Jimi Kuusimo

HARHAANJOHTAVA INFORMAATIO JA AVOIMIEN LÄHTEIDEN TIEDUSTELU DIGITAALISISSA INFORMAATIOYMPÄRISTÖISSÄ: TEKOÄLYN UHAT JA MAHDOLLISUUDET

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2024

TIIVISTELMÄ

Kuusimo, Jimi
Harhaanjohtava informaatio ja avoimien lähteiden tiedustelu digitaalisissa informaatioympäristöissä: Tekoälyn uhat ja mahdollisuudet
Jyväskylä: Jyväskylän yliopisto, 2024, 44 s.
Tietojärjestelmätiede, kandidaatintutkielma
Ohjaaja: Siitonen, Valtteri

Tutkielmassa tarkasteltiin harhaanjohtavan informaation (sisältäen mis- ja disinformaation) vaikutusta sekä tekoälyn uhkia ja mahdollisuuksia avoimien lähteiden tiedustelun toimintaan digitaalisissa informaatioympäristöissä, kuten internetissä ja sosiaalisessa mediassa. Tutkielma toteutettiin kuvailevana kirjallisuuskatsauksena, jossa avattiin laajasti kontekstia ja ympäristöä informaatiolle, informaatioympäristöille sekä avoimien lähteiden tiedustelun toiminnalle ja tarkasteltiin niiden yhteistoimintaa tekoälyn kanssa. Tutkielman tarkoitus oli selvittää tekoälyn luomia uhkia ja mahdollisuuksia harhaanjohtavan informaation käsittelyyn avoimien lähteiden tiedustelun toiminnassa. Laaja katsaus kirjallisuuteen osoitti, että avoimien lähteiden tiedustelutoiminta vastaa uusiin tekoälyllä levitettyihin harhaanjohtavan informaation muotoihin samalla tavalla, kuin se vastaa muuhunkin harhaanjohtavaan informaatioon. Tämä ei tarkoita, että toimintaan ei olisi tullut muutoksia, vaan että samoja keinoja käytetään edelleen, mutta niitä voidaan tehostaa tekoälyavusteisesti. Tämä tarkoittaa laajempaa ulottuvuutta ja tehokkaampaa toimintaa, mutta myös uusia keinoja varmistaa informaation lähde ja todenperäisyys. Myös disinformaation levittäjät ovat löytäneet uusia tapoja disinformaation tuottamiseen ja levittämiseen tekoälyavusteisesti. Tiedustelutoiminnan tulee olla tietoinen ja pyrkiä vastaamaan tähän tekoälyn avulla tuotettuun ja levitettyyn disinformaatioon omalla toiminnallaan ja hyödyntää tekoälyn luomia mahdollisuuksia informaatioympäristön luotettavuuden varmistamisessa. Tutkielman tulokset osoittavat harhaanjohtavan informaation olevan uhka ja riskitekijä nyt ja tulevaisuudessa ja osoittaa tarpeen ongelman kokoaikaiseen ja jatkuvaan tutkimiseen, jotta informaatioympäristöt voidaan pitää tulevaisuudessakin luotettavina informaation lähteinä.

Asiasanat: OSINT, avoimien lähteiden tiedustelu, disinformaatio, misinformaatio, synteettinen media, tekoäly

ABSTRACT

Kuusimo, Jimi
Misleading information and open-source intelligence in digital information environments: Threats and opportunities of artificial intelligence
Jyväskylä: University of Jyväskylä, 2024, 44 pp.
Information Systems Science, Bachelor's Thesis
Supervisor: Siitonen, Valtteri

This thesis examines the impact of misleading information (including misinformation and disinformation) and the threats and opportunities of artificial intelligence on open-source intelligence in digital information environments like internet and social media. The thesis was conducted as a descriptive literature review and describes broadly the context and environment for information, information environments, and open-source intelligence and how they function in conjunction with artificial intelligence. The purpose of the thesis was to investigate the opportunities and threats the use of artificial intelligence poses for handling misleading information in open-source intelligence. An extensive review of the literature showed that open-source intelligence is responding to new forms of misleading information disseminated and produced by artificial intelligence in the same way it responds to other forms of misleading information. This is not to say that intelligence operations have not made use of the new technologies. Open-source intelligence uses the same proven methods that have worked before but uses artificial intelligence to enhance its ability to respond and process misleading information. This means a broader reach and more efficient operations, but also new ways to ensure the source and veracity of information. Disseminators of disinformation have also found new ways to produce and disseminate disinformation using artificial intelligence. Intelligence operations must be aware of and seek to respond to this disinformation generated and disseminated by artificial intelligence through their activities and exploit the potential of artificial intelligence to ensure the reliability of the information environment. The results of this thesis show that misleading information is a threat and a risk factor now and in the future, and shows a need for full-time and continuous research into the problem to ensure the integrity of information environments going forward.

Keywords: OSINT, open-source intelligence, disinformation, misinformation, synthetic media, artificial intelligence

KUVIOT

TAULUKOT

SISÄLLYS

1 JOHDANTO

Informaation arvo muuttuvissa ja digitalisoituvissa yhteiskunnissa ei vähene. Tämä ei tarkoita, että informaatio olisi arvokkaampaa nyt kuin se oli aiemmin, mutta sen arvo ei ole vähentynyt tai tule vähenemään. Tämä väite on totta niin kauan, kun on mahdollista erotella harhaanjohtava informaatio todellisesta ja luotettavasta informaatiosta. Uudet informaation levittämiseen ja luomiseen käytetyt tekoälytyökalut muokkaavat tätä informaation varmistamisen prosessia. Tässä tutkielmassa keskitytään tarkastelemaan harhaanjohtavan informaation vaikutuksia digitaalisiin informaatioympäristöihin erityisesti painottaen tekoälyn luomia uhkia ja mahdollisuuksia. Eli toisin sanoen, miten harhaanjohtava informaatio voidaan erotella luotettavasta informaatiosta tekoälyn muokkaamissa digitaalisissa informaatioympäristöissä.

Sanastokeskuksen (2022, s. 8) mukaan informaatioympäristö ”koostuu informaatiosta ja informaatiojärjestelmistä sekä informaatiota tuottavista, hyödyntävistä ja vastaanottavista ihmisistä ja organisaatioista.” Erilaiset digitaaliset informaatioympäristöt, kuten tässä tutkielmassa erityisesti tarkasteltava internet ja sosiaalinen media, ovat viime vuosien aikana kokeneet mullistuksen synteettisen median muodossa. Synteettinen media tarkoittaa mitä tahansa mediaa, kuten videoita, kuvia, audiota tai tekstiä, minkä on luonut tai sitä on muokannut tekoäly (Hyska, 2023). Tässä tutkielmassa tarkastellaan digitaalisten informaatioympäristöjen sisällä erityisesti internetiä ja sosiaalista mediaa ympäristönä avoimien lähteiden tiedustelun (engl. open-source intelligence, OSINT) toiminnalle, koska tekoäly ja erilaiset synteettisen median muodot, kuten syvävääreennökset (engl. deepfakes), ovat erityisesti keskittyneet näihin ympäristöihin.

OSINT on tiedustelutoimintaa, joka tapahtuu avoimissa ja julkisissa ympäristöissä, kuten internetissä tai mediassa (Jardines, 2016). Avoimet informaatioympäristöt toimivat ympäristönä sekä OSINTille että erilaisille informaatiovaikuttamisen muodoille. Informaatiovaikuttaminen tarkoittaa ihmisten mielipiteisiin ja asenteisiin vaikuttamaan pyrkivää systemaattista viestintää (Valtioneuvoston kanslia, 2016, s. 13). Erityisesti harhaanjohtavan informaation levittäminen jotain poliittisia tai sotilaallisia tarkoituksia kohti on näkyvä informaatiovaikuttamisen muoto.

Tutkielmassa etsitään vastausta havaittuun ongelmaan informaation varmistamisesta todelliseksi ja luotettavaksi informaatioksi, jolla on arvoa digitaalisissa informaatioympäristöissä ja OSINTin toiminnassa sekä tekoälyn vaikutuksia näissä ulottuvuuksissa. Tämä tarkoittaa tutustumista informaation, avoimien tietolähteiden ja informaatioympäristöjen sekä OSINTin toimintaan ja toimintaympäristöihin ja kuinka nämä kaikki yhdessä uusien teknologioiden kanssa muokkaavat digitaalisia informaatioympäristöjä. Tutkielman näkökulmaksi valikoitui OSINT, koska se on kaiken muun toiminnan keskiössä ja sen toiminta on luonnostaan rajoittunut avoimiin tietolähteisiin ja informaatioympäristöihin.

Yleistyvät informaatiovaikuttamisen yritykset, helpottunut harhaanjohtavan informaation luominen ja jakaminen tekoälyn avulla sekä sen uudet muodot muuttavat informaatioympäristöjä ja luovat pohjan tämän tutkielman tarpeellisuudelle. Nämä ongelmat ovat vasta viime vuosina kehittyneet vakavasti otettaviksi riskitekijöiksi informaatioympäristöjen integriteetille. Samalla ne avaavat oven OSINTin, harhaanjohtavan informaation sekä tekoälyn yhteiselle tutkimiselle. Nämä riskitekijät muodostavat tutkielman tarkoituksen ja motivaation ja johtavat tutkielman tutkimuskysymyksiin:

1. Miten tekoälyn avulla luotu ja levitetty harhaanjohtava informaatio vaikuttaa OSINT-toimintaan?
2. Miten tekoälyä voi hyödyntää osana harhaanjohtavan informaation käsittelyä OSINT-toiminnassa?

Tutkimuskysymykset kuvaavat tekoälyn luomia mahdollisuuksia ja uhkia tiedustelutoiminnalle käsitellessä harhaanjohtavaa informaatiota. Näitä vaikutuksia tutkiessa on tärkeää tiedostaa, että ne eivät tapahdu tyhjiössä, joten tutkielmassa avataan laajasti kontekstia ja tekijöitä OSINTin ja harhaanjohtavan informaation ympärillä. Aiheen tutkimisella pyritään luomaan ajankohtainen katsaus nykyaikaiseen OSINT-toimintaan ja sen metodeihin sekä tarkastella harhaanjohtavan informaation ja tekoälyn vaikutusta näihin. Tarkoituksena on luoda myöhempää tutkimusta ja tarkastelua varten otanta nykyhetkestä, jotta tulevaisuudessa on mahdollista tarkastella toiminnan kehitystä nopeasti muuttuvissa digitaalisissa informaatioympäristöissä.

Tutkielman tutkimusmenetelmänä on järjestelmällisesti toteutettu kuvaileva kirjallisuuskatsaus, jossa tutustutaan laajasti aiheita käsittelevään kirjallisuuteen ja vastataan tutkimuskysymyksiin niistä luodun synteesin perusteella. Tämä valikoitui tutkimusmenetelmäksi, koska sen avulla voidaan tehdä luotettava laaja katsaus OSINTin, harhaanjohtavan informaation ja tekoälyn yhdistettyyn toimintaympäristöön. Tutkielma suunniteltiin alkavan vain harhaanjohtavan informaation ja OSINTin tutkimisella, mutta tekoäly valikoitui mukaan aiherajaukseen, koska sillä on vahva ja peruuttamaton asema nykyaikaisissa digitaalisissa informaatioympäristöissä. Lähdekirjallisuutta valittiin eri kriteerein riippuen aiheesta: informaatiota käsitellessä tarvittiin tukeva tieteellinen pohja, OSINTia tarkastellessa luotettavuutta ja tunnettuja toimintamalleja ja tekoälyä tarkastellessa oli tarve tarkastella uutta informaatiota uusista teknologioista, painottaen kuitenkin lähteiden tieteellistä pohjaa. Tekoäly ja sen erilaiset sovellukset

muuttuvat nopeasti, joten jo muutamia vuosia vanhat artikkelit voivat sisältää vanhentunutta informaatiota ja jo ylitettyjä ongelmia.

Tutkielmassa tarkastellut aiheet ovat yksinään laajalti tutkittuja, kuten synteettinen media, syväväärennökset ja harhaanjohtava informaatio. Viimeaikaiset muutokset informaatioympäristössä vaativat kuitenkin jatkuvasti uusia tilannetta kartoittavia katsauksia ja päivittyviä tutkimuksia. Uudet teknologiat tarkoittavat uusia haasteita, joten on tärkeää pitää kirjaa informaatioympäristöjen ja teknologian muutoksen vaiheista, jotta voidaan valmistautua tulevaan. OSINTia on tutkittu lukuisista näkökulmista, mutta sen kokemuksia tekoälyllä luotuun harhaanjohtavaan informaatio ei ole kartoitettu laajoilla katsauksilla. Samoin tekoälyä osana OSINT-toimintaa tutkivat tutkimukset ovat kohdistuneet pitkälti yksittäisten teknologioiden tutkimiseen, eikä laajoihin monia erilaisia tapoja yhdistäviin tutkimuksiin.

Tutkielmassa käytettyjä lähteitä etsittiin monista tietokannoista, kuten JYKDOKista ja Google Scholarista. Hakusanoina toimivat kaikki OSINTin ympärillä olevat, tekoälyä käsittelevät sekä harhaanjohtavaan informaatioon yhdistettävät termit ja niiden synonyymit. Erityisesti OSINTia käsitellessä jouduttiin astumaan tieteellisen ympäristön ulkopuolelle ja tarkasteltiin tieteellisissä lähteissä käytettyä pohjakirjallisuutta. Valtiollisten toimijoiden suorittamana OSINTin nykyaikaiset työkalut ja menetelmät ovat, kenties hieman paradoksaalisesti, yleisesti ei-avointa informaatiota. Nykyisin on erkaantunut myös yksityisiä tiedusteluorganisaatioita, jotka ylläpitävät julkisia kertomuksia käyttämistään resursseista ja metodeista, joihin tässä tutkielmassa turvaudutaan tutkiessa OSINTin työtapoja.

Tutkielman lähdekirjallisuutena toimii pääasiassa aiempi aihetta käsittelevä kirjallisuus, toimitetut teokset sekä vertaisarvioidut artikkelit. Lähteiden luotettavuutta arvioitiin esimerkiksi tarkastelemalla niille annettua JUFO-tasoluokitusta, tutkimalla kirjottajien tai toimittajien aiempia tuotoksia ja asemaa käsitellyn aiheen diskursseissa sekä artikkeleissa käytettyjä lähteitä. Lähteitä suljettiin myös ulkopuolelle monilla kriteereillä, esimerkiksi jos tuotos näytti selkeää puolueellisuutta tai kannanottoa erityisesti harhaanjohtavan informaation ympärillä sekä tarkastelemalla tekstin ja kirjoittajan objektiivisuutta, sponsoreita tai ilmoitettuja tai ymmärrettyjä eturistiriitoja. Uusien tekoälyä käsittelevien artikkelien kanssa viittauksien määrä ei ollut luotettava laadun arviointikriteeri, mutta laatu pyrittiin aina tarkistamaan yhdistelemällä muita aiemmin mainittuja tapoja. Lähempään laaduntarkkailuun tutkielmassa valikoitui lähdekirjallisuus tutkimalla niiden otsikoita, tiivistelmiä, löydettyjä tuloksia sekä mahdollista pohdintaa tulosten ympärillä.

Tutkielma on jaettu kolmen pääotsikon alle ja etenee (harhaanjohtavan) informaation ja OSINT-toiminnan tarkastelun jälkeen niitä yhdistävään osioon, missä tutkitaan harhaanjohtavan informaation vaikutuksia OSINT-toimintaan digitaalisissa informaatioympäristöissä. Tekoälyn vaikutuksia avataan informaation käsittelyn yhteydessä, mutta siihen palataan tarkemmin viimeisessä osiossa. Ensimmäisen pääotsikon alla määritetään informaation, avoimien tietoläh-

teiden ja digitaalisten ympäristöjen määritteitä ja niissä vaikuttavia tekijöitä. Toisessa pääotsikossa tutustutaan OSINTiin, sen toimintaympäristöihin sekä sen asemaan osana laajempaa tiedustelusykliä. Kolmannen pääotsikon sisällä tutkitaan harhaanjohtavan informaation yleisiä vaikutuksia väestöön ja tiedustelun suorittajiin, tarkastellaan miten OSINT vastaa tiedustelutoimintana tekoälyllä levitettyyn harhaanjohtavaan informaatioon ja kuinka tekoälyä voi käyttää osana OSINT-toimintaa.

2 HARHAANJOHTAVA INFORMAATIO AVOIMISSA TIETOLÄHTEISSÄ

Harhaanjohtavaa informaatiota tarkastellessa on tärkeää erottaa erilaiset informaation ilmenemismuodot ja määritellä miten niitä tarkastellaan tutkielman sisällä sekä erilaisissa informaatioympäristöissä, erityisesti digitaalisissa avoimissa tietolähteissä. Tässä osiossa tarkastellaan informaation sekä avoimien tietolähteiden käsitteitä ja asetetaan niitä määrittävät rajaukset ja määritelmät. Lisäksi tarkastellaan erilaisia informaatioympäristöjä muokkaavia tekijöitä, informaation leviämistä sosiaalisessa mediassa, leviämiseen vaikuttavia tekijöitä sekä synteettisen median vaikutusta informaatioympäristöihin.

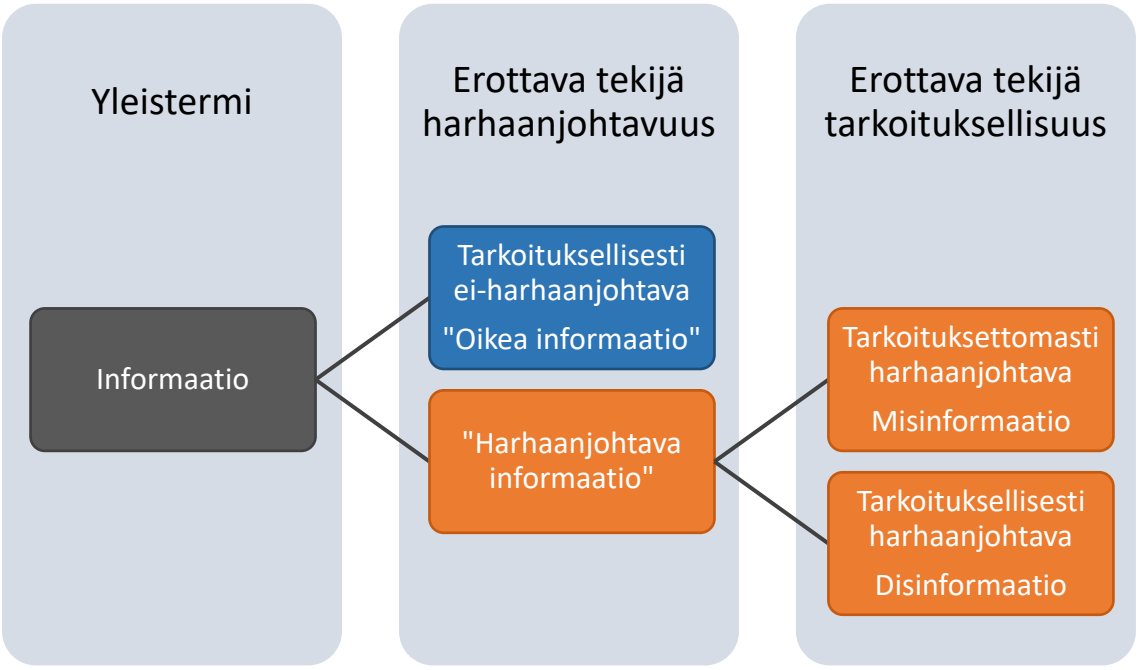
2.1 Informaation määritelmä ja muodot

Informaation jako eri muotoihin on vahvasti semanttisuuteen, eli ilmaisun merkityksellisyyteen, liittyvä ongelma ja liittyy käsitteiden määrittelyssä käytettyihin termeihin ja niiden merkityksiin. Tieteellisessä ympäristössä jako perustuu käytännön tarpeeseen varmistaa tarkoituksenmukainen kommunikaatio erilaisia tutkimuksia ja tarkkaa käsittelyä varten. Tästä huolimatta informaatiosta sekä erityisesti misinformaatiosta ja disinformaatiosta käytetään vaihtelevia määritelmiä. Tässä osiossa tutustutaan informaation eri määritelmiin ja asetetaan tutkielmassa käytetyt määritelmät.

Informaation eri muodoille on annettu lukemattomia eri määritelmiä (Dretske, 1982; Fallis, 2015; Fetzer, 2004; Floridi 2005; Fox, 1983; Scarantino & Piccini, 2010). Tutkittu lähdekirjallisuus viittaa misinformaation ja disinformaation käsitteisiin yksinkertaisin määritelmin: misinformaatio on väärää informaatiota ja disinformaatio on tarkoituksellisesti levitettyä tai tuotettua väärää informaatiota (esimerkiksi Bastick, 2021; Venema, 2024). Tässä tutkielmassa tarkastellaan informaatiota kuitenkin Søen (2019) määrittelemän mallin mukaan.

Søe (2019) erottaa informaation muodot totuusarvosta ja määrittelee informaation luonnetta määritteleviksi parametreiksi harhaanjohtavuuden (engl. misleadingness) sekä tarkoituksellisuuden (engl. intentionality). Tarkoituksena huomion siirtämisessä on se, että määritelmä antaa tilaa esimerkiksi oikeasti väärän informaation ja ironian tai sarkasmin tarkastelun välille (Søe, 2019). Tämä on tärkeää tarkastellessa erityisesti internetissä ja sosiaalisessa mediassa leviävää informaatiota.

Søe (2019) esittelee lähtökohdaksi yleistermin ”informaatio”, joka kuvaa yleisesti kaikkea informaatiota sen luonteesta tai parametreista huolimatta. Informaatioon viitataan, kun puhutaan yleisesti mistä tahansa informaatiosta, mikä on myös tässä tutkielmassa käytetty määritelmä. Informaation lisäksi määritellään tarkoituksellisesti ei-harhaanjohtava informaatio (vastaisuudessa ”oikea informaatio”), tarkoituksettomasti harhaanjohtava informaatio (vastaisuudessa misinformaatio) sekä tarkoituksellisesti harhaanjohtava informaatio (vastaisuudessa disinformaatio) (Søe, 2019). Tutkielman sisällä voidaan viitata sekä mis- että disinformaatioon yhteisellä termillä ”harhaanjohtava informaatio” silloin, kun informaation harhaanjohtavuuden tarkoituksellisuudella ei ole merkitystä tai se ei ole tiedossa. Informaation ulottuvuuksia visualisoidaan kuviossa 1.



KUVIO 1 Informaation eri muodot (Søe, 2019 mukaan muotoiltu)

Søen määrittelemä malli valikoitui tutkielmassa käytettäväksi määritelmäksi, koska sen avulla voidaan huomioida tarkemmin kontekstuaaliset merkitykset informaation jakamisessa. Esimerkiksi on tärkeää huomioida, että harhaanjohtava informaatio voi olla totuudellista informaatiota, joka on harhaanjohtavaa esimerkiksi informaation puutteellisuuden takia. Samoin esimerkiksi ironia, satiiri tai vitsit voivat olla väärää informaatiota, mutta eivät ole harhaanjohtavia, koska ne

ovat tarkoitettu huomioitavan vitseinä, mikä tekee niistä tarkoituksellisesti ei-harhaanjohtavia (Søe, 2019; Fallis, 2015).

Tärkeät käsitteet tutkielmassa informaation tarkasteluun ovat siis informaatio, oikea informaatio, misinformaatio, disinformaatio sekä harhaanjohtava informaatio, joiden määrittelevinä parametreina ovat harhaanjohtavuus ja tarkoituksellisuus. Lisäksi termiä ”tieto” käytetään tarkoittamaan jotain käyttötarkoitusta kohti jäsenneltyä ja jalostettua informaatiota, kuten tiedustelutieto. Informaatiolle voidaan määritellä muitakin muotoja, kuten propaganda, valeuutiset tai malinformaatio (Wardle & Derakhshan, 2017). Tutkielmassa näitä tarkastellaan kuitenkin Søen määrittelemien käsitteiden sisällä.

Määritelmistä huolimatta ero informaation eri muotojen välillä on vaikeaa muodostaa sosiaalisissa ja avoimissa ympäristöissä. Eron tekeminen harhaanjohtavan ja ei-harhaanjohtavan välillä on haastavaa ja tarkoituksellisuuden arviointi jopa mahdotonta. Erotus mis- ja disinformaation välille voidaan tehdä vain perustuen havaittuun ja todennettuun tietoon levittäjän motiiveista, vaikka informaatio itsessään tiedettäisiinkin harhaanjohtavaksi informaatioksi.

2.2 Informaatio avoimissa tietolähteissä

Avoimet tietolähteet on laaja käsite, joka koostuu erilaisista informaation jakelukanavista ja tietolähteistä. Avoimille tietolähteille ei ole yksittäistä määritelmää, koska niitä hyödyntävät tahot määrittelevät niitä eri tavalla. Seuraavaksi asetetaan tässä tutkielmassa käytetty määritelmä avoimelle informaatiolle sekä avoimille tietolähteille ja avataan niitä erilaisten esimerkkien avulla.

Avoin informaatio tarkoittaa julkisesti saatavilla olevaa informaatiota, mitä voidaan laillisesti hankkia keräämällä tai ostamalla (Ministry of Defence, 2023, s. 81; Jardines, 2016, s. 5; Kivimäki, 2024, s. 283). H. Gibson (2017, s. 81) painottaa, että termiä ”avoin” ei tule sekoittaa termiin ”ilmainen”, eikä avoimuus myöskään tarkoita helppoa pääsyä informaatioon. Tämä määritelmä ei siis sulje ulos informaatiota, jota voi kerätä pyytämällä, tarkkailemalla tai luomalla käyttäjätilin, kunhan toiminta on laillista. Avoimet tietolähteet tarkoittavat nimensä mukaisesti tietolähteitä avoimelle informaatiolle. Lukemisen helpottamiseksi termiä ”tietolähde” käytetään termin ”lähde” sijasta viitatessa informaation lähteisiin, ellei tämä ole muuten selvää kontekstissa.

Nykyaikaisesti isoimpiin avoimen informaation jakelukanaviin kuuluvat internet sekä sosiaalinen media (Clark, 2013; Jardines, 2016). Tutkielmassa tarkasteltavat internet ja sosiaalinen media toimivat sekä ympäristöinä että ympäristöjen sisällä jakelukanavina tai tietolähteinä informaatiolle. Internet tarjoaa paljon informaatiota lukemattomista eri lähteistä sisältäen esimerkiksi foorumeita, wikisivustoja ja tietokantoja sekä kirjoja, blogeja ja videoita (H. Gibson, 2017). World Wide Web ja internet ovat erillisiä käsitteitä, mutta tutkielman sisällä termiä internet käytetään kattamaan kaikki erilaiset internetissä olevat palvelut ja toiminnot, mukaan lukien WWW, ellei ole tarpeellista painottaa WWW:n ominaisuuksia.

WWW voidaan jakaa monin eri tavoin, kuten esimerkiksi pintaverkkoon (engl. surface web), syväverkkoon (engl. deep web) sekä pimeään verkkoon (engl. dark web) (Basheer & Alkhatib, 2021; H. Gibson, 2017). Pintaverkko tarkoittaa erilaisten hakukoneiden (esim. Google ja Bing) indeksoimia verkkosivuja (Basheer & Alkhatib, 2021), joihin pääsee suoraan näiden hakukoneiden avulla käyttämällä normaaleja selaimia (esim. Google Chrome tai Microsoft Edge). Suurin osa WWW:n informaatiosta on kuitenkin syvemmällä kuin pintaverkossa sijaitsevat informaation lähteet (Bergman, 2001).

Syväverkko ja pimeä verkko muodostavat indeksoimattoman osan internetistä, jota hakukoneet eivät voi palauttaa vastauksena hakukyselyyn (Basheer & Alkhatib, 2021; H. Gibson, 2017; Price & Sherman, 2001). Pimeä verkko tarjoaa ideaalisen alustan rikollisuudelle (Kalpakis ym., 2017, s. 113), mutta sen käyttäminen itsessään ei ole laitonta. Vaikka syväverkko ja pimeä verkko ovat enemmän piilossa ja jopa salaisia verrattuna pintaverkkoon, ovat ne silti yleisesti avoimia tietolähteitä informaatiolle. Yhdessä pintaverkon kanssa ne muodostavat maailman laajimman jakelukanavan informaatiolle, WWW:n.

Internet ja WWW eivät ole kuitenkaan ainoita avoimen informaation jakelukanavia. Taulukko 1 kuvaa tarkemmin erilaisia avoimen informaation jakelukanavia ja esimerkkejä tietolähteistä, joita eri tahot tuovat esille. Jakelukanaviin kuuluvat esimerkiksi internet, media, kaupalliset kanavat, julkinen informaatio, "harmaa kirjallisuus" sekä havainnoimalla, raportoimalla tai henkilölähteistä saatu informaatio (Clark, 2013; H. Gibson, 2017; S. D. Gibson, 2007; Jardines, 2016; NATO, 2002; Office of the Director of National Intelligence [ODNI], 2013). Tutkielmassa keskitytään pitkälti internetiin ja sosiaaliseen mediaan informaation jakelukanavina ja tietolähteinä, koska ne toimivat huomattavina ympäristöinä sekä OSINT-toiminnalle sekä tekoälyn luomille uhkille ja mahdollisuuksille. Sosiaalinen media on myös toiminut mahdollistajana laajemmalle disinformaation leviämiselle digitaalisessa maailmassa (Kivimäki, 2024).

Avoimen informaation määritelmä on edelleen laaja ja jopa ongelmallinen (S. D. Gibson, 2013; Hatfield, 2023) ja jättää paljon tulkinnan varaan, mutta on tämän tutkielman kontekstiin riittävä. Avoimen informaation ala muuttuu kuitenkin koko ajan ja on muuttunut monesti eri aikakausina. Painokone, radio ja televisio mullistivat informaation levittämisen aikanaan, kuten myös internet, WWW ja sosiaalinen media myöhemmin. Avoimen informaation määrä lisääntyy kaikkialla, ja uudet informaatioympäristöjä muokkaavat tekijät tekevät avoimen informaation määrittämisestä ja varmistamisesta yhä haasteellisempaa.

TAULUKKO 1 Informaation jakelukanavia ja lähteitä (Clark, 2013; H. Gibson, 2017; S. D. Gibson, 2007; Jardines, 2016; NATO, 2002; ODNI, 2013 mukaan kerätty)

Jakelukanava	**Esimerkkejä tietolähteistä**
Internet ja WWW Kattaa pintaverkon, syväverkon ja pimeän verkon. Sisältää myös paljon tietolähteitä muista jakelukanavista.	- Tietokannat, wikit - Sosiaalisen median palvelut - Video-, audio- ja kuvamateriaali - Nettisivut, blogit, foorumit - RSS syötteet, - Metadata, API
Media Kattaa erilaiset median muodot, joilla saavutetaan laajasti erilaisia yleisöjä. Kanavana voi toimia tietokoneet, televisio, radio tai painettu materiaali.	- Televisio-ohjelmat - Radio-ohjelmat - Aikakausi- ja sanomalehdet - Kirjat, pelit, elokuvat, sarjakuvat - Uutiset
Kaupalliset kanavat Yleensä tuotteet tai palvelut, joita jokin taho tarjoaa kaupallisesti vastiketta vastaan. Data ja informaatio sekä niiden käyttöön liittyvä tietämys ja taito voivat olla kaupallisia kohteita.	- Kaupalliset tietokannat - Satelliittikuvat ja muu data - Kaupalliset palvelut (Oxford Analytica) - Kartoituspalvelut - Riskianalyysit ja -mittaukset
Julkinen informaatio ja dokumentit Valtioiden, yritysten ja muiden toimijoiden tuottama julkinen data ja informaatio esimerkiksi selonteoista, esityksistä tai ehdotuksista.	- Asiakirjat ja raportit - Budjetit, demografiat - Lait ja lakiehdotukset - Tilinpäätökset, vuosikatsaukset - Tutkinnat - Lehdistötilaisuudet - Turvallisuusohjeet
”Harmaa kirjallisuus” Kattaa muut kuin kaupalliset julkaisut, joita yleensä levitetään rajallisesti jollekin tietylle yleisölle, esimerkiksi jonkun organisaation sisällä. Sisältää paljon tieteellistä, poliittista, sosioekonomista tai sotilaallista informaatiota. (Clark, 2013; Jardines, 2016; NATO, 2002)	- Työmuistiot - Tiedotteet - Opinnäytetyöt - Kongressijulkaisut - Patentit - Uutiskirjeet - Tuote-esitteet - Tekniset raportit - Tutkimukset
Havainnointi ja henkilölähteet Julkisesti toimivat asiantuntijat, harrastajat, ammattilaiset, erilaiset järjestöt ja muut toimijat kuten myös julkisesti havainnoitavat tapahtumat, ilmiöt ja ympäristön muutokset.	- Katutaide, graffitit - Lentokonetarkkailijat - Harrasteradistit - Journalistit - Pakolaiset - Museot - Kansalaisjärjestöt (Punainen Risti)

2.3 Digitaalisia informaatioympäristöjä muokkaavat tekijät

Informaatioympäristöissä ja erityisesti sosiaalisessa mediassa vaikuttaa lukematon määrä erilaisia tekijöitä ja toimijoita, jotka muokkaavat ympäristöjen sisältöä ja kokemusta niiden käyttämisestä. Tässä osiossa tutustutaan erilaisiin informaatioympäristöjä muokkaaviin tekijöihin sekä niiden vaikutuksiin harhaanjohtavan informaation leviämisessä. Näitä ympäristöjä muokkaavia tekijöitä ovat esimerkiksi sosiaalisen median automattiset botit sekä erilaiset uudet teknologiat, kuten viime vuosina aina vahvemmin keskusteluun noussut generatiivinen tekoäly ja sen erilaiset sovellukset, kuten syväväärennökset.

Uudet teknologiset innovaatiot avaavat uusia mahdollisuuksia disinformaation levittämiselle erityisesti sosiaalisessa mediassa (Kivimäki, 2024). Erityisen tunnettu tapa harhaanjohtavan informaation levittämiselle sosiaalisessa mediassa ovat erilaiset botit. Boteilla tarkoitetaan täysi- tai puoliautomaattisia ohjelmia ja algoritmeja, jotka vuorovaikuttavat ihmisten kanssa, tuottavat tai jakavat sisältöä ja yrittävät jäljitellä ihmisten käyttäytymistä (Ferrara ym., 2016). Bottien toimintaa on mahdollista tehostaa käyttämällä tekoälyä, tehden niistä tehokkaampia toiminnassaan ja vaikeampia tunnistaa (Arcos & Arribas, 2024). Botit eivät kuitenkaan ole ainoa syy harhaanjohtavan informaation leviämiselle.

Vosoughi ja muut (2018) kertovat harhaanjohtavan informaation leviävän helpommin ja nopeammin kuin oikea informaatio ja osoittavan ihmisten olevan syy tähän ilmiöön. Heidän mukaansa botit kiihdyttävät sekä oikean että harhaanjohtavan informaation leviämistä yhtä paljon, mutta ihmiset levittävät aktiivisemmin harhaanjohtavaa informaatiota. Lu ja muut (2022) kertovat syyksi tälle toiminnalle ihmisten tavan jakaa informaatiota tunneperäisistä syistä. Yleinen ahdistus voi toimia laukaisimena misinformaation jakamiselle ja toiveikkaat ja ennalta misinformoituja uskomuksia ylläpitävät henkilöt jakavat misinformaatiota myös helpommin eteenpäin (Lu ym., 2022). Ihmisten levittämää misinformaatiota voidaan kuitenkin vähentää lisäämällä joukkoistetun faktantarkistuksen toimenpiteitä (Drolsbach & Pröllochs, 2023).

Bottien lisäksi keskusteluun nykyaikaisesta internet-ilmapiiristä sitoutuu myös uudet tekoälypohjaiset sovellukset, jotka täyttivät lyhyessä ajassa keskustelun harhaanjohtavan informaation tuottamisesta ja levittämisestä. Tekoälyn erilaisia sovelluksia on lukemattomia, joiden takana on kymmeniä erilaisia teknologioita. Termillä tekoäly viitataan tutkielmassa kaikkiin tekoälyn eri kategorioihin, alakategorioihin ja käytettyihin teknologioihin, kuten koneoppimiseen, syväoppimiseen tai suuriin kielimalleihin. Poikkeuksena ovat erikseen nimetyt sovellukset, kuten syväväärennökset sekä generatiiviset sisällöntuotantosovellukset (vastaisuudessa tekoälysovellukset).

Harhaanjohtavan informaation kontekstissa nämä tekoälysovellukset sisältävät ison riskin erityisesti misinformaation levittämiselle. Sovellukset kuten ChatGPT, Midjourney tai ElevenLabs voivat luoda autenttisen tuntuista, mutta todellisuudessa synteettisesti luotua teksti tai audiovisuaalista sisältöä pelkän tekstimuotoisen kehotteen ja niille syötetyn datan perusteella (Gregory, 2023).

Tämä tarkoittaa, että näiden sovellusten tuottama sisältö on vain niin luotettavaa, kuin niihin syötetty data-aineisto on, riippumatta niiden näennäisesti älykkäästä toiminnasta.

Dienin mukaan (2023) nämä tekoälysovellukset eivät sisällä minkäänlaista päättelykykyä tilanteisiin, joihin niitä ei ole erikseen ohjelmoitu hallitsemaan. Hänen mukaansa nämä tekoälysovellukset eivät kykene hallitsemaan luotettavasti kysymyksiä tai ongelmia, joita niille syötetty data-aineisto ei ole käsitellyt ja ratkaissut. Hänen mukaansa sovellus voi luoda epäjohdonmukaisia, mutta luotettavilta vaikuttavia synteettisiä vastauksia kysymyksiin, joihin se ei voi tietää vastausta, koska sille syötetty data-aineisto ei ole käsitellyt asiaa. Tekoälysovellusten tuottamasta harhaanjohtavasta, puutteellisesta tai virheellisestä informaatiosta käytetään nimitystä "hallusinaatio" (Dien, 2023; Monteith ym., 2023). Tämän tekoälyn luoman synteettisen informaation leviämistä voi nopeuttaa informaation paikkansapitävyyden kyseenalaistamattomuus (Monteith ym., 2023).

Tekoälysovellukset voivat luoda todennäköisiä ja mahdollisia väitteitä ilman konkreettisia perusteita väitteiden takana, luoden illuusion väitteen paikkansapitävyydestä (Monteith ym., 2023; Sobieszek & Price, 2022). Tekoälysovelluksella luodusta informaatiosta on vaikeaa tarkistaa lähteiden todenperäisyyttä, sillä tekoälymallit eivät yleisesti erottele käyttämiään lähteitä data-aineistosta. Tekstiä luovat sovellukset yleisesti syntetisoivat vastauksen monista eri lähteistä (Dien, 2023), eivätkä tarjoa informaatiota siitä, onko sen tarjoama vastaus varmasti luotettava (Monteith ym., 2023; Sobieszek & Price, 2022) tai koottu luotettavista lähteistä (van Dis ym., 2023). Sovellukset voivat myös vaihtaa vastaustaan aiemmin kysyttyihin identtisiin kysymyksiin tai luoda synteettisiä lähteitä tukemaan annettuja vastauksia (Dien, 2023).

Näitä tekoälysovelluksia käytetään yleisesti persoonalliseen käyttöön, ilman tarkoitusta levittää harhaanjohtavaa informaatiota. Sobieszek ja Price (2022) kumoavatkin pelkoa tekoälysovellusten käytöstä kohti pahantahtoisia toimia ja uskovat tekoälysovellusten luomien häiriötilanteiden johtuvan käyttäjien välinpitämättömyydestä misinformaatiota kohtaan ja sovelluksista itsestään. Toisaalta he tuovat esille mahdollisuuden syntyvästä noidankehästä, missä ihmiset käyttävät tekoälyllä luotua virheellistä sisältöä aineistona kirjoittamiseen, jota myöhemmin käytetään sovelluksien uutena data-aineistona, ja näin syntyvä kierre lopulta saastuttaa informaatioympäristöt. Tekoälyä voi kuitenkin käyttää myös toimintaan, jolla yritetään aiheuttaa häiriötilanteita tai levittää disinformaatiota.

Syväväärennökset tarkoittavat tekoälyllä luotua aidolta tuntuvaa synteettistä mediaa, joka on muokattu olemassa olevan audiovisuaalisen materiaalin perusteella (Ienca, 2023; Venema, 2024). Syväväärennöksien yleisin muoto on synteettiset ja muokatut videot, missä henkilön piirteet vaihdetaan toisen henkilön piirteisiin, esimerkiksi vaihtamalla tai muokkaamalla kasvoja (Groh ym., 2021; Mirsky & Lee, 2021). Visuaalinen harhaanjohtava informaatio koetaan yleisesti luotettavammaksi kuin pelkkä tekstimuotoinen (Hameleers ym., 2020) ja voidaan kokea tällaisena voimakkaammin ja pitkäkestoisemmin (Weikmann &

Lecheler, 2023b), joten visuaaliset syvävääreennökset ovat riskitekijä informaatioympäristöissä.

Toisaalta syvävääreennöksillä voidaan jäljitellä henkilöiden ääntä ja puhetta ja voivat olla vaarallisia nimenomaan visuaalisuuden puutteen takia (Weikmann & Lecheler, 2023a, s. 10). Syvävääreennöksien potentiaalisia haittoja ovat esimerkiksi harhaanjohtavan informaation levittäminen, ihmisten yksityisyyden loukkaaminen tai väärennetyn todistusaineiston luominen (Ienca, 2023). Syvävääreennöksien realistinen luonne tekee niiden leviämisestä misinformaationa helppoa, koska ihmiset eivät tajua sisällön olevan synteettisesti luotua (Venema, 2024).

Botit, synteettinen media, tekoälysovellukset ja syvävääreennökset luovat paljon potentiaalisia ongelmia harhaanjohtavan informaation leviämiseen. Teknologisista uhkakuvista huolimatta ihmiset ovat kuitenkin vielä pääsyyllisiä harhaanjohtavan informaation leviämiseen sekä sosiaalisen median ympäristössä että tekoälysovellusten luoman synteettisen median kanssa. Harhaanjohtava informaatio on haasteellista tunnistaa, mutta sen käsittely ja varmistaminen on oleellinen osa OSINTin toimintaa.

3 OSINT JA SEN TOIMINTAYMPÄRISTÖT

Tiedustelutieto tarkoittaa analysoitua ja jalostettua informaatiota, jonka tarkoitus on olla hyödyllistä päätöksentekijöille (Goldman & Maret, 2016). Tämä tiedustelun tuottama tieto on monen tekijän summa ja koostuu tiedustelun toimintaympäristössä vaikuttavista toimijoista, määritellystä tiedustelutarpeesta sekä informaation käsittelystä tiedustelutiedoksi. Näiden lisäksi tässä osiossa tutustutaan tiedustelun ja OSINTin toimintaan, tiedustelua määrittäviin ja rajoittaviin tekijöihin sekä tiedusteluprosessiin.

3.1 OSINT

Tässä osiossa tutustutaan OSINTin määritelmään ja toimintaan. Yleisesti tiedustelutoiminta on informaation käyttämistä jonkin tiedustelutarpeen tai tiedustelupyynnön (engl. request for information) täyttämiseksi. Tästä informaatiosta muuttuu tiedustelutietoa, kun sitä kerätään ja käytetään jonkin tiedustelutarpeen täyttämiseksi, eli se on kerätty, prosessoitu ja kohdennettu täyttämään asetetut tai ymmärretyt päätöksentekijöiden tarpeet (Goldman & Maret, 2016; Lowenthal, 2019).

Modernista medianäkyvyydestään huolimatta tiedustelu on vanha ilmiö ja monet historialliset tekstit alleviivaavat tiedustelun käyttöä osana poliittista ja sotilaallista toimintaa. Sun Tzu painottaa informaation keräämisen ja hyödyntämisen tärkeyttä jo 400 vuotta ennen ajanlaskun alkua historiallisessa teoksessa Sodankäynnin taito, missä hän omistaa kokonaisen luvun vakoojien käytön tärkeydelle informaation keräämisen välineenä. Tiedustelu ja tiedustelutoiminta ovat kehittyneet yhteiskuntien mukana ja muotoutuneet vastaamaan päätöksentekijöiden tarpeisiin aikakausien määrittämällä tavalla aina historiasta nykyaikaan asti ja eteenpäin.

OSINT ei myöskään ole uusi tiedustelun muoto (Andrew, 2018; Block, 2023), vaikka nykyisin tietomäärältään laajin avoimen informaation jakelukanava on

nykyaikainen internet (Clark, 2013). OSINT-toiminnalle on vaikeaa asettaa historiallista alkupäivämäärää, koska on mahdotonta määritellä, mikä oli eri ajanjaksoina nykyisellä OSINTin määritelmällä julkista ja avointa informaatiota (Block, 2023). Nykyaikaisessa muodossaan, omana tiedustelulajinaan ja osana muita määriteltyjä tiedustelulajeja, OSINT voidaan määritellä tiedustelutiedoksi, jota tuotetaan aiemmin määritellyistä avoimista tietolähteistä, jaetaan oikealle yleisölle sekä hyödynnetään tiedustelutarpeen täyttämiseksi (ODNI, 2013, s. 46). Jardines (2016) laajentaa tätä määritelmää jakamalla sen viiteen avainvaatimukseen, joiden perusteella OSINT-toimintaa voidaan määritellä tarkemmin.

Jardinesin (2016) asettamat kaksi ensimmäistä avainvaatimusta koskevat informaation avointa luonnetta. Ensimmäinen vaatimus määrittelee yksinkertaisesti jo aiemmin määritellyn avoimen informaation käyttämisen osana tiedustelutoimintaa. Toinen vaatimus jatkaa tätä määritelmää, painottaen avoimen informaation vaatimusta olla laillisesti kerätty. Informaation keräämiseen ja käsittelyyn ei saa olla mitään laillisia esteitä, eikä informaation keräämisessä saa käyttää laittomia keinoja, kuten varastamista, hakkerointia, sosiaalista manipulointia (engl. social engineering) tai henkilön yksityisyyden loukkaamista (Jardines, 2016).

Jardinesin (2016) kolme jälkimmäistä avainvaatimusta määrittävät informaation keräämistä ja sen käyttöä osana tiedustelutoimintaa. Kolmas vaatimus määrittää, että kerätty informaatio tulee tarkistaa luotettavaksi ja ei-harhaanjohtavaksi. Tähän syynä on neljäs vaatimus, joka määrittelee avoimet tietolähteet välikäden tietolähteiksi (Jardines, 2016). Informaation on siis kerännyt, järjestellyt ja julkaissut tuntematon toimija, joka seuraa omia motivaatioitaan tälle toiminnalle (Jardines, 2016). Verrattuna muihin tiedustelulajeihin OSINTissa painottuu tiedottomuus tietolähteen luotettavuudesta informaation keräämisessä. Jardinesin (2016) viides vaatimus OSINTin märittelylle on sen tarve täyttää jokin tiedustelutarve. Hänen mukaansa tämä vaatimus erottaa OSINTin esimerkiksi avoimien lähteiden tutkinnasta, johon tutustutaan myöhemmin tutkielmassa.

OSINT-toiminnassa käytettyä informaatiota kuvaa myös sen suuri määrä (Clark, 2013; Jardines, 2016). Clarkin (2013) mukaan tämä määrä voi olla ongelma, sillä oleellisen ja hyvän materiaalin löytäminen voi osoittautua vaikeaksi. Hänen mukaansa informaation järjestely, todentaminen ja varmistus ovat haasteita, jotka korostuvat erityisesti internetissä suuren informaatiomäärän takia. Tiedustelua suorittavan analyytikon on mahdotonta käyttää hyödyksi kaikkea saatavilla olevaa materiaalia (Clark, 2013), mutta tekoälysovellukset ovat mahdollinen ratkaisu tähän ongelmaan. Tekoälyn käyttämiseen osana tiedustelutoimintaa palataan myöhemmin tutkielmassa.

OSINTia ei kuvaa pelkästään kerätyn informaation määrä, vaan sen kattama osuus lopullisesta tiedustelutiedosta: arvioiden mukaan 80 % käytetystä tiedustelutiedosta tulee avoimista tietolähteistä (S. D. Gibson, 2013; Hulnick, 2010). Tästä huolimatta avoimista tietolähteistä tuotettu tiedustelutieto mielletään toissijaiseksi verrattuna muiden tiedustelulajien tuottamaan tiedustelutie-

toon (Clark, 2013; S. D. Gibson, 2013). Analyytikot tunnustavat salaisen tiedustelutiedon huomattavasti luotettavammaksi kuin avoimista tietolähteistä kerätyn identtisen tiedustelutiedon (Pedersen & Jansen, 2019).

Avoimet tietolähteet nähdään lisäksi uhkina, koska kuka tahansa voi käyttää niistä löytyvää informaatiota ilman erityistä lupaa (Clark, 2013). OSINT on tiedustelulajina kenen tahansa käytettävissä, koska avoimet tietolähteet ja avoin informaatio ovat kenen tahansa tarkasteltavissa ja hyödynnettävissä. Jardines (2016, s. 7) painottaa kuitenkin, että tiedustelutiedon arvoa ei mitata sen perusteella, miten haastavaa se on kerätä, vaan tarkastelemalla miten hyvin se vastaa tiedustelutarpeeseen. Tähän tiedustelutarpeeseen vastataan käyttämällä muitakin mahdollisia tiedustelulajeja yhdessä OSINTin kanssa, joten tiedustelulajit täydentävät toisiaan ja ovat osa suurempaa kokonaisuutta. Tiedustelulajit tukevat toisiaan, toimivat osittain samoilla alueilla ja vaikuttavat samoissa ympäristöissä.

3.2 OSINTin toimintaympäristöt

OSINTin toimintaympäristöön vaikuttaa monet erilaiset tekijät, kuten tiedustelua suorittavat ja hyödyntävät tahot, muut tiedustelulajit sekä muut ympäristössä toimivat tekijät. Tässä osiossa tutustutaan kaikkeen tähän ja tarkastellaan OSINTin toimintaympäristön vaikutusta tiedustelutoimintaan. OSINT on tiedustelutoimintana laaja-alainen ja koskettaa kaikkia muita tiedustelulajeja, mutta voi tästä huolimatta olla yksinään riittämätön täyttämään tiedustelupyynnön tarpeet.

OSINT tunnetaan tiedustelulajina ensimmäisenä keinona, jolla pyritään vastaamaan tiedustelupyyntöön, koska sen tuottamaa tiedustelutietoa on helppo kerätä, käsitellä ja levittää (Clark, 2013; Jardines, 2016). Tämän lisäksi se tunnetaan mahdollisesti viimeisenä keinona ja jopa kaikkina keinoina näiden välissä (Mercado, 2004, s. 49). OSINT on siis läsnä koko tiedusteluprosessissa ja sillä pyritään hahmottamaan aukkoja, mitä muut tiedustelulajit voivat täyttää (Jardines, 2016) tai pyritään täyttämään aukkoja, mitä muut tiedustelulajit jättivät (Clark, 2013). Jardinesin (2016, s. 9) mukaan OSINT voi täydentää muita tiedustelulajeja eri tavoin, esimerkiksi:

- Henkilötiedustelussa (engl. human intelligence, HUMINT) OSINT-toiminnalla voidaan etsiä sopivia henkilöitä agenteiksi.
- Signaalitiedustelussa (engl. signals intelligence, SIGINT) OSINT voi tarjota teknisiä tunnistetietoja seuratusta tietoliikenneinfrastruktuurista.
- Mittaus- ja tunnusmerkkitiedustelua (engl. measurement and signature intelligence, MASINT) OSINT voi auttaa tarjoamalla avoimien tietolähteiden informaatiota liittyen esimerkiksi seismologisiin tapahtumiin.
- Geospatiaalista tiedustelua (engl. geospatial intelligence, GEOINT) OSINT voi täydentää esimerkiksi kaupallisten kuvantamislähteiden avulla.

OSINT voi kerätä tiedustelutietoa kaikista luetelluista tiedustelulajeista sekä muista luettelemattomista tiedustelulajeista. OSINTia määrittää sen laillinen toiminta, eikä se vaadi tiedustelun suorittajalta toimivaltuuksia, minkä takia sitä voidaan hyödyntää osana lukuisia organisaatioita ja tarkoituksia. Clark (2013) alleviivaa näihin tarkoituksiin poliittiset, ekonomiset, sotilaalliset sekä kilpailulliset tarpeet muiden ohella. OSINTia toiminnassaan hyödyntää esimerkiksi sotilasorganisaatiot, tiedustelupalvelut, lainvalvontaviranomaiset, terroristiorganisaatiot, yksityiset organisaatiot sekä uutisorganisaatiot (Clark, 2013; Higgins, 2022; Jardines, 2016; Ramwell ym., 2017). Avoimen informaation ympärille on muodostunut myös yhteisöjä, jotka käyttävät avointa informaatiota ja OSINT-menetelmiä osana tutkivaa journalismia ja tutkimushankkeita. Esimerkiksi Bellingcat on tällainen yksityinen tutkintaorganisaatio, joka käyttää OSINT-menetelmiä vastatakseen tiedontarpeeseen (Higgins, 2022). Yksityiset tutkintaorganisaatiot voivat erottaa toimintansa tiedustelusta, mutta käyttävät silti pitkälti samoja metodeja informaation keräämiseen, käsittelyyn ja hyödyntämiseen (Higgins, 2022). Yksityiset tutkintaorganisaatiot ovat esimerkki avoimien tietolähteiden hyödyllisyydestä osana tutkimustoimintaa sekä tiedustelutoimintaa.

Yksityiset tutkintaorganisaatiot sekä tiedustelun suorittajat ovat ensiaskel uusien tapahtumien, konfliktien ja muutosten todellisen luonteen tutkimisessa ja varmentamisessa. Informaatio uusien tapahtumien ympärillä on kuitenkin ristiriitaista, sekavaa ja vaikeasti varmistettavaa. Hernández-Escayolan (2024) mukaan erilaiset faktantarkistajat (engl. fact-checkers) ja faktojen tarkistuksen ympärille syntyneet organisaatiot ovat tärkeä tekijä journalismin, avoimen informaation sekä johdettuna myös OSINTin toimintaympäristössä. Näiden faktantarkistajien päätehtävänä on hallita harhaanjohtavasta informaatiosta syntyviä häiriötiloja ja ylläpitää informaatioympäristöjen luotettavuutta, mutta ovat yleisesti tärkeä tekijä OSINTin toimintaympäristössä faktojen tarkistamiseen ja tiedustelutiedon varmistamiseen.

Faktantarkistajien tehtävänä ei ole vain tunnistaa harhaanjohtava informaatio, vaan pyrkiä estämään harhaanjohtavan informaation leviämistä kertomalla miksi jokin asia on harhaanjohtavaa (Weikmann & Lecheler, 2023a). Hernández-Escayolan (2024) mukaan faktantarkastajien on oltava läpinäkyviä, puolueettomia, reiluja ja avoimia toimijoita pysyäkseen luotettavina tahoina informaatioympäristöissä. Hänen mukaansa väestön tulee tietää, ketä faktantarkistajat ovat ja miten he toimivat, jotta heidän varmistama informaatio voidaan nähdä luotettavana. Harhaanjohtavan informaation levittäjät yrittävät kuitenkin murentaa faktantarkistajien luotettavuutta luomalla sekaannusta heidän toimintansa ympärille (Hernández-Escayola, 2024). Tämä voi johtaa tilanteeseen, missä julkiset henkilöt ja poliitikot käyttävät hyödyksi aiemmin esiteltyjen syväväärennöksien olemassaoloa ja faktantarkistajien murentuvaa luotettavuutta. Esimerkiksi poliitikot voivat väittää faktantarkistajien oikeaksi varmistamaa videota syväväärennökseksi ja pyrkiä näin väistämään vastuuta videolla tapahtuvista asioista. Syväväärennöksien käyttäminen mahdollisen kiistettävyyden rakentamisessa tunnetaan nimellä ”valehtelijan osinko” (engl. liar’s dividend) (Citron & Chesney, 2019).

Faktantarkistusprojektit ylettyvät vain murto-osaan leviävästä harhaanjohtavasta informaatiosta (Higgins, 2022), eikä faktoista olla aina edes samaa mieltä (Vinhas & Bastos, 2022). Luotettavan kuvan varmistamiseksi faktantarkistajien tulee toimia varovaisesti ja ottaa kantaa vain väitteisiin, joihin heillä on varmat todisteet (Hernández-Escayola, 2024, s. 245), joten tiedustelua suorittavien tahojen tulee tehdä oma työnsä informaation varmistamiseksi. OSINTin ympäristössä toimii siis muut tiedustelulajit, avoimien lähteiden tutkintaa suorittavat toimijat sekä faktantarkistajat, jotka kaikki käyttävät omia prosessejaan ja metodejaan informaation käsittelyyn ja varmistamiseen.

3.3 OSINT osana tiedusteluprosessia

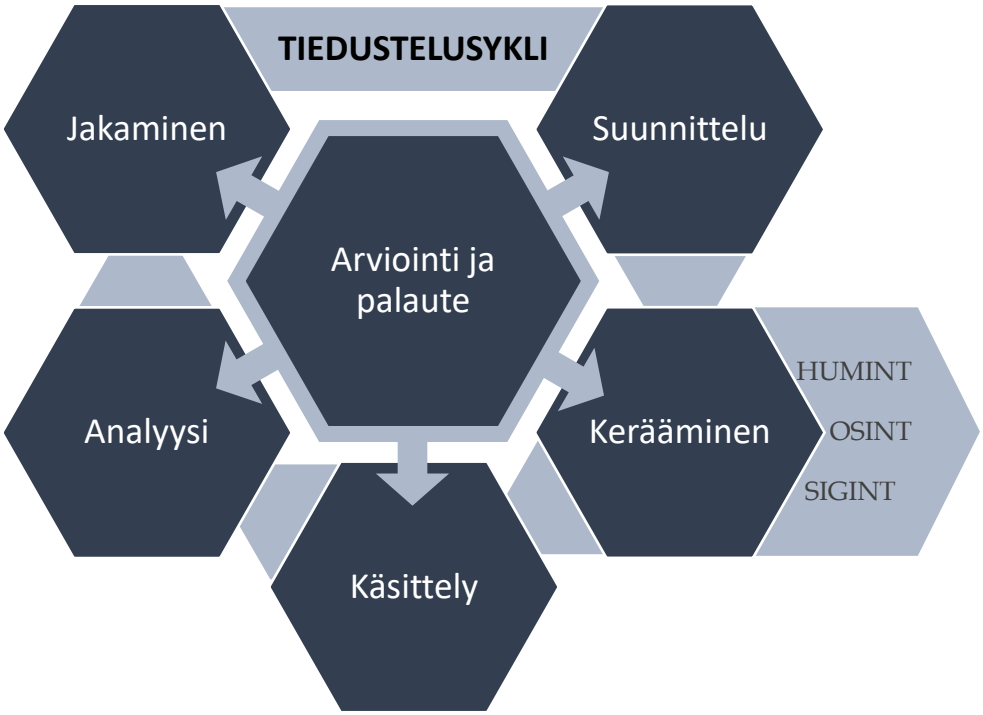
"Whoever is best at gathering (and exploiting) relevant information tends to win." (Lowenthal & Clark, 2016, s. 1). Lowenthal ja Clark (2016, s. 1) alleviivaavat informaation käsittelyn tärkeyttä tiedustelutoiminnassa pelkän informaation keräämisen lisäksi. He jatkavat, että ilman tarkkaa kerätyn informaation analyysiprosessia, missä annetaan konteksti informaatiolle ja tarkistetaan mitä informaatiosta puuttuu, on pelkkä informaation kerääminen hyödytöntä. Eri organisaatiot ja tiedustelun suorittajat käyttävät erilaisia prosesseja, mutta ne sisältävät yleisesti samat vakiintuneet askeleet. Tässä osiossa esitellään tiedustelusykli (engl. intelligence cycle), internetlähteiden lähdeanalyysi sekä yksittäisen informaation ja lähteen luotettavuuden analysointimalli.

NATO (2002) kuvaa OSINTin perustaksi kaikelle muulle tiedustelutoiminnalle ja korostaa sen roolia kontekstin luomisessa kerätylle informaatiolle. OSINT on tärkeä osa koko tiedusteluprosessia ja täyttää tiedustelutarpeen yhteistoimin muiden tiedustelulajien kanssa. Kuviossa 2 kuvattu tiedustelusykli kuvaa tiedustelussa tapahtuvat eri toimet ja muodostaa viitekehyksen kaikelle tiedustelutoiminnalle, missä eri tiedustelulajit toimivat yhdessä kohti määriteltyjä tavoitteita (Marzell, 2017). Esitelty tiedustelusykli koostuu viidestä erillisestä vaiheesta, jotka sisältävät erilaisia toimia (Joint Chiefs of Staff, 2013; Marzell, 2017; Ministry of Defence, 2023; Omand, 2013):

- Suunnitteluvaiheessa määritellään tiedustelutarpeet, suoritettavan tiedustelun rakenne, valmistellaan keräyssuunnitelma ja lähetetään tiedustelupyyntö tiedustelun suorittajille. Vaiheessa päätetään, mitä tiedustelutietoa tarvitaan, kuinka nopeasti ja mihin tarpeisiin.
- Keräämisvaiheessa seurataan määriteltyä keräyssuunnitelmaa. Vaiheessa kerätään tiedustelupyynnön kannalta oleellista informaatiota eri tiedustelulajeilla esimerkiksi OSINTin, HUMINTin tai SIGINTin keinoin.
- Käsittelyvaiheessa käsitellään keräysvaiheessa kerättyä informaatiota ja tarkistetaan sen oleellisuus. Tämä tarkoittaa informaation muokkaamista käytettävään muotoon ja sen asettamista kontekstiin. Esimerkiksi

vieraan kielen kääntämistä ja kulttuuristen piirteiden avaamista ymmärrettävään muotoon.

- Analyysivaiheessa arvioidaan tiedustelutiedon vastaavuus lähetettyyn tiedustelupyyntöön. Analyysiin sisältyy tiedustelutiedon luotettavuuden, oikeellisuuden sekä merkityksellisyyden arviointi.
- Jakamisvaiheessa uusi tiedustelutieto jaetaan ja levitetään päätöksentekijöille ja muille tarpeellisille tahoille. Tämä voi tarkoittaa esimerkiksi suullisia selontekoja, kirjoitettuja raportteja, kuvamateriaalia tai karttoja.
- Koko syklin ajan toteutetaan tiedustelutoiminnan arviointia ja annetaan palautetta. Tämä tarkoittaa tiedustelutiedon arvioimista kohti alkuperäisiä tiedustelutarpeita sekä syklin aloittamista uudelleen varustettuna uudella tietämyksellä ja tarkennetuilla vaatimuksilla



KUVIO 2 Tiedustelusykli (Joint Chiefs of Staff, 2013, s. I-6; Marzell, 2017, s. 39; Ministry of Defence, 2023, s. 37–38 mukaan muotoiltu)

Vaiheet eivät nimestään huolimatta etene askelittain, vaan vaiheet toimivat samanaikaisesti sekä osittain päällekkäin (Ministry of Defence, 2023) ja tiettyjä vaiheita voidaan jopa ohittaa tarpeen vaatiessa (Joint Chiefs of Staff, 2013). Keräysvaiheessa tiedustelusyklin rakennetta voidaan soveltaa tiedustelulajien sisällä tapahtuvassa informaation kokoamisessa tai käyttää jotain erillistä prosessia. Esitellyn mallin lisäksi on lukuisia muita malleja, kuten esimerkiksi TCPED-malli,

joka seuraa pitkälti samoja vaiheita eri nimillä (Ministry of Defence, 2023). Tiedustelusyklistä irrallaan NATO (2002) kuvaa internetlähteiden todennuksen ja lähdeanalyysin askeleet seuraavanlaisesti:

- Informaation tarkkuuden vahvistaminen, joka tarkistetaan muita tietolähteitä vastaan.
- Tietolähteen luotettavuuden ja auktoriteetin arviointi, eli tuoko tietolähde esille itsestään selkeän ja todellisen kuvan?
- Ajankohtaisuuden tarkastaminen, eli onko tarjottu informaatio varmistettavissa ajankohtaiseksi?
- Tietolähteen objektiivisuuden analysointi, eli onko tietolähde todennettavasti jonkun ryhmän edustaja?
- Asiaankuuluvuuden tutkiminen, eli onko informaatio relevanttia täytettävään tiedustelutarpeeseen?

Osana tiedusteluprosessia tarkastellaan myös informaation ja informaation lähteen luotettavuutta ja annetaan niille alustava arviointi. Taulukko 2 kuvaa NATOn tiedustelutiedon arviointiasteikkoa (myös amiraliteettikoodi, engl. admiralty code), missä arvioidaan informaation lähteen luotettavuus asteikolla A–F sekä itse informaation oikeellisuus asteikolla 1–6 (Ministry of Defence, 2023; Department of the Army, 2006). Arviolla luodaan alustava kuva informaation käyttämisestä osana tiedusteluprosessia. Esimerkiksi luokitus F2 tarkoittaa lähteen olevan mahdotonta luokitella ja informaatio itsessään on loogista ja vastaa muuta aiheesta saatua informaatiota, mutta on vahvistamatonta. Luokitus A4 tarkoittaa informaation tulevan luotettavasta lähteestä, mutta informaatio itsessään ei vaikuta loogiselta eikä ole vahvistettu, mutta on kuviteltujen mahdollisuuksien sisällä.

Tiedustelulajien yhteistoiminta isommassa syklissä, tiedustelulajien sisällä toimiva prosessi, internetlähteiden luotettavuuden arviointi sekä yksittäisen informaation ja lähteen arviointi ovat osa OSINTin sisäistä ja välittömässä toimintaympäristössä tapahtuvaa toimintaa. Informaation kerääminen, käsittely ja hyödyntäminen ovat kuitenkin monimutkainen ja aikaa vievä prosessi. Seuraavaksi tarkastellaan, miten digitaalisissa informaatioympäristöissä leviävä harhaanjohtava informaatio ja synteettinen media vaikuttavat näihin prosesseihin.

TAULUKKO 2 Tietolähteen ja informaation luotettavuuden arviointiasteikko (Ministry of Defence, 2023, s. 60; Department of the Army, 2006, s. B-1 & B-2 mukaan muotoiltu)

Tietolähteen luotettavuuden arviointi		
A	Luotettava	Ei epäilystä autenttisuudesta, luotettavuudesta tai pätevyydestä. Historiallisesti luotettava tietolähde
B	Yleensä luotettava	Pieni epäilys autenttisuudesta, luotettavuudesta tai pätevyydestä. Historiallisesti yleensä luotettava tietolähde.
C	Melko luotettava	Epäilys autenttisuudesta, luotettavuudesta tai pätevyydestä. On tuottanut pätevää informaatiota aiemmin.
D	Harvoin luotettava	Huomattava epäilys autenttisuudesta, luotettavuudesta tai pätevyydestä. On tuottanut pätevää informaatiota aiemmin.
E	Epäluotettava	Puutteellinen autenttisuus, luotettavuus tai pätevyys. Historiallisesti epäluotettava tietolähde.
F	Ei arvioitavissa	Ei perusteita tietolähteen luotettavuuden arviointiin.
Informaation luotettavuuden arviointi		
1	Vahvistettu	Vahvistettu muista riippumattomista tietolähteistä, looginen itsessään, yhdenmukainen aiheellisen muun informaation kanssa.
2	Todennäköisesti totta	Ei vahvistettu, looginen itsessään, yhdenmukainen muun aiheellisen informaation kanssa.
3	Mahdollisesti totta	Ei vahvistettu, kohtuullisen looginen, sopii yhteen muun aiheellisen informaation kanssa.
4	Epävarma	Ei vahvistettu, mahdollista mutta ei loogista, ei muuta informaatiota aiheesta.
5	Epätodennäköinen	Ei vahvistettu, ei loogista itsessään, ristiriidassa muun aiheellisen informaation kanssa.
6	Ei arvioitavissa	Ei perusteita informaation luotettavuuden arviointiin.

4 HARHAANJOHTAVA INFORMAATIO OSINTIN TOIMINTAYMPÄRISTÖSSÄ

Mediassa leviävät kauhukuvat uusista teknologioista ja niiden vaikutuksesta digitaalisiin informaatioympäristöihin ruokkivat keskustelua tulevaisuudesta: tekoäly tulee luomaan kaiken internetin sisällön, kaikki sosiaalisen median käyttäjät tulevat olemaan botteja ja syvävääreennöksien takia edes videoihin ei voi enää luottaa. Sensaationaalisista kauhukuvista välittämättä uudet teknologiat tulevat muokkaamaan informaatioympäristöjä. Miten harhaanjohtava informaatio vaikuttaa väestöön, ihmisiin sekä tiedustelua suorittaviin henkilöihin? Miten synteettinen media ja tekoäly vaikuttaa OSINT-toimintaan? Miten OSINT-toiminnassa voidaan käyttää tekoälyä harhaanjohtavan informaation käsittelyyn? Näihin kysymyksiin tutustutaan tässä osiossa.

4.1 Digitaaliset informaatioympäristöt ja harhaanjohtava informaatio

Jotta harhaanjohtava informaatio voi vaikuttaa ihmisiin, tulee ihmisten jotenkin nähdä tai kokea tämä levitetty informaatio. Tämä avaa etenkin internetin ja sosiaalisen median erilaisille informaatiovaikuttamisen muodoille. Samalla pelikentällä toimii myös OSINT, joten on tärkeää tiedostaa, miten disinformaatiota pyritään levittämään, millaisiin ongelmiin ympäristössä törmätään ja miten ongelmiin voidaan tiedustelutoimintana vastata. Disinformaation ainoa päämäärä ei ole mielipiteisiin vaikuttaminen tai häiriötilanteiden luominen väestössä ja yksilötasolla. Disinformaatiota levittämällä tiedustelun suorittajaan voidaan yrittää vaikuttaa esimerkiksi tuhlaamalla resursseja, peittelemällä muuta toimintaa tai johtamalla harhaan. Tässä osiossa tutustutaan harhaanjohtavan informaation vaikutuksiin sekä väestössä että OSINT-toiminnassa.

Laajoissa informaatioympäristöissä, kuten sosiaalisessa mediassa, voi disinformaatio nopeasti muuttua leviäväksi misinformaatioksi. Henkilö voi hel-

posti jakaa hänelle tarkoituksellisesti levitettyä disinformaatiota eteenpäin misinformaationa, jos tämä luulee informaation olevan oikeaa informaatiota. Tämä kuvaa informaatiovaikuttamisen toimintamekanismia, missä misinformaatio leviää pahaa-aavistamattoman väestön keskuudessa. Valtioneuvoston kanslian (2016, s. 13) mukaan informaatiovaikuttaminen tarkoittaa väestön ja päätöksentekijöiden asenteisiin, mielipiteisiin ja käyttäytymiseen vaikuttamaan pyrkivää häiriöllistä viestintää ja toimintaa. Toiminnan tarkoitus on vaikuttaa yhteiskunnan toimintakykyyn esimerkiksi harhaanjohtavaa informaatiota levittämällä (Valtioneuvoston kanslia, 2016).

Sosiaalisen median palvelut ja niiden käyttö muokkaavat informaatiovaikuttamisen taistelukenttää. Viimeisen vuosikymmenen sisään mahtuu monia esimerkkejä informaatiovaikuttamisesta sekä disinformaatiokampanjoista, jotka ovat levinneet pääosin sosiaalisessa mediassa. Yhdysvaltojen presidentinvaalit ja Yhdistyneen kuningaskunnan Brexit-äänestys vuonna 2016 sekä Ukrainan ja Venäjän välinen sota erityisesti sen kärjistyessä vuonna 2022 ovat luoneet massiivisia harhaanjohtavan informaation aaltoja (Pherson ym., 2024). Näiden lisäksi esimerkiksi COVID-19 pandemian ympärillä liikkuva valtioiden sponsoroima harhaanjohtava informaatio lisäsi taudin leviämistä (Lin ym., 2022).

Altistuminen harhaanjohtavalle informaatiolle voi heikentää ihmisten kykyä erottaa oikea informaatio harhaanjohtavasta. Pennycook ja muut (2018) osoittavat, että jopa yksittäiset altistumiset harhaanjohtavalle informaatiolle voi muuttaa henkilöiden kykyä hahmottaa informaation luotettavuutta. Tämä tapahtuu, vaikka faktantarkistajat olisivat merkinneet informaation harhaanjohtavaksi tai informaatio olisi vastoin henkilön ennalta ylläpitämiä uskomuksia (Pennycook ym., 2018). Aiemmin tutkielmassa Lu ja muut (2022) toivat esille informaation jakamisen tunneperäisistä syistä. Saling ja muut (2021) jatkavat tätä näkemystä, painottaen että jopa aktiivisesti faktoja tarkistavat henkilöt voivat jakaa harhaanjohtavaa informaatiota eteenpäin, vaikka olisivatkin tietoisia sen mahdollisesta harhaanjohtavuudesta.

Faktojen tarkistuksesta ja informaation oikaisusta huolimatta henkilöt luottavat edelleen ainakin osittain informaatioon, minkä he tietävät harhaanjohtavaksi (Lewandowsky ym., 2017). Harhaanjohtavan informaation leviämiseen vaikuttaa myös misinformaation levittäjien itsevarmuus. Misinformoituja mielipiteitä ylläpitävät henkilöt ovat usein itsevarmempia uskomuksiinsa kuin oikein informoidut (Kuklinski ym., 2000). Vaikka henkilöt eivät olisikaan itsevarmasti misinformoituja, ovat ihmiset silti yleisesti liian itsevarmoja kykyynsä huomata harhaanjohtava informaatio. Lyonsin ja muiden (2021) mukaan ihmiset pitävät itseään keskimääräistä parempina harhaanjohtavan informaation tunnistamisessa, mutta tämä itsevarmuus ei korreloi tutkimuksen tulosten kanssa. Tutkimuksen mukaan huonoiten suoriutuivat kaikista itsevarmimmat henkilöt (Lyons ym., 2021).

Ihmiset ovat siis helposti harhaan johdettavissa ja alttiita harhaanjohtavan informaation vaikutuksille. Sosiaalinen media tekeekin oikean informaation ja manipulointiyritysten erottamisen toisistaan vaikeaksi (Hernández-Escayola,

2024). Cialdinin (2014) mukaan näkemyksemme hyväksyttävästä käyttäytymisestä mukautuvat sen perusteella, mitä muut miettivät oikeana, muodostaen "sosiaalisen hyväksyttävyyden" (engl. social proof) toiminnalle. Sosiaalisessa mediassa tämän sosiaalisen hyväksyttävyyden muodostaa sisällön ”viraalius”, eli tykkäyksien ja jakamisien muodostama sisällön leviäminen (Kivimäki, 2024, s. 287). Tämä yleinen hyväksyntä ja viraalius lisää luottamusta leviävään informaatioon riippumatta sen harhaanjohtavuudesta (Luo ym., 2022).

Suosiota keränneen harhaanjohtavan informaation tarkastaminen internetistä saattaa kuitenkin usein jopa nostaa uskomusta harhaanjohtavaan informaatioon (Aslett ym., 2023). Omatoimisen faktantarkistuksen toimimattomuus voi johtua vahvistusvinoumasta (engl. confirmation bias), missä henkilöt etsivät omia näkemyksiään vahvistavaa informaatiota (Roberts & Qahri-Saremi, 2023). Bastickin (2021) mukaan harhaanjohtava informaatio voi vaikuttaa henkilöiden tiedostamattomaan käyttäytymiseen ja asenteisiin jopa lyhyellä altistuksella. Bastickin tutkimuksen tulokset eivät ole suoraan linkitettävissä suuriin disinformaatiokampanjoihin, mutta osoittaa altistumisen mahdollisesti vaikuttavan henkilöiden tiedottomaan toimintaan.

Tiedustelua suorittavat toimijat ovat tietoisia näistä ongelmista, mutta eivät immuuneja niille. Lisäksi tiedustelutoiminnassa informaation arviointia voi haitata esimerkiksi jonkin tiedustelulajin tuottaman tiedon suosiminen tai hyljeksiminen, tunneperäiset syyt kuten ylpeys, ajankohtaisvinouma (engl. recency bias) tai ”vividness weighting”, eli jonkun tietolähteen suosiminen pelkästään sen aiheuttaman suuren reaktion takia (Clark & Mitchell, 2019). Kivimäki (2024, s. 290) nostaa esille tämän suuren reaktion aiheuttamaa vaikutusta jopa informaatiota analysoivaan henkilöön. Hänen mukaansa analyytikoiden tulee kiinnittää erityistä huomiota materiaaliin, joka herättää vahvoja tunteita tai reaktioita, koska se voi osoittaa jotain tarkoitusta varten luotua disinformaatiota.

Tämä tarkoitus voi olla esimerkiksi hämäystoimintaa (engl. deception). Tiedustelun suorittajat ovat koko ajan tietoisia hämäyksen mahdollisuudesta (Wirtz, 2024). Hämäystä on esimerkiksi tiedustelulle syötetty puutteellinen oikea informaatio tai disinformaatio, tarkoituksena esimerkiksi hallita vastustajan tietämystä oikeasta tilanteesta tai ohjata vastustaja keskittymään muualle (Clark & Mitchell, 2019; Wirtz, 2024). Hämäystoiminta itsessään perustuu tiedustelun tuottamalle tiedolle vastustajan uskomuksista, suunnitelmista ja odotuksista, johon hämäystoiminnalla vastataan (Clark & Mitchell, 2019; Wirtz, 2024). Informaatiovaikuttaminen voi toimia osana hämäystoimintaa, kun harhaanjohtavalla informaatiolla pyritään vaikuttamaan päätöksentekijöiden toimintaan ja käyttäytymiseen. Hämäystoiminnalla johdatetaan vastustaja keskittymään heille tarjottuun realistiseen ja uskottavaan informaatioon, eikä siihen mikä heiltä puuttuu (Wirtz, 2024, s. 33). Tiedustelutoiminnan näkökulmasta tiedusteluun on käytössä vain rajallinen määrä resursseja tiedustelutarpeen täyttämiseksi (Ministry of Defence, 2023, s. 44), joten hämäystoiminnan mahdollisuutta tulee epäillä, jos näitä resursseja tulee käyttää paljon uuden informaation varmistamiseksi (Clark &

Mitchell, 2019). Sosiaalinen media luo suurimman kanavan myös hämäystoiminnalle (Clark & Mitchell, 2019) joten sen tarjoamat mahdollisuudet, kuten automaatio ja kohdennettu mainonta, ovat hämäystoiminnan käytettävissä.

Sosiaalinen media avaa oven henkilökohtaisesti kohdennetun ja räätälöidyn disinformaation levittämiselle (Ienca, 2023). Disinformaation levittäjät voivat ostaa sosiaalisen median alustoilta näkyvyyttä disinformaatiota levittäville käyttäjätileille sekä hyödyntää mainostamisen kohdentamiseen tarkoitettuja palveluita (Ienca, 2024; Kivimäki, 2024). Lisänäkyvyyttä ja viraaliutta erilaisille disinformaation levittämisen toimenpiteille voi saada käyttämällä aiemmin tutkielmassa mainittuja botteja (Kivimäki, 2024, s. 285), hyödyntämällä niin kutsuttuja "trollitehtaita" (engl. troll farms) (Lukito, 2024) tai molempia. Tämä yhdistettynä aiemmin esille tuotuun näkemykseen asioiden sosiaalisesta hyväksymisestä niiden suosion perusteella voi helposti vaikuttaa ihmisten mielipiteisiin sekä tiedustelutoiminnan sisällä että sen ulkopuolella.

Jardinesin (2016, s. 6) mukaan OSINTin ympäristössä informaation sensurointi tai harhaanjohtavan informaation levittäminen on tuloksetonta, koska OSINT tukeutuu toiminnassaan valtavaan määrään erilaisia lähteitä informaation tarkistamiseksi. Informaation arviointi ja tarkistaminen luotettavaksi voi kuitenkin osoittautua haastavaksi jopa OSINT-toiminnassa. Hämäystoiminta, faktojen tarkistamisessa kohdatut ongelmat sekä harhaanjohtavan informaation aiheuttamat vinoutumat ja vääristymät päätöksentekoprosessissa vaikeuttavat OSINT-toimintaa. Tämä voi vahvistua entisestään synteettisen informaation ja median täyttäessä digitaalisia informaatioympäristöjä.

4.2 Uusien disinformaation muotojen vaikutus OSINT-toimintaan

OSINT on tiedustelulajina uniikki sen käytössä olevan informaation määrän takia. Ongelma ei ole informaation kerääminen, vaan informaation määrän takia sen käsittely. Myös synteettinen media lisää harhaanjohtavan informaation leviämisen muotoja sosiaalisessa mediassa. Tekoälysovellukset kykenevät luomaan aidolta näyttäviä, jopa valokuvamaisia kuvia asioista, ihmisistä tai tapahtumista, mitä ei ole olemassa tai ei ole ikinä tapahtunut. Tässä osiossa tarkastellaan tekoälyllä tuotetun tai levitetyn disinformaation sekä synteettisen median vaikutusta OSINT-toimintaan.

Tekoälyä voi käyttää disinformaation levittämiseen monin eri tavoin. Tekoälyllä voidaan kohdentaa ja räätälöidä disinformaatiota yksittäisille henkilöille, automatisoida disinformaation laaja levittäminen informaatiovaikuttamisena, ajaa keskusteluihin synteettistä haitallista liikennettä sekä häiritä informaation lähteitä esimerkiksi täyttämällä ne turhalla tai häiritsevällä informaatiolla (Government Communications Headquarters [GCHQ], 2021). Lisäksi tekoälyllä voidaan tuottaa syväväärennöksiä sekä muuta audiovisuaalista tai tekstimuotoista synteettistä materiaalia.

Audiovisuaalisen materiaalin käyttäminen osana disinformaatiota ei ole uusi keksintö ja voi tapahtua monin eri tavoin, esimerkiksi vaihtamalla asioiden kontekstia, muokkaamalla sisältöä tai näyttämällä vain valittuja osia sisällöstä. Väärään kontekstiin liitetty audiovisuaalinen materiaali, jopa ilman mitään manipulaatiota tai muokkausta, on yksinkertainen tapa levittää uskottavaa disinformaatiota (Kivimäki, 2024; Weikmann & Lecheler, 2023a). Tämä voi tarkoittaa esimerkiksi vanhasta konfliktista otettujen kuvien liitämistä uuden konfliktin keskusteluihin kertomatta kuvan alkuperää. Aiemmin esitelty visuaalisen disinformaation koettu luotettavuus ei välttämättä riipu kontekstin oikeellisuudesta, kunhan se näyttää uskottavalta.

Tekoälyllä luotu harhaanjohtava informaatio ulottuu kuitenkin pidemmälle kuin pelkkä kuvien muokkaaminen tai kontekstin vaihtaminen. Syvävääränrennökset ja synteettinen media muokkaavat informaatioympäristöjä ja voivat olla osana informaatiovaikuttamista tai hämäystoimintaa. Yhdistettynä aiemmin esitettyyn faktojen tarkistamisen ongelmallisuuteen tai esiteltyyn valehtelijan osinko -ilmiöön voi synteettisen median aiheuttama epävarmuus ja epätietoisuus koetusta ympäristöstä aiheuttaa häiriötilanteita. Pelkkä synteettisen median olemassaolo aiheuttaa ongelmia informaatioympäristöissä. Syvävääränrennöksiä voi käyttää tapahtuneiden asioiden vääristämiseen (Weikmann & Lecheler, 2023a), tekosyynä niiden kiistämiseen (Citron & Chesney, 2019) sekä faktantarkistajien luotettavuuden heikentämiseen ja tuloksien kiistämiseen (Higgins, 2022).

Tästä huolimatta syvävääränrennökset eivät ole vielä aihe huolelle ja ovat tulevaisuuden ongelmia (Higgins, 2022; Weikmann & Lecheler, 2023a). Audiovisualinen materiaali ei ole informaatioympäristössä vapaasti liikkuva riippumaton yksilö, vaan osa isompaa informaatiorakennetta (Higgins, 2022). Synteettistä informaatiota voidaan tarkastella saman prosessin kautta kuin muutakin informaatiota, eli kyseenalaistamalla informaatio ja sen ympärillä olevat ulottuvuudet (Higgins, 2022). Esimerkiksi aineiston sisältö, konteksti, levittäjä tai levittäjän motivaatiot voidaan kyseenalaistaa. Aiemmin esitellyt tiedustelussa käytetyt prosessit informaation luotettavuuden arviointiin toimivat edelleen, vaikka informaatio olisikin synteettistä.

OSINTin näkökulmasta syvävääränrennetyn sisällön käsittely ei eroa muusta levitetystä harhaanjohtavasta mediasta. Median luotettavuuden arvioinnissa voidaan käyttää esimerkiksi aiemmin esiteltyä amiraliteettikoodia tietolähteen ja informaation arvioimiseksi, vaikka informaatiota ei etukäteen tiedetäkään synteettiseksi. NATOn julkaisema vuoden 2002 käsikirja OSINT-toimintaan alleviivaa Joe Barkerin kehittämät nettisivujen luokitteluun käytettävät viisi tärkeää peruskysymystä: kuka, mitä, missä, milloin ja miksi (NATO, 2002, s. 26). Osana tiedusteluprosessia vielä 2020-luvulla käytetään samoja kysymyksiä, mutta kyseenalaistetaan kerätty informaatio myös kysymällä ”miten”, jotka muodostavat yhdessä kuusi tiedustelun peruskysymystä (Higgins, 2022; Ministry of Defence, 2023). Tämä ei tarkoita, että keinoja ja tapoja ei ole kyseenalaistettu aiemmin kysymällä ”miten”, vaan että keinojen kyseenalaistamisen arvo on noussut uusien harhaanjohtavan informaation levityskeinojen mukana. Näiden kysymysten perusteella pyritään arvioimaan tietolähteen luotettavuus ja käyttökelpoisuus

OSINT-toimintaan sekä levitetyn materiaalin taustatietoja ja kontekstia ympäristössä. Kysymyksillä voidaan kyseenalaistaa lisäksi informaation laajempi oikean maailman ympäristö ja konteksti.

Segondin (2024) mukaan harhaanjohtavan informaatio voidaan tunnistaa analysoimalla sen sisältöä, sen leviämistä tai kumpaakin. Tämä tarkoittaa jo odotetusti informaation kyseenalaistamista, taustojen tutkimista ja ympäristön ja kontekstin tarkastelua. Hernández-Escayolan (2024) mukaan informaatiota käsitellessä voidaan kyseenalaistaa käytetty kieli, tarjotut todisteet sekä viitatut lähteet. Sensaationaalinen kielenkäyttö tai puutteelliset todisteet tai lähteet kertovat informaation luotettavuudesta (Hernández-Escayola, 2024). Tiivistetysti, sosiaalisessa mediassa leviävää disinformaatiota voidaan tunnistaa ja varmistaa kyseenalaistamalla informaation sisältö, sen asema ympäristössä ja sitä ympäröivä konteksti riippumatta siitä, onko se synteettistä vai ei.

Synteettinen media ei siis itsessään ole vaarallisempaa kuin muu disinformaatio (Weikmann & Lecheler, 2023a), mutta sen uskotaan muodostuvan poliittiseksi sotakeinoksi tulevaisuudessa (Paterson & Hanley, 2020). Weikmannin ja Lechelerin (2023a) mukaan teknologisesta kehityksestä huolimatta vaarallisin visuaalisen disinformaation muoto on aiemmin mainittu väärään kontekstiin liitetty audiovisuaalinen materiaali. He nostavat mahdolliseksi ongelmaksi myös pintaväärennökset (engl. shallowfakes), jotka ovat minimaalisesti muokattuja videoita esimerkiksi nopeuttamalla tai hidastamalla videon nopeutta.

Audiovisuaalisen median luotettavuutta voidaan tarkastella tutkimalla siihen liitettyä metadataa, joka yksinkertaistetusti kuvaa dataa datasta. Tutkielmassa tämä tarkoittaa esimerkiksi informaatiota, jonka puhelin tallentaa tiedostoon otetun kuvan kanssa tai käsittelyohjelma tallentaa muokattuun tiedostoon. Metadata voi sisältää esimerkiksi informaatiota siitä missä, milloin ja kuinka jokin materiaali on tuotettu (Kivimäki, 2024). Metadatasta voi esimerkiksi päätellä, jos tiedosto on otettu eri päivänä kuin sen väitetysti kuvaama tapahtuma tapahtui (Padilha ym., 2022).

Pelkkään metadataan ei kuitenkaan voi luottaa. Useat sosiaalisen median palvelut irrottavat metadatan palveluun ladatusta mediasta (Kivimäki, 2024, s. 290; Venema, 2024, s. 181) ja metadataa on mahdollista muokata tai väärentää (Padilha ym., 2022). Tulevaisuuden ratkaisu luotettavaan vesileimaan voi olla esimerkiksi lohkoketjujen käyttäminen median varmistamisessa (Higgins, 2022, s. 209; Venema, 2024, s. 185). Tulevaisuutta odotellessa jaetun sisällön todenperäisyys tulee vahvistaa muita keinoja käyttämällä riippumatta siitä, onko tutkitussa sisällössä metadata vai ei. Hyödyksi voidaan käyttää esimerkiksi aiemmin mainittuja faktantarkistajia, kyseenalaistamista tai käyttämällä muita työkaluja sisällön vahvistamiseksi. Taulukossa 3 tarkastellaan erilaisia esimerkkejä työkaluista ja niiden käyttötarkoituksia. Työkaluja on moniin eri tarpeisiin ja niitä voi hyödyntää levitetyn disinformaation narratiivin purkamisessa tai osoittamalla epäkohtia harhaanjohtavan informaation ja levittäjän väitteiden välillä. Esimerkiksi voidaan etsiä epäkohtia väitetyn tapahtuman ajankohdasta käyttäen aurinko- tai varjoanalyysiä, satelliittikuvia tai liikennetietoja.

TAULUKKO 3 Avoimien lähteiden tutkinnan työkaluja (Bellingcat, 2023 mukaan kerätty)

Avoimien lähteiden tutkinnan työkaluja	
Aurinko- ja varjoanalyysi	Arvio kuvattujen tapahtumien vuorokaudenajasta perustuen varjojen pituuteen tai auringon sijaintiin ja oletettuun sijaintiin maapallolla.
Käänteinen kuvahaku	Videoiden ja kuvien etsiminen muualta internetistä. Mahdollisuus löytää alkuperäinen lähde tai tunnistaa kuvattuja asioita tai esineitä.
Kasvojentunnistus	Kasvojen tunnistamiseen ja etsimiseen käytettäviä ohjelmistoja. Mahdollisuus tunnistaa synteettisesti luotuja kasvoja tai verrata kasvoja laajoihin tietokantoihin.
Metadatan tarkistaminen	Videoiden ja kuvien metadatan tarkastaminen ja analysointi. Mahdollista tunnistaa manipuloitua sisältöä ja tarkistaa median tunnusmerkkejä.
Sosiaalisen median analytiikka	Analytiikan tarkasteluun tai sisällön hallintaan työkaluja. Mahdollisuus tarkastella yksittäisiä kohteita, tilejä, julkaisuja jne.
Henkilöiden yhteistietojen tarkastaminen	Ihmisten yhteystietojen etsiminen ja tarkastaminen. Sisältää esimerkiksi puhelinnumeroita, käyttäjänimiä, sähköposteja ja osoitteita.
Kartta-, satelliitti- ja sijaintitiedot	Esimerkiksi kartat, satelliittikuvat, karttapalvelut, maamerkkien tunnistaminen tai historialliset sijaintitiedot. Mahdollistaa jopa live-aikaisen karttojen tarkastelun.
Maa-, meri- ja ilmaliikenne	Liikennetietoja ja -karttoja, jotka päivittyvät jopa liveajassa. Mahdollisuus tarkastella yksittäisten kulkuneuvojen, kuten lentokoneiden tai junien, tietoja, reittejä ja liikennehistoriaa.
Arkistot	Arkistoitua sisältöä esimerkiksi internetistä poistetuista tai päivitetyistä sivuista, videoista tai tiedostoista. Mahdollisuus tarkastella tilannekatsauksia eri ajoilta.

Digitaalisissa informaatioympäristöissä esiintyvän disinformaation käsittely OSINTissa koostuu monista eri tekijöistä. Disinformaation tunnistaminen voi olla haasteellisempaa syvävääreennöksien ja synteettisen median lisääntyessä, mutta käsittelyn toimenpiteet itsessään eivät ole muuttunut teknologioiden kehittymisen takia. Riippumatta median tai informaation synteettisestä luonteesta, ei pelkästään näennäisesti luotettava ja aidon näköinen harhaanjohtava informaatio ole vielä vaarallista OSINTin toiminnalle, kunhan informaation alkuperä, konteksti ja asema ympäristössä kyseenalaistetaan ja niitä tarkastellaan kriittisesti. Isoin ongelma OSINT-toiminnan kannalta on edelleen avoimien tietolähteiden sisältämän (harhaanjohtavan) informaation määrä sekä sen tehokas käsittely. Tähän ongelmaan etsitään vastausta samoista teknologioista, joilla myös harhaanjohtavaa informaatiota tuotetaan ja levitetään.

4.3 Tekoäly osana OSINT-toimintaa

Kehittyvät teknologiat tarkoittavat uudenlaisia tapoja luoda ja levittää disinformaatiota, mutta mahdollistavat ne myös uusia mahdollisuuksia OSINT-toiminnalle. Erityisesti suurien informaatiomäärien käsittelyyn soveltuvat tekoälysovellukset sekä informaation varmistamisessa ja arvioinnissa käytettävät teknologiat tasapainottavat kamppailua harhaanjohtavan informaation leviämistä vastaan. Tässä osiossa tutustutaan siihen, miten tekoälyä voi hyödyntää osana OSINTia ja esitetään käyttöä avaavia esimerkkejä.

GCHQ (2021), Yhdistyneessä kuningaskunnassa toimiva tiedustelu- ja turvallisuuspalvelu, painottaa tekoälyn käyttämisen osana tiedustelutoimintaa olevan perustana Yhdistyneen kuningaskunnan turvallisuudelle. Erityisesti GCHQ (2021) painottaa tekoälytyökalujen tärkeyttä lisääntyneen ja monimuotoistuneen informaation sekä tekoälyn mahdollistamia uhkia vastaan puolustautumisessa. Tekoälytyökaluja voidaan käyttää faktantarkistukseen, tunnistamaan syvävääreennettyä ja manipuloitua sisältöä sekä tunnistamaan ja estämään bottiverkostoja ja trollitehtaita (Arcos & Arribas, 2024; GCHQ, 2021; Segond, 2024).

Tekoälyavustus on myös tärkeä osa digitaalisia valvontatyökaluja, joilla voidaan seurata informaation leviämistä, leviämisen nopeutta sekä tekstin tai median sisältöä (Segond, 2024). Tämä tarkoittaa esimerkiksi erilaisia poiminta-, analyysi- ja visualisointityökaluja tai argumentin tarkistustyökaluja, joilla voidaan luoda kerätyn aineiston perusteella vastaväitteet leviävään harhaanjohtavaan informaatioon (Arcos & Arribas, 2024). Informaation sisällön tutkiminen ja varmistaminen tarkoittaa käytännössä informaation aseman kyseenalaistamista tekoälyn avulla, eli jo tunnettua informaation ja tietolähteen luotettavuuden automaattista arviointia ja kyseenalaistamista.

Informaation leviämistä tarkastellessa voidaan käydä läpi aiemmin esitellyt kuusi peruskysymystä tekoälyavusteisesti. Esimerkiksi selvitetään missä ja milloin asia tapahtui tai kuka on informaation alkuperäinen lähde. Segondin (2024) mukaan tekoäly voi automaattisesti vertailla eri lähteitä ja etsiä reitin alkuperäiseen lähteeseen ja sen julkaisuaikaan. Lisäksi tekoälytyökaluilla voidaan pitää kirjaa eri lähteistä ja arvioida niiden luotettavuutta, seurata informaation levittäjinä toimivia tilejä tai seurata yleisesti erilaisia aiheita ja teemoja (Segond, 2024).

Harhaanjohtavan informaation sisällön tarkastelu tekoälyavusteisesti pyrkii vastamaan aiemmin alleviivattuihin kysymyksiin ja löytämään epäkohtia kerrotun informaation ja todellisuuden välillä automaattisesti. Tekoälyllä voidaan esimerkiksi tutkia harhaanjohtavan informaation ympärillä olevaa keskustelua ja kielenkäyttöä sekä tekstin tai median sisältöä (Küçük & Can, 2020; Nandwani & Verma, 2021) tai luoda vertailua varten aiheittain erillisiä ryhmiä (Segond, 2024). Tekoälytyökaluilla syvävääreennettyjen videoiden tunnistaminen tapahtuu analysoimalla yksittäisiä videon kuvia, mistä tarkkaillaan videolla näkyvien henkilöiden ilmeitä, muutoksia tunnusmerkeissä tai viitteitä synteettisesti luotuihin kasvoihin ja vertaamalla tuloksia kontrollitapaukseen (Groh ym., 2021; Segond, 2024; Wang & Dantcheva, 2020). Syvävääreennetyn audion tarkastelussa voidaan

tekoälyllä verrata tarkasteltavaa audiota varmennetusti oikeaan audioon tapahtumasta tai puheesta ja etsiä manipuloimisen tai syvävääräntämisen jälkiä (Segond, 2024). Tämä voi tarkoittaa erilaisten ympäristön äänien tunnistamista, erottelua ja tarkkailua tekoälyn avulla (Mesaros ym., 2021) tai puheessa henkilön sanavalintojen, tunnetilan tai äänensävyn vertailua todennettuun audioon (Segond, 2024).

Leviämisen ja sisällön analysoimisen lisäksi yksinkertaisia tapoja käyttää tekoälyä osana tiedustelutoimintaa on informaation kerääminen ja laajojen informaatiomassojen käsittely (GCHQ, 2021; Higgins, 2022; Ministry of Defence, 2023). Suurista informaatiomassoista voidaan etsiä harhaanjohtavaa informaatiota sekä kartoittaa disinformaatioaktiviteetteja tekoälyavusteisesti. Kivimäki (2024, s. 291) painottaa, että tiedustelun näkökulmasta kiinnostava kysymys ei välttämättä ole informaation todenperäisyys, vaan minkä takia disinformaatio on edes ensikädessä olemassa. Kyseenalaistamalla syyt disinformaation olemassaolon ja disinformaatioaktiviteettien takana on mahdollista tarkastella harhaanjohtavan informaation olemassaolon tarkoitusta laajemmassa kokonaiskuvassa, ja ymmärtää kuka hyötyisi disinformaation levittämisestä (Kivimäki, 2024, s. 291). Tämä voi tarkoittaa esimerkiksi informaatiovaikuttamisen tai hämäystoiminnan takana olevien toimijoiden tunnistamista.

Syvävääräennöksien ja synteettisen median tunnistaminen on kuitenkin pitkälti tilanne, missä disinformaatiota levittävät tahot toimivat koko ajan askeleen edellä. Kun keino syvävääräennöksien luotettavaan tunnistamiseen löydetään, niin synteettistä mediaa luovat ohjelmistot käyttävät hyödyksi tätä tietoa ja tuottavat entistä aidomman näköisiä syvävääräennöksiä (Venema, 2024, s. 181). Tämä tekee esimerkiksi tiedustelupalveluista haluttomia jakamaan syvävääräennöksien tunnistamiseen käytettyjä metodeja (Venema, 2024, s. 181). Tämä voi olla rajoittava tekijä tulevaisuudessa syvävääräennöksien tunnistamisen tutkintaan, koska täysin avoin tutkimus ei välttämättä ole mahdollista. Veneman (2024) mukaan synteettisen median maailma kehittyy äärimmäisen nopeasti ja keinot informaatioympäristön suojaamiseksi jäävät hyödyttömiksi nopeasti. Tulevaisuudessa tekoälyllä uskotaan kuitenkin olevan tärkeä asema myös ennakoivassa toiminnassa disinformaatioaktiviteettien tunnistamisessa, ennen kuin niistä muodostuu ongelmia (Arcos & Arribas, 2024; Ministry of Defence, 2023).

Tekoälyllä on tärkeä rooli osana OSINT-toimintaa, eikä pelkästään osana harhaanjohtavan informaation ja synteettisen median luomista ja levittämistä. Tekoälyä voi käyttää osana useita OSINTin prosesseja, esimerkiksi informaation keräämisessä, harhaanjohtavan informaation tunnistamisessa, tiedustelutiedon käsittelyssä, faktojen tarkistuksessa sekä disinformaatiota levittävien toimijoiden seuraamisessa. Erityisesti tekoälytyökalut voivat auttaa suurien informaatiomassojen käsittelyssä sekä manuaalisten vaiheiden tehostamisessa. Aiemmin määritellyt OSINTissa käytettävät kysymykset informaation varmistamiseksi ja luotettavuuden takaamiseksi vaativat edelleen vastauksia, jotka on mahdollista hoitaa tekoälyä hyödyntämällä automaattisesti.

5 YHTEENVETO JA POHDINTA

Tutkielmassa etsittiin vastausta OSINTin asemaan tekoälyn ja harhaanjohtavan informaatio keskiössä. Tämä tarkoitti perehtymistä informaatioon ja digitaalisia informaatioympäristöjä muokkaaviin tekijöihin, tiedustelutoimintaan ja sen prosesseihin, harhaanjohtavan informaation vaikutuksiin digitaalisissa informaatioympäristöissä sekä tekoälyn vaikutusta disinformaation levittämiseen ja osana OSINT-toimintaa. Tutkielmassa etsittiin vastausta ongelmaan informaation varmistamisesta todelliseksi ja luotettavaksi informaatioksi OSINT-toiminnassa, ja tutkielman tarkoituksena oli luoda otanta nykyhetken tilanteesta, jotta tulevaisuudessa on mahdollista tarkastella nopeasti kehittyvää toimintaa tekoälyn, harhaanjohtavan informaation sekä OSINTin ympärillä.

Tutkimalla harhaanjohtavan informaation vaikutuksia huomattiin, että tekoälyn ja synteettisen median kasvava läsnäolo informaatioympäristöissä sekä osana harhaanjohtavan informaation levittämistä voivat vaikuttaa suuresti ihmisiin yksilö- ja väestötasolla, esimerkiksi informaatiovaikuttamisena sosiaalisessa mediassa, mutta vaikutus OSINTin sisäiseen toimintaan voi olla mitätön. Tämä johtuu siitä, että OSINTin toiminta on jo pitkälti vakiintunutta ja uudet työskentelymetodit tarvitsevat aikaa vakiintuakseen. Disinformaation levittäjät sekä OSINT hyödyntävät tekoälyä toiminnassaan, mutta vanhat toimintatavat ohjaavat näiden uusien teknologioiden käyttöä.

Tämä tarkoittaa laajempaa disinformaation levittämistä ja laajempaa informaation käsittelyä tekoälyn voimin. Toisin sanoen, tekoälyä käytetään osana OSINT-toimintaa vastaamaan samoihin kysymyksiin, joita tietolähteiden luotettavuuden arviointiin käytettiin jo 20 vuotta sitten, mutta huomattavasti laajemmalla skaalalla. Tähän skaalaan vastaa myös helpompi disinformaation tuottaminen ja levittäminen. Tekoälyavusteiset botit voivat levittää luotettavalta näyttävää harhaanjohtavaa synteettistä mediaa nopeasti ja ovat yhä vaikeampia erottaa boteiksi OSINTin keinoin. Tämä teknologinen kisa muuttuu kissa ja hiiri -leikiksi, missä disinformaatiota levittävät tahot ovat aina askeleen edellä, koska teknologiset kehitykset molemmin puolin ruokkivat lisää teknologista kehitty-

mistä, joka suosii ensiksi toimivaa osapuolta, eli levittäjiä. Tekoälyllä on tulevaisuudessa mahdollisuus toimia myös ennaltaehkäisevästi, joten leikin voittajaa ei ole vielä päätetty.

Tutkielmassa etsittiin vastausta kahteen tutkimuskysymykseen:

1. Miten tekoälyn avulla luotu ja levitetty harhaanjohtava informaatio vaikuttaa OSINT-toimintaan?
2. Miten tekoälyä voi hyödyntää osana harhaanjohtavan informaation käsittelyä OSINT-toiminnassa?

Tutkielman tulokset osoittavat vastauksena ensimmäiseen tutkimuskysymykseen, että jopa kehittyvissä informaatioympäristöissä OSINTin toiminnassa vastataan harhaanjohtavaan informaatioon vastaamalla samoihin kysymyksiin, kuin jo ennen tekoälyn käyttämistä osana disinformaation luomista ja levittämistä. Tämä tarkoittaa vastaamista kuuteen peruskysymykseen: kuka, mitä, missä, milloin, miksi ja miten. Informaation tai median synteettisyys ei muuta toimintatapoja, eli se mitä tehdään, on edelleen sama, mutta se miten asiat tehdään voivat erota huomattavasti. Kysymyksiin voidaan etsiä vastauksia uusilla työkaluilla, kuten automaattisella varjoanalyysillä tai kasvojentunnistusteknologialla. Jos joku näyttää epäilyttävältä tai liian hyvältä ollakseen totta, aiheuttaa kohua ja reaktioita väestön keskuudessa, asiaa tutkitaan ja perehdytään sen alkuperään ja ympärillä olevaan kontekstiin informaation varmistamiseksi tai paljastamiseksi. Informaation kyseenalaistaminen ja ymmärrys disinformaation luomisesta ja levittämiseen käytetyistä teknologioista ovat avainasemassa edelleen.

Vastauksena toiseen tutkimuskysymykseen, ehkä jopa jo odotetusti, tekoälyn käyttäminen osana tiedustelutoimintaa ei ole muokannut OSINTin toimintaa huomattavasti. Tekoälyn käyttäminen disinformaation levittämiseen ja OSINTissa sen tunnistamiseen palauttaa asiat tietynlaiseen tasapainoon, koska molemmat tapahtuvat suuremmalla skaalalla. OSINT toimii edelleen käyttäen samoja keinoja, mutta hyödyntää manuaalisen työn sijasta tekoälytyökaluja. Tekoälyä voi käyttää esimerkiksi informaation keräämiseen ja käsittelyyn, faktantarkistukseen, synteettisen median tunnistamiseen, disinformaation lähteiden etsimiseen sekä osana digitaalisia valvontatyökaluja. Lisäksi tekoälyavusteisesti voidaan etsiä vastauksia jo aiemmin esitettyihin peruskysymyksiin seuraamalla informaation leviämistä ja sisältöä. Tekoäly ei ole syrjäyttänyt ihmistä vielä, koska ihminen on tarpeellinen vielä erityisesti lopputuotteen arvioinnissa, hyödyntämisessä ja päätöksenteossa.

Mitä nämä löydökset tarkoittavat, jos mikään ei ole merkityksellisesti muuttunut? Ne tarkoittavat, että faktantarkistajien, tiedustelun suorittajien sekä väestön on oltava tietoisia uusista harhaanjohtavan informaation muodoista sekä teknologioista, jolla sitä niitä tuotetaan. Tämä tarkoittaa, että kaikkea liian otollista, huomiota herättävää ja jopa aiempia uskomuksia täydentävää informaatiota on tutkittava kriittisesti, vaikka sen synteettisestä luonteesta ei voi olla varma. Teknologisesti on pyrittävä pysymään disinformaation levittäjien kantapäillä, jotta synteettinen media ja tekoälyllä levitetty disinformaatio eivät voita leikkiä.

Synteettinen media ei ole vielä tämän päivän huoli, mutta odottaminen kunnes asiasta muodostuu ongelma ei ole toimiva ratkaisu. Tällä hetkellä synteettinen media ei ole itsessään haitallista OSINT-toiminnalle, kunhan sen olemassaolo tiedostetaan ja tunnistetaan. Jopa tekoälyllä luodun ja levitetyn disinformaation tunnistaminen alkaa informaation aseman ja ympäristön kyseenalaistamisella. Tämä tarkoittaa yleensä informaation alkuperäisen lähteen selvittämistä, mikä voi kertoa paljon informaation olemassaolon syistä, ja tietämällä syy on mahdollista tarkastella disinformaation olemassaolon tarkoitusta laajemmassa kontekstissa. Informaation varmistamiseksi on keksittävä myös uusia toimintatapoja ja varmistuskeinoja riippumatta sen synteettisestä luonteesta, esimerkiksi hyödyntämällä aiemmin esiteltyjä lohkoketjuja, jotta syvävääreennetyt tai väärään kontekstiin liitetyt kuvat voidaan varmistaa helposti luotettavaksi tai harhaanjohtavaksi.

Tutkielman tarkoituksena ja tavoitteena oli tallentaa tämänhetkinen tilanne tekoälyn asemasta osana disinformaation levittämistä ja osana OSINTin toimintaa. Teknologia kehittyy nopeasti ja tämän päivän ongelmat voivat olla huomenna jo unohdettu tai vastoin jopa huonontuneet, mutta pohja tulevaisuuden tutkimuksille rakennetaan nykyisten ongelmien ja keinojen tarkastelulla. Harhaanjohtava informaatio ottaa kuitenkin monia muotoja, ja tutkielman näennäisesti tarkka rajaus oli sekä liian laaja, että liian suppea kattamaan nämä kaikki. Tutkielman tarkastelusta karsiutui pois monia tärkeitä ja käsittelyä tarvitsevia aiheita, kuten pikaisesti mainitut pintavääreennökset tai audiosyvävääreennökset.

Erityisesti lisää tutkimusta ansaitsee tutkielmassa esiteltyjen henkilökohtaisten vaikutusten, kuten päätöksenteon vääristymien, sosiaalisen hyväksyttävyyden sekä vividness weighting -ilmiön tutkiminen tiedustelua suorittavaan henkilöön sekä muihin faktantarkistajiin. Miten jatkuva altistus harhaanjohtavalle informaatiolle vaikuttaa tiedustelun suorittajaan, vaikka he ovatkin täysin tietoisia informaation harhaanjohtavuudesta? Aiheuttaako harhaanjohtava informaatio muutoksia henkilöiden maailmankuvaan, ja miten tämä eroaa riippuen tietoisuudesta informaation harhaanjohtavuudesta?

LÄHTEET

Andrew, C. (2018). *The secret world: A history of intelligence*. Yale University Press.

Arcos, R., Arribas, C. M. (2024). Anticipatory approaches to disinformation, warning and supporting technologies. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 401–416). Routledge. https://doi.org/10.4324/9781003190363

Aslett, K., Sanderson, Z., Godel, W., Persily, N., Nagler, J., & Tucker, J. A. (2023). Online searches to evaluate misinformation can increase its perceived veracity. *Nature*, *625*, 548–556. https://doi.org/10.1038/s41586-023-06883-y

Basheer, R., & Alkhatib, B. (2021). Threats from the dark: A review over dark web investigation research for cyber threat intelligence. *Journal of Computer Networks and Communications, 2021*. https://doi.org/10.1155/2021/1302999

Bastick, Z. (2021). Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior*, *116*. https://doi.org/10.1016/j.chb.2020.106633

Bellingcat. (2023). *Bellingcat's online investigation toolkit* [laskutaulukko]. Haettu 17.4.2024 osoitteesta https://docs.google.com/spreadsheets/d/18rtqh8EG2q1xBo2cLNyhIDuK9jrPGwYr9DI2UncoqJQ/edit#gid=930747607

Bergman, M. K. (2001). White Paper: The deep web: Surfacing hidden value. *The Journal of Electronic Publishing*, *7*(1). https://doi.org/10.3998/3336451.0007.104

Block, L. (2023). The long history of OSINT. *Journal of Intelligence History*, 1–15. https://doi.org/10.1080/16161262.2023.2224091

Cialdini, R. B. (2014). *Influence: Science and practice* (5. uud. painos). Pearson.

Citron, D. K., & Chesney, R. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, *107*(6), 1753–1820. https://doi.org/https://doi.org/10.15779/Z38RV0D15J

Clark, R. M. (2013). *Intelligence collection*. Sage.

Clark, R. M., & Mitchell, W. L. (2019). *Deception: Counterdeception and counterintelligence*. Sage.

Department of the Army. (2006). *Human Intelligence Collector Operations* (FM 2-22.3). https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/fm2_22x3.pdf

Dien, J. (2023). Editorial: Generative artificial intelligence as a plagiarism problem. *Biological Psychology*, *181*. https://doi.org/10.1016/j.biopsycho.2023.108621

Dretske, F. I. (1982). *Knowledge and the flow of information*. MIT Press.

Drolsbach, C. P., & Pröllochs, N. (2023). Diffusion of community fact-checked misinformation on Twitter. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–22. https://doi.org/10.1145/3610058

Fallis, D. (2015). What is disinformation? *Library Trends*, *63*(3), 401–426. https://doi.org/10.1353/lib.2015.0014

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of Social Bots. *Communications of the ACM*, *59*(7), 96–104. https://doi.org/10.1145/2818717

Fetzer, J. H. (2004). Information: Does it have to be true? *Minds and Machines*, *14*(2), 223–229. https://doi.org/10.1023/b:mind.0000021682.61365.56

Floridi, L. (2005). Is semantic information meaningful data? *Philosophy and Phenomenological Research*, *70*(2), 351–370. https://doi.org/10.1111/j.1933-1592.2005.tb00531.x

Fox, C. J. (1983). *Information and misinformation: An investigation of the notions of information, misinformation, informing, and misinforming*. Greenwood Press.

Gibson, H. (2017). Acquisition and preparation of data for OSINT investigations. Teoksessa B. Akhgar, P. S. Bayerl & F. Sampson (toim.), *Open source intelligence investigation: From strategy to implementation* (s. 69–93). Springer International Publishing. https://doi.org/10.1007/978-3-319-47671-1

Gibson, S. D. (2007). *Open source intelligence (OSINT): A contemporary intelligence lifeline* [väitöskirja, Cranfield University]. Cranfield CERES. https://dspace.lib.cranfield.ac.uk/handle/1826/6524

Gibson, S. D. (2013). Open source intelligence. Teoksessa R. Dover, M. S. Goodman & C. Hillebrand (toim.), *Routledge companion to intelligence studies* (s. 123–131). Routledge. https://doi.org/10.4324/9780203762721

Goldman, J., & Maret, S. (2016). *Intelligence and information policy for national security: Key terms and concepts*. Rowman & Littlefield.

Government Communications Headquarters. (2021). *The Ethics of Artificial Intelligence: Pioneering a New National Security*. https://www.gchq.gov.uk/files/GCHQAIPaper.pdf

Gregory, S. (2023). Fortify the truth: How to defend human rights in an age of Deepfakes and Generative AI. *Journal of Human Rights Practice*, *15*(3), 702–714. https://doi.org/10.1093/jhuman/huad035

Groh, M., Epstein, Z., Firestone, C., & Picard, R. (2021). Deepfake detection by human crowds, machines, and machine-informed crowds. *Proceedings of*

the National Academy of Sciences, *119*(1). https://doi.org/10.1073/pnas.2110013119

Hameleers, M., Powell, T. E., Van Der Meer, T. G. L. A., & Bos, L. (2020). A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political Communication*, *37*(2), 281–301. https://doi.org/10.1080/10584609.2019.1674979

Hatfield, J. M. (2023). There is no such thing as open source intelligence. *International Journal of Intelligence and CounterIntelligence*, *37*(2), 397–418. https://doi.org/10.1080/08850607.2023.2172367

Hernández-Escayola, P. (2024). Journalistic approaches to information sources, fact-checking, verification, and detached reporting. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 239–249). Routledge. https://doi.org/10.4324/9781003190363

Higgins, E. (2022). *We are Bellingcat: An intelligence agency for the people*. Bloomsbury publishing.

Hulnick, A. S. (2010). The dilemma of open sources intelligence: Is OSINT really intelligence? Teoksessa L. K. Johnson (toim.), *The Oxford handbook of national security intelligence* (s.229–241). Oxford University Press.

Hyska, M. (2023). The politics of past and future: Synthetic media, showing, and telling. *Philosophical Studies*. https://doi.org/10.1007/s11098-023-02062-x

Ienca, M. (2023). On artificial intelligence and manipulation. *Topoi*, *42*(3), 833–842. https://doi.org/10.1007/s11245-023-09940-3

Jardines., E. A. (2016). Open source intelligence. Teoksessa M. M. Lowenthal & R. M. Clark (toim.), *The five disciplines of intelligence collection* (s. 5–43). Sage.

Joint Chiefs of Staff. (2013). *Joint publication 2-0: Joint intelligence*. https://irp.fas.org/doddir/dod/jp2_0.pdf

Kalpakis, G., Tsikrika, T., Cunningham, N., Iliou, C., Vrochidis, S., Middleton, J., Kompatsiaris, I. (2017). OSINT and the dark web. Teoksessa B. Akhgar, P. S. Bayerl & F. Sampson (toim.), *Open source intelligence investigation: From strategy to implementation* (s. 111–132). Springer International Publishing. https://doi.org/10.1007/978-3-319-47671-1

Kivimäki, V.-P. (2024). Open-source information for intelligence purposes: The challenge of disinformation. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 283–294). Routledge. https://doi.org/10.4324/9781003190363

Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, *62*(3), 790–816. https://doi.org/10.1111/0022-3816.00033

Küçük, D., & Can, F. (2020). Stance detection: A survey. *ACM Computing Surveys*, *53*(1), 1–37. https://doi.org/10.1145/3369026

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "Post-Truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lin, T.-H., Chang, M.-C., Chang, C.-C., & Chou, Y.-H. (2022). Government-sponsored disinformation and the severity of respiratory infection epidemics including COVID-19: A global analysis, 2001–2020. *Social Science & Medicine*, *296*. https://doi.org/10.1016/j.socscimed.2022.114744

Lowenthal, M. M. (2019). *Intelligence: From secrets to policy* (8. uud. painos). Sage.

Lowenthal, M. M. & Clark, R. M. (2016). *The five disciplines of intelligence collection*. Sage.

Lu, X., Vijaykumar, S., Jin, Y., & Rogerson, D. (2022). Think before you share: Beliefs and emotions that shaped COVID-19 (mis)information vetting and sharing intentions among WhatsApp users in the United Kingdom. *Telematics and Informatics*, *67*. https://doi.org/10.1016/j.tele.2021.101750

Lukito, J. (2024) Digital disinformation, electoral interference, and systemic distrust. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 122–134). Routledge. https://doi.org/10.4324/9781003190363

Luo, M., Hancock, J. T., & Markowitz, D. M. (2022). Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues. *Communication Research*, *49*(2), 171–195. https://doi.org/10.1177/0093650220921321

Lyons, B. A., Montgomery, J. M., Guess, A. M., Nyhan, B., & Reifler, J. (2021). Overconfidence in news judgments is associated with false news susceptibility. *Proceedings of the National Academy of Sciences*, *118*(23). https://doi.org/10.1073/pnas.2019527118

Marzell, L. (2017). OSINT as part of the strategic national security landscape. Teoksessa B. Akhgar, P. S. Bayerl & F. Sampson (toim.), *Open source intelligence investigation: From strategy to implementation* (s. 33–55). Springer International Publishing. https://doi.org/10.1007/978-3-319-47671-1

Mercado, S. C., (2004). Sailing the sea of OSINT in the Information Age. *Studies in Intelligence, 48*(3). https://www.cia.gov/resources/csi/static/Sailing-the-Sea-OSINT.pdf

Mesaros, A., Heittola, T., Virtanen, T., & Plumbley, M. D. (2021). Sound event detection: A tutorial. *IEEE Signal Processing Magazine*, *38*(5), 67–83. https://doi.org/10.1109/msp.2021.3090678

Ministry of Defence. (2023). *Joint doctrine publication 2-00: Intelligence, counter-intelligence and security support to joint operations* (4. uud. painos).

https://assets.publishing.service.gov.uk/media/653a4b0780884d0013f71bb0/JDP_2_00_Ed_4_web.pdf

Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. *ACM Computing Surveys*, *54*(1), 1–41. https://doi.org/10.1145/3425780

Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & Bauer, M. (2023). Artificial intelligence and increasing misinformation. *The British Journal of Psychiatry*, *224*(2), 33–35. https://doi.org/10.1192/bjp.2023.136

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*. https://doi.org/10.1007/s13278-021-00776-6

NATO. (2002). *NATO open source intelligence handbook v1.2.* https://archive.org/details/NATOOSINTHandbookV1.2/mode/2up

Office of the Director of National Intelligence. (2013). *U.S. National Intelligence: An Overview 2013*. https://www.dni.gov/files/documents/USNI%202013%20Overview_web.pdf

Omand, D. (2013). The cycle of intelligence. Teoksessa R. Dover, M. S. Goodman & C. Hillebrand (toim.), *Routledge companion to intelligence studies* (s. 59–70). Routledge. https://doi.org/10.4324/9780203762721

Padilha, R., Salem, T., Workman, S., Andaló, F. A., Rocha, A., & Jacobs, N. (2022). Content-aware detection of temporal metadata manipulation. *IEEE Transactions on Information Forensics and Security*, *17*, 1316–1327. https://doi.org/10.1109/tifs.2022.3159154

Paterson, T., & Hanley, L. (2020). Political warfare in the digital age: Cyber subversion, information operations and 'deep fakes.' *Australian Journal of International Affairs*, *74*(4), 439–454. https://doi.org/10.1080/10357718.2020.1734772

Pedersen, T., & Jansen, P. T. (2019). Seduced by secrecy – perplexed by complexity: Effects of secret vs open-source on intelligence credibility and analytic confidence. *Intelligence and National Security*, *34*(6), 881–898. https://doi.org/10.1080/02684527.2019.1628453

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, *147*(12), 1865–1880. https://doi.org/10.1037/xge0000465

Pherson, R. H., Labriny, D., DiOrio, A. (2024) Historical disinformation practices: Learning from the Russians. The challenge of disinformation. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 59–83). Routledge. https://doi.org/10.4324/9781003190363

Price, G., & Sherman, C. (2001). *The invisible web: Uncovering information sources search engines can't see*. Information Today.

Ramwell, S., Day, T., Gibson, H. (2017). Use cases and best practices for LEAs. Teoksessa B. Akhgar, P. S. Bayerl & F. Sampson (toim.), *Open source intelligence investigation: From strategy to implementation* (s. 197–211). Springer International Publishing. https://doi.org/10.1007/978-3-319-47671-1

Roberts, N., & Qahri-Saremi, H. (2023). Tragedy, truth, and technology: The 3T theory of social media-driven misinformation. *Journal of the Association for Information Systems*, *24*(5), 1358–1377. https://doi.org/10.17705/1jais.00831

Saling, L. L., Mallal, D., Scholer, F., Skelton, R., & Spina, D. (2021). No one is immune to misinformation: An investigation of misinformation sharing by subscribers to a fact-checking newsletter. *PLoS ONE*, *16*(8). https://doi.org/10.1371/journal.pone.0255702

Sanastokeskus. (2022). *Informaatiovaikuttamisen käsitteitä*. https://www.huoltovarmuuskeskus.fi/files/08b491d7e81367aadf73960423421cfb897135f1/informaatiovaikuttamisen-kasitteita.pdf

Scarantino, A., & Piccini, G. (2010). Information without truth. *Metaphilosophy*, *41*(3), 313–330. https://doi.org/10.1111/j.1467-9973.2010.01632.x

Segond, F. (2024). AI technologies to support detection of information manipulation on social networks and online media: A quick overview. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 417–425). Routledge. https://doi.org/10.4324/9781003190363

Sobieszek, A., & Price, T. (2022). Playing games with AIs: The limits of GPT-3 and similar large language models. *Minds and Machines*, *32*(2), 341–364. https://doi.org/10.1007/s11023-022-09602-0

Søe, S. O. (2019). A unified account of information, misinformation, and disinformation. *Synthese*, *198*(6), 5929–5949. https://doi.org/10.1007/s11229-019-02444-x

Valtioneuvoston kanslia. (2016). *Avoimesti, rohkeasti ja yhdessä: Valtiohallinnon viestintäsuositus*. https://vnk.fi/documents/10616/3541383/Valtionhallinnon-viestintasuositus-2016.pdf

van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023, helmikuu 3). *ChatGPT: Five priorities for research*. Nature. https://www.nature.com/articles/d41586-023-00288-7

Venema, A. E. (2024). Deepfake disinformation: How digital deception and synthetic media threaten national security. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 175–191). Routledge. https://doi.org/10.4324/9781003190363

Vinhas, O., & Bastos, M. (2022). Fact-checking misinformation: Eight notes on consensus reality. *Journalism Studies*, *23*(4), 448–468. https://doi.org/10.1080/1461670x.2022.2031259

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wang, Y., & Dantcheva, A. (2020). A video is worth more than 1000 lies. comparing 3DCNN approaches for detecting deepfakes. *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 515–519. https://doi.org/10.1109/fg47880.2020.00089

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe report. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Weikmann, T., & Lecheler, S. (2023a). Cutting through the hype: Understanding the implications of deepfakes for the fact-checking actor-network. *Digital Journalism*. https://doi.org/10.1080/21670811.2023.2194665

Weikmann, T., & Lecheler, S. (2023b). Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society*, *25*(12), 3696–3713. https://doi.org/10.1177/14614448221141648

Wirtz, J. J. (2024) Military deception and perception management. Teoksessa R. Arcos, I. Chiru & C. Ivan (toim.), *Routledge handbook of disinformation and national security* (s. 31–41). Routledge. https://doi.org/10.4324/9781003190363