focal length	5.2 mm
Aperture	1.8
Exposure	1/100
ISO	50
Flash	No Flash

Kuva 1.4 Exif-dataa itse ottamastani kuvasta ImgOps.com -palvelussa.

Kaikista kuvista ei enää saa sijaintitietoja, vaikka ne olisi kameran asetuksissa kytketty päälle. Sijaintitiedot ovat yleensä koordinaatti muodossa. Voi olla, että laitteen valmistajat salaavat näitä tietoja, etteivät ne päädy vääriin käsiin, kuten kyberrikollisille. Muita mahdollisesti hyödyllisiä tietoja kyllä löytää käyttämällä osintframework.com sivustolta löydettyjä työkaluja. Kuvassa 1.4 on joitain tietoja kuvasta, joka on otettu Samsung älypuhelimella. Käytetty palvelu on nimeltään ImgOps, joka kokoaa erilaisia kuvankäsittelytyökaluja kuten efix-datan tutkimiseen ja niin sanotulle käänteiselle kuvahaulle tarkoitettuja työkaluja yhteen.

Google Hacking Database

Filters | Reset All

Show 15

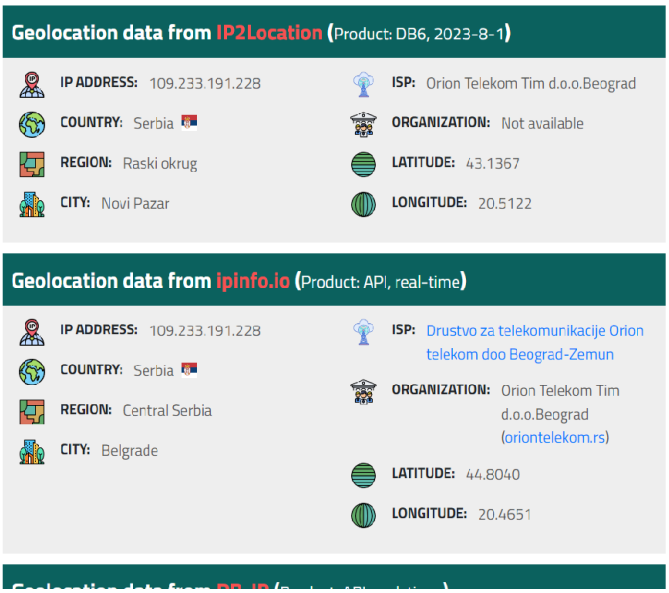
Quick Search webcam

Date Added	Dork	Category	Author
2023-11-07	intitle:"Webcam" inurl:WebCam.htm	Various Online Devices	s Thakur
2023-01-31	intitle:"Index of /webcam/"	Files Containing Juicy Info	Shuvrosayar Das
2022-06-17	inurl:webcam site:skylinewebcams.com inurl:roma	Various Online Devices	Simone Gasparato
2021-11-09	intitle:"webcamXP" inurl:8080	Various Online Devices	Krishna Agarwal
2021-10-19	intitle:"webcamXP 5" inurl:admin.html	Various Online Devices	César Hernández Obispo
2021-09-29	intitle:"webcam" "login"	Pages Containing Login Portals	Yash Singh
2021-09-15	intitle:"yawcam" "It's a webcam!" "user" "pass"	Various Online Devices	Mugdha Peter Bansode
2021-08-20	inurl:/multi.html intitle:webcam	Various Online Devices	Anmol K Sachan
2021-05-28	intitle:"webcamxp" "Flash JPEG Stream"	Various Online Devices	Anmol K Sachan
2021-05-25	inurl:mobile.html intitle:webcamXP	Various Online Devices	Anmol K Sachan
2021-04-30	intitle:"Web Client" inurl:"webcamera.html"	Various Online Devices	J. Igor Melo

Kuva 1.5 GHDB-palvelun tuloksia hakusanalla ”webcam”

Toinen hyvin yleinen ja tehokas työkalu avointen lähteiden tiedusteluun on nimeltään GHDB (Google Hacking Database/Google Dorks). GHDB käytännössä hakukyselyjen indeksipalvelu, jota hyödynnetään julkisesti saatavilla olevan tiedon löytämiseksi. Palvelusta voi hakea oikeastaan ihan mitä tahansa, esimerkiksi haavoittuvaisia webbikameroita. Kuten kuvasta 1.5 voi havaita, hakusana ”webcam” antaa paljon tuloksia hakukyselyille, joita käyttäen on mahdollista löytää webbikameroita internetistä syöttämällä hakukyselyn google-hakuun. Voisi kuvitella, että tämä palvelu olisi enemmän hyödyksi yksityisille toimijoille, kuten journalisteille ja kyberrikollisille valtion turvallisuusviranomaisteen sijaan, koska palvelua voi käyttää tehokkaasti muun muassa haavoittuvaisien IoT-laitteiden tai tietovuodoissa paljastuneille käyttäjätunnuksien ja salasanojen löytämiseen.

Osintframework.com sivustolta löytyy myös muita hyödyllisiä työkaluja, kuten työkaluja IP-osoitteiden analysointiin. IP-osoite voi toimia yksilöivänä tietona GDPR:n mukaan [7]. IP-osoitteen avulla voi muun muassa määrittää laitteen sijainnin vaihtelevalla tarkkuudella. Tähän hyödyllisiä työkaluja ovat muun muassa IP2Location.com tai iplocation.net.



Kuva 1.6 kuvakaappaus iplocation.net palvelusta

Kuvassa 1.6 on hakutulos GHDB:n avulla löydetyn webbikameran IP-osoitteelle. Kuten kuvasta näkyy, webbikameran sijainnille tarjotaan useita eri sijainteja. Sijainti kerrotaan koordinaattien muodossa, jotka voi syöttää esimerkiksi Google Mapsiin paikantaakseen kyseisen webbikameran. Yhtä hyvin tätä palvelua voi käyttää minkä tahansa muun IoT-laitteen paikantamiseen. Oleellista tässä on se, että OSINT-operaatiossa voi käyttä paljon

erilaisia työkaluja ja tuloksia yhdistelemällä pienistäkin tiedonrippeistä voidaan saada jotain hyödyllistä tiedustelutietoa irti.

3 Yksityisyyden suoja ja GDPR

Tänä päivänä ihmiset luovuttavat paljon tietoa itsestään erilaisille internetissä toimiville palveluille. Toisinaan tieto voi olla arkaluonteista tai yksilöivää ja siten myös mahdollisesti arvokasta kolmansille osapuolille, esimerkiksi suoramarkkinoinnin tai tiedustelun kannalta. Siksi eri puolilla maailmaa kuten Euroopassa ja Yhdysvalloissa on kehitetty ja asetettu erilaisia kansalaisten yksityisyyden suojaa ja tietosuojaa parantavia lainsäädäntöjä. Tietosuoja-asetukset täytyy huomioida kaikenlaisessa liiketoiminnassa ja muussakin toiminnassa, jossa käsitellään yksityishenkilöiden tietoja. Eräs tunnetuin ja tiukoin esimerkki tällaisesta lainsäädännöstä on GDPR [11].

GDPR (General Data Protection Law) on Euroopan Unionin kehittämä lainsäädäntö, jonka tavoite on parantaa EU:n kansalaisten ja EU:ssa asuvien ihmisten yksityisyyden suojaa ja turvallisuutta. Asetus tuli voimaan toukokuussa 2018. Lain tavoite on suojata henkilötietoja internettiä käyttäessä. [11]

Internet-käyttäjistä kerätään henkilötietoja muun muassa silloin kun käyttäjä tekee ostoksia verkkokaupassa, käyttää sosiaalisen median palveluja ja myös silloin, kun hän asioi verkkopankissa tai muissa tunnistautumista vaadittavissa palveluissa. Myös internet-sivun avaamisen yhteydessä välittyy yksilöiviä tietoja, kuten käyttäjän IP-osoite ja päätelaitteen teknisiä tietoja kuten esimerkiksi näytön koko.

GDPR koskee niin yksityisiä, kuin myös julkisia palveluita ja organisaatiota ja se on voimassa myös *EU:n ulkopuolella* olevilla yrityksillä ja organisaatiolla, silloin kun ne käsittelevät EU:n kansalaisen tai EU:ssa asuvan henkilön tietoja [11]. Kyseisen tietosuoja-asetuksen rikkojille voidaan määrätä tuntuvia sakkoja ja uhreilla on mahdollisuus haastaa rikkojat oikeuteen [11].

Euroopan Unioni haluaa GDPR:n avulla ottaa tiukan linjan kansalaistensa tietosuojaan ja tietoturvaan, varsinkin kun tänä päivänä ihmiset luovuttavat hyvin paljon henkilötietoja ja muuta arkaluonteista ja arvokasta tietoa erilaisilla internetissä toimiville yrityksille [11]. Tämä on hyvin perusteltua, sillä ovathan tietomurrot ja niistä koituvat seuraamukset olleet melko paljon uutisten otsikoissa Suomessakin, kuten psykoterapiakeskus Vastaamon tietomurto on osoittanut.[3] Koska GDPR:n rikkojille voi olla luvassa ankara rangaistus, se on hyvä keino

[3] https://yle.fi/t/18-318971/fi

ennalta-ehkäistä tai ainakin saattaa tekijät vastuuseen, mikäli henkilötietojen keräyksestä koituu suoraan tai epäsuoraan harmia ihmisille.

GDPR tuo mukanaan paljon tärkeitä määritelmiä, joita pohditaan myöhemmin avointen lähteiden tiedustelun yksityisyyden suojan ongelmien kautta. Tietosuoja-asetuksen mukaan *henkilötiedot* (engl. personal data) ovat sellaista tietoa, joka viittaa johonkin henkilöön, tai jonka avulla hänet voidaan suoraan tai epäsuorasti yksilöidä. Kun henkilötietoja käsitellään, eli kun puhutaan henkilötietojen käsittelystä (engl. data processing) on kyseessä kaikki toimenpiteet, jossa kerätään, nauhoitetaan, organisoidaan, jäsennellään, säilytetään tai käytetään henkilötietoja. *Rekisterinpitäjä* (engl. data controller) on henkilö, joka päättää, miksi ja miten henkilötietoja kerätään. Hän voi olla tietoja keräävän yrityksen tai organisaation johtaja tai työntekijä. *Henkilötietojen käsittelijällä* (engl. data processor) tarkoitetaan kolmatta osapuolta, joka käsittelee henkilötietoja rekisterinpitäjän sijaan. [11]

GDPR:ssä on seitsemän tietosuojaa koskevaa periaatetta [11]:

1. Tietojen käsittelyn täytyy olla lainmukaista, reilua ja läpinäkyvää käsittelyn kohteelle
2. Tietoja käsitellään vain niihin tarkoituksiin, joista on kerrottu käsittelyn kohteelle
3. On kerättävä ja käsiteltävä vain minimimäärä tietoja
4. Tiedot täytyy pitää tarkkoina ja ajantasaisina
5. Tietoja saa säilyttää vain niin kauan, kun niitä tarvitaan aiemmin määriteltyyn käyttötarkoitusta varten
6. Käsittely täytyy toteuttaa niin, että asianmukainen turvallisuus, luottamuksellisuus ja eheys ei vaarannu, esim. salauksen avulla
7. Tilivelvollisuus: *rekisterinpitäjä* on vastuussa näiden sääntöjen noudattamisesta

Avointen lähteiden tiedustelun tapauksessa on hyvin selvää, että menetelmä rikkoo GDPR-asetusta useammassakin kohdassa. Kuvitellaan tilanne, missä yksityinen yritys päättää ottaa OSINT-menetelmiä käyttöön tutkiakseen ihmisten ostoskäyttäytymistä verkossa. Yritystä kiinnostaa tietää, että mitä tuotteita selataan ja kuka niitä selaa sen verkkokaupassa. Tätä varten yritys on kehittänyt suunnitelman OSINT-operaatiolle, joka mukailee luvussa 2.2

määriteltyä tiedustelusykliä. Tässä tapauksessa suunnitteluvaiheen tunnisteina käytettäisiin esimerkiksi verkkokauppaan tallennettuja tietoja, kuten sähköpostiosoitetta ja yhteystietoja, joihin lukeutuu asiakkaan koko nimi, osoite ja puhelinnumero. Tavoite on yksilöidä verkkokaupan käyttäjä ja pitää kirjaa mitä tuotteita hän selaa, jotta tuotteita voitaisiin myydä hänelle tehokkaammin esimerkiksi kohdistetun kampanjoinnin avulla. Käyttäjältä ei kysytä lupaa missään vaiheessa. Käyttäjän vierailun tiedot tallennettaan tietokantaan, josta ne voidaan myöhemmin hakea käsiteltäväksi. Tietoja ei salata, eikä niillä ole asetettu mitään aikarajaa, kuinka kauan niitä pidetään tietokannassa. Yritys suunnittelee myös myyvänsä tietoja niistä kiinnostuneille kolmansille osapuolille.

Tässä kuvitteellisessa esimerkkitapauksessa yritys rikkoo hyvin montaa kohtaa GDPR:ssä. Tietojen kerääminen tapahtuu ilman käyttäjän suostumusta, eikä yritys edes kerro aikeisteen asiakkaallensa. Yritys ei myöskään kerää vain minimi määrää tietoja, mutta myös tarpeetonta tietoa siihen nähden, että tavoite on vain tehostaa myyntiä käyttäjälle. Esimerkiksi sähköpostiosoite riittäisi varsin hyvin kohdennetulle mainonnalle, mutta yritys kerää myös ylimääräisiä henkilötietoja. Henkilötietoja ei myöskään salata, mikä rikkoo GDPR:n periaatteiden kohtaa 6. Pahimmassa tapauksessa nämä tiedot voisivat päätyä vaikka kyberrikollisten käsiin tietomurron yhteydessä ja siitä voisi koitua vakavia seurauksia asiakkaalle.

Tällainen kuvitteellisessa esimerkissä tapahtuva laiton toiminta ei välttämättä rajoitu vain asiakkaisiin. Myös työntekijät saattavat olla epäoikeudenmukaisen ja laittoman tiedonkeruun- ja käsittelyn kohteena. Lokakuussa 2019 pikamuotijätti H&M jäi kiinni työntekijöihinsä kohdistuneesta laittomasta tiedonkeruusta- ja käsittelystä. Tapaus sijoittui Saksan Nurembergiin ja H&M saikin 35 miljoonan euron sakon rikkomuksesta, GDPR:n artikloihin 5 ja 6 vedoten. Laitonta tiedonkeruuta oli harjoitettu ainakin vuodesta 2014 lähtien. Tiedot olivat monessa tapauksessa arkaluonteisia, kuten yksityiselämää koskevia tietoja, tietoa työntekijän uskonnollisesta vakaumuksesta sekä yksityisiä terveystietoja. Tiedot oli kerätty muun muassa työntekijöiden ja esihenkilöiden välisien haastattelujen kautta, ja tietoja käytettiin esimerkiksi työntekijän kyvykkyyttä uralla etenemiseen arvioidessa. Kaikki tiedot tallennettiin verkkolevyille, eikä työntekijöille kerrottu mitä tiedoilla tehdään, vaikka GDPR:ssä nimenomaan määrätään näin kertomaan. Tietojen avulla työntekijöistä tehtiin yksityiskohtaisia profiileja, joita käytettiin esimerkiksi edellä mainituissa arvioinneissa. [12]

4 Avointen lähteiden tiedustelun vaikutus tavalliseen internet-käyttäjään

OSINT-menetelmiä voi yhtä hyvin käyttää rikollisten tai terroristien lisäksi myös muihinkin kohteisiin. Yksityiselle sektorilla onkin jo alettu kehittämään ja otettu käyttöön ratkaisuja, joiden tarkoitus on saada avoimesta datasta kaupallista hyötyä [15]. Kohteena voi olla myös yrityksen työntekijät, kuten luvun kolme H&M esimerkkitapauksessa kävi ilmi. Oikeastaan kaikki internetin käyttäjät voivat syystä tai toisesta joutua OSINT-menetelmien kohteeksi, mikäli joku organisaatio voisi siitä hyötyä, eikä sitä vastaan voi kovin helposti suojautua.

Suojautuminen OSINT-menetelmiä vastaan voi olla nimittäin hyvin hankalaa. Esimerkiksi viime aikoina on kehitetty algoritmi, joka pystyy yksilöivästi tunnistamaan 99,98 % amerikkalaisista käyttäen julkista dataa anonymisoiduista tietokannoista [16]. Ajatuksena tässä on se, että nimenomaan suuria määriä varsin mitättömiltä vaikuttavia tietoja yhdistelemällä voidaan luoda kokonaiskuva, josta voi saada paljon informaatiota irti. OSINT-menetelmiä voi näin ollen käyttää myös ihmisten uudelleentunnistamisessa sellaisissa tapauksissa, jossa alkuperäinen data on anonymisoitu [16]. Tässä kohtaa on aiheellista pohtia, onko tiedon anonymisointi enää tarpeeksi järeä toimenpide internet-käyttäjien yksityisyydensuojan turvaamiseksi. Yleisesti ottaen mitä vähemmän julkisesti avointa dataa ihmisestä on saatavilla, sitä hankalampaa on kohdentaa avointen lähteiden tiedustelumenetelmiä häneen. Siinä tapauksessa paras keino suojautua on olla julkaisematta mitään yksilöiviä tai arkaluonteisia tietoja itsestään mihinkään palveluun. Hyvä tapa vähentää niin sanottua hyökkäyspinta-alaa voisi olla esimerkiksi käyttämällä nimimerkkiä oikean nimen sijaan [6], asettamalla sosiaalisen median tilit yksityiseksi rajaten julkaisuja näkevää yleisöä sekä pysymällä kokonaan erossa tietyistä sosiaalisen median palveluista kuten Facebookista.

4.1 Laillisuus

Avointen lähteiden tiedustelun laillisuus riippuu pitkälti siitä, että kuka tiedustelua harjoittaa. Yksityisiin yrityksiin, yksityishenkilöihin ja muihin organisaatioihin pätevät eri säännöt kuin turvallisuusviranomaisiin kuten Suojelupoliisiin tai muuhun virkavaltaan. Turvallisuusviranomaisilla on siis lain mukaan vapaat kädet käyttää erilaisia tiedustelumenetelmiä, mikäli siihen on syytä [13,14]. Sen sijaan kaikki muut kuin edellä mainitut toimijat voivat mahdollisesti ja myös melko todennäköisesti rikkoa lakia avointen

lähteiden tiedustelumenetelmiä käyttäessä. Kuten kolmannessa luvussa kävikin jo ilmi, OSINT-operaatiot voivat olla ristiriidassa muun muassa GDPR:n tai jonkun toisen tietosuoja-asetuksen kanssa muun muassa siksi, että näissä operaatiossa kerätään henkilörekisteriä ilman kohteen suostumusta.

Tosin myös viranomaisten harjoittamassa tiedustelussa on harmaa alue. Vaikka viranomaisilla olisi valtuudet kohdentaa avointen lähteiden tiedustelua tai jotain muuta vakoilua yksittäiseen rikoksesta epäiltyyn henkilöön tai ryhmään, laajemmat OSINT-haut kuten massavalvonta on vieläkin kiistanalainen aihe [6].

Toinen kiistanalainen aihe liittyy avointen lähteiden informaatioon. Jos tarkastellaan esimerkiksi melkeinpä mistä tahansa sosiaalisesta mediasta löytyviä käyttäjien julkaisuja, ne voidaan periaatteessa luokitella avoimeksi ja julkiseksi. Mutta useimmat sosiaalisen median palvelut vaativat käyttäjältä rekisteröintiä ja sen yhteydessä tapahtuvaa käyttöehtojen hyväksyntää. Joissain tapauksissa vain rekisteröityneet käyttäjät pääsevät selaamaan muiden julkaisua ja näin ollen julkaisut eivät olekaan enää käytännössä kaikille avoimia. Tämä sotii tavallaan koko *avointen* lähteiden tiedustelun ideaa vastaan. Sen lisäksi jotkut palvelut, kuten Meta (Facebook) itseasiassa kieltää tyypillisesti OSINT-menetelmissä käytettyjen automaattisten työkalujen käytön. [6] Facebook myös yrittää aktiivisesti estää niin kutsuttua datan haravointia (web scraping) palveluissaan.[4]

4.2 Eettisyys

Avointen lähteiden tiedustelun eettiset ongelmat liittyvät lähinnä ihmisten yksityisyyden suojaan. Arveluttavasta tiedonkeruusta voi koitua esimerkiksi sosiaalista haittaa kuten mainehaittaa tiedustelun kohteelle [16]. Esimerkiksi vuonna 2012 Saudi-Arabialainen bloggaaja ja aktivisti Hamza Kashgari joutui pulaan Twitter julkaisustaan, kun hän väitetysti loukkasi profeetta Muhammedia. Hän yritti paeta Saudi-Arabiasta Uuteen-Seelantiin, mutta jäi Malesiassa viranomaisten haltuun, josta hänet lähettiin takaisin Saudi-Arabiaan [3]. Hamza oli tunnettu aktivisti kotimaassaan, joten voisi spekuloida, että hänen sosiaalisen median tilejään valvottiin tarkasti viranomaisten puolesta. Tällaisesta poliittisesta vainoamisesta onkin ollut jo paljon keskustelua monien tutkijoiden joukossa aiheen ajankohtaisuuden vuoksi [13].

[4] https://about.fb.com/news/2022/09/deterring-scraping-by-protecting-facebook-identifiers/

Yksi suuri ongelma avointen lähteiden tiedustelun eettisyydessä on se, että onko kerätty tieto ylipäätään varteenotettavaa [14]. Avoimista lähteistä, kuten sosiaalisesta mediasta tai terroristien viestintäkanavista kerätty tieto voi olla harhaanjohtavaa ja sen validiteetti on usein hankalaa varmistaa [14]. Jos tiedustelun kohde on tietoinen siitä, että häntä vakoillaan, hän saattaa yrittää harhauttaa tiedustelijoita esimerkiksi harhaanjohtavilla julkaisuilla. Tämän vuoksi OSINT-menetelmistä kerätty tieto ei pitäisi yksinomaan riittää kenenkään tuomitsemiseen oikeudessa, vaikka tällaisiakin tapauksia löytyy. Esimerkiksi vuonna 2010 vihainen matkustaja Robin Hoodin lentokentällä Englannissa purki turhautumistaan Twitterissä myöhästyneestä lennosta uhkailemalla räjäyttävänsä lentokentän [4]. Kyseinen julkaisu oli ilmeisesti tarkoitettu vitsiksi, mutta poliisi otti sen hyvin vakavasti ja twiitin kirjoittaja pidätettiin sen toimesta. Hän sai twiitistään tuomion, joka myöhemmin kuitenkin peruttiin, mutta kirjoittaja ei kuitenkaan säästynyt muilta seuraamuksilta, sillä hän menetti silloisen työpaikkansa [4].

Eikä pitäisi myöskään pitää itsestään selvänä sitä, että some-kirjoituksen takana on juuri se henkilö, joka tilin omistaa. On kuitenkin mahdollista, että tili on kaapattu jonkun vihamielisen tahon toimesta, jonka intresseissä voisi olla pilata tilin omistajan maine.

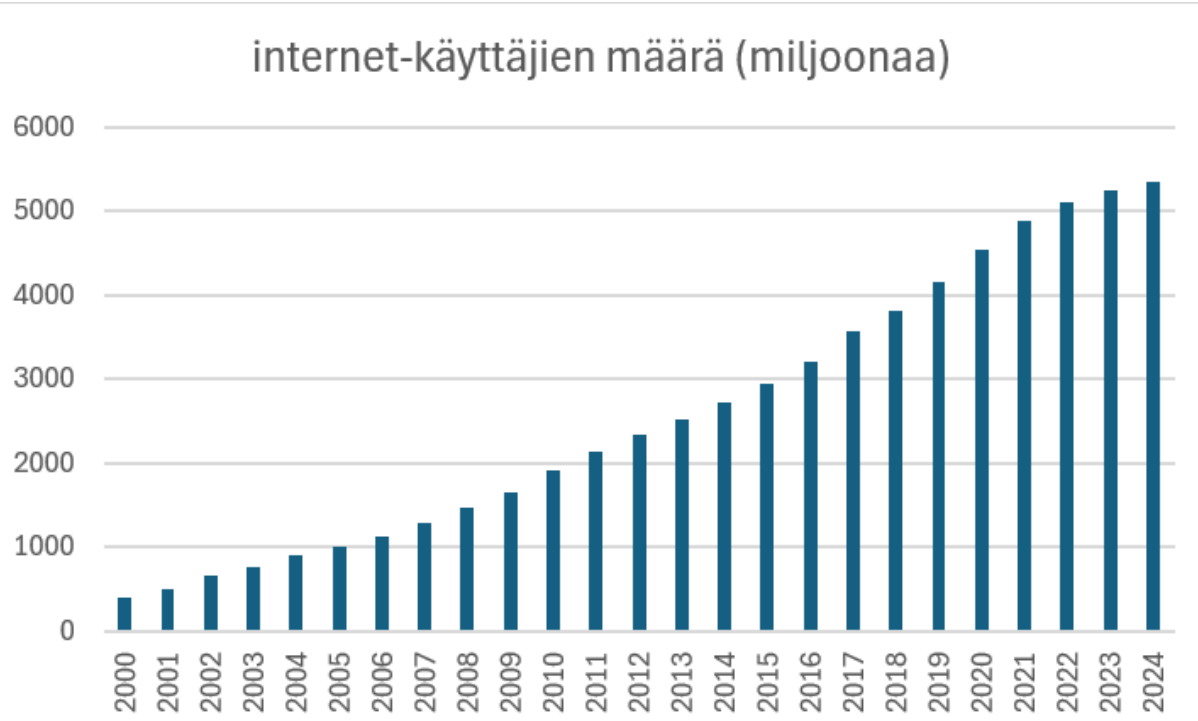
OSINT-menetelmät voivat toimia työkaluina myös massavalvonnalle. Yleensä ihmisten massavalvontaa harjoitetaan esimerkiksi valvontakameraverkkoja hyödyntäen, mutta esimerkiksi Kiinassa on jo yhdistetty valvontakameraverkoilla valvonta myös avointen lähteiden tiedusteluun [17]. Tätä kutsutaan toisinaan digitaaliseksi autoritaarisuudeksi (engl. digital authoritarianism). Kiinassa tarkkaillaan sosiaalisen median lisäksi myös kansalaisten ostokäyttäytymistä [17]. Esimerkiksi kiinalaiset teknologiajätit kuten Tencent ja Ant Group, jotka keräävät dataa yli miljardin kiinalaisen käyttäytymisestä, ostoksista ja viestinnästä, ovat käytännössä maan hallinnon kontrollissa. Nämä yhtiöt ovat velvollisia raportoimaan viranomaisille kaiken mitä he pyytävät [17]. Tässä kohtaa voikin jo sanoa, että avointen lähteiden tiedustelua käytetään ihan tavalliseen internet-käyttäjään ja sillä on hyvin vakavia seurauksia käyttäjien yksityisyyden suojalle. Kiinan tapauksessa tavallisille internet-käyttäjille ei ole juuri ollenkaan yksityisyyden suojaa.

Avointen lähteiden tiedustelu voi siis vaikuttaa tavallisiin internet-käyttäjiin muun muassa silloin, kun auktoritaarinen hallinto valvoo kansalaisiaan ja mahdollisesti käyttää tiedustelutietoja poliittisiin vainoihin tai sortamiseen [17]. Sen lisäksi yksityisien yrityksien

harjoittama OSINT-tiedustelu voi antaa yrityksille kaupallista hyötyä avoimesta datasta, jota ihan tavalliset internet-käyttäjät tuottavat internettiin [15].

5 Yhteenveto ja pohdintaa

Tässä kandidaatintutkielmassa perehdyttiin avointen lähteiden tiedusteluun erilaisien näkökulmien kautta. Avointen lähteiden tiedustelu sisältää siis paljon erilaisia tiedonkeruumenetelmiä, jossa hyödynnetään julkisista lähteistä kerättyä dataa, jota onkin runsaasti tarjolla nyt internetin-aikakaudella. Avoimen datan määrä tulee todennäköisesti vielä lisääntymään, kun internet vakiintuu ihmisten käyttöön eri puolilla maailmaa, jonka seurauksena myös sosiaalisen median palvelujen ja keskustelupalstojen käyttäjämäärät kasvavat. Tämän myötä myös kyseessä olevan tiedustelumenetelmän suosio tulee todennäköisesti kasvamaan vielä entisestään niin viranomaisten kuin myös yksityisien toimijoiden joukossa. Kuvassa 1.7 on havainnollistettu internet-käyttäjien määrän kehitystä viimeisten 24 vuoden aikana.



Kuva 1.7 internet-käyttäjien kokonaismäärän kehitys vuosina 2000–2024 [19]

Johdannossa esiteltiin kaksi tutkimuskysymästä, joihin lähdettiin etsimään vastauksia erilaisia lähteitä käyttämällä. Ensimmäinen tutkimuskysymys oli:

”**Mikä selittää avointen lähteiden tiedustelumenetelmän suosion tänä päivänä?”**

Tutkimalla avointen lähteiden tiedustelun teknisestä ja käytännönläheisestä näkökulmasta ja sen tyypillisiä käyttötarkoituksia, käyttökohteita ja muita ominaisuuksia tutkielmassa

päädyttiin siihen tulokseen, että avointen lähteiden tiedustelun suosiota selittää ennen kaikkea se, että menetelmässä hyödynnettävän avoimen datan määrä on lisääntynyt hyvin paljon internetin vallankumouksen ja WWW:n käytön yleistyttyä. Eli tiedustelutietoa on tarjolla hyvin paljon, ja sitä voidaan käsitellä automaattisesti erilaisia työkaluja hyödyntäen. Lisäksi menetelmä on osoittautunut monella tavalla halvemmaksi sekä joissain tapauksissa tehokkaammaksi kuin muut tiedustelumenetelmät. Myös turvallisuusuhat ja niiden kehittyminen on siirtynyt yhä enemmän ja enemmän internettiin, mikä tarkoittaa sitä, että juuri tällaisia internettiä hyödyntäviä tiedustelumetodeja tarvitaan enemmän. Toisaalta suosiota selittää myös se, että näitä menetelmiä käytetään myös muiden kuin turvallisuusviranomaisten toimesta, kuten yksityisten yritysten, jotka tavoittelevat kaupallista hyötyä, tai tutkivien journalistien. jotka yrittävät kerätä tietoa tutkimuksiinsa. Selkeästi menetelmä on hyvin monikäyttöinen ja myös kustannustehokas, niin kuin aiemmissa luvuissa on käynyt ilmi.

Toinen tutkimuskysymys muodostui siitä oletuksesta, että OSINT-menetelmät ovat jo laajassa käytössä, mikä osoittautui oikeaksi ensimmäisen tutkimuskysymyksen ratkettua. Toinen tutkimuskysymys oli:

”**Miten avointen lähteiden tiedustelu vaikuttaa tavallisen internet-käyttäjän yksityisyyteen?**”

OSINT-menetelmiä käytetään tyypillisesti valtion viranomaisten toimesta, kuten aiemmissa luvuissa on käynyt ilmi. Viranomaisilla kuten Suojelupoliisilla ja Poliisilla on lupa käyttää erilaisia tiedustelumenetelmiä työssään silloin kun niiden käyttö on perustelua. Niissä tapauksissa tiedustelun kohde ei ole tavallinen internet-käyttäjä, vaan rikoksesta epäilty henkilö, joten tiedustelumenetelmien käyttö on hyvin perusteltua. Yksityisten toimijoiden, kuten yrityksien ja journalistien käsissä OSINT-menetelmien käyttö voi olla laillisesti kyseenalaista muun muassa GDPR:n tai muiden tietosuoja-asetuksien kautta katsottuna sekä eettisesti arveluttavaa, varsinkin kun kyseessä olevia menetelmiä vastaan on todellisuudessa hyvin hankala suojautua ja niiden käyttö väistämättä loukkaa ihmisten yksityisyyden suojaa jollakin tasolla.

Etenkin luvussa 4 esiteltyjen esimerkkitapausten valossa on hyvin selkeää, että avointen lähteiden tiedustelun kohteena voi olla käytännössä kuka tahansa. Tiedustelun kohteena voi olla esimerkiksi yrityksen asiakkaiden lisäksi sen omat työntekijät. Suuret yritykset voivat tehdä yhteistyötä auktoritaarisen hallinnon kanssa, jolloin seuraamukset ovat tavallisten

internet-käyttäjien kannalta vakavia. Avointen lähteiden tiedustelu antaa mahdollisuuden kansalaisten massavalvontaan joko sellaisenaan, tai yhdistettynä muihin tiedustelumenetelmiin kuten Kiinan esimerkissä kävi ilmi. Tämä vaarantaa tavallisen-internetkäyttäjän yksityisyyden suojan, sillä käyttäjän julkaisemien tietojen perusteella hänestä voidaan tehdä profiili, jota voidaan käyttää kaupallisissa tarkoituksessa esimerkiksi myymällä tietoja eteenpäin niistä kiinnostuneille kolmansille osapuolille. Käyttäjän tietoja voi myös käyttää auktoritaarisen hallinnon toimesta vainoamiseen tai sortamiseen.

Vaikka avointen lähteiden tiedustelusta herää huolia sen eettisyydestä ja laillisuudesta, on tässä menetelmässä kieltämättä myös paljon hyötyä yhteiskuntarauhan turvaamiseksi. Näiden menetelmien avulla on pystytty virkavallan käsissä muun muassa ratkomaan murhamysteereitä, jossa tekijät ovat kerskailleet teoillaan sosiaalisessa mediassa sekä havaitsemaan ihmiskauppaan viittaavaa toimintaa [18]. Sen lisäksi menetelmillä on paljastettu myös laittomia sotilasoperaatioita, kuten luvussa 2 kerrottiin.

Kaiken kaikkiaan avointen lähteiden tiedustelulla pitäisi olla paikka viranomaisten arsenaalissa rikoksia ja terrorismia torjuessa. Kuitenkin tätä menetelmää pitäisi käyttää eettisesti ja lain puitteissa. OSINT-menetelmillä saatu tiedustelutieto pitäisi pystyä yhdistämään muihin tietoihin kattavan kokonaiskuvan luomiseksi. On selvää, että tavallisten internet-käyttäjien massavalvonta on eettisesti hyvin arveluttavaa sillä se loukkaa myös sellaisten ihmisten yksityisyyden suojaa, ketkä eivät ole syyllistyneet rikokseen, eli silloin tiedustelu on tavallaan perusteetonta ja moraalisesti väärin. GDPR:n avulla voi säädellä yksityisen toimijoiden aikeita käyttää tätä menetelmää, mutta turvallisuusviranomaisten tapauksessa lainsäädäntöä olisi hyvä tarkastella uudelleen, jotta massavalvontaa ei sovellettaisi esimerkiksi ihmisoikeuksia polkevien auktoritaaristen hallintojen kuten Kiinan, Venäjän tai tiettyjen arabimaiden toimesta.

Lähteet

[1] Best Jr., Richard A. & Cumming, Alfred. (2007). Open source intelligence (OSINT): Issues for Congress, Congressional Research Center (CRS), ss. 1-10

[2] Koops, Bert-Jaap & Hoepman, Jaap-Henk & Leenes, Ronald. (2013). Open-source intelligence and privacy by design. Computer Law & Security Review. 29. ss. 676–688. 10.1016/j.clsr.2013.09.005.

[3] Eijkman, Quirine & Weggemans, Daan. (2013). Open Source Intelligence and Privacy Dilemmas: Is it Time to Reassess State Accountability?. Security and Human Rights. 23. ss. 285-296 10.1163/18750230-99900033.

[4] Trottier, Daniel. (2015). Open source intelligence, social media and law enforcement: Visions, constraints and critiques. European Journal of Cultural Studies. 18. ss. 530-547. 10.1177/1367549415577396.

[5] Finlex telekuuntelu, televalvonta ja tekninen tarkkailu, Poliisilaki 872/2011 (viitattu 1.4.2024), https://www.finlex.fi/fi/laki/ajantasa/2011/20110872

[6] Böhm, Isabelle & Lolagar, Samuel. (2021). Open source intelligence: Introduction, legal, and ethical considerations. International Cybersecurity Law Review. 2. 10.1365/s43439-021-00042-7 ss. 318-334

[7] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

[8] Selfie soldiers: Russia checks into Ukraine, Vice News, 17.6.2015 https://video.vice.com/en_us/video/selfie-soldiers-russia-checks-into-ukraine/55ba5014018008e821c71e52

[9] Ünver, Hamid Akın. (2018). Digital Open Source Intelligence and International Security: A Primer. 10.13140/RG.2.2.16242.56000 ss. 1-20.

[10] Akhgar, Babak & Bayerl, Petra & Sampson, Fraser. (2016). Open Source Intelligence Investigation: From Strategy to Implementation. 10.1007/978-3-319-47671-1 ss. 7-74.

[11] Wolford, Ben. What is GDPR, the EU's new data protection law? GDPR.eu, 7.11.2018 (viitattu 1.4.2024) https://gdpr.eu/what-is-gdpr/

[12] European Data Protection Board, Hamburg Commissioner Fines H&M 35.3 Million Euro for Data Protection Violations in Service Centre, 2.10.2020. (viitattu 15.4.2024) https://www.edpb.europa.eu/news/national-news/2020/hamburg-commissioner-fines-hm-353-million-euro-data-protection-violations_en

[13] Lakomy, Miron. (2023). Open-source intelligence and research on online terrorist communication: Identifying ethical and security dilemmas. 10.1177/17506352231166322.

[14] Staniforth, Andrew. (2016). Police Use of Open Source Intelligence: The Longer Arm of Law. 10.1007/978-3-319-47671-1_3. ss. 21-31

[15] Hassan, Nihad & Hijazi, Rami. (2018). The Evolution of Open Source Intelligence. 10.1007/978-1-4842-3213-2_1 ss. 1-17.

[16] Pastor-Galindo, Javier & Nespoli, Pantaleone & Gomez Marmol, Felix & Martinez Perez, Gregorio. (2020). The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2965257 ss. 10283-10302.

[17] Qiang, Xiao (2021). Chinese Digital Authoritarianism. The Project on Middle East Political Science partnered with Stanford University's Center for Democracy, Development and the Rule of Law and its Global Digital Policy Incubator ss. 35-41

[18] Mateescu, Alexandra & Brunton, Douglas & Rosenblat, Alex & Patton, Desmond, Gold, Zachary & Boyd, Danah. (2015). Social Media Surveillance and Law Enforcement ss. 1-2

[19] Simon Kemp - 5 Billion Social Media Users, DataReportal 31.1.2024 (viitattu 12.4.2024) https://datareportal.com/reports/digital-2024-deep-dive-5-billion-social-media-users