Joel Juusti

IDENTITEETIN PEITTÄMINEN JA SUOJAAMINEN TEHTÄESSÄ AVOINTEN LÄHTEIDEN TIEDUSTELUA VERKOSSA

JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2023

TIIVISTELMÄ

Juusti, Joel
Identiteetin peittäminen ja suojaaminen tehtäessä avointen lähteiden tiedustelua verkossa
Jyväskylä: Jyväskylän yliopisto, 2023, 97 s.
Kyberturvallisuus, pro gradu -tutkielma
Ohjaaja: Hämäläinen, Timo

Ukrainan sota on nostanut esille verkossa toteutettavan avointen lähteiden tiedustelun ja globaalin digitalisaation ansiosta kenen tahansa on mahdollista hankkia tietoa verkon avoimista lähteistä. Verkossa toimiminen jättää kuitenkin jälkeensä jälkiä, joiden avulla käyttäjän identiteetti voidaan yrittää selvittää. Tutkimuksen tavoitteena oli selvittää, millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää toteutettaessa avointen lähteiden tiedustelua, ja miten näitä jalanjälkiä voidaan peittää tehokkaasti. Lisäksi tavoitteena oli selvittää, kykeneekö tutkijan kehittämää teoreettista mallia hyödyntämällä peittämään ja suojaamaan käyttäjän identiteetin tehokkaasti teknisestä näkökulmasta, kun tehdään avointen lähteiden tiedustelua verkossa. Tutkimus toteutettiin laadullisena tutkimuksena, jonka aineisto kerättiin kirjallisuuskatsauksen, teemahaastatteluiden sekä empiiristen testien avulla. Tutkimuksessa aineiston analyysimenetelmänä käytettiin teorialähtöistä sisällönanalyysia. Merkittävimmiksi passiivisiksi digitaalisiksi jalanjäljiksi osoittautuivat IP-osoite, selaimen- ja web-palveluiden keräämä data sekä MAC-osoite. Turvallisimmaksi ratkaisuksi jälkien peittämisessä osoittautui teknisten ratkaisujen välttäminen tiedonhankinnassa tai passiivisten työkalujen ja lähteiden hyödyntäminen. Merkittävimpiin passiivisiin digitaalisiin jalanjälkiin tulee ottaa kantaa tutkimuksessa kehitetyn mallin mukaisesti, mikäli tiedonhankinta täytyy verkossa teknisesti kuitenkin toteuttaa. Valittaessa kontrollit passiivisten digitaalisten jalanjälkien peittämiseksi, tulisi valitut kontrollit suhteuttaa tapauskohtaiseen riskiarvioon. Lisäksi peitettäessä digitaalisia jalanjälkiä, ei jalanjäljistä muodostettu kokonaisuus saisi erottua anomaliana muista käyttäjistä tai liikenteestä. Tutkimuksessa kehitetyn Maksimaalisen digitaalisen turvallisuuden mallin nähtiin soveltuvan viitekehykseksi passiivisten digitaalisten jalanjälkien peittämisessä, ja haastatellut asiantuntijat kertoivat itse käyttävänsä samoja menetelmiä suojautuessaan verkossa. Suurimmiksi haasteiksi nähtiin mallin ymmärtämiseen vaadittava korkea tietotaito sekä mallin käytäntöön soveltaminen. Käytännön ratkaisun toteuttamisen nähtiin vaativan merkittävästi osaamista, sillä sovellettaessa mallia käytäntöön, korostuvat myös aktiiviset digitaaliset jalanjäljet. Vaikka käyttäjän käyttämät tekniset ratkaisut olisivatkin toimivat, voi käyttäjä omalla toiminnallaan paljastaa identiteettinsä verkossa.

Asiasanat: digitaaliset jalanjäljet, digitaalisten jalanjälkien peittäminen, identiteetti, avointen lähteiden tiedustelu

ABSTRACT

Juusti, Joel
Masking and protecting identity when conducting open-source intelligence online.
Jyväskylä: University of Jyväskylä, 2023, 97 pp.
Cyber Security, Master's Thesis
Supervisor: Hämäläinen, Timo

The war in Ukraine has brought attention to open source intelligence gathering conducted online and the global digitalization, which enables anyone to obtain information from open sources on the Internet. However, operating online leaves traces that can be used to attempt to identify a user's identity. The objective of the research was to determine the types of passive digital footprints that remain on the internet when conducting open source intelligence and how these footprints can be effectively concealed. Additionally, the aim was to investigate whether the researcher's developed theoretical model could effectively hide and protect a user's identity from a technical perspective when conducting open source intelligence online. The study was conducted as a qualitative research, and the data was collected through literature review, thematic interviews, and empirical tests. The data analysis method employed in the research was theory-driven content analysis. The most significant passive digital footprints were found to be IP address, data collected by the browser and web services, and MAC address. The safest solution for concealing these traces proved to be avoiding technical solutions in information gathering and instead utilizing passive tools and sources. Regarding the most significant passive digital footprints, the research recommended taking a stance according to the developed model if technical information gathering is still required online. When selecting controls to conceal passive digital footprints, the chosen controls should be tailored to the specific risk assessment. Furthermore, when concealing digital footprints, the resulting ensemble should not stand out as an anomaly among other users or traffic. The developed Maximum Digital Security model in the research was seen as suitable framework for covering passive digital footprints, and the interviewed experts reported using the same methods to protect themselves online. The main challenges were perceived to be the high level of expertise required to understand the model and the application of the model in practice. Implementing practical solutions was seen as demanding significant expertise, as the application of the model also emphasizes active digital footprints. Even if the user's chosen technical solutions were effective, the user can still expose their identity through their own actions online.

Keywords: digital footprints, covering digital footprints, identity, open-source intelligence gathering

KUVIOT

SISÄLLYS

1 JOHDANTO

Vuonna 2023 on arvioitu, että internetiä käyttää hieman yli viisi miljardia ihmistä (Statista, 2023). Informaatioteknologian käyttö lisääntyykin jatkuvasti maailman digitalisoituessa ja tieto muuttuu entistä enemmän digitaaliseen muotoon. Tiedon muuttuessa digitaaliseksi ja siirtyessä verkkoon, myös tiedonhankinnan merkitys verkon avoimista lähteistä kasvaa. Avointen lähteiden tiedustelu, eli OSINT-tiedustelu verkossa, ei ole enää vain valtiollisten toimijoiden yksinoikeus, vaan sen harjoittaminen on mahdollista kenelle tahansa. Toimiessaan verkossa ja etenkin toteuttaessaan avointen lähteiden tiedustelua, tulisi myös suojautumiseen kiinnittää huomiota.

Tutkimuksessa pyritään selvittämään kuvailevan kirjallisuuskatsauksen avulla, millaisia passiivisia digitaalisia jalanjälkiä henkilö jättää toimiessaan verkossa ja miten näitä jälkiä voidaan peittää. Vastausten pohjalta kehitetään teoreettinen malli, jota soveltamalla kuka tahansa voi peittää passiivisia digitaalisia jalanjälkiään suojatakseen identiteettinsä verkossa. Teoreettisen mallin toimivuus varmistetaan asiantuntijahaastatteluiden avulla ja empiirisillä PoC (Proof of Concept) -henkisillä testeillä. Tämä pro gradu -tutkielma toteutettiin ilman toimeksiantoa, ja aihe valikoitui tutkijan omasta mielenkiinnosta ja halusta perehtyä aiheeseen syvällisemmin. Tutkimuksella pyritään vastaamaan seuraaviin kysymyksiin:

Tutkimuksen pääkysymys:

1. Kykeneekö tutkijan kehittämää teoreettista mallia hyödyntämällä peittämään ja suojaamaan käyttäjän identiteetin tehokkaasti teknisestä näkökulmasta, tehtäessä avointen lähteiden tiedustelua verkossa?

Tutkimuksen apukysymykset:

1. Millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää tehtäessä avointen lähteiden tiedustelua?

2. Miten passiivisia digitaalisia jalanjälkiä voidaan peittää tehokkaasti verkossa?

Tutkimus on ajankohtainen, sillä tiedon siirryttyä verkkoon OSINT-tiedustelusta on tullut merkittävä osa muidenkin kuin valtiollisten toimijoiden toimintaa. On olemassa yrityksiä, jotka hankkivat, käsittelevät tai jopa tuottavat tietoa. Pörssiyhtiöt hankkivat tietoa kilpailijoistaan ja toimintaympäristöstä pärjätäkseen markkinoilla. Tiedonhankinnan merkitys myös yksittäisissä työtehtävissä kasvaa jatkuvasti. OSINT-tiedustelu on esimerkiksi merkittävä osa tietoturvatestaajien päivittäistä työtä (Li, 2021, s. 61–62). Ukrainan sodan alkamisen jälkeen OSINT-tiedustelu on nostanut profiiliaan myös Suomessa. Lehdistö on alkanut käyttämään entistä enemmän OSINT-tiedusteluun liittyvää tietoa, lähteitä ja asiantuntijoita uutisoinnissaan (Iltalehti, 2022a). Ukrainan sodan seuraaminen on mahdollista lähes reaaliajassa kenelle tahansa sosiaalisen median alustojen avulla. Kotimaisen tiedusteluosaamisen kysyntä on kasvussa, tiedustelun tutkimus ja opetus kehittyvät jatkuvasti, ja niiden tarve tunnustetaan yhä paremmin.

Tutkimus on tarpeellinen, sillä toteutettaessa kuvailevaa kirjallisuuskatsausta, ei tutkija löytänyt yhtäkään tutkimusta, joka käsittelisi OSINT-tiedustelua verkossa turvallisuuden näkökulmasta. Ulkomaalaisessa kirjallisuudessa aihetta on käsitelty paljon, mutta teokset perustuvat kirjoittajien asiantuntemukseen eivätkä tieteelliseen tutkimukseen. Lisäksi kirjallisuudessa käsitellyt kokonaisuudet ovat usein keskenään ristiriitaisia eivätkä käsittele aihetta kokonaisvaltaisesti. Ne keskittyvät usein johonkin spesifiin osa-alueeseen digitaalisiin jalanjälkiin liittyen. Tutkimusta tehtäessä selvisi myös, että digitaalisten jalanjälkien peittämiseksi on olemassa kaupallisia tuotteita, joita voidaan soveltaa digitaalisten jalanjälkien peittämiseksi, kuten ZeroTier (2023) sekä turvallisuus edellä suunniteltuja käyttöjärjestelmiä. Näissä ongelmana voidaan kuitenkin nähdä se, että yleensä toimija joutuu jättämään merkittävästi tietoja tuotteen tarjoajalle, ne voivat olla "tavalliselle käyttäjälle" haastavia käyttää, eivät lisää käyttäjän ymmärrystä digitaalisista jalanjäljistä tai tuotteet maksavat, jolloin ne eivät välttämättä ole kaikkien saatavilla.

Ihmisoikeuskeskus ja YK:n ihmisoikeustoimisto (2022) ovat luoneet käytännöllisen ohjeen siitä, kuinka OSINT-tiedustelu tulee toteuttaa, jotta tieto on käyttökelpoista esimerkiksi kansainvälisen oikeuden näkökulmasta. Samanlainen ohje löytyy Bellingcatin (2022) tekemänä Ukrainan sotaan liittyen. Molemmat ohjeet korostavat turvallisuuden merkitystä ja suojautumista. Mallit määrittävät minimimenetelmät ja standardit suojautumista varten korkealla, abstraktilla tasolla. Niiden ongelmaksi voidaan nähdä se, että mallit eivät ole kokonaisvaltaisia, vaan niissä määritellään "minimi" toimenpiteet, jotka pitäisi ottaa huomioon suojautumiseen liittyen. Ne eivät kerro lukijalle yksityiskohtaisesti, mitkä ovat kyseisten menetelmien vahvuudet ja heikkoudet sekä miksi tietyllä tavalla tulisi menetellä (Human Rights Center & United Nations Human Rights, 2022, s. 31–34; Bellingcat, 2022, s. 10–13.) Bellingcatin ohjeiden taustalla voidaan nähdä olevan omakohtaista kokemusta vaikutusyrityksistä. Suomalaisen tutkijan autettua Bellingcatiä selvittämään Malaysian Airlinesin lennon MH17 alasampumista,

hänen sähköpostiinsa yritettiin tunkeutua (Iltasanomat, 2021). Toinen Bellingcatiin liittyvä taho on joutunut Venäjän etsintäkuuluttamaksi (Yle, 2022). Mikäli nämä OSINT-tiedusteluun erikoistuneet ammattilaiset paljastuvat verkossa ja joutuvat vaikutusyritysten kohteeksi, mikä takaa muille toimijoille sen, ettei näin tapahtuisi heille? Tutkimuksen tarkoituksena onkin lisätä kaikkien OSINT-tiedustelun parissa toimivien ymmärrystä passiivisista digitaalisista jalanjäljistä ja passiivisten digitaalisten jalanjälkien peittämisestä. Kaikilla OSINT-tiedustelua tekevillä tahoilla ei tietotekninen osaaminen välttämättä ole korkealla tasolla, jolloin he saattavat paljastaa todellisen henkilöllisyytensä ja altistaa itsensä erilaisille vaikutusyrityksille. Operaatioturvallisuudella, jota voidaan käyttää lyhenteillä OPSEC tai OPTU, tarkoitetaan informaatioita, joita pyritään pitämään salassa vastustajalta, jotta operaation tavoitteita ja toimintaa voidaan suojella (Puolustusvoimat, 2020, s. 48–49). Digitaaliset jalanjäljet voidaan nähdä merkittävänä osana operaatioturvallisuutta, kun tehdään avointen lähteiden tiedustelua verkossa. Digitaalisten jalanjälkien jättäminen voi paljastaa toimijan aikeita, menetelmiä, taktiikoita ja strategioita. Digitaalisten jalanjälkien peittäminen ja suojaaminen ovat siten keskeisiä osa-alueita, joita tulisi ottaa huomioon OSINT-tiedustelua tehdessä.

Tutkimuksen avulla voidaan lisätä tietoisuutta passiivisista digitaalisista jalanjäljistä ja niiden peittämisestä OSINT-tiedustelun yhteydessä. Tutkimuksen tuloksia voidaan hyödyntää niin yksityisten henkilöiden kuin organisaatioidenkin toiminnassa verkossa. Lisääntynyt tietoisuus digitaalisista jalanjäljistä ja niiden peittämisestä voi auttaa suojaamaan henkilöiden ja organisaatioiden identiteettiä verkossa sekä vähentämään altistumista vaikutusyrityksille ja tietomurroille. Tuloksena kehitetty teoreettinen malli voi tarjota käytännön ohjeita ja menetelmiä passiivisten digitaalisten jalanjälkien peittämiseen, joka voi auttaa sekä aloittelevia että kokeneita OSINT-tiedustelijoita suojautumaan verkossa. Mallin avulla kuka tahansa voi lisätä turvallisuuttaan ja parantaa tietosuojaansa verkossa, kun he tekevät avointen lähteiden tiedustelua.

1.1 Tutkimusmenetelmät ja tutkimuksen rakenne

Tutkimus on kvalitatiivinen, eli laadullinen. Tutkimuksessa käytetään myös konstruktiivista tutkimusotetta teoreettisen mallin kokonaisuuksien testaamiseen käytännössä. Kirjallisuuskatsauksen metodina käytetään kuvailevaa kirjallisuuskatsausta. Kirjallisuuskatsauksen muoto on narratiivinen. Narratiivisessa kirjallisuuskatsauksessa epäyhtenäistä tietoa järjestetään jatkuvaksi tapahtumaksi. Kirjallisuuskatsauksessa pyritäänkin rakentamaan usean lähteen pohjalta yhtenäinen kuvaus käsiteltävästä aiheesta. Kirjallisuuskatsauksen toimitustapana on käytetty yleiskatsausta. Kyse on laajemmasta prosessista, jonka tarkoituksena on tiivistää aiemmin tehtyä kirjallisuutta (Salminen, 2011, s. 7–8). Kirjallisuuskatsauksen pääasialliset lähteet ovat asiantuntijoiden kirjoittamaa kirjallisuutta. Näitä lähteitä ovat aihealueeseen liittyvät kirjat, artikkelit, tutkimukset, ohjeet ja muut asiakirjat. Esimerkkejä asioiden konkretisoimiseksi, lukuelämyksen

parantamiseksi ja tietoaukkojen täydentämiseksi paikataan Internet-lähteillä, kuten uutisartikkeleilla ja alan tunnustettujen toimijoiden Internet-julkaisuilla. Kirjallisuuskatsauksen avulla pyritään vastaamaan tutkimuksen alakysymyksiin ja luomaan tutkijan oma konstruktio, teoreettinen malli identiteetin peittämiseksi ja suojaamiseksi toimittaessa verkossa.

Tutkimuksen empiirisessä osassa aineistonkeruumenetelmänä käytetään asiantuntijahaastatteluja ja testataan tutkijan konstruktiivista mallia empiirisesti PoC (Proof of Concept) -hengessä. Testien avulla pyritään testaamaan haastatteluissa ilmenneitä asioita käytännössä ja luomaan lukijalle esimerkkejä siitä, miksi tietyt teoriakokonaisuudet mallissa tulee huomioida. Haastattelut toteutetaan teema- eli puolistrukturoituina haastatteluina. Puolistrukturoiduissa haastatteluissa haastateltavat saavat vastata samoihin kysymyksiin vapaasti, omin sanoin, valmiiden vastausvaihtoehtojen sijaan (Eskola & Suoranta, 1998, s. 63). Näin varmistetaan, että asiantuntijoilla on mahdollisuus lisätä huomioitaan aiheeseen liittyen, mikäli tutkijalta on jäänyt jokin merkittävä asia huomioimatta. Empiiristä testaamista varten tutkijan kehittämää mallia sovelletaan käytäntöön ja tehdään havaintoja mallin mukaisista kokonaisuuksista. Testien tuloksia verrataan toisiinsa ja selvitetään mallin eri kokonaisuuksien toimivuutta käytännössä. Tutkimuksen aineiston analyysimenetelmänä käytetään teorialähtöistä analyysiä. Analyysissä nojataan tutkijan konstruktiivisen tutkimuksen avulla kehittämään teoreettiseen malliin käyttäjän identiteetin peittämiseksi ja suojaamiseksi (Tuomi & Sarajärvi, 2018, s. 107–108.)

Tutkimuksen teorialuvuissa käsitellään kattavasti digitaaliset jalanjäljet, menetelmät jälkien peittämiseksi ja esitellään teoreettinen malli, jonka avulla digitaalisia jalanjälkiä peittämällä käyttäjän identiteetti voidaan suojata. Mallia käytetään tutkimuksen viitekehyksenä. Neljännessä luvussa esitellään tutkimuksen toteutus ja tutkimusmenetelmät. Viidennessä luvussa käsitellään tutkimuksen tulokset, johtopäätökset ja pohdinta. Viimeisessä luvussa käsitellään mahdollisia jatkotutkimusaiheita ja tarkastellaan tutkimuksen eettisyyttä sekä luotettavuutta.

1.2 Tutkimuksen rajaus ja etiikka

Bazzell (2018) määrittelee OSINT-tiedustelun tarkoittavan tiedustelua ja tiedonhankintaa, jossa tieto kerätään julkisista, avoimista lähteistä. Viranomaisille se voi tarkoittaa ilmakuvien tai uutisten seuraamista vieraista valtioista. Asianajajalle se voi tarkoittaa julkisten rekisterien ja asiakirjojen pyytämistä viranomaisilta. Useimmille ihmisille se on kuitenkin julkisen tiedon hankintaa internetin välityksellä. Samanlainen tulkinta on myös Yhdysvaltojen puolustusministeriöllä (Bazzell, 2018, s. IV; Hassan & Hijazi, 2018, s. 2). Tässä tutkimuksessa käsitellään digitaalisia jälkiä, joita jää OSINT-tiedustelusta verkossa, eli internetissä.

Useat teokset, jotka käsittelevät teknisestä näkökulmasta digitaalisia jälkiä ja niiden peittämistä, alkavat usein "vastuuvapauslausekkeella", jossa pohditaan sitä, kuinka teosten oppeja voidaan käyttää väärin. Shavers & Bair (2016)

käsittelevät kirjassaan poliisin digitaalista forensiikkaa. He vertaavat digitaalisiin jälkiin liittyvää teknistä osaamista autolla ajamiseen ja korostavat yksilön vastuuta. Autokoulussa opetetaan ajamaan autoa, mutta ei opeteta ajamaan ylinopeutta tai kuinka ajoneuvolla paetaan viranomaisia (Shavers & Bair, 2016, s. vi.)

Tutkimuksen rajaamiseksi ja eettisyyden lisäämiseksi tässä tutkimuksessa tarkastellaan passiivisia digitaalisia jälkiä ja teknisiä ratkaisuja niiden peittämiseksi. Tutkijan kehittämä malli on tekninen viitekehys passiivisten digitaalisten jälkien peittämiseksi. Käyttäjän tekemät ratkaisut, virheet ja niiden välttämisen käsittely jätetään mahdollisimman vähäiseksi. Teoreettisen mallin hyödyntäminen käytännössä jätetään lukijan vastuulle. Shaversin ja Bairin (2016) mukaan usein suurin heikkous digitaalisten jälkien peittämisessä on itse käyttäjä ja hänen tekemänsä virheet, teknologisten ratkaisujen sijaan. Kaikki tutkimuksessa käsitellyt asiat perustuvatkin julkisiin lähteisiin.

Tutkimuksen tärkeimpänä tuotteena on tekninen viitekehys, malli passiivisten digitaalisten jälkien peittämiseksi ja identiteetin suojaamiseksi verkossa. Tietyt ohjelmat tai ohjelmistoratkaisut eivät välttämättä kestä aikaa. Teoreettisen mallin voidaan kuitenkin nähdä kestävän paremmin aikaa. Mallin testaamista varten kehitetään kuitenkin esimerkkiratkaisu, jossa käytetään valittuja ohjelmia. Ohjelmien valinta on rajattu niin, että malli rakennetaan Windows-käyttöjärjestelmän päälle. Windows-käyttöjärjestelmän voidaan katsoa olevan suurimmalle osalle ihmisistä tuttu, jolloin mallin soveltamisen voidaan nähdä olevan kaikille yleisesti helpompaa. Ohjelmina digitaalisten jälkien peittämiseksi ja suojaamiseksi pyritään käyttämään ja suosimaan ilmaisia, avoimen lähdekoodin ohjelmia, lukuun ottamatta maksullista VPN-palvelua. Avoimen lähdekoodin ohjelmien voidaan katsoa olevan turvallisia siinä mielessä, että niiden toiminnallisuutta on mahdollisuus tarkastella ja käyttäjä voi varmistua siitä, mitä ohjelma oikeasti tekee. Ilmaisilla ohjelmilla pyritään siihen, että mallin soveltaminen olisi kenelle tahansa mahdollista, lähtökohdasta tai taustasta riippumatta.

Tutkimuksen tarkoituksena ei ole antaa menetelmiä ja keinoja siihen, miten digitaaliset jäljet peitetään kotimaisilta viranomaisilta. Tutkimuksen tarkoituksena on antaa kokonaisvaltainen malli, teknisestä ja teoreettisesta näkökulmasta lukijalle, jota soveltamalla lukija voi suojautua muilta verkosta kohdistuvilta uhilta. Niinpä tässä tutkimuksessa ei käsitellä kattavasti rahaliikenteen peittämistä ja laitteiden hankkimista. Ne käsitellään tarvittavilta osin, jotta pystytään välttämään ylimääräisten tietojen antaminen kolmansille osapuolille. Digitaalisia jälkiä ja niiden peittämistä käsitellään vain tavanomaisen tietokoneen näkökulmasta, ja muut käyttöön soveltuvat laitteet suljetaan tämän tarkastelun ulkopuolelle, sillä niihin liittyy omia ominaispiirteitään ja ne aiheuttavat kyseisille laitteille tyypillisiä digitaalisia jalanjälkiä.

2 DIGITAALISET JALANJÄLJET JA NIIDEN PEITTÄMINEN

Christenssonin (2014) mukaan digitaalinen jalanjälki on datasta muodostuva jälki, joka syntyy käyttäjän käyttäessä internetiä. Jälkiä jää esimerkiksi käytettäessä internetissä verkkosivuja, lähetettäessä sähköpostia tai käyttäjän itse laittaessa dataa internetpalveluihin. Christensson jakaa jäljen aktiiviseen ja passiiviseen jalanjälkeen. Aktiivinen jalanjälki tarkoittaa sitä dataa, jonka käyttäjä itse tietoisesti laittaa internetiin, kuten kuvat tai blogikirjoitukset. Passiivisella jalanjäljellä tarkoitetaan dataa, jonka käyttäjä jättää tarkoituksettomasti internetiin ja jonka jokin taho tai toiminnallisuus kerää käyttäjästä. Tämä data voi olla esimerkiksi evästeitä ja muita selaimen tallentamia tietoja (Christensson, 2014.) Tässä tutkimuksessa tarkastellaan passiivisia digitaalisia jalanjälkiä ja digitaalisten jalanjälkien peittämisellä tarkoitetaan tässä tutkimuksessa niitä menetelmiä, joiden avulla käyttäjä voi peittää, suojata, muuttaa tai millä tahansa tavalla häivyttää digitaalisia jalanjälkiään.

2.1 Aiempi tutkimus

Ayed (2011) käsitteli tutkimuksessaan teknistä lähestymistapaa digitaalisen identiteetin aiheuttamien riskien vähentämiseksi kirjallisuuskatsauksen avulla. Hänen mukaansa digitaalisen identiteetin muodostamiseen liittyy merkittäviä riskejä, joita ovat muun muassa identiteettivarkaus ja tietovuodot. Ayed ehdottaakin XML-pohjaisen metadata dokumentin käyttöä. Dokumentti lisäisi metadatan keräämiseen läpinäkyvyyttä sekä käyttäjän mahdollisuuksia hallita omaa digitaalista identiteettiään. Fan, Chow ja Xu (2014) sekä Abramson ja Aha (2013) osoittivat tutkimuksissaan, kuinka on mahdollista tunnistaa henkilö pelkästään hänen selaimessa suorittamiensa toimintojen perusteella (Ayed, 2011; Fan, Chow & Xu, 2014; Abramson & Aha, 2013.)

Jamal ja Zain (2022) tutkivat digitaalisten jalanjälkien merkitystä, niiden liittymistä verkkorikollisuuteen ja parhaita menetelmiä digitaalisten jalanjälkien peittämiseksi. Heidän mukaansa suurimmat hyötyjät digitaalisista jalanjäljistä ovat kaupalliset toimijat, kuten mainostajat ja palveluntarjoajat. He korostavat sitä, että monet kyberrikokset liittyvät uhrien digitaalisiin jalanjälkiin. Tällaisia rikoksia ovat esimerkiksi tietojenkalastelu ja identiteettivarkauden avulla tehty petos. Parhaina käytäntöinä digitaalisten jalanjälkien peittämiseksi yksilön näkökulmasta he pitävät harkittua tiedon jakamista sosiaalisessa mediassa ja luotettujen verkkosivujen käyttämistä, jotka käyttävät HTTPS-protokollaa tietoliikenteen salaamiseksi. Ranakoti, Yadav, Apurva, Tomer ja Roy (2017) tutkivat deep web:iä ja tekniikoita sekä tapoja, joiden avulla internetin käyttäjä voi suojata yksityisyyttään verkossa anonymiteetin avulla. Suurimpana uhkana yksilöille he pitävät verkkorikollisia. Parhaan anonymiteetin saavuttamiseksi verkossa tulee

heidän mukaansa ottaa huomioon evästeet, käyttää Tor-selainta, proxy-palvelimia tai VPN-palvelua, peittää kamera, ottaa huomioon käytettävä selain, käytettävä mainosten blokkaajaa ja internetiä tulisi käyttää virtuaalikoneella, joka käyttää turvallista käyttöjärjestelmää (Jamal & Zain, 2022; Ranakoti, Yadav, Apurva, Tomer & Roy, 2017.)

Juan, Shimin, Jun, Bin ja Lei (2021) tutkivat Tor-liikenteen tunnistamista koneoppimisen avulla. He tutkivat onko mahdollista tunnistaa Tor-liikenne tietovirrasta, jossa on mukana muutakin liikennettä. Tutkimuksen johtopäätöksenä oli, että käyttämällä Random-Forest ja KNN algoritmeja, voidaan asiakkaiden Tor-liikenne tunnistaa tehokkaasti. Vlajic, Madani ja Nquyen (2017) tutkivat Tor-selaimen kykyä tarjota anonymiteettiä käyttäjälle. Tutkimuksen mukaan käytettäessä Tor-selainta oletusasetuksilla, ei se tarjoa suojaa neljää yleisimmin käytettyä tekniikkaa vastaan, joilla käyttäjää seurataan verkossa. Tor-selaimen käytön lisäksi tulee heidän mukaansa tehdä myös muita toimenpiteitä, aidon anonymiteetin saavuttamiseksi verkossa. Tutkimuksen mukaan verkkosivujen tekniikat seurata käyttäjää jaetaan IP-perustaiseen seurantaan, eväste seurantaan, URL seurantaan ja välimuisti perustaiseen seurantaan. Tutkimuksen mukaan vaikka käytettäisiin Tor-selainta, pystyvät verkkosivut seuraamaan käyttäjää näiden menetelmien avulla. Tor-selain vaihtaa reititys reittiä kymmenen minuutin välein, mutta keskimääräinen verkkosivuilla käytetty aika on noin neljä ja puoli minuuttia. Näin verkkosivut ehtivätkin keräämään tietoja käyttäjästä. Käyttäjän tulisikin vaihtaa manuaalisesti Tor:in käyttämä reititys muutamien verkkosivun klikkausten välein, klikkaamalla ”New Tor Circuit for this Site”, ”New Identity” ominaisuuksia selaimessa ja muuttamalla selaimen lisäasetukset estämään näitä tekniikoita käyttämällä selaimen ”about:config” ominaisuutta (Juan, Shimin, Jun, Bin & Lei, 2021; Vlajic, Madani & Nquyen, 2017.)

Bernardos, Zuniga, ja O’Hanlon (2015) tutkivat muuttumattomia osoitteita, jotka jäävät verkkoon digitaalisena jalanjälkenä. Heidän mukaansa verkkolaitteiden verkkokortissa sijaitseva, muuttumaton MAC-osoite on haasteellinen yksityisyyden näkökulmasta. MAC-osoite mahdollistaa laitteen yksilöinnin ja se usein sisällytetään metadataan. Ratkaisuksi he ehdottavat MAC-osoitteen satunnaistamista, jolloin osoite muuttuisi aika ajoin. MAC-osoitteen satunnaistamista testattiin tutkimuksessa kolmella erilaisella empiirisellä kokeella. Tulokset osoittivat, että käytäntö lisäisi yksityisyyttä, mutta sen käytössä tulisi huomioida myös muut ympäristön ominaisuudet tehokkuuden lisäämiseksi (Bernardos, Zuniga, & O’Hanlon, 2015.)

2.2 Identiteetti

Identiteetti tarkoittaa niiden piirteiden summaa, joiden avulla henkilö on mahdollista tunnistaa niin, ettei häntä voi sekoittaa toiseen henkilöön. Henkilökohtainen ja digitaalinen identiteetti eroavat toisistaan. Henkilökohtainen identiteetti muodostuu henkilön fyysisistä piirteistä, kun taas digitaalinen identiteetti muodostuu henkilön itsensä jättämistä ja luomista tiedoista, joiden paikkaansa

pitävyys voi vaihdella faktasta fiktioon (Shavers & Bair, 2016, s. 187–188.) Digitaalinen identiteetti tarkoittaa ainutlaatuista identiteettiä, jonka avulla henkilö voidaan tunnistaa digitaalisesti. Se muodostuu lukuisista piirteistä tai ominaisuuksista, joita henkilöstä kerätään verkossa. Nämä piirteet yhdistämällä saadaan luotua ainutlaatuinen kokonaisuus, digitaalinen identiteetti, jonka avulla käyttäjä voidaan tunnistaa. Tämän tiedon keräämiseksi ja henkilön tunnistamiseksi, ei välttämättä tarvitse tietää esimerkiksi henkilön viestien sisältöä. Digitaalinen identiteetti on mahdollista muodostaa pelkästään keräämällä henkilöstä riittävästi metadataa. Metadata tarkoittaa tietoa tiedosta. Esimerkiksi viestin sisältö voidaan nähdä tietona, mutta tiedot siitä kuka lähetti viestin, mihin kellonaikaan, mistä sijainnista ja kenelle ovat metadataa (Ayed, 2011, s. 607–608.)

2.2.1 Identiteettien erottaminen

Mitnickin (2017) mukaan käyttäjän tulee luoda uusi digitaalinen identiteetti ja erotettava oikea identiteetti, uudesta digitaalisesta identiteetistä ollakseen anonyymi verkossa (Mitnick & Vamosi, 2017, s. 46). Tämä tarkoittaa sitä, että käyttäjä ei voi jakaa verkossa mitään sellaista tietoa aktiivisesti tai passiivisesti, josta hänet olisi mahdollista tunnistaa fyysisessä maailmassa. Jotta käyttäjä kykenee keräämään dataa verkossa, tulee hänen kuitenkin usein rekisteröityä erinäisten palveluiden käyttäjäksi (Hassan & Hijazi, 2018).

Oikean identiteetin piilottamiseksi tulee käyttäjän luoda alias sähköposti, alias puhelinnumero, alias kotiosoite ja käyttää anonyymeja maksumenetelmiä, jotka eivät ole seurattavissa käyttäjään. Nämä ovat yleisimmät tiedot, joita henkilöltä kysytään, hänen rekisteröityessään erilaisten verkkopalvelujen käyttäjäksi. Usein rekisteröinti tulee myös vahvistaa, joko puhelimeen tai sähköpostiin saapuvasta viestistä. Verkosta löytyy palveluntarjoajia, jotka eivät vaadi tunnistautumista sähköpostin luomiseksi. Tiettyjen palveluiden käyttämiseksi voi olla tarpeellista luoda suurten palveluntarjoajien sähköpostitili, kuten Hotmail tai Gmail-tili. Tässä tapauksessa tulee olla tietoinen palvelun ehdoista ja tili ei saa olla jäljitettävissä käyttäjään. Myöskään käytettäessä konetta ei tiliin saisi olla kirjautuneena, sillä muuten käyttäjästä kerätään dataa digitaalisen identiteetin rakentamiseksi. Alias puhelinnumeron saa helposti hankkimalla prepaid liittymän. Toinen vaihtoehto on hankkia virtuaalinen puhelinnumero verkosta. Huomioitavaa puhelinnumeroissa on kuitenkin se, että käyttäjän ei ole tarkoitus käyttää näitä numeroita viestintään, vaan vain erilaisten tilien luomiseen. Kotiosoitteen piilottamiseksi on kaksi menetelmää. Toinen menetelmä on hankkia itselleen maksullinen postilokero, mikäli posti halutaan oikeasti vastaanottaa ja toinen mahdollisuus on käyttää osoitetta, jota ei ole olemassa. Tällaisen osoitteen toimimisen varmistamiseksi kannattaa osoite valita uudelta asuin alueelta, mutta muuttaa osoitteen numerot sellaisiksi, että ne eivät vastaa mitään oikeaa osoitetta. Maksutietojen piilottamiseksi parhaat menetelmät ovat virtuaalivaluuttojen käyttö anonyymisti tai kertakäyttöiset prepaid luottokortit (Bazzell, 2016a.)

Näitä alias tietoja ei tule missään nimessä antaa tai kertoa kenellekään. Esimerkiksi annettaessa puhelinnumero ystävälle, tallentaa ystävä puhelinnumeron

todennäköisesti puhelimeensa. Mikäli hän on kirjautuneena esimerkiksi Googlen tiliin, Google kerää ja tallentaa numeron sekä siihen yhdistetyn nimen. Bazzell pitääkin minimi nyrkkisääntönä sitä, että mikäli kyseessä ei ole viranomainen tai pankkia vastaava palvelu, ei koskaan tulisi samalla kertaa tarjota kahta oikeaa yhteystietoa palvelulle. Jos annetaan oikea nimi, ei anneta oikeaa kotiosoitetta ja jos annetaan kotiosoite, ei anneta oikeaa nimeä. Suurimmat syyt näin radikaaleille menetelmille henkilöllisyyden suojaamiseksi ovat tietovuodot ja palveluiden liiketoiminta. Usein tietovuotojen takia verkosta löytyvät yhteystiedot, ovat juuri näitä samoja tietoja, joita henkilön tulee antaa palveluille rekisteröintien yhteydessä (Bazzell, 2016a, s. 329 & 352–353). Antaessa tietojaan palveluille, käyttäjä luovuttaa tietonsa ulkopuolisen toimijan vastuulle. Hyvä esimerkki valtavasta tietovuodosta Suomessa on psykoterapiakeskus Vastaamo, jossa arvioiden mukaan yli 30000 ihmisen mielenterveystiedot vuotivat verkkoon tietomurron yhteydessä (Iltalehti, 2022b). Mikäli Vastaamo olisi säilyttänyt Bazzellin periaatteen mukaisesti henkilötiedot ja terveystiedot erillään, olisi selvitty merkittävästi vähemmillä vahingoilla. Verkkopalveluiden käyttäjien tietojen myyminen muille yrityksille ja erityisesti mainostajille on suurta liiketoimintaa, eikä käyttäjä välttämättä edes ole itse tästä tietoinen. Yritysten myydessä käyttäjän tietoja, ei käyttäjä voi olla itse täysin varma siitä, minne kaikkialle hänen tietojaan on välitetty (Bazzell, 2016a.)

"If you're not paying for the product then you're the product" – Andrew Lewis

2.2.2 Metadata

Metadata tarkoittaa dataa, datasta. Se on usein kuvailevaa ja näkyviltä piilotettua tietoa, joka kuuluu usein esimerkiksi osaksi erilaisia digitaalisia tiedostoja. Se voi olla osana millaisia tiedostoja tahansa, kuten kuvat, videot ja erilaiset tekstitiedostot. Tällaista tietoa voivat olla esimerkiksi käyttäjän nimi, tiedoston koko, sijainti, luomisen ajankohta, luomiseen käytetyt laitteet ja kommentit. Kuvissa tätä metadataa kutsutaan EXIF dataksi (Hassan & Hijazi, 2016, s. 46–47.) Sähköposti viestit sisältävät metadataa, vaikka viestien sisältö olisikin salattu. Tällaista dataa voi olla IP-osoite, laitteen MAC-osoite, tieto siitä, kuka on lähettänyt viestin, kenelle ja milloin. Metadataa muodostuu myös muunlaisesta viestinnästä, kuten puheluista ja tekstiviesteistä (Mitnick & Vamosi, 2017, s. 39–42.)

Shavers & Bair (2016) korostavatkin metadatan merkitystä poliisin digitaalisessa forensiikassa. He mainitsevat sen yhdeksi parhaista tavoista kerätä tietoa ja todistusaineistoa käyttäjistä, varsinkin silloin kun käyttäjä on pyrkinyt salaamaan identiteettiään ja tietoliikennettään (Shavers & Bair, 2016.) Usein myös verkkopalvelut keräävät käyttäjästä metadataa luodakseen käyttäjälle digitaalisen identiteetin mainonnan kohdistamista varten. Metadatan poistamiseksi on useita menetelmiä. Usein laitteissa, jotka tuottavat digitaalisia tiedostoja, on mahdollista säätää asetuksia niin, että käyttäjä voi itse määrittää sen millaista metadataa tiedostoihin tallennetaan. Toinen menetelmä on tarkastella tiedostojen

metadataa erilaisilla selainpohjaisilla työkaluilla. Samoilla työkaluilla onnistuu usein myös metadatan poistaminen tiedostoista (Bazzell, 2016a, s. 274.)

2.2.3 Viestintä

Mitnickin (2017) mukaan, vaikka sähköisestä viestinnästä onkin lakeja, jotka suojaavat viestien yksityisyyttä, monet ilmaiset sähköpostin tarjoajat lukevat ja skannaavat sähköpostiviestien sisältöjä kohdistaakseen mainontaansa. Käyttäjä usein hyväksyy tämän tiedostamattaan palvelun ehdoissa, lukematta ehtoja sen tarkemmin. Myös käytettäessä työnantajan tarjoamaa sähköpostia, saattaa työnantajalla olla pääsy käyttäjän sähköposteihin. Vaikka useimmat palveluntarjoajat salaavatkin viestit lähetyksen aikana, eivät viestit välttämättä ole salattuja lähettäjän ja vastaanottajan sähköposti kansioissa. Parhaita käytäntöjä sähköpostiviestien salaamiseksi ja yksityisyyden säilyttämiseksi ovat yksityisyyttä arvostavan palvelun valinta, tilin yhteiskäyttö ja asymmetrisen Pretty Good Privacy (PGP) salauksen käyttäminen. Palvelun valinnassa tulee kiinnittää huomiota palvelun ehtoihin. Tehokas menettely viestien kaappaamisen estämiseen on yhteisen tilin käyttäminen, lähettämättä yhtäkään sähköpostiviestiä. Tämä onnistuu luomalla viestit luonnokset kansioon ja käymällä vuorotellen lukemassa viestit luonnokset kansiosta. PGP on asymmetrinen salausmenetelmä, jonka avulla viestit on mahdollista itse salata. Salaamalla viestit itse, voidaan estää palveluntarjoajaa lukemasta tai skannaamasta sähköpostikansioissa olevia viestejä. Asymmetrisessä salauksessa molemmat viestien osapuolet luovat kaksi avainta. Julkisen ja yksityisen avaimen. Julkiset avaimet ovat julkisesti saatavilla ja henkilön julkisen avaimen avulla viestit voidaan salata ja lähettää vastaanottajalle. Vain vastaanottaja pystyy avaamaan nämä viestit salaisella yksityisellä avaimellaan. Jos joku näkee salatun viestin tai kaappaa sen, hän ei kykene sitä lukemaan ja purkamaan salausta ilman vastaanottajan yksityistä avainta. Viestinnässä tulisi käyttää palveluita, jotka käyttävät päästä-päähän salausta. Tämä tarkoittaa sitä, että viesti säilyy salattuna, kunnes se saavuttaa vastaanottajan (Mitnick & Vamosi, 2017, s. 32–48 & 118.)

Muut suositellut viestintä menetelmät ovat VoIP ja muut verkkoyhteyttä käyttävät ohjelmat. VoIP eli Voice Over Internet Protocol mahdollistaa puheluiden tekemisen suoraan tietokoneelta. VoIP-palveluista on usein myös mahdollista saada itselleen anonyymi puhelinnumero. Toinen suositeltava tapa on käyttää kommunikointiin ohjelmia, jotka arvostavat yksityisyyttä ja käyttävät viestintään verkkoyhteyttä, normaalien puheluiden sijaan. Nämä palvelut vaativat usein puhelinnumeron. Puhelinnumero näihin palveluihin on mahdollista saada anonyymisti virtuaalisena numerona tai prepaid-liittymän avulla (Bazzell, 2016a, s. 156-190.)

2.2.4 Laitteiden ja ohjelmistojen valinta sekä käyttö

Flown (2021) mukaan tietokone sisältää valtavan määrän käyttäjän henkilökohtaista ja arvokasta informaatiota, jonka perusteella käyttäjä on mahdollista myös

tunnistaa. Jos käyttäjä haluaa olla anonyymi, tulee käyttäjän hankkia itselleen käyttötarkoitukseen sopiva tietokone, jolla hän ei ole missään vaiheessa tehnyt mitään henkilökohtaisia aktiviteetteja, käyttänyt henkilökohtaisia tilejä tai julkaissut informaatiota, joka olisi yhdistettävissä käyttäjän oikeaan identiteettiin. Tietokoneen valinnassa olisi myös hyvä kiinnittää huomiota siihen, ettei sen käyttäminen vaadi minkään suuren palveluntarjoajan tilin käyttämistä. Mikäli käyttäjä käyttää konetta tunnistettavalla tilillä kirjautuneena esimerkiksi googlen palveluihin, kerää google jatkuvasti tietoa käyttäjän toiminnasta (Flow, 2021, s. 6.)

Kovalevyjen elinikä on rajallinen ja ne hajoavat tietyn ajan kuluttua, jolloin data katoaa. Lisäksi OSINT-tiedustelussa voi olla tarpeellista tallentaa dataa tietokoneen ulkopuolelle. Hyviä ratkaisuja ovat ulkoiset, fyysiset muistit tai pilvipalvelut, jotka tarjoavat datan säilöntään mahdollisuuksia. Tärkeintä säilönnässä on se, että data on kahdennettuna, jossakin muualla ongelmatilanteita varten ja se, että data on luotettavasti salattu tallennustavasta riippumatta. Mikäli USB-muistitikkujen käyttäminen epäilyttää, voi käyttäjä käyttää tiedon tallentamiseen Cd-levyjä tai SD muistikortteja. Näiden etuna on se, että ne eivät ole aktiivisia ja näin kykene suorittamaan koneessa mitään ilman käyttäjän lupaa (Bazzell, 2016b.)

Ostettaessa uusia laitteita, jää usein laitteen myyjälle laitteita koskevia tunnistetietoja haltuun, joiden avulla laitteiden ja ostajan yksilöllinen tunnistaminen on mahdollista. Mikäli haluaa hankkia itselleen jäljittämättömät laitteet ja mahdollisimman edullisesti, se onnistuu ostamalla käytetyt laitteet, käyttämällä anonyymia identiteettiä. Luodakseen verkossa tarvittavia anonyymeja tilejä, käyttäjä tarvitsee myös ylimääräisen puhelimen, jossa on mahdollista käyttää prepaid liittymää (Bazzell, 2016a, s. 156–162.) Laitteita valittaessa on myös hyvä tarkastaa mihin asti palveluntarjoaja on luvannut tukea laitteiden ohjelmistopäivityksiä, joilla paikataan havaittuja haavoittuvuuksia. Yleensä erityisesti tiettyihin vanhempiin puhelinmalleihin tuki lopetetaan tietyssä vaiheessa laitteiden elinkaarta.

Digitaalisten jalanjälkien peittämiseksi ja identiteetin suojaamiseksi tulee luottaa tarkasti valittuihin ohjelmiin. Tarvittavia ohjelmistoja valittaessa tulee kiinnittää huomiota ohjelmien palveluehtoihin, tietosuojakäytäntöön ja siihen, että palvelut olisivat avoimen lähdekoodin ohjelmia. Avoimen lähdekoodin ja voittoa tavoittelemattomien tahojen tuottamat ohjelmat ovat usein turvallisimpia, sillä niiden toimintaperiaatetta on kenen tahansa mahdollista tarkastella. Käytettäessä ohjelmaa, joka ei ole avoimen lähdekoodin ohjelma, täytyy sokeasti luottaa palveluntarjoajaan. Se että palveluntarjoaja on nimekäs brändi, ei automaattisesti tarkoita turvallista ja luotettavaa ohjelmaa (Mitnick & Vamosi, 2017, s. 64.) Ilmaisissa ohjelmissa mainosten näkeminen on yleensä hyvä asia, sillä silloin ohjelman tarjoaja suuremmalla todennäköisyydellä tekee voittoa mainonnalla, eikä käyttäjän tiedoilla (Bazzell, 2016b, s. 135).

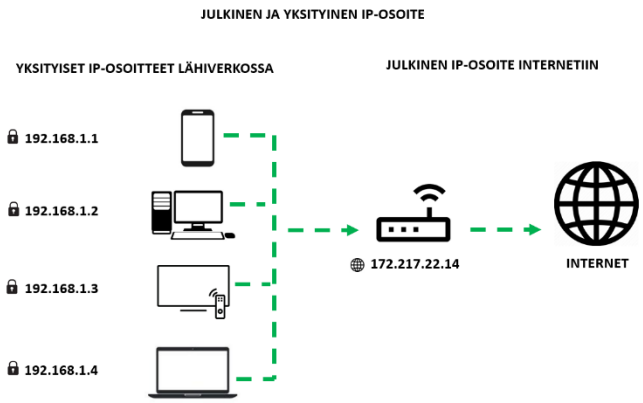
2.3 IP-osoite

IP-osoite on uniikki osoite, jolla laite yhdistetään internetiin. IP-osoitteen avulla laitteet voidaan tunnistaa ja ne pystyvät kommunikoimaan toistensa kanssa internetissä. IP-osoite on ainutlaatuinen jokaiselle laitteelle tietyssä verkossa. Maailmassa on kaksi erilaista IP-osoite protokollaa. IPv4- ja IPv6 osoitteet. IPv6-osoitteet otettiin käyttöön, sillä IPv4-osoitteet alkoivat loppumaan. Aikanaan IPv4-osoitteita luotaessa, ei osattu ennakoida internetin valtavaa kasvua. Esimerkit IPv4- ja IPv6-osoitteista:

IPv4-osoite: 192.168.1.1

IPv6-osoite: 2a00:1190:0001:0002:0003:0000:0000:ffff

Yhdistyttäessä internetiin käytetään, joko staattista tai dynaamista IP-osoitetta. Staattinen tarkoittaa, että IP-osoite pysyy samana eikä muutu ajan kuluessa. Dynaaminen osoite tarkoittaa sitä, että IP-osoite muuttuu jokaisella kerralla, kun käyttäjä muodostaa yhteyden internetiin. IP-osoitteita on kahta tyyppiä. Julkinen ja yksityinen. Julkinen osoite sallii suoran yhteyden internetiin, kun taas yksityinen osoite jaetaan esimerkiksi kotiverkon jokaiselle laitteelle ja niiden paljastuminen ulkoverkkoon pyritään estämään. Esimerkiksi kotiverkolla voi olla yksi julkinen IP-osoite, joka on määritetty reitittimelle ja jokaiselle laitteelle kotiverkossa on määritetty yksityinen IP-osoite. Näin reitittimen julkinen IP-osoite näkyy internetiin, ja kotiverkon yksityiset IP-osoitteet näkyvät vain kotiverkon sisällä (Hassan & Hijazi, 2018, s. 52–53.)



KUVIO 1 Julkinen ja yksityinen IP-osoite

2.3.1 ”War driving”

Flown (2021) mukaan pyrittäessä olemaan anonyymi, tulisi aina olettaa, että oma käytössä oleva fyysinen IP-osoite on jonkun ulkopuolisen tahon tiedossa. Flown mukaan war driving on verkkorikollisten tapa peittää oma fyysinen IP-

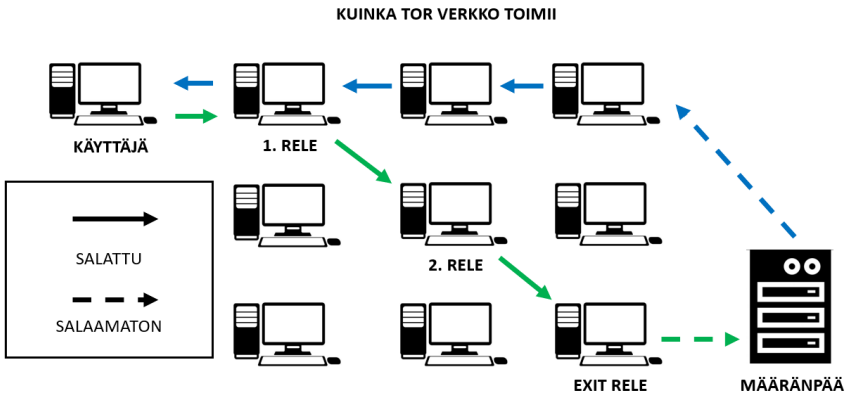
osoitteensa. Ilmiössä on kyse siitä, että rikollinen ajelee ajoneuvolla ympäri kaupunkia etsiäkseen julkisia, avoimia Wi-Fi verkkoja. Käytettäessä jonkun muun omistamaa tai tarjoamaa verkkoa, johtaa IP-osoite paljastuessaan kyseiseen sijaintiin (Flow, 2021, s. 5–6.) Lähes kaikki verkkosivut tallentavat vierailijoiden IP-osoitteet, muun metadatan ohessa. Internet yhteyden palveluntarjoaja saattaa myös tallentaa IP-osoitteeseen liittyvää dataa ja esimerkiksi sähköpostiosoitteen luonnin yhteydessä tallennetaan käyttäjän IP-osoite (Hassan & Hijazi, 2018, s. 54). IP-osoitteen tietäminen helpottaa hyökkääjää kohdistamaan hyökkäyksiä käyttäjään. Kun IP-osoite tiedetään, kuka tahansa pystyy myös selvittämään verkossa kyseisen yhteyden palveluntarjoajan, mahdollisen sijainnin ja lukemattomasti muuta hyödyllistä informaatiota.

Mitnickin (2017) mukaan parhaat menetelmät fyysisen IP-osoitteen piilottamiseksi ovat jonkun muun tarjoaman julkisen Wi-Fi verkon tai anonyymin prepaid hotspot yhteyden käyttäminen. Wi-Fi verkkoa käytettäessä tulisi kuitenkin huomioida, että käytetään vain HTTPS-protokollaa käyttäviä verkkosivuja ja käytössä on VPN, jolloin liikenne ei ole luettavissa selkokielisenä, jos joku muu samassa verkossa onnistuu kaappaamaan tietoliikennettä. Molemmissa menetelmissä korostetaan myös epäjohdonmukaisuutta sijainnin valinnassa. Yhteyttä ei saisi käyttää liian usein samasta sijainnista, koska vaikka käyttäjä ei olisi pääteltävissä IP-osoitteesta, voi sijainti paljastua ja käyttäjä löytyä kyseisestä sijainnista. Hotspotin vahvuutena on se, että käyttäjä voi hallinnoida ja sulkea muut käyttäjät verkon ulkopuolelle. Tässäkin tapauksessa hotspot tulee olla anonyymisti hankittu (Mitnick & Vamosi, 2017, s. 114.)

2.3.2 The Onion Router (Tor)

Tor on Firefox-selaimesta kehitetty selain käyttäjän ja vastaanottajan IP-osoitteen peittämiseksi käytettäessä selainta tai lähetettäessä sähköpostia. Sen vahvuutena on helppo käyttöisyys ja se, että se on ilmainen sekä kenen tahansa käytettävissä. Tor reitittää käyttäjän internet liikenteen vähintään kolmen satunnaisesti valitun solmun läpi verkossa ja dataan lisätään yksi kerros salausta jokaisessa solmussa. Viimeisen solmun ja vastaanottajan välillä liikennettä ei ole kuitenkaan salattu ja tässä välissä onkin epäilty olevan mahdollisuus kaapata tietoliikennettä sekä lukea sitä selkokielisenä. Jokainen solmu tietää vain edellisen solmun, josta liikenne on vastaanotettu. Tor:in heikkoutena voidaan nähdä se, että solmuja ylläpitävät vapaaehtoiset, jotka ovat antaneet laitteensa Tor-verkon solmuiksi. Näin ei voida luotettavasti tietää kenen omistamien solmujen läpi liikenne kulkee ja onkin epäilty, että joissakin tilanteissa Tor-verkon salaus on saatettu jo onnistua murtamaan verkon haavoittuvuuksien takia. Koska Tor-liikenne kulkee myös useamman solmun läpi, on Tor muihin selaimiin verrattuna hidas käyttöinen. Muiksi negatiivisiksi puoliksi voidaan lukea se, että Tor:in käyttö voi näkyä ulospäin, jolloin esimerkiksi internetin palveluntarjoaja voi yrittää estää sen käyttämisen, Tor on tietyissä valtioissa laiton ja se salaa vain web-selauksen, eikä muuta tietoliikennettä. Tor-selaimeen ei myöskään saa asentaa mitään

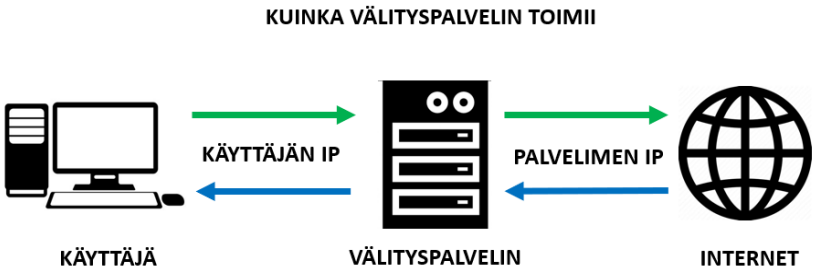
selainlaajennoksia, sillä ne saattavat aiheuttaa oikean IP-osoitteen paljastumisen (Shavers & Bair, 2016, s. 11–19; Tor, 2023.)



KUVIO 2 Kuinka Tor-verkko toimii

2.3.3 Välityspalvelimet (Proxy)

Välityspalvelimet ovat välittäjiä käyttäjän tietokoneen ja internetin välillä. Yritykset käyttävät välityspalvelimia turvallisuus kerrosten lisäämiseksi, niin että yrityksen sisäinen verkko saadaan eroteltua internetistä sekä sisällön suodattamiseen. Välityspalvelimia on erilaisia. Yleisin on web välityspalvelin, jota käytetään verkon resurssien hakemiseksi. Niitä voidaan käyttää myös käyttäjän IP-osoitteen peittämiseksi, jolloin käyttäjän IP-osoite korvataan välityspalvelimen IP-osoitteella. Välityspalvelimien käyttäminen ei ole kuitenkaan suositeltavaa, sillä niiden alkuperä ja ylläpitäjä voivat olla epäluotettavia, eikä tietoliikenne välttämättä ole salattu, kuten esimerkiksi Tor-selainta käytettäessä. Ilmaiset välityspalvelimet kykenevät tarkastelemaan käyttäjän tietoliikennettä ja lokittamaan sitä sekä saattavat myös ladata käyttäjän selaimeen haitallista sisältöä (Hassan & Hijazi, 2018, s. 66–67.)



KUVIO 3 Kuinka välityspalvelin toimii

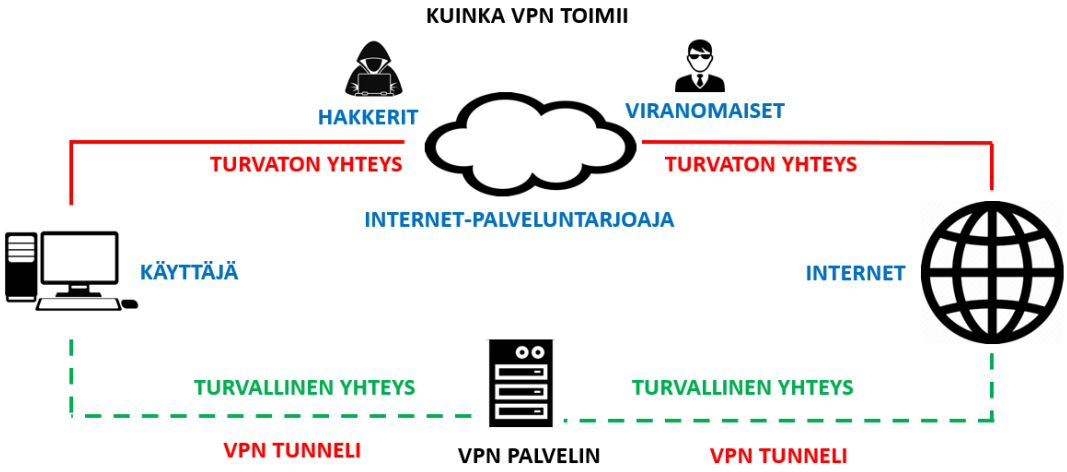
2.3.4 Virtual private network (VPN)

VPN reitittää kaiken verkkoliikenteen turvallisen ”tunnelin” läpi. Liikenne menee VPN-palvelimelle ja kaikki data salataan käyttäjän laitteen sekä palvelimen välillä. Näin salatun liikenteen tarkastelusta ei ole hyötyä ulkopuoliselle ennen kuin liikenne saapuu VPN-palvelimelle. Liikenteen saavuttua palvelimelle, VPN-palvelin välittää liikenteen eteenpäin määränpäähänsä, näin lisäten vastaanottajalle haastetta jäljittää ja seurata lähettäjää. Näin VPN esimerkiksi auttaa käyttäjää peittämään julkisen IP-osoitteensa (Bazzell, 2016a, s. 93–94.)

Mitnickin mukaan (2017) VPN:ää käytettäessä on hyvä huomioida, että VPN tarjoaa yksityisyyttä, mutta ei tee automaattisesti käyttäjästä anonyymia. Ainut tapa olla anonyymi VPN:ää käytettäessä, on olla anonyymi alusta lähtien ja ostaa palvelu alias yhteys- ja maksutiedoilla, kuten prepaid luottokortilla. Parhaat VPN-palvelut ovat maksullisia. Ilmaiset palvelut usein rajoittavat palveluiden nopeutta ja saattavat tehdä voittoa käyttäjien tiedoilla. Yleensä VPN-palveluntarjoajilla on palvelimia useissa eri maissa, jolloin käyttäjä kykenee valitsemaan minkä maan IP-osoitetta haluaa käyttää. Tämä on kätevää sillä monet maat rajoittavat tietoliikennettä ja palveluiden käyttöä IP-osoitteiden perusteella. VPN:än käyttö voikin mahdollistaa täysin uusiin palveluihin ja verkkoihin pääsyn (Mitnick & Vamosi, 2017, s. 114–116.) Flown (2021) mukaan useimmat VPN-palveluntarjoajat väittävät, etteivät ne kerää lokitietoja käyttäjästä. Hänen mukaansa tämä on kuitenkin käytännössä lähes mahdotonta ja ainakin jonkinlaisen metadatan kerääminen lokien muodossa on näiden palveluiden luotettavan toiminnan kannalta välttämätöntä. VPN-palveluita käytettäessä korostuukin palveluntarjoajan valinta ja palveluntarjoajan tietosuojakäytäntö (Flow, 2021, s. 4–5.)

Hassanin ja Hijazin (2018) mukaan VPN-palveluntarjoajaa valittaessa korostuu maa, josta VPN-palveluntarjoaja toimii ja sen maan lainsäädäntö tietojenkäsittelyyn liittyen. Toinen merkittävä tekijä on se, että palveluntarjoajan tulisi omistaa omat DNS-palvelimensa ja tukea suojautumista DNS-vuodoilta. Palveluntarjoajan tulisi tukea OpenVPN ohjelmaa, joka mahdollistaa kenen tahansa tarkastella ja auditoida palvelu mahdollisten sisäänrakennettujen ”takaovien” varalta, jotka voisivat mahdollistaa ulkopuolisen tahon päästä käsiksi salattuun liikenteeseen. Palvelun olisi hyvä hyväksyä anonyymit maksumenetelmät ja olla tilattavissa ilman tunnistautumista sekä sen tulisi tukea useita laitteita samalla tilauksella (Hassan & Hijazi, 2018, s. 65–66.)

Käytettäessä yksityisyyttä ja anonymiteettiä tarjoavia palveluita tulee huomioida, että yksittäistä palvelua käytettäessä yhteys voi vahingollisesti vuotaa ja paljastaa oikean IP-osoitteen. Näin voi tapahtua, kun osaa tietokoneen liikenteestä ei vahingollisesti ohjatakkaan turvalliseen kanavaan, vaan se ohjautuu internet-palveluntarjoajan palvelimille, mahdollistaen heidän tarkastella ja lokittaa tietoliikennettä, vaikka käytettäisiinkin esimerkiksi VPN:ää. Tätä kutsutaan DNS-vuodoksi. Usein useamman menetelmän samanaikainen käyttö tarjoaakin parhaan lopputuloksen. Esimerkiksi käyttämällä VPN:ää, voidaan piilottaa internet-palveluntarjoajalta Tor-selaimen käyttäminen (Hassan & Hijazi, 2018, s. 67.)



KUVIO 4 Kuinka VPN toimii

2.4 Media Access Control (MAC)

Kytkimissä protokollana toimii ARP. Sen tehtävänä on lähiverkoissa selvittää IP-osoitetta vastaava MAC-osoite ja näin jakaa paketteja oikeisiin laitteisiin lähiverkon sisällä. Ennen liikennöintiä lähetetään verkkoon ARP-kysely, johon on liitetty haluttu IP-osoite. Se kone, jolla on haluttu IP-osoite, lähettää vastausviestissä oman MAC-osoitteensa. MAC-osoite löytyy verkkoon kytkettävien laitteiden verkkokortista. Se on laitteen fyysisesti yksilöivä osoite. Osoite koostuu kuudesta, kaksi numeroisesta heksadesimaaliluvusta. Kolme ensimmäistä lukua ovat valmistajan kertova etuliite ja kolme viimeistä laitteen yksilöivä luku (Huuskonen, 2020.) MAC-osoite näyttää tältä:

3A:34:52:C4:69:B8

Ethernet lähiverkon laitteet kytketään usein yhteen kytkimen avulla. Kytkintä ja reititintä ei tule sekoittaa toisiinsa. Reititin on kiinnostunut IP-osoitteista. Kytkin taas tunnistaa laitteiden MAC-osoitteita, mutta ei IP-osoitteita. Kytkimessä sijaitsee MAC-osoitetaulu. MAC-osoitetaulua käytetään määrittämään mistä kytkimen portista liikenne välitetään eteenpäin (Pearson, 2022.)

MAC-osoitteen muuttaminen tulee tehdä välittömästi, koneen ensimmäisen käynnistyksen yhteydessä, ennen muita toimenpiteitä. Tämä johtuu siitä, että jokainen laite on yksilöitävissä MAC-osoitteen avulla ja MAC-osoite tallennetaan usein monien toimenpiteiden yhteydessä metadataan. Esimerkiksi MAC-osoite saattaa näkyä sähköpostin metadatassa tai reitittimen lokitiedoissa. MAC-osoitteen peittämiseen on useampiakin keinoja. Yksi keino on muuttaa osoite käyttöjärjestelmän konfiguroinnista käsin. Se miten osoite muutetaan, vaihtelee käyttöjärjestelmien mukaan. Toinen menetelmä on käyttää tietokoneessa ulkoista verkkoadapteria, jolla on oma MAC-osoite, jolloin ei käytetä lainkaan tietokoneen oman verkkokortin MAC-osoitetta. On olemassa myös ohjelmia, joiden avulla

osoitetta voidaan väärentää. Virtuaalikoneiden luonti, käyttö ja hallinta ovat yksi kätevä keino peittää aitoa MAC-osoitetta. Käytetystä MAC-osoitteesta tulisi varmistua aina ennen internetin käyttöä (Mitnick & Vamosi, 2017, s. 101–102 & 117–121.)

2.5 Selain

Mitnickin (2017) mukaan käytettäessä internetiä selaimella, verkkosivu kerää selaimelta tietoja käyttäjästä. Tällaisia tietoja voivat olla esimerkiksi muun muassa tietokoneella käytetty kieli, näppäimistön kieli, käyttöjärjestelmä versioineen, millaisia ohjelmia käyttäjä on asentanut ja jopa mitä ohjelmia käyttäjän kone on ajamassa. Näiden tietojen pohjalta verkkosivujen ja palvelujen on mahdollista rakentaa käyttäjälle digitaalista identiteettiä, sillä kerättyään riittävästi dataa, voidaan käyttäjä yksilöidä ainutlaatuisesta kokonaisuudesta, jonka verkkosivu on käyttäjästä kyennyt luomaan kerätyn datan avulla. Tätä kutsutaan selaimen sormenjäljeksi (engl. browser fingerprint) (Mitnick & Vamosi, 2017, s. 42–44.)

Usein verkkosivut pyrkivät seuraamaan käyttäjää yhdistämällä selaimessa tehdyt toimenpiteet käyttäjän IP-osoitteeseen. Vaikka käyttäjä vaihtaisi säännöllisesti IP-osoitetta, kykenee verkkosivu seuraamaan käyttäjän toimia, piilottamalla istunto-id:n verkkosivun linkkeihin. Tämä tarkoittaa sitä, että aina käyttäjän klikatessa linkkiä, löytyy linkistä istunto-id, jonka avulla verkkosivu kykenee tunnistamaan käyttäjän. Estääkseen istunto-id:n avulla tehtävän seuraamisen, tulee käyttäjän avata verkkosivu säännöllisesti uudestaan (Vlajic, Madani & Nquyen, 2017, s. 111.)

Evästeet ovat pieniä tekstitiedostoja, jotka yleensä tallennetaan käyttäjän tietokoneen selaimeen. Evästeet pitävät sisällään tietoa, jonka avulla verkkosivut voivat tunnistaa käyttäjän muista käyttäjistä. Tällaista tietoa voivat olla esimerkiksi verkkosivun nimi ja käyttäjä-ID. Evästeitä on karkeasti kahdenlaisia. Istuntoevästeet (engl. session cookies) ja pysyvät evästeet (engl. persistent cookies). Istuntoevästeet tallennetaan tilapäisesti käyttäjän selaimeen ja poistetaan sen jälkeen, kun käyttäjä sulkee selaimen tai kirjautuu ulos palvelusta. Näitä evästeitä verkkosivut käyttävät esimerkiksi käyttäjän ostoskorin sisällön muistamiseen tai tiedon säilyttämiseksi saman verkkosivun eri sivujen välillä. Pysyviä evästeitä on kahdenlaisia. Flash-evästeet (engl. flash cookies) ja ikuisuus evästeet (engl. ever cookies). Pysyvät evästeet ovat pidempi ikäisiä, kuin muut evästeet. Niiden avulla seurataan käyttäjän toimia eri verkkosivujen välillä. Flash-evästeet tallennetaan käyttäjän kovalevylle, eivätkä ne katoa poistettaessa evästeitä selaimesta. Niitä hyödynnetään adobe flash-lisäosan avulla ja ne mahdollistavat esimerkiksi median toiston tallentamisen tiettyyn kohtaan, seuraavaa kertaa varten. Ikuisuus evästeet ovat JavaScript pohjaisia ja voivat selvitä, vaikka aiemmin mainitut evästeet poistettaisiin laitteelta. Evästeet voidaan myös jakaa ensimmäisen ja kolmannen osapuolen evästeisiin. Ensimmäisen osapuolen evästeillä pyritään seuraamaan käyttäjää tietyn verkkosivun sisällä, kun taas kolmansien osapuolten evästeillä käyttäjää seurataan eri sivustojen välillä. Kolmannen osapuolen

evästeet usein sisältävätkin tietoja käyttäjästä. Nykyään löytyy kuitenkin selain- ja haittaohjelmien torjuntaohjelmia, jotka voivat havaita ja poistaa nämäkin evästeet (Hassan & Hijazi, 2018.)

Vaikka evästeet olisikin kytketty pois päältä, on verkkosivujen siitä huolimatta mahdollista yksilöidä käyttäjä verkossa. Mikäli käyttäjä käyttää vielä jossakin palvelussa oikeaa henkilöllisyyttään on selaimen keräämä tieto yhdistettävissä käyttäjän oikeaan identiteettiin. Tekniikat käyttäjän seuraamiseksi ja yksilöimiseksi ovat Script-pohjaiset tekniikat ja canvas tekniikka. JavaScript tekniikassa käyttäjän selaimeen ladataan Script, joka kerää teknistä informaatiota käyttäjän selaimesta ja laitteesta. Tätä tietoa käytetään myöhemmin käyttäjän identifioimiseksi ja seuraamiseksi samaan tapaan kuin IP-osoitetta. Canvas on HTML-elementti, jota käytetään erilaisten kuvioiden piirtämiseen. Se piirtää näkymättömän kuvan käyttäjän selaimeen, joka kerää teknistätietoa käyttäjän selaimesta ja käyttöjärjestelmästä. Tämän avulla käyttäjän toimia voidaan seurata eri verkkosivujen välillä. Canvas tekniikkaa käyttäjien seuraamiseksi käyttävät erityisesti mainostajat (Hassan & Hijazi, 2018, s. 55–56.)

Muita selaimen käyttämiä tekniikoita käyttäjän seuraamiseksi ovat Katri Ahlgrenin (2019) mukaan web bug -jäljitteet, superevästeet, HSTS ja ETag. Jäljitteet ovat evästeitä huomaamattomampi menetelmä käyttäjien seuraamiseksi. Ne ovat pikselin kokoisia GIF tai PNG-kuvia, scriptejä tai elementtejä, jotka näkyvät vain sivun lähdekoodissa. Superevästeet eivät tallennu selaimeen, kuten HTTP-evästeet, vaan javascriptin avulla evästeiden dataa tallennetaan eri puolille selainta, jolloin se voidaan poistamisen jälkeen luoda uudelleen muualta muistista. HSTS:n avulla verkkosivun ylläpitäjä voi asettaa verkkopalvelimen sallimaan vain HTTPS-yhteydet, HTTP-yhteyksien sijaan. Myös tätä tekniikkaa voidaan käyttää käyttäjän seuraamiseksi. ETag:in avulla verkkopalvelimet ja käyttäjän tietokone voivat tarkastella, onko käyttäjän pyytämä sivu muuttunut käyttäjän viime vierailun jälkeen. Mikäli ETag osoittaa, että sivu ei ole muuttunut, ladataan sivu tai elementti käyttäjän välimuistista. ETag:it mahdollistavatkin seurata käyttäjän vierailuja sivuilla, vaikka käyttäjän IP-osoite vaihtuisi vierailujen välillä (Ahlgren, 2019, s. 17–29.)

2.5.1 Selain ja selaimen koventaminen

Hassanin ja Hijazin (2018) mukaan Firefox on ainut aidosti avoimen lähdekoodin selain, kun puhutaan yleisistä, ”normaaleista” selaimista (Hassan & Hijazi, 2018, s. 59). Bazzellin (2018) mukaan Firefoxin etuina on sen muokattavuus, yksityisyys verrattuna muihin selaimiin ja kyky asentaa siihen kolmansien osapuolten ohjelmistolaajennoksia, jotka tekevät turvallisesta käytöstä entistä helpompaa. Bazzell mainitsee, että muun muassa Tor-selain on rakennettu Firefoxin pohjalta. Firefox mahdollistaa normaaleista selainasetuksista ominaisuuden valinnan, jonka avuilla selain hyväksyy vain HTTPS-protokollaa käyttävät verkkosivut. Tämä tarkoittaa sitä, että mikäli tietoliikenne kaapataan, ei liikenne ole kaappaajan luettavissa selkokielisenä. Mikäli kyseinen asetus ei ole valittuna, voi selain ohjata käyttäjän HTTP-protokollaa käyttäville sivuille, jolloin liikenne on

selkokielistä ja hyökkääjän luettavissa. Toinen hyödyllinen asetus on, asettaa ”Älä seuraa -signaalin” lähetys päälle selaimen asetuksista. Tärkeää on huomioida, että kaikki verkkosivut eivät kuitenkaan noudata tätä pyyntöä. Firefoxin voi halutessaan myös konfiguroida todella yksityiseksi ja mieleisekseen yksityiskohtaisempien lisäasetusten avulla, kirjoittamalla verkko-osoite kenttään Firefoxissa ”about:config”. Asioita, joihin asetuksissa tulee kiinnittää huomiota ovat tavat, joilla selain voi kerätä tietoa käyttäjästä. Näitä tietoja voivat olla esimerkiksi sivuhistorian, sijainnin ja laitteen virrantason seuraaminen selaimen toimesta (Bazzell, 2018, s. 5–7.) Brave väittää tietosuojakäytännössään, heidän selaimensa olevan valmiiksi yksityinen selain ja se saattaakin soveltua heikomman ATK-taustan omaaville, jolloin kaikkia asetuksia ei tarvitse itse konfiguroida (Brave, 2023).

2.5.2 Selainlaajennokset ja siivousohjelmistot

Selainlaajennokset ovat pieniä ohjelmia, jotka toimivat selaimessa ja tarjoavat käyttäjälle jonkin spesifin toiminnallisuuden. OSINT-tiedustelussa spesifien ohjelmien käyttö voi korostua ja eri selaimiin onkin ladattavissa erilaisia selainlaajennoksia, eivätkä kaikki laajennokset ole tarjolla kaikissa selaimissa. Valittaessa selainlaajennoksia tulisi käyttäjän perehtyä laajennosten luotettavuuteen, koska ne voivat myös kerätä käyttäjän yksityisiä ja henkilökohtaisia tietoja, yksityisyyden tarjoamisen sijaan. Firefoxin selainlaajennoksista löytyy laajennoksia, jotka mahdollistavat aiemmin mainittujen, käyttäjää seuraavien menetelmien, kuten evästeiden ja JavaScriptien estämisen sekä hallinnan. Muita hyviä toiminnallisuuksia yksityisyyden kannalta ovat laajennokset, jotka peittävät verkkosivuilta User-Agent tiedon. User-Agent on metadataa, joka kertoo verkkosivulle millaista selainta tai mahdollisesti käyttöjärjestelmää käyttäjä käyttää. Suositeltavaa on myös käyttää laajennosta, joka pakottaa selaimen käyttämään HTTPS-protokollaa, jos tätä toiminnallisuutta ei löydy vaihtoehtona selaimen asetuksista. Mikäli jälkien piilottamisen sijaan haluaakin piilotella näkyvillä, löytyy myös laajennoksia, jotka mahdollistavat ylimääräisen liikenteen luomisen mainostajien ja seuraajien hämäämiseksi. Lisäksi Jotkin VPN:ät ja haittaohjelmien torjuntaohjelmat pyrkivät torjumaan jo tänä päivänä selainten seuraamistekniikoita (Bazzell, 2018, s. 8.)

Tietokoneen käyttö aiheuttaa ja jättää ”roskaa” tietokoneen kovalevylle päivittäisessä käytössä ja tämä voi viedä tietokoneelta tarpeettomasti esimerkiksi muistia tärkeämmältä käytöltä. Tällaisen ”roskan” ja koneelle tallennettujen tiedostojen hallintaan on olemassa erikseen tarkoitettuja siivousohjelmia, joiden avulla kone voidaan pitää ”puhtaana”. Nämä ohjelmat mahdollistavat myös evästeiden poistamisen laitteelta. Osa siivousohjelmista mahdollistaa valinnan siitä, mitä ohjelmia käynnistetään tietokoneen käynnistyksen yhteydessä ja näin voidaan esimerkiksi estää Flash- ja Java-ohjelmien käynnistyminen kokonaan. Usein kaikkien näiden aiemmin mainittujen toimintojen estäminen voi aiheuttaa sen, että sivut eivät ole normaalisti käytettävissä. Nämä menetelmät kuitenkin

tarjoavat monipuolisen kyvyn hallita ja valita käytettävät evästeet sekä menetelmät, jotka selaimelle sallitaan (Bazzell, 2018, s. 3–4.)

2.5.3 Hakukoneet

Suuret hakukone yritykset seuraavat käyttäjää monitoroimalla käyttäjän verkko aktiviteetteja kohdistaakseen mainontaa (Hassan & Hijazi, 2018, s. 138). Mitnickin (2017) mukaan, vaikka käyttäjä poistaisi selaushistoriansa laitteelta, tallentavat suuret hakukoneyritykset käyttäjän historian omille pilvipalvelimilleen ja Yhdysvalloista on esimerkkejä, kuinka viranomaiset ovat ilmestyneet yksityishenkilön ovelle hänen haettuaan hakukoneilla epämääräisiä asioita. Vaikka käyttäjä ei olisikaan kirjautunut hakukone-palveluntarjoajan tilille, tehdään seuraaminen yhdistämällä hakuhistoria käyttäjän IP-osoitteeseen. Tätä voidaan estää käyttämällä yksityisyyttä arvostavia hakukoneita, joita on lukuisia. DuckDuckGo on vakio hakukoneena Firefoxissa ja lupaa, että käyttäjän IP-osoitteita ei lokiteta. DuckDuckGo ylläpitää myös omaa Tor-relettä, joka mahdollistaa hakukoneen käytön myös Tor-verkossa, ilman suurempaa hidastumista. Koska DuckDuckGo ei seuraa käyttäjää, ei se myöskään suodata käyttäjän hakutuloksia aiempien hakujen perusteella (Mitnick & Vamosi, 2017, s. 78–80.) Startpage on yksityisyyttä arvostava hakukone, joka tarjoaa vain Googlen hakukoneiden tuloksia, kun taas DuckDuckGo:n tulokset koostuvat useammista hakukoneista (Bazzell, 2018, s. 89). Muita yksityisyyttä arvostavia hakukoneita ovat: Qwant, Oscobo, Swisscows, Privatelee, Gigablast ja Gibiru (Hassan & Hijazi, 2018, s. 140).

2.6 Haittaohjelmat ja haavoittuvuudet

Perlroth (2021) kuvaa kirjassaan, kuinka nollapäivähaavoittuvuuksien etsiminen ja myyminen on valtioiden tiedustelupalveluiden ja verkkorikollisorganisaatioiden välistä kilpailua. Nollapäivähaavoittuvuus tarkoittaa järjestelmän tai ohjelman vaarantavaa haavoittuvuutta, jonka joku ulkopuolinen taho on löytänyt, mutta joka ei ole vielä palveluntarjoajan tiedossa ja näin siihen ei ole olemassa päivitystä ongelman korjaamiseksi. Perlrothin mukaan näitä nollapäivähaavoittuvuuksia hyödyntämällä valtiot ja rikollisjärjestöt pyrkivät hankkimaan luvattomasti tietoa digitaalisessa ympäristössä ja pahimmillaan jopa ottamaan uhrien laitteita hallintaansa. Usein käyttäjälle tarjottavien ohjelmistojen päivitystiedostot ovatkin päivitystiedostoja haavoittuvuuksien korjaamiseksi, ohjelmistojen kehittämisen lisäksi (Perlroth, 2021.) Muita uhkia ovat erilaiset haittaohjelmat, jotka saattavat varastaa käyttäjän dataa, muuttaa tai estää niihin pääsyn toteuttamalla haitallisen toiminnallisuuden uhrin laitteessa. Erilaisia haittaohjelmia ja haittaohjelma tyyppejä on useita. Erilaisia haittaohjelma tyyppejä ovat esimerkiksi kiristyshaittaohjelmat, vakoiluhaittaohjelmat, mainosohjelmat, virukset, madot, pelotteluohjelmat ja rootkitit (Hassan & Hijazi, 2018, s. 22).

2.6.1 Järjestelmän suojaus

Käyttäjien laitteilla on usein lukemattomia ohjelmia, joiden ajan tasalla pitäminen ja hallinta voi olla käyttäjälle haastavaa. Osa ohjelmista vastaanottaa automaattisia päivityksiä, mutta osassa ohjelmista tällaista ominaisuutta ei välttämättä ole. Usein tämä altistaakin käyttäjän hyökkäyksille, sillä usein aiemmin havaittuihin haavoittuvuuksiin on ehditty myös ulkopuolisten tahojen toimesta kehittämään hyödyntämismenetelmiä, joilla käyttäjän haavoittuville järjestelmille kyetään aiheuttamaan haittaa ja pahimmillaan ulkopuolinen voi etänä ottaa käyttäjän tietokoneen hallintaansa. Suojatakseen itsensä ulkopuolisilta hyökkääjiltä, tulisi käyttäjän pitää huolta, että ohjelmistot ovat ajan tasalla. Tämä onnistuu parhaiten käyttämällä päivitysohjelmistoa, joka päivittää ohjelmia automaattisesti ja kertoo käyttäjälle ohjelmien päivitysten ja versioiden tilanteen (Bazzell, 2016b, s. 16–18.)

Virustentorjuntaohjelmat tarkkailevat jatkuvasti kaikkia käyttäjän aktiviteetteja tietokoneella, kuten tiedostojen avaamista, ohjelmien käynnistämistä ja tiedostojen lataamista internetistä verraten näitä tunnettuihin uhkiin. Käytettäessä virustentorjuntaohjelmaa tietokoneella, on tärkeää, että koneelle on asennettu vain yksi virustentorjuntaohjelma. Mikäli koneella on useampi ohjelma, alkavat ohjelmat ”taistella” järjestelmästä, eivätkä ne toimi ihanteellisesti (Bazzell, 2016b, s. 14.) Hassanin ja Hijazin (2018) mukaan hyvästä virustentorjuntaohjelmasta löytyy seuraavat piirteet: sisäänrakennettu palomuuri, ohjelma skannaa sähköpostit mahdollisten tietojenkalastelujen varalta, päivittyy automaattisesti, kykenee tunnistamaan kehittyneet haittaohjelmat eikä kuluta merkittävästi tietokoneen resursseja. Mikäli virustentorjuntaohjelma ei pidä sisällään palomuuria, tulisi käyttäjän varmistua, että tietokoneella on toimiva palomuuri. Palomuurin tehtävä on tarkkailla ja kontrolloida sisään ja ulospäin kulkevaa verkkoliikennettä, paljastaen mahdolliset verkkorikolliset ja haittaohjelmat (Hassan & Hijazi, 2018, s. 31.)

Mikään ohjelma ei kykene suojaamaan tietokonetta kaikilta uhilta. Mikäli ohjelma väittää niin, se on todennäköisesti itse jonkinlainen haittaohjelma. Virustentorjuntaohjelman keskittyessä pääsääntöisesti viruksiin, tunnistamalla ne tunnettujen uhkien perusteella, pystyy haittaohjelmien torjuntaohjelma tunnistamaan erilaisia uhkia. Haittaohjelmat ja verkkorikollisten käyttämät hyökkäysmenetelmät kehittyvät jatkuvasti. Haittaohjelmien torjuntaohjelma auttaa havaitsemaan ne uhat, joita tavanomainen virustentorjuntaohjelma ei havaitse. Saavuttaakseen maksimaalisen suojan, tulee käyttäjällä olla sekä virustentorjuntaohjelma, että haittaohjelmien torjuntaohjelma. Nämä ohjelmat auttavat nykyään myös jo estämään haitallisia evästeitä (Hassan & Hijazi, 2018, s. 33.)

2.6.2 Käyttöjärjestelmät ja käyttöjärjestelmien koventaminen

Riippumatta siitä, mitä ohjelmia käyttäjä lataa tietokoneelle, tulisi käyttäjän ensimmäiseksi varmistua siitä, että käyttöjärjestelmä on turvallinen. Suurimmat uhat, jotka uhkaavat käyttöjärjestelmää ovat haittaohjelmat ja verkkorikollisten

kyky saada laite hallintaansa etänä. Näitä uhkia voidaan pienentää koventamalla käyttöjärjestelmää. Tämä tarkoittaa sitä, että muutetaan käyttöjärjestelmän asetuksia niin, että hyökkääminen muuttuu haastavammaksi ja käyttöjärjestelmä turvallisemmaksi. Näitä menetelmiä voivat olla esimerkiksi automaattisten päivitysten päälle laittaminen, sellaisen käyttäjätunnuksen käyttäminen, jolla ei ole kaikkia järjestelmänvalvojan oikeuksia, vahvojen salasanojen käyttö, etätuen poistaminen käytöstä, piilotettujen tiedostojen tekeminen näkyviksi, salasanan asettaminen BIOS:siin/UEFI:in sekä tarpeettomien palveluiden ja porttien poistaminen käytöstä. Osaan käyttöjärjestelmistä on myös saatavissa niin sanottuja jäädytysohjelmia, jotka kykenevät tallentamaan tietokoneen ”tilan”. Tällaista ohjelmaa käyttämällä esimerkiksi haittaohjelman iskiessä, käyttäjä voi perua vahingot vain käynnistämällä tietokoneen uudestaan (Hassan & Hijazi, 2018, s. 33–38.)

Olemassa on myös käyttöjärjestelmiä, jotka on suunniteltu tarjoamaan merkittävää yksityisyyttä ja jopa anonymiteettiä. Tällaisia käyttöjärjestelmiä ovat esimerkiksi Tails, Qubes ja Whonix. Shaversin ja Bairin (2016) mukaan Tails on käyttöjärjestelmä, joka on tehty tarjoamaan maksimaalista anonymiteettiä. Tails:iä on mahdollista käyttää lataamatta sitä lainkaan isäntäkoneelle. Tämä onnistuu USB-tikun, DVD-levyn tai virtuaalikoneen avulla. Tails:in merkittävimpänä ominaisuutena voidaan pitää sitä, että se ei tallenna ja jätä lainkaan jälkiä isäntätietokoneeseen. Se ei koske lainkaan tietokoneen kovalevyyn ja pyyhkii käyttämänsä muistin sammutuksen yhteydessä. Edes USB-tikulta käytettäessä Tails ei tallenna mitään, ellei käyttäjä erikseen niin määritä. Tails myös käyttää Tor:ia salatakseen kaiken liikenteensä, kun taas normaalisti, pelkkä Tor:in käyttö salaisi vain selain liikenteen. Lisäksi Tails tarjoaa merkittävästi myös muita yksityisyyttä lisääviä palveluita. Se esimerkiksi automaattisesti väärentää käyttäjän MAC-osoitteen (Shavers & Bair, 2016, s. 29–35; Tails, 2023.)Muita yksityisyyteen tähtääviä käyttöjärjestelmiä ovat Qubes ja Whonix. Qubesin ominaispiirre on se, että se käyttää virtualisointiteknologiaa erottaakseen eri ohjelmat ja prosessit toisistaan. Jokainen ohjelma suoritetaan omassa ”laatikossaan”, joka on erillinen virtuaalikone. Tämä vaikeuttaa koko järjestelmän vaarantumista esimerkiksi verkkohyökkäyksen tapahtuessa. Qubes tukee myös Tor-yhteyksiä (Qubes, 2023.) Whonix on käyttöjärjestelmä, jota suoritetaan virtuaalikoneen sisällä ja joka reitittää kaiken liikenteen Tor:in kautta. Se käyttää kahta virtuaalikonetta. Toista käyttäjälle ja toista internet-yhteydelle, joka lisää kerroksen turvallisuuteen. Whonix sisältää myös palomuurin ja muita yksityisyyttä helpottavia toiminnallisuuksia (Whonix, 2023.)

2.6.3 Virtuaalikoneet

Virtuaalikoneet mahdollistavat vieraiden virtuaalisten käyttöjärjestelmien käyttämisen tietokoneella, jossa on jo isäntäkäyttöjärjestelmä. Ne käyttäytyvät ja ovat kuin oma, erillinen tietokoneensa. Ne kuitenkin käyttävät isäntäkoneen fyysisiä resursseja kuten prosessoria ja muistia. Virtuaalikoneet lisäävät käyttäjälle turvallisuutta, sillä ne ovat yleensä täysin erotettu muusta isäntäkoneen käyttöjärjestelmästä. Esimerkiksi mikäli virtuaalikoneeseen tarttuu haittaohjelma,

voidaan virtuaalikone poistaa ja korvata helposti uudella, ilman että suurempaa vahinkoa on tapahtunut. Virtuaalikoneista voidaan myös helposti luoda klooneja, jolloin koneen menettäminen ei käytännössä aiheuta lainkaan harmia tai vaivaa. Virtuaalikoneella voidaan myös lisätä yksityisyyttä, sillä ne usein mahdollistavat digitaalisten jalanjälkien väärentämisen. Esimerkiksi virtuaalikoneelle voidaan aina luoda erillinen MAC-osoite ja ne voidaan poistaa välittömästi käytön jälkeen. OSINT-tiedustelun kannalta virtuaalikoneet ovatkin käteviä, sillä ne mahdollistavat erilaisten ”ympäristöjen” luonnin ja hallinnan helppokäyttöisesti. Virtuaalikoneet voidaan esimerkiksi tallentaa aina johonkin tiettyyn hetkeen sulkemisen yhteydessä (Hassan & Hijazi, 2018, s. 86–87.)

Virtuaalikoneiden käyttö lisää turvallisuuteen ja yksityisyyteen kerroksia. Niiden avulla voidaan estää tietojen tallentuminen isäntäkoneen kovalevylle. Virtuaalikoneet voidaan myös tallentaa ulkoiselle laitteelle, jolloin niistä ei löydy mitään jälkiä isäntäkoneesta. Virtuaalikoneet myös yleensä vaativat käyttöjärjestelmien tapaan salasanaa, jolloin pääsynhallintaan tulee automaattisesti yksi kerros lisää. Virtuaalikoneiden käyttäminen myös tarjoaa mahdollisuuden käyttää erilaisia käyttöjärjestelmiä erilaisiin tarkoituksiin. Näin voidaan lisätä tietojen paljastumisen estämiseksi volyymia lohkomalla käyttöä, sillä jokainen virtuaalikone on oma käyttöjärjestelmä kokonaisuutensa (Shavers & Bair, 2016, s. 165–168.)

2.7 Tietojen säilyttäminen ja fyysiset salaus ratkaisut

Kaikki aiemmin mainitut menetelmät ovat turhia, mikäli käyttäjän tietokone varastetaan tai kaapataan käyttäjältä, jolloin pahimmillaan kaappaaja voi saada käyttäjän koneen haltuunsa niin, että tietokoneelle on kirjauduttu sisään ja kaappaajalla on välittömästi pääsy koneelle (Bazzell, 2018.) Toinen huomioitava asia on se, että käyttöjärjestelmien tavanomaiset salasanat, joita kysytään tietokoneen käynnistyksen yhteydessä ovat ohitettavissa tähän tarkoitukseen suunniteltujen ohjelmien avulla, jotka voidaan suorittaa USB-tikulta tietokoneen käynnistämisen yhteydessä (Shavers & Bair, 2016.)

2.7.1 Salaus, siivoaminen ja steganografia

Turvallisin ratkaisu suojata tietokone ja tiedostot ulkopuoliselta käytöltä on käyttää kolmannen osapuolen tarjoamaa avoimen lähdekoodin salausohjelmistoa, joka tarjoaa modernit salausmenetelmät. Saatavilla on ilmaisia ohjelmia, jotka mahdollistavat kovalevyn ja tiedostojen salauksen modernilla AES-salausalgoritmilla. Salaamalla koko kovalevy, estetään tietokoneen käyttäminen ulkopuolisilta, mikäli tietokone varastetaan niin, että se on suljettuna. Kyseiset ohjelmat myös mahdollistavat tiedostojen salaamisen auki olevassa tietokoneessa, jolloin tiedostojen joutuessa vääriin käsiin, niitä eivät ulkopuoliset pysty hyödyntämään. Joissakin ohjelmissa on myös mahdollisuus piilottaa useampia käyttöjärjestelmiä

eri salasanojen taakse. Tämä tarkoittaa sitä, että mikäli käyttäjää kiristetään luovuttamaan salasana kovalevyn avaamiseksi, voi käyttäjä luovuttaa salasanan, joka johtaa ”naamioituun” käyttöjärjestelmään, jota kiristäjä voi luulla kohde käyttöjärjestelmäksi, vaikka kohde käyttöjärjestelmä olisi edelleen piilossa eri salasanan takana (Shavers & Bair, 2016, s. 164–165.)

Käyttäjän poistaessa tietoa tietokoneeltaan ei tieto oikeasti välittömästi katoa minnekään, vaan se säilyy edelleen kovalevyllä. Tämä mahdollistaa ”poistettujen” tiedostojen palauttamisen kovalevyltä, vaikka käyttäjä luulisi niiden olevan poistettu. Tämä kannattaa huomioida myös laitteen hävittämisen yhteydessä. Käyttöjärjestelmät pitävät osoittimien avulla kirjaa siitä missä tiedostot sijaitsevat kovalevyllä. Käytännössä jokaiselle käyttäjän kansiolle ja tiedostolle tietokoneella on tällainen osoitin, joka kertoo käyttöjärjestelmälle mistä tiedosto alkaa ja mihin se loppuu kovalevyllä. Käyttäjän poistaessa tiedostoja, hän käytännössä poistaa käyttöjärjestelmän osoittimen ja käyttöjärjestelmä merkitsee kyseiset alueet kovalevyllä vapaaksi tulevia tallennuksia varten. Data kuitenkin säilyy kovalevyllä siihen asti, kunnes käyttöjärjestelmä kirjoittaa datan päälle uutta dataa. Tähän asti tiedostot ovat yleensä myös vielä palautettavissa kovalevyltä. Aiemmin mainitut siivousohjelmat mahdollistavat kuitenkin tällaisten vapaaksi merkittyjen kovalevyn alueiden ylikirjoituksen milloin tahansa, tietojen palauttamisen vaikeuttamiseksi (Bazzell, 2016b, 80–81.)

Mikäli tiedostojen ja datan salaaminen ei tule kysymykseen on mahdollisuus hyödyntää steganografiaa. Steganografia tarkoittaa datan piilottamista, salauksen sijaan. Yksinkertaisimmillaan steganografia voi olla sitä, että normaaliin word-tiedostoon kirjoitetaan tekstiä valkoisilla kirjaimilla, jolloin teksti ei näy tiedostossa. Tiedostojen, kuten word:in metadatan kommenttikenttään voidaan piilottaa tekstiä tai tiedostojen tyypit voidaan vaihtaa niin, että ne eivät suostu aukeamaan. Tietyt ohjelmat myös mahdollistavat tietojen piilottamisen tiedostoihin, kuten kuviin niin, että niiden havaitseminen visuaalisesti on lähes mahdotonta (Shavers & Brett, 2016, s. 115–124.)

2.7.2 Kannettavat ohjelmat ja fyysiset turvallisuusratkaisut

Kannettavat ohjelmat ovat ohjelmia, joita ei tarvitse ladata tietokoneelle, niiden suorittamiseksi. Kannettavat ohjelmat on mahdollista suorittaa ulkoisesta muistista, kuten USB-tikulta, jolloin ne jättävät minimaalisesti jälkiä tietokoneelle. Näitä käytettäessä tulee kuitenkin huolehtia ulkoisena muistina käytetyn laitteen suojauksesta. Tätä varten markkinoilta löytyykin muisteja, joihin on rakennettu ulkoinen pääsykoodi ja salaus. Tästä huolimatta muistien sisälle tallennetut tiedostot tulisi vielä salata itse erikseen mahdollisuuksien mukaan (Shavers & Bair, 2016, s. 163–164.)

Vaikka koneen kovalevy olisi salattu, mutta kone on avoinna ja käytössä, on riskinä koneen joutuminen kaapatuksi kesken käytön. Tämän takia tuleekin kiinnittää huomiota kykyyn sulkea kone kesken käytön. Mitnick (2017) kertoo, kuinka tarjolla on ohjelmia, joiden avulla tietokoneen voi kytkeä bluetooth-yhteydellä toiseen laitteeseen. Mikäli laitteiden välinen bluetooth-yhteyden väli

kasvaa liian suureksi, lukitsee kone itsensä. Tällainen ominaisuus löytyy muun muassa Windowsista oletuksena (Mitnick & Vamosi, 2017, s. 181.) Muut menetelmät koneen sammuttamiseksi tällaisessa tilanteessa ovat koneen käyttäminen ilman akkua, niin että kone on jatkuvasti virtajohdossa, pikanäppäimen määrittäminen koneen sulkemiseksi ja tällaiseen käyttöön tarkoitetut USB-laitteet, joiden poistaminen sulkee tietokoneen (Shavers & Bair, 2016, s. 170.) Muita fyysistä turvallisuutta lisääviä menetelmiä ovat tietokoneen kameran peittäminen ja tietokoneen sammuttaminen siksi aikaa, kun sitä ei käytetä. Sammuttamista Bazzell perustelee sillä, että sammutettuun koneeseen on haastavaa hyökätä ja tietokoneen kovalevy ei kulu (Bazzell, 2016b, s. 13.)

2.7.3 Salasanat ja salasanojen hallinta

Mitnickin (2017) mukaan yksi parhaista menetelmistä salasanojen hallintaan on kirjoittaa salasanat ylös. Tällä hän tarkoittaa sitä, että kirjoittaa pitkän ja vaikean salasanan alun ylös, niin että alun nähtyään muistaa itse koko salasanan. Hänen mukaansa, vaikka joku näkisi salasanan alun, ei se paljasta näkijälle tarvittavaa tietoa salasanasta (Mitnick & Vamosi, 2017, s. 21.) Bazzellin (2016b) mukaan salasanojen vahvuudessa ratkaisee kaksi asiaa. Ensimmäinen on se, että onko käyttäjän salasana arvattavissa ja löytyykö se jo joltakin tietovuotoon liittyvältä listalta, jolloin salasana on verkkorikollisten arvattavissa. Toinen ratkaiseva tekijä on salasanan pituus, joka määrittää sitä kauanko tietokoneelta laskennallisesti menee arvata käyttäjän salasana. Mitä pidempi salasana on, sitä kauemmin salasanan arvaaminen kestää. Bazzell suosittelee käyttämään kaikkiin eri palveluihin tai ainakin erityyppisiin palveluihin eri salasanoja, kuitenkin niin, että ne muistuttavat toisiaan ja ovatkin näin helposti muistettavissa. Esimerkiksi seuraavasti:

Rahaan liittyvissä palveluissa merkin ”$” käyttäminen: ProGradu$@2023$!!

Sosiaalisenmedian palveluissa merkin ”?” käyttäminen: ProGradu?@2023?!!

Verkkokauppa palveluissa merkin ”%” käyttäminen: ProGradu%@2023%!!

Mikäli haluaa käyttää salasanojen hallintaohjelmistoa, Bazzell suosittelee käyttämään avoimen lähdekoodin ohjelmistoa, joka ei tallenna salasanoja internetiin tai synkronoi niitä internetin kautta. Hänen mukaansa on hyvä sallia ohjelman luoda käyttäjälle uniikit salasanat. Usein luotettavimpia salasanan hallintaohjelmia ovat ohjelmat, joiden tekniikka perustuu siihen, että ne säilyttävät salasanat käyttäjän laitteella. Useimpien salasanan hallintaohjelmistojen salasana käyttäjän tulee kuitenkin muistaa ja salasana ei ole palautettavissa edes pyydettäessä maksimaalisen turvallisuuden takaamiseksi. Muita hyviä käytäntöjä salasanoihin liittyen on estää tietokonetta ja selainta tallentamasta salasanoja, käyttää kaksivaiheista tunnistautumista, ottaa tilin ja salasanan palauttamiseen liittyvät toiminnot pois käytöstä, jos mahdollista ja poistaa ”luotetut laitteet” asetus käytöstä (Bazzell, 2016b, s. 42-66.)

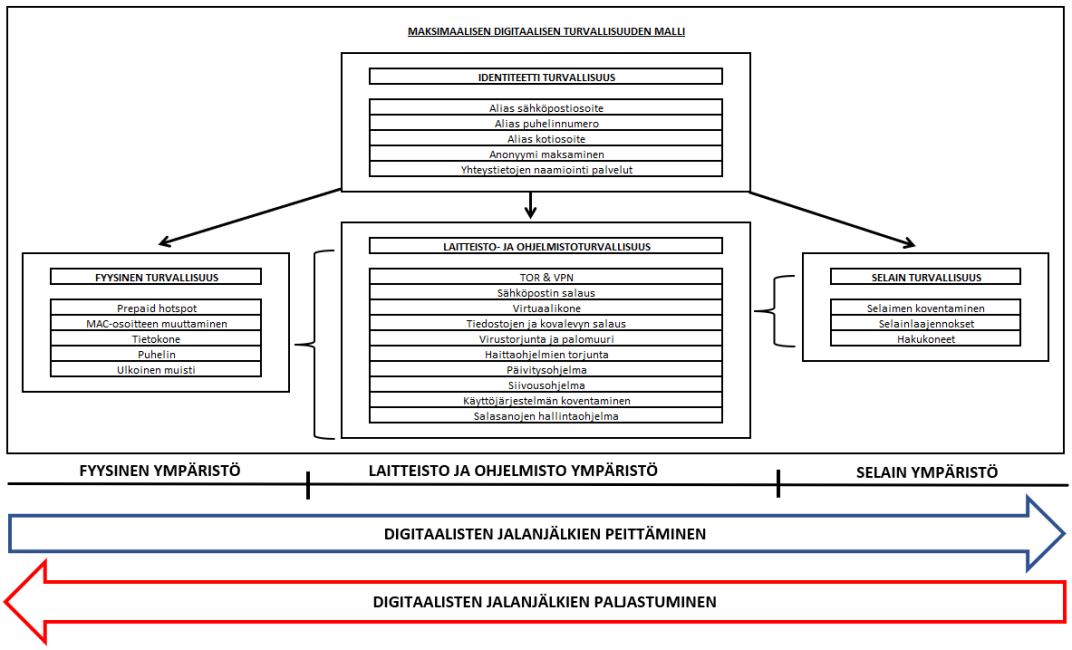
3 MAKSIMAALISEN DIGITAALISEN TURVALLISUUDEN MALLI

Tutkimuksen viitekehyksenä käytetään tutkijan itsensä kehittämää ”Maksimaalisen digitaalisen turvallisuuden” - mallia. Malli on kehitetty aiheeseen liittyvän kirjallisuuden pohjalta ja se on tehty siitä näkökulmasta, että se soveltuisi parhaiten käytettäväksi tehtäessä OSINT-tiedustelua verkossa. Mallin ulkopuolelle on jätetty ohjelmien suorittaminen ulkoisesta muistista, proxyjen eli välityspalvelinten käyttäminen ja ympäristön rakentaminen etäpalvelimelle. Näin on pyritty vähentämään mallin kompleksisuutta ja lisäämään käytettävyyttä. Malli ei näin olekaan muita vaihtoehtoja ja menetelmiä poissulkeva. Mallin on tarkoitus tarjota viitekehys, jota soveltamalla taustasta riippumatta OSINT-tiedustelua verkossa suorittavien tahojen on mahdollista rakentaa käyttämänsä ympäristö riittävän anonyymiksi ja turvalliseksi teknisestä näkökulmasta. Anonyymius on kovaa työtä ja vaatii jatkuvasti ylimääräisiä panoksia sekä ponnisteluja. Malli ei siis sovikaan kaikille sellaisenaan, vaan tarkoituksena on, että jokainen voi valita teoreettisesta mallista itselleen parhaiten sopivat kokonaisuudet osaksi tietokoneen sekä verkon käyttöä ja toteuttaa ne käytännössä itse parhaaksi katsomallaan tavalla.

Tässä kappaleessa esitellään tutkimuksessa käytetty malli. Mallin yhteydessä esitellään esimerkkiohjelmia tiettyjen toiminnallisuuksien saavuttamiseksi. Esimerkeiksi on pyritty valitsemaan avoimen lähdekoodin ohjelmia läpinäkyvyyden ja yksityisyyden takaamiseksi. Ohjelmiksi on valittu ilmaisia ohjelmia, jotta mallin soveltaminen olisi mahdollisimman helppoa ja mahdollista kenelle tahansa taustasta riippumatta. Mallissa ainoana maksullisena ohjelmana käytetään maksullista VPN:ää. Ladattaessa ohjelmia koneelle tulisi jokaisen perehtyä ohjelmien tietosuojakäytäntöihin. Ohjelmia ladatessa tulee varmistua siitä, että ohjelmat ladataan luotettavasta lähteestä. Mikäli ohjelmien alkuperä ja luotettavuus epäilyttävät, voi ohjelmat tarkastaa tiedostojen tarkastukseen tarkoitettujen palvelujen avulla. Yksi tällainen palvelu on Virustotal, joka mahdollistaa tiedostojen tarkastamisen verkkosivuillaan. Tarjoamalla palvelulle tarkasteltavan tiedoston, tarkastaa se tiedoston skannaamalla sen yli 70 virustentorjuntaohjelman läpi (Virustotal, 2023.) Ohjelmat tulisi myös ladata vain HTTPS-protokollaa hyväksi käyttäen, jotta tiedostojen peukalointi ei olisi ulkopuolisten tahojen osalta mahdollista latausvaiheen aikana. Tiedoston eheyden voi tarkastaa tarkistussumman (engl. checksum) avulla. Usein palveluntarjoajat tarjoavat verkkosivuillaan tarkistussumman ladattaviin ohjelmiin liittyen. Ladattuaan käyttäjän on mahdollista selvittää tiedostostaan tarkistussumma ja verrata tätä palveluntarjoajan verkkosivulla ilmoittamaan summaan. Mikäli summat vastaavat toisiaan, on ladattu tiedosto, sama tiedosto, jonka palveluntarjoaja lupaa verkkosivuillaan (Hickey & Arcuri, 2020, s. 29–31.)

Maksimaalisen digitaalisen turvallisuuden malli jaetaan neljään kokonaisuuteen, joita ovat: 1. Identiteetti turvallisuus, 2. Fyysinen turvallisuus, 3 Laitteisto- ja ohjelmistoturvallisuus ja 4. Selain turvallisuus. Identiteetti

turvallisuuden tarkoituksena on tarjota käyttäjälle anonyymi identiteetti OSINT-tiedusteluun verkossa ja se on läsnä sekä tulee huomioida kaikissa muissa mallin vaiheissa. Fyysinen turvallisuus ottaa kantaa niihin ratkaisuihin, joihin voidaan vaikuttaa fyysisesti tai jotka liittyvät fyysiseen turvallisuuteen. Fyysinen turvallisuus liittää käyttäjän fyysisestä ympäristöstä digitaaliseen ympäristöön. Laitteisto- ja ohjelmistoturvallisuus kattaa ne menetelmät, joissa ohjelmisto tai laitteisto ratkaisuilla voidaan lisätä digitaalista yksityisyyttä ja turvallisuutta. Selain turvallisuus kattaa ne menetelmät, jotka selaimessa voidaan tehdä yksityisyyden ja turvallisuuden lisäämiseksi selaimen näkökulmasta. Fyysinen turvallisuus, laitteisto- ja ohjelmistoturvallisuus sekä selain turvallisuus nähdään mallissa lineaarisena ketjuna. Ketjussa digitaalisten jalanjälkien peittäminen täytyy aloittaa huomioimalla fyysinen turvallisuus ja siihen liittyvät ratkaisut. Tämän jälkeen tulee kiinnittää huomiota laitteisto- ja ohjelmistoturvallisuuteen sekä lopuksi selain turvallisuuteen. Digitaalisten jalanjälkien paljastuminen kulkee mallissa ketjua vastapäivään ja alkaa selain turvallisuudesta. Identiteetti turvallisuus vaikuttaa ja tulee huomioida mallin kaikissa kokonaisuuksissa.



KUVIO 5 Maksimaalisen digitaalisen turvallisuuden malli

3.1 Identiteetti turvallisuus

Mallin ”sydämenä” toimii identiteetti turvallisuus. Jotta mallissa voidaan pysyä anonyymina, tulee oikea identiteetti erottaa uudesta digitaalisesta identiteetistä kokonaan ja pitää nämä toisistaan erillään mallin kaikissa vaiheissa. Jotta käyttäjä voi olla anonyymi verkossa, tulee hänen pystyä tarjoamaan palveluille sähköpostiosoite, puhelinnumero, kotiosoite ja maksuvälineet.

Mallissa suositellaan käyttämään kahta eri sähköpostiosoitetta. Gmail tai Hotmail osoite vaaditaan usein tiettyjen palveluiden käyttämiseksi. Jotta anonyymin Gmail-osoitteen luominen on mahdollista, tarvitaan myös toinen sähköpostiosoite, jolla Gmail-osoitteen luominen varmennetaan. Tutanota mahdollistaa salatun sähköpostiosoitteen luomisen ilmaiseksi, ilman että palveluun tarvitsee tunnistautua tai tarjota henkilötietoja itsestään (Tutanota, 2023). Tutanota sähköpostiosoite kannattaa hankkia itselleen, ennen kuin alkaa määrittämään uutta tietokonetta OSINT-käyttöön, sillä useat nykyiset koneet vaativat Windows-käyttöjärjestelmää käytettäessä kirjautumaan koneelle jonkin suuren palveluntarjoajan, kuten Microsoftin sähköpostiosoitteen avulla tai luomaan sellaisen ensimmäisen käytön yhteydessä. Tutanota sähköpostiosoitetta luodessa, tulisi sähköpostiosoite luoda mielellään avoimessa Wi-Fi verkossa ja MAC-osoite vaihdettuna. Näin varmistutaan siitä, etteivät oikea IP-osoite ja MAC-osoite tallennu metadataan sähköpostin luonnin yhteydessä. Gmail-osoitteen luonnin jälkeen, tulee varmistua siitä, että Gmail-osoitteeseen ei olla tarpeettomasti kirjautuneena, käyttäjän tietojen keräämisen vähentämiseksi.

Mallissa puhelinnumerona käytetään prepaid numeroa, joka on hankittu anonyymisti. Numeroa ei ole tarkoitettu yhteydenpitoon tai viestimiseen. Numeron ainut tarkoitus on mahdollistaa vahvistusviestien vastaanottaminen, tilien luonnin yhteydessä. Eettisin ratkaisu mallissa alias kotiosoitteeksi on vuokrata postilokero tai osoite itselleen verkon avulla. Näiden alias yhteystietojen avulla varmistutaan siitä, että tehtäessä OSINT-tiedustelua verkossa, ei tarjota omia henkilökohtaisia tietoja palveluntarjoajille ja tilitiedot eivät ole yhdistettävissä käyttäjään. Parhaat menetelmät anonyymiin maksamiseen ovat prepaid-luottokortit ja virtuaalivaluuttojen käyttäminen, niin että niiden alkuperä voidaan häivyttää. Mallissa maksamiseen voidaan kuitenkin käyttää omaa maksukorttia, mikäli tarjotaan muut yhteystiedot alias yhteystietoina. Näin maksaminen on mahdollisimman helppoa, käytännöllistä ja palveluntarjoajien ei ole mahdollista yhdistää korttia oikeaan henkilöllisyyteen. Alias henkilöllisyyttä käytettäessä verkossa tulee kuitenkin huomioida aina se, että ei käytä kenenkään oikean ihmisen henkilöllisyyttä ja kaikki maksaminen tapahtuu aina käyttäjän omaa pääomaa käyttämällä.

Mallissa käytetään myös palveluita, jotka tarjoavat yhteystietojen naamiointia. Esimerkkinä käytetyistä palveluista ovat 33mail ja Blur. 33mail on sähköpostin välityspalvelu. Palveluun rekisteröityessä palvelu pyytää sähköpostiosoitetta, johon kaikki palvelun kautta välitetyt viestit toimitetaan. Tarjoamalla palvelulle sähköpostiosoite prograd u@tutanota.com, välitetään kaikki välitysosoitteeseen tulleet sähköpostit kyseiseen osoitteeseen. Sitten palvelu luo käyttäjälle sähköpostiosoitteen ilman etuliitettä. Tällainen osoite voisi olla esimerkiksi @ProGradu.33mail.com. Käyttäjä voi lisätä kyseisen välitysosoitteen etuliitteeksi mitä tahansa ja viestit välitetään aina haluttuun sähköpostiosoitteeseen, ilman että käyttäjän tarvitsee tarjota tai kertoa kenellekään oikeaa osoitetta. Välitysosoitteet voi myös poistaa, jolloin haluttua postia voidaan hallita (33MAIL, 2023.) Toinen yhteystietojen naamiointiin sopiva palvelu on Blur. Blur:ista on tarjolla ilmainen, että maksullinen versio. Palvelu on erittäin kätevä ja monipuolinen,

sillä se kykenee tarjoamaan maksullisena versiona monipuolisesti palveluita, yksityisyyden lisäämiseksi. Blur mahdollistaa muun muassa puhelinnumeroiden, sähköpostien ja maksukorttien naamioimisen. Lisäksi sen avulla on mahdollista luoda ja hallita salasanoja sekä estää verkkosivuja seuraamasta käyttäjää (Ironvest, 2023.)

3.2 Fyysinen turvallisuus

Mallin fyysisen turvallisuuden kokonaisuus voidaan nähdä rajapintana käyttäjän fyysisen ja digitaalisen ympäristön välissä. Tässä kokonaisuudessa tulee ottaa huomioon ne fyysiset tekijät, jotka voivat vaarantaa käyttäjän digitaalisen turvallisuuden. Näitä tekijöitä ovat käytettävä yhteys, käytetyt laitteet ja verkkokortin määrittämä MAC-osoite. Mallissa yhteytenä käytetään prepaid hotspot yhteyttä, joka on hankittu anonyymisti. Hotspot on OSINT-tiedustelijalle käytännöllisempi ratkaisu, kuin jatkuva avointen Wi-Fi verkkojen etsiminen. Näin mahdollistetaan se, että käyttäjän IP-osoitteen paljastuessa, käyttäjän identiteetin paljastuminen olisi mahdollisimman epätodennäköistä. Hotspot yhteyttä käytettäessä tulee kuitenkin huomioida, että myös sen sijainnin on mahdollista paljastua ja käyttöä kannattaakin toteuttaa useammista eri paikoista, epäsäännöllisesti yksityisyyden maksimoimiseksi.

Tietokoneena mallissa käytetään käyttötarkoitukseen uutena hankittua kannettavaa tietokonetta, jolla ei ole tehty mitään omia henkilökohtaisia toimenpiteitä tai käytetty käyttäjään yhdistettävissä olevia tilejä. Mikäli tietokone on hankittu käytettynä, tulee varmistua siitä, että kone on puhdistettu ja nollattu asiaankuuluvalla tavalla. Kovalevyn tarkastaminen on mahdollista toteuttaa TestDisk ohjelman avulla. TestDisk on kuitenkin komentorivipohjainen työkalu, jonka käyttäminen voi olla osalle haastavaa. Vaihtoehtoinen ja helpompi käyttöinen ohjelma on Recuva. Se ei kuitenkaan ole avoimen lähdekoodin ohjelma. Esimerkissä tietokoneena käytetään kannettavaa, jossa on 16 gigatavun kokoinen keskusmuisti ja teratavun kokoinen kovalevy. Riittävä keskusmuisti mahdollistaa useampien ohjelmien samanaikaisen suorittamisen sujuvasti ja riittävä kovalevyn koko tarjoaa mahdollisuuden luoda useampia virtuaalikoneita. Tämä on tärkeää sillä virtuaalikoneet käyttävät ja vaativat toimiakseen tietokoneen laitteisto resursseja, kuten muistia. Koneen käyttöönotossa tulee myös huomioida, mitä asetuksia koneeseen määrittää. Esimerkiksi käyttöjärjestelmän kielen tai näppäimistön kielen valinta voivat näkyä myöhemmin selaimen keräämissä tiedoissa.

Mallissa MAC-osoite muutetaan käyttöjärjestelmän asetuksien avulla. Näin vähennetään käyttäjän tarvetta hankkia uutta verkkoadapteria. Lisäksi varmistutaan siitä, ettei koneen MAC-osoite tallennu mihinkään metadataan oikeana, tehtäessä tulevia toimenpiteitä. MAC-osoitteen vaihtaminen ja väärentäminen on mahdollista myös virtuaalikoneen avulla, mutta saadaksemme ympäristön toimimaan, tulee meidän tehdä toimenpiteitä myös isäntäkäyttöjärjestelmän

avulla. Windowsissa MAC-osoitteen vaihtaminen onnistuu rekisterieditorin avulla.

Prepaid liittymää varten mallissa tarvitaan puhelin. Mallin esimerkissä käytetään vanhaa älypuhelinta, jonka tarkoituksena on vain vastaanottaa vahvistusviestit tilienluonnin yhteydessä, prepaid numeroon. Ulkoisena muistina mallissa käytetään iStorage:n neljän gigatavun USB-tikkua, joka on salattu numeerisella PIN-koodilla ja tikun sisältö salataan AES-salausalgoritmin avulla (iStorage, 2023). Mikäli tikun luotettavuus epäilyttää, voi sisällön salaukseen lisätä suojauskerroksia myös itse tietokoneella, ennen tallentamista tikulle tai käyttää tallennuslaitteena SD-korttia.

3.3 Laitteisto- ja ohjelmistoturvallisuus

Mallin laitteisto- ja ohjelmistoturvallisuus kattaa ne kokonaisuudet, joissa käyttäjän tietokoneelle ladatuilla ohjelmilla lisätään yksityisyyttä ja turvallisuutta käyttäjälle. Ensimmäinen ladattava ohjelma on Tor-selain. Tor-selaimen salatun yhteyden avulla ladataan muut tarpeelliset ohjelmat, jolloin palveluntarjoajat kykenevät tallentamaan mahdollisimman vähän tietoja käyttäjästä. Tor-selaimen merkitys mallissa onkin suurin juuri ympäristön rakentamisen aikana.

3.3.1 Käyttöjärjestelmän koventaminen

Windowsia ei ole tarkoitettu turvalliseksi ja anonyymiksi käyttöjärjestelmäksi käyttäjän näkökulmasta. Tails, Qubes ja Whonix ovat käyttöjärjestelmiä, jotka on suunniteltu käyttäjän anonymiteettiä ajatellen, mutta niiden käyttäminen vaatii perehtyneisyyttä, jonka takia isäntäkäyttöjärjestelmän esimerkkinä käytetään Windows-käyttöjärjestelmää. Windowsiin voidaan lisätä turvallisuutta konfiguroinnin avulla. Mallin esimerkissä noudatetaan soveltaen Hassanin ja Hijazin (2018) OSINT-käyttöön tarkoitettuun Windows-käyttöjärjestelmään suosittelemia konfigurointeja. Näitä konfigurointeja ovat:

- Automaattisten päivitysten asettaminen päälle käyttöjärjestelmään.

- Windowsin ohjelmien päivityksistä varmistuminen ja säännöllinen päivittäminen.

- Tietokoneen lukitseminen vahvemmin, kuin Windowsin tunnistautumiseen luottaen. (Mallissa tämä voidaan toteuttaa luomalla pikavalinta näppäimet sammuttamiselle ja salaamalla kovalevy VeraCrypt ohjelman avulla)

- Vähemmän oikeuden käyttäjätilien käyttäminen mahdollisuuksien mukaan.

- User Account Control (UAC) ilmoitusten määrittäminen maksimiin. Tämä aiheuttaa sen, että tietokone ilmoittaa käyttäjälle aina kun jokin ohjelma yrittää asentaa ohjelmia tai yrittää tehdä muutoksia tietokoneeseen.

- Etätuen (Remote Assistance) kytkeminen pois päältä. Tämän toiminnallisuuden päällä oleminen voi mahdollistaa hyökkääjän päästä etänä käsiksi koneeseen.

- Piilotettujen tiedostojen tekeminen näkyviksi. Jotkin haitalliset ohjelmat käyttävät Windowsin ominaisuuksia piilottaakseen tiedostojaan.

- Salasanan asettaminen BIOS:iin/UEFI:in, jolla varmistetaan, että ulkopuoliset tahot eivät kykene ajamaan ohjelmia ulkoisesta muistista tietokoneen käynnistämisen yhteydessä. Hickeyn & Arcurin (2020) mukaan kaikki kovalevyn salausohjelmat eivät salaa BIOS/UEFI:a, jolloin niille on tarpeellista määrittää oma salasana. Salasanan asettaminen BIOS:siin/UEFI:in on tärkeää, mikäli tietokone ei ole jatkuvasti käyttäjän valvonnassa ja toimitaan toimintaympäristössä, jossa ulkopuolinen taho voi päästä tietokoneeseen käsiksi. Chaos Computer Club:in tapahtumassa esiteltiin PoC (engl. Proof of Consept) tyyppisesti käytännön esimerkki, miten hyökkääjä kykenee toteuttamaan evil maid hyökkäyksen, vain hyödyntämällä BIOS:ia. Evil maid -hyökkäys tarkoittaa hyökkäystä, jossa hyökkääjä pääsee fyysisesti käsiksi valvomattomaan laitteeseen ja tekee siihen muutoksia huomaamattomasti (Kmille, 2022.)

- Ylimääräisten porttien, protokollien ja palveluiden kytkeminen pois päältä.

- Windowsin ohjelmien, kuten kameran käyttöoikeuksien säätäminen tarpeen mukaan minimiin. Erityisesti sijainnin käyttöoikeus tulisi kytkeä kaikilta ohjelmilta pois päältä. Muita huomioita ovat esimerkiksi kieli ja näppäimistön käyttämä kieli, jotka voivat näkyä verkkosivujen keräämissä tiedoissa.

3.3.2 VPN

Ensimmäinen Tor-selaimen avulla ladattava ohjelma on VPN, jotta voidaan lisätä salaukseen lisää kerroksia mahdollisimman nopeasti. Mallin testaamiseen tarkoitetussa esimerkissä käytettävä VPN on avoimen lähdekoodin ProtonVPN:än maksullinen versio. Maksullisella versiolla varmistutaan siitä, ettei palveluntarjoaja rajoita verkkoyhteyden nopeutta ja vähennetään todennäköisyyttä sille, että käyttäjän tietoja käytettäisiin väärin. ProtonVPN lupaa tietosuojakäytännössään, että se ei lokita käyttäjän tietoliikennettä. ProtonVPN on kotoisin Sveitsistä, jossa on tunnetusti tiukka lainsäädäntö yksityisyyteen ja tietojenkäsittelyyn liittyen.

Sveitsin laki muun muassa velvoittaa ilmoittamaan henkilölle, mikäli hän on joutunut valvonnan kohteeksi ja hänen tietojaan on luovutettu viranomaisille. ProtonVPN tukee myös OpenVPN:ää, lupaa estää DNS-vuodot ja omistaa omat DNS-palvelimensa. Näin palvelu siis mahdollistaa myös käyttäjän estää oman internet-palveluntarjoajansa lokittaa käyttäjän tietoliikennettä (ProtonVPN, 2023.)

3.3.3 Salaus ja salasanat

Salaus jaetaan mallissa kahteen kokonaisuuteen, jotka ovat tiedostojen ja kovalevyn salaus sekä viestinnän salaus. Mallissa kovalevy ja arkaluontoiset tiedostot tulee salata ulkopuolisen käytön estämiseksi. Tämä turvaa myös käyttäjän datan, mikäli laite varastetaan. Kovalevyn ja tiedostojen salaus voidaan toteuttaa mallissa esimerkiksi ilmaisella, avoimen lähdekoodin ohjelmalla VeraCrypt. VeraCrypt mahdollistaa AES-salausalgoritmin ja SHA-512 HASH-algoritmin avulla koko kovalevyn salaamisen. Mikäli ulkopuolinen taho yrittää avata konetta, ei avaaminen onnistu ilman oikean salasanan tarjoamista koneelle ja kovalevy pysyy salattuna. VeraCrypt mahdollistaa myös yksittäisten tiedostojen salaamisen auki olevassa tietokoneessa luomalla salatun virtuaalisen levyn. VeraCrypt mahdollistaa myös ulkoisten muistien, kuten USB-tikun salaamisen. VeraCrypt:in avulla voidaan myös piilottaa tietokoneeseen käyttöjärjestelmiä. Tämä tarkoittaa sitä, että koneen avaamisen yhteydessä, kysyttäessä salasanaa, voidaan käynnistää toinen vaihtoehtoinen käyttöjärjestelmä tarjoamalla siihen liittyvä salasana (VeraCrypt, 2023.)

Vaikka Tutanota perustuukin avoimeen lähdekoodiin, lupaa päästä-päähän salauksen, tukee kaksivaiheista tunnistautumista ja anonymiteettiä, voi lisäturvallisuus viestintään joissakin tapauksissa olla tarpeen (Tutanota, 2023). OpenPGP on maailman laajimmin käytetty sähköpostien päästä-päähän salausstandardi, joka on Internet Engineering Task Forcen (IETF) ehdottama RFC4880 standardi. Kleopatra on sertifikaattien hallintaohjelma ja graafinen käyttöliittymä. Se mahdollistaa salauksessa käytettävien yksityisten ja julkisten avainten hallinnan sekä viestien salaamisen sekä salauksen purun. Kleopatraa käyttämällä voidaan varmistua sisällön pysymisestä yksityisenä, riippumatta käytettävästä sähköpostin-palveluntarjoajasta. Salatessa sähköpostiviestejä on hyvä kuitenkin muistaa, että sähköpostien metadataan on mahdollista jättää paljon dataa käytävästä viestinnästä (OpenPGP, 2023.)

OSINT-tiedustelua tehtäessä verkossa, saattaa tiedustelijalla olla lukuisia tilejä käytössään eri palveluihin. Näiden tilien hallintaan tarvitaan mallissa salasanan hallintaohjelma. Luotettava avoimen lähdekoodin salasanan hallintaohjelma, jota mallissa voidaan käyttää, on KeePassXC. KeePassXC on avoimen lähdekoodin ohjelma, joka luo salatun tiedoston, joka koostuu salasanoista ja käyttäjätunnuksista, suojaten tiedoston vahvalla salasanalla. Ohjelma kykenee myös luomaan käyttäjälle vahvoja salasanoja. Salattuja tietoja ei tallenneta tai synkronoida verkkoon, ellei käyttäjä erikseen halua tallentaa salattuja tietoja pilveen (KeePassXC, 2023.)

3.3.4 Virustentorjunta, palomuuri ja haittaohjelmien torjunta

Tietokoneen suojaamiseksi, mallissa käytetään virustentorjunta- ja haittaohjelmien torjuntaohjelmaa sekä palomuuria. Windows 10 ja 11 tarjoavat Microsoft:in oman Windows defender:in, joka pitää sisällään virustentorjuntaohjelman ja palomuurin. Esimerkissämme käytämme ympäristössämme siis Microsoftin tarjoamaa Windows defenderiä. Haittapuolena defenderin käytössä on se, että se tukee Microsoftin omia sovelluksia. Se ei siis välttämättä tarjoa kaksista suojaa esimerkiksi käytettäessä Firefox-selainta. Siksi käytämmekin vaihtoehtoisia selaimia mallissa pääsääntöisesti virtuaalikoneiden sisällä (Microsoft, 2023.)

Turvallisuuden lisäämiseksi mallin käytännön esimerkissä käytetään haittaohjelmien torjuntaohjelmana Malwarebytes ohjelmaa. Malwarebytes mahdollistaa laitteella olevien tiedostojen tarkastamisen skannaamalla. Ohjelma lupaa toimia jo valmiiksi asennettujen virustentorjuntaohjelmien kanssa. Malwarebytes ei ole avoimen lähdekoodin ohjelma ja tätä perustellaan sillä, että mikäli se olisi avoin, tarjoaisi se hyökkääjille keinon selvittää kuinka ohjelma on ohitettavissa. Ohjelman maksullinen versio on myös saatavilla ilmaisena kokeilu versiona (Malwarebytes, 2023.) Käytettäessä näitä ohjelmia, tulee kuitenkin varmistua, että ohjelmat ovat päällä, ne on konfiguroitu oikein, eikä niiden yhteiskäyttö aiheuta kompromisseja turvallisuuteen. Esimerkiksi käytettäessä Malwarebytes ohjelmaa, se saattaa korvata osan Windows Defender:in toiminnallisuuksista.

3.3.5 Päivitys- ja siivousohjelmat

Turvallisuuden takaamiseksi mallissamme käytetään päivitysohjelmaa, joka auttaa huolehtimaan, että tietokoneelle asennetut ohjelmat pysyvät ajan tasalla eivätkä ole haavoittuvia tunnettujen haavoittuvuuksien toimesta. Mikäli tietokoneen ohjelmien ajantasaisuudesta ei pidetä huolta, eivät virustentorjunta-, palomuuri- ja haittaohjelmien torjuntaohjelmatkaan välttämättä pysty suojaamaan käyttäjää luotettavasti. Mallin esimerkissä päivitysohjelmana käytetään Patch My PC Home Updater ohjelmaa. Ohjelma on ilmainen, muttei valitettavasti perustu avoimeen lähdekoodiin. Se tukee kuitenkin yli kolmensadan ohjelman päivitysten seurantaa ja on erittäin helppokäyttöinen (Patchmypc, 2023.)

Jotta voidaan olla varmoja, että tietokoneen kovalevyltä on kadonnut poistettu tieto, voidaan siivousohjelman avulla kovalevyn vapaaksi merkityt alueet uudelleenkirjoittaa. Siivousohjelmat myös mahdollistavat muistiin kertyneen ”roskan” poistamisen sekä evästeiden poistamisen ja hallinnan. Mallissa käytetään siivousohjelmana avoimen lähdekoodin siivousohjelmaa BleachBit, joka mahdollistaa kovalevyn uudelleen kirjoittamisen ja evästeiden poistamisen (BleachBit, 2023). CCleaner siivousohjelma oli aiemmin suosittu, mutta vuonna 2017 ohjelmasta paljastui merkittävä haavoittuvuus. Ohjelma ei myöskään ole avoimen lähdekoodin ohjelma ja sitä on monesti kritisoitu sekä syytetty siitä, että se kerää käyttäjien tietoja (Bleedingcomputer, 2020; CCleaner, 2023.) BleachBit:in käytössä tulee kuitenkin huomioida se, että se on erittäin aggressiivinen siivousohjelma, jonka huolimaton käyttö voi aiheuttaa käyttäjälle tärkeiden asioiden

poistamisen. Sitä käytettäessä tuleekin olla tarkka, ettei poista mitään itselleen tärkeää (Bazzell, 2018, s. 382).

3.3.6 Virtuaalikoneet

OSINT-tiedustelun käytännöllisyyden ja turvallisuuden maksimoimiseksi tulee mallissa käyttää kaikkeen toimintaan virtuaalikonetta, kun ympäristö on saatu rakennettua valmiiksi. Virtuaalikone suojaa käyttäjän isäntäkäyttöjärjestelmää uhilta, joita käyttäjään kohdistuu käytettäessä selainta OSINT-tiedusteluun. Virtuaalikone myös mahdollistaa OSINT-tiedustelijalle useiden erilaisten ympäristöjen rakentamisen, eri käyttötarkoituksiin. Luotaessa virtuaalikoneita, voi käyttäjä määrittää koneille uuden MAC-osoitteen. Virtuaalikoneet myös mahdollistavat koneiden helpon hallinnan. Koneet on mahdollista tallentaa tiettyyn vaiheeseen ennen sulkemista, niiden poistaminen ja luominen on helppoa sekä koneita voidaan kloonata varmuuskopioiksi, mikäli käytetty kone esimerkiksi saastuu haittaohjelmalla.

Mallin esimerkissä virtuaalisointi ohjelmana käytetään ilmaista, avoimen lähdekoodin VirtualBox ohjelmaa (VirtualBox, 2023). Virtualbox ohjelmaan voidaan luoda esimerkiksi neljä virtuaalikonetta, joiden käyttöjärjestelmät ovat Kali Linux, Whonix ja Tails. Näin saadaan käyttöä hajautettua eri virtuaalikoneille ja koneet tarjoavat esimerkin siitä, kuinka voidaan käyttää tehtäviin parhaiten soveltuvia käyttöjärjestelmiä. Tails:iä voidaan käyttää sähköpostiviestintään ja toimenpiteisiin, joissa ei tarvitse tallentaa tietoa sekä toiminta on toteutettavissa Tor-selaimen avulla. Tails salaa automaattisesti kaiken liikenteen Tor:in avulla ja väärentää MAC-osoitteen. Se ei tallenna mitään tietoa normaalissa tilanteessa tietokoneen kiinteään muistiin. Virtuaalikone kuitenkin mahdollistaa Tails:in kätevämmän käytön, sillä virtuaalikoneita voidaan tallentaa tiettyyn tilaan ennen sammuttamista, jolloin Tails:iä ei tarvitse joka kerta erikseen määrittää, vaan sen poistaminen on tarpeellista vain hätätilanteessa. Ulkopuolisten ohjelmien lataaminen voi vaarantaa Tails:in turvallisuuden, joten se ei sovellukaan parhaalla mahdollisella tavalla tiedonhankintaan. Toisena ja kolmantena virtuaalikoneena käytetään esimerkissä Whonix-käyttöjärjestelmää, jonka avulla käytetään selainta OSINT-tiedusteluun. Whonix on kehitetty niin, että se pysyy eristettynä isäntäjärjestelmästä, eikä lataa mitään isäntälaitteen IP-osoitteita internetiin. Whonixiin kuuluu kaksi virtuaalikonetta, gateway ja workstation. Gateway pitää huolen, että kaikki Whonixin tietoliikenne välitetään Tor-verkon kautta. Workstation virtuaalikoneessa käyttäjä suorittaa haluamiaan ohjelmia ja tekee aktiviteettejaan. Näin saadaan lisättyä turvallisuuteen ja yksityisyyteen yksi kerros lisää. Neljäntenä virtuaalikoneena käytetään Kali Linux-käyttöjärjestelmää. Kali sisältää ja siihen on ladattavissa helposti OSINT-tiedusteluun soveltuvia työkaluja verkosta. Kali mahdollistaa OSINT-tiedustelun tehokkaasti komentorivipohjaisten työkalujen avulla. Kalissa tietoliikennettä ei automaattisesti kuitenkaan salata Tor-verkon avulla (Kali Linux, 2023.)

3.4 Selain turvallisuus

Riippumatta käytettävästä selaimesta, on selaimen turvallisuutta ja yksityisyyttä mahdollista parantaa muuttamalla selaimen asetuksia. Mallin esimerkissä käytetään kahta erilaista selainta. Selaimet ovat Tor ja Firefox. Tor mahdollistaa verkon selaamisen anonyymisti, mutta toimii hitaasti ja osa verkkosivuista estää Tor-selaimen käytön. Tor-selaimeen ei myöskään ole suositeltavaa tehdä muutoksia tai ladata selainlaajennuksia, jotta selaimen tarjoama turvallisuus ei vaarannu. Tor on Firefox-pohjainen selain ja Firefox-selaimesta onkin mahdollista saada korkealla tasolla yksityisyyttä arvostava selain konfiguroimalla selaimen asetukset oikein. Firefox-selaimesta on mahdollista luoda myös varmuuskopio tiedosto, jonka voi esimerkiksi tallentaa USB-tikulle, jolloin Firefox-selainta ei tarvitse jokaiselle koneelle ja käyttöjärjestelmälle konfiguroida erikseen. Tiedosto pitää sisällään tallennetut ja ladatut kirjanmerkit, selainlaajennokset sekä selaimen asetukset. Firefox mahdollistaa vakio-ominaisuutena yksityisen selaimen käyttämisen painamalla Crtl + Shift + P selaimen avaamisen jälkeen. Tämä avaa uuden yksityisen selainikkunan Firefox:iin. Firefox lupaa, ettei se tallenna yksityistä selainta käytettäessä sivuhistoriaa, evästeitä, väliaikaisia tiedostoja ja hakuja sekä lupaa estää verkkosivujen seurantaa.

3.4.1 Selaimen perusasetukset

Firefox-selaimen perusasetukset löytyvät ”settings” valinnan takaa selaimessa. Perusasetukset, joita mallin esimerkissä käytämme ovat:

- Oletushakukoneen asettaminen yksityisyyttä arvostavaksi hakukoneeksi Startpage.
- Asetetaan ”Älä seuraa signaali” lähetettäväksi vierailtaville sivuille.
- Asetetaan Firefox oletusselaimeksi.
- Estetään sivuhistorian muistaminen ja sivustojen seuraaminen.
- Määritetään evästeet poistettavaksi selaimen sulkemisen yhteydessä.
- Kielletään tunnistustietojen, kuten salasanojen muistaminen selaimessa.
- Estetään osoitekentän ehdotukset, kuten sivuhistoria.
- Kielletään tarpeettomat käyttöoikeudet selaimelta, kuten sijainti, kamera ja mikrofoni.

- Laitetaan päälle HTTPS-only moodi, sallimaan vain verkkosivut, jotka käyttävät HTTPS-protokollaa.

- Estetään vaarallinen sisältö, lataukset ja varoitetaan haitallisista ohjelmista.

- Estetään Firefoxin datan keräys ja käyttö ulkopuolisten toimesta.

3.4.2 Selaimen lisäasetukset

Firefoxin muokattavuus mahdollistaa yksityisyyden maksimoimiseksi lisäasetusten muokkaamisen. Näihin asetuksiin on mahdollista päästä käsiksi kirjoittamalla selaimen osoitekenttään ”about:config” ja hakemalla osoitetta. Selain voi varoittaa sivun olevan haitallinen, mutta kyseessä on vain selaimen sivu, jolla lisäasetuksia voidaan hallita. Lisäasetuksissa tulee sivun hakukenttään syöttää haluttu asetus, sen saamiseksi näkyviin. Näitä asetuksia oikein käyttämällä ulkopuolisten tahojen on haastavampi seurata käyttäjän toimenpiteitä. Mallin esimerkissä muutetut asetukset ovat:

Browser.formfill.enable → false
Services.sync.prefs.sync.browser.formfill.enable → false

Nämä asetukset estävät verkkosivuja ehdottamasta käyttäjälle syötteitä kysyttäessä käyttäjän syötettä verkkosivuilla ja muistamasta käyttäjän aiempia syötteitä verkkosivulle.

Browser.cache.disk.enable → false
Browser.cache.disk_cache_ssl → false
Browser.cache.offline.enable → false

Näillä asetuksilla estetään selainta käyttämästä kovalevyä tietojen tallentamiseen.

Dom.event.clipboardevents.enabled → false

Jotkut sivut pyytävät ilmoituksen siitä, jos käyttäjä kopioi sivulta tekstiä tai kuvia. Tämä asetus estää verkkosivuja saamasta kyseistä tietoa.

Geo.enabled → false

Tämä asetus estää Firefoxia jakamasta käyttäjän sijaintia.

Network.cookie.lifetimePolicy → 2

Tämä asetus määrittää milloin selaimen käyttämät evästeet vanhenevat. Asetuksella ”2” evästeet vanhenevat, kun selain suljetaan.

Privacy.trackingprotection.enabled → true

Tämä asetus estää verkkosivuja seuraamasta käyttäjää.

Browser.safebrowsing.phishing.enabled → false
Browser.safebrowsing.malware.enabled → false

Nämä asetukset estävät Googlea seuraamasta ja skannaamasta käyttäjän näkymiä verkossa, ”haittaohjelmien” varalta.

Media.navigator.enabled → false
Dom.battery.enabled → false

Nämä asetukset estävät verkkosivuja näkemästä onko käyttäjän mikrofoni ja kamera päällä sekä mikä on käyttäjän akun varauksen määrä, jolloin näitä ei voida käyttää tietoina käyttäjän identifioimiseen.

Extensions.pocket.enabled → false

Pocket palvelu on selaimeen tarkoitettu ominaisuus, joka mahdollistaa sisällön, kuten videoiden tallentamisen. Tämä asetus poistaa toiminnon käytöstä.

Media.peerconnection.enabled → false
Media.peerconnection.turn.disable → true
Media.peerconnection.use_document_iceservers → false
Media.peerconnection.video.enabled → false

Näiden asetusten tarkoituksena on estää IP-osoitteen vuotaminen WebRTC-haavoittuvuuden takia. Tämä voi aiheuttaa IP-osoitteen vuotamisen, vaikka käytössä olisi VPN-palvelu.

3.4.3 Selainlaajennokset evästeiden hallintaan

Selaimiin on ladattavissa pieniä ohjelmia, joita kutsutaan selainlaajennoksiksi. Laajennosten tehtävä on toteuttaa selaimessa jokin tietty toiminnallisuus. Käyttämällä oikein valittuja ja käyttötarkoitukseen sopivia laajennoksia, voidaan lisätä selaimen turvallisuutta käyttäjän näkökulmasta. Suositeltavaa on käyttää selainlaajennoksia, jotka auttavat estämään ja hallitsemaan evästeitä sekä käyttäjän seuraamista verkkosivujen toimesta. Mikäli selain ei tue omista asetuksistaan HTTPS-protokollan automaattista valintaa, on tätä varten hyvä ladata oma selainlaajennoksensa. Myös verkkosivujen keräämää metadataa voidaan väärentää selainlaajennosten avulla.

Mallin esimerkissä evästeiden ja verkkosivujen seuraamisen hallitsemiseksi ja estämiseksi käytetään laajennoksia uBlock Origin ja Privacy Beaver. Nämä ohjelmat tarjoavat yhdessä parhaan suojan. uBlock mahdollistaa verkkosivujen

käyttämien, seuraamiseen tarkoitettujen tekniikoiden, kuten evästeiden hallinnan ja estää myös mainokset. Sen toiminta kuitenkin perustuu siihen, että se määrittelee sallitut ja ei sallitut ratkaisut vertaamalla niitä tunnettuihin listoihin, jotka määrittävät tuleeko kyseinen toiminto estää. Privacy Badger käyttää heuristiikkaa estämään verkkosivuja seuraamasta käyttäjää, mutta ei estä mainoksia. Tämä tarkoittaa käytännössä sitä, että Privacy Badger paikkaa niissä tilanteissa, joissa uBlock ei kykene estämään käyttäjän seuraamista. Vakiona nämä molemmat laajennokset toimivat itsenäisesti taustalla, mutta uBlock:ista on mahdollista asettaa kehittyneempi tila päälle, joka mahdollistaa käyttäjää hallitsemaan mitä sallitaan ja mitä estetään. Tämä onnistuu menemällä uBlock:in asetuksiin ja valitsemalla kohta ”I am an advanced user”. Tämä luo laajennokseen taulukon, jonka avulla käyttäjä voi itse hallita laajennosta (uBlock Origin, 2023; Privacy Badger, 2023.)

Muita mallin esimerkissä käytettäviä laajennoksia ovat Random User-Agent ja Smart HTTPS sekä ylimääräistä liikennettä generoivat laajennokset. User-Agent laajennos mahdollistaa väärentää verkkosivujen keräämään metadataan tietoja järjestelmästä, kuten käytetyn selaimen ja käyttöjärjestelmän. HTTPS-laajennos varmistaa, että käytetään vain HTTPS-protokollaa käyttäviä sivuja, ellei käyttäjä muuta määritä. Näin estetään sellaisten sivujen käyttäminen, joilta käyttäjän tietoliikenteen lukeminen selkokielisenä voisi olla mahdollista.

4 TUTKIMUKSEN TOTEUTUS

Tämän luvun kappaleissa 4.1 ja 4.2 käsitellään tutkimuksen tavoitteet, tutkimuskysymykset, näkökulma, rajaus sekä kuvataan tutkimusmenetelmän teoreettisia taustoja. Kappaleessa 4.3 käsitellään tutkimuksen toteutus vaihe vaiheelta, sisältäen tiedonhankinnan, haastateltavien kohderyhmän, haastatteluiden toteutuksen, aineiston käsittelyn ja analyysin. Lopussa käydään läpi myös tutkimuksen tietoturvaan ja tietosuojaan liittyviä asioita.

4.1 Tutkimustavoitteet ja -kysymykset

Tutkimuksen tavoitteena oli selvittää millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää tehtäessä avointen lähteiden tiedustelua, miten näitä jälkiä voidaan peittää sekä kehittää teoreettinen malli digitaalisten jalanjälkien peittämiseksi toteutettaessa avointen lähteiden tiedustelua verkossa. Passiivisten digitaalisten jalanjälkien peittämiseksi ei aiemmin ole tutkimuksen avulla kehitetty teoreettista mallia, joka hahmottaisi kokonaisuudet, jotka tulisi huomioida suojatessa identiteettiä verkossa. Tämä näkökulma tarjoaakin tilaisuuden tuottaa uutta tietoa tutkittavasta ilmiöstä sekä mahdollisuuden luoda teoreettinen malli, jonka avulla on mahdollista hahmottaa aiheeseen liittyviä laajoja kokonaisuuksia. Näiden tutkimustavoitteiden pohjalta tutkimukselle asetettiin kolme tutkimuskysymystä:

Tutkimuksen pääkysymys:

1. Kykeneekö tutkijan kehittämää teoreettista mallia hyödyntämällä peittämään ja suojaamaan käyttäjän identiteetin tehokkaasti teknisestä näkökulmasta, tehtäessä avointen lähteiden tiedustelua verkossa?

Tutkimuksen apukysymykset:
1. Millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää tehtäessä avointen lähteiden tiedustelua?
2. Miten passiivisia digitaalisia jalanjälkiä voidaan peittää tehokkaasti verkossa?

Apukysymyksillä pyritään selvittämään kirjallisuuden ja asiantuntijahaastatteluiden avulla millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää tehtäessä avointen lähteiden tiedustelua verkossa ja miten näitä jälkiä voidaan peittää. Tutkija on kehittänyt kirjallisuuden pohjalta teoreettisen mallin passiivisten digitaalisten jalanjälkien peittämiseksi. Pääkysymyksellä pyritäänkin selvittämään asiantuntijahaastatteluiden ja PoC (engl. Proof of Consept) tyyppisten testien avulla, kykeneekö tutkijan teoreettista mallia hyödyntämällä suojaamaan käyttäjän identiteetin teknisestä näkökulmasta, toteutettaessa avointen lähteiden tiedustelua verkossa.

4.2 Tutkimusmenetelmä

Tutkimus on laadullinen eli kvalitatiivinen. Laadullinen tutkimus on empiiristä ja siinä korostuu aineiston keräämis- ja analyysimetodit sekä niiden selostaminen uskottavasti lukijalle, antaen lukijan mahdollisuuden arvioida tutkimusta ja sen tulosten uskottavuutta. Laadullisella tutkimuksella pyritään kuvaamaan ja ymmärtämään tutkittavan kohteen laatua, ominaisuuksia ja merkityksiä kokonaisvaltaisesti, kun taas määrällinen tutkimus painottuu tutkittavan ilmiön selittämiseen ja nojaa aineistossaan tilastoihin sekä numeroihin (Tuomi & Sarajärvi, 2018, s. 19-23.) Tutkimuksen menetelmäsuuntaukseksi valikoitui laadullinen tutkimus, sillä sen katsottiin soveltuvan parhaiten tutkittavan ilmiön ja ongelman tutkimiseen.

Tutkimuksessa käytetään myös konstruktiivista tutkimusotetta, jonka avulla tutkijan kehittämä ”Maksimaalisen digitaalisen turvallisuuden malli” kehitettiin. Aiempaan kirjallisuuteen ja tutkimukseen perehdyttäessä ei löytynyt vastaavanlaista teoreettista mallia, jolla olisi pyritty hahmottamaan mahdollisimman yksinkertaisesti, laajojen teknisten kokonaisuuksien vaikutusta käyttäjän yksityisyyteen ja suojaan verkossa. Tähän ongelmaan pyrittiin kehittämään teoreettinen malli, joka hahmottaisi käyttäjälle niitä kokonaisuuksia, jotka hänen tulisi huomioida toimiessaan mahdollisimman anonyymisti ja turvallisesti verkossa. Konstruktiivisella tutkimusotteella pyritään ratkaisemaan reaalimaailman ongelma, tuottamalla uusi konstruktio. Konstruktioille on tyypillistä, että ne eivät ole löydettyjä, vaan ne kehitetään ja keksitään. Konstruktiivinen tutkimus on kokeellista ja konstruktiota tulisi tarkastella kokonaisuutena, jolla yritetään havainnollistaa, jalostaa, testata aikaisempaa teoriaa tai luoda uusi teoria. Uusi konstruktio pyritään luomaan sekä teorian, että käytännön näkökulmasta (Lukka, 2001.)

4.2.1 Tutkimuksen rajaus

Tutkimus rajattiin käsittelemään vain passiivisia digitaalisia jalanjälkiä, sillä katsottiin, että niiden peittämiseen kyetään vaikuttamaan teknisillä ratkaisuilla. Tämä mahdollisti tutkimuksen keskittymisen teknisiin näkökulmiin, sosiaalisten näkökulmien sijaan. Käyttäjän tekemät ratkaisut, virheet ja niiden välttämisen käsitteleminen jätettiin mahdollisimman vähäiseksi. Käyttäjätilien luomisen käsitteleminen katsottiin kuitenkin välttämättömäksi osaksi kokonaisuutta. Tutkimuksen konstruktioksi valikoitiin teoreettinen malli, sillä sen voidaan nähdä kestävän paremmin aikaa, kuin tietyillä ohjelmilla ja ratkaisuilla toteutetun käytännön ratkaisun. Teoreettisen mallin hyödyntäminen käytäntöön jätetäänkin lukijan vastuulle ja tutkimuksessa esitetyt ohjelmat edustavatkin esimerkkejä, joita voidaan käyttää konstruktion soveltamiseen käytäntöön. Tarkasteltavaksi laitteistoksi valittiin kannettava tietokone, sillä sen nähtiin soveltuvan liikuteltavuutensa kannalta hyvin aiheeseen. Lisäksi Shavers ja Bair (2016) kuvaavat miten erilaisista laitteista, kuten puhelimesta jää erilaisia passiivisia digitaalisia

jalanjälkiä (Shavers & Bair, 2016). Valitsemalla tarkasteluun kannettava tietokone, kyettiin rajaamaan tarkasteltavien digitaalisten jalanjälkien määrää. Digitaalisista jalanjäljistä pyrittiin käsittelemään kattavasti kaikki merkittävimmät passiiviset digitaaliset jalanjäljet sekä niiden peittämiseen soveltuvat käytännöt. Tutkimuksen rajaamiseksi ja eettisyyden lisäämiseksi anonyymiä maksamista, viestintää sekä laitteiden hankintaa ei tarkasteltu, kuin siltä laajuudelta, mikä on tarpeellista tietojen piilottamiseksi palveluidentarjoajilta. Lisäksi avointen lähteiden tiedustelu verkossa liitettiin aiheeseen rajauksen, eettisyyden ja ajankohtaisuuden lisäämiseksi. Näin kyettiin jättämään kyberrikoksiin liittyvät ominaispiirteet, kuten kohdelaitteisiin tunkeutuminen käsittelyn ulkopuolelle. Konstruktion testauksessa käytettävässä esimerkissä käytetyiksi ohjelmiksi pyrittiin valitsemaan ilmaisia, avoimen lähdekoodin ohjelmia. Näin pyrittiin luomaan lukijalle kuva siitä, mitä käytettävien ohjelmien valinnassa tulisi huomioida ja, että mallin toteuttaminen on mahdollista myös pienin kustannuksin. Tutkimuksen tutkimuskysymykset pyrittiin luomaan niin, että ne rajaisivat tutkimusta ja tukisivat toisiaan.

4.2.2 Aineistonkeruumenetelmä

Tutkimuksen kirjallisuuskatsauksen metodina käytettiin kuvailevaa kirjallisuuskatsausta ja kirjallisuuskatsauksen muotona narratiivista kirjallisuuskatsausta. Narratiivisessa kirjallisuuskatsauksessa epäyhtenäistä tietoa järjestetään jatkuvaksi tapahtumaksi. Kirjallisuuskatsauksessa pyrittiin rakentamaan usean lähteen pohjalta yhtenäinen kuvaus käsiteltävästä aiheesta. Kirjallisuuskatsauksen toimitustapana käytettiin yleiskatsausta. Yleiskatsauksessa on kyse laajemmasta prosessista, jonka tarkoituksena on tiivistää aiemmin tehtyä kirjallisuutta (Salminen, 2011, s. 7-8.) Kirjallisuuskatsauksen pääsääntöiset lähteet ovat asiantuntija tahojen kirjoittamaa kirjallisuutta. Näitä lähteitä ovat aihealueeseen liittyvät kirjat, artikkelit, tutkimukset, ohjeet ja muut asiakirjat. Esimerkkejä asioiden konkretisoimiseksi, lukuelämyksen parantamiseksi ja tietoaukkojen täydentämiseksi paikattiin internet-lähteillä, kuten uutisartikkeleilla ja alan tunnustettujen toimijoiden internetjulkaisuilla. Kirjallisuuskatsauksen avulla pyrittiin osittain vastaamaan tutkimuksen alakysymyksiin ja luomaan tutkijan oma konstruktio, teoreettinen malli identiteetin peittämiseksi ja suojaamiseksi toimittaessa verkossa.

Tutkimuksen empiirisen osan aineistonkeruumenetelmänä käytettiin asiantuntijahaastatteluja, joilla pyrittiin vastaamaan kaikkiin tutkimuksen tutkimuskysymyksiin. Haastattelut valikoituivat aineistonkeruumenetelmäksi, sillä niiden avulla pyrittiin ymmärtämään haastavaa aihetta sekä täydentämään kirjallisuuskatsauksen avulla luotua kokonaisuutta asiantuntijoiden käsityksillä ja kokemuksilla. Aiheen haastavuuden takia haastattelut sopivat loistavasti aineistonkeruumenetelmäksi, sillä ne mahdollistivat aineistonkeräämisen joustavasti. Tuomi ja Sarajärvi (2018) toteavatkin haastattelujen etuna joustavuuden, sillä ne mahdollistavat kysymysten tarkentamisen, väärinkäsitysten oikaisemisen ja haastattelun toteutuksen muokkaamisen tarpeen mukaan (Tuomi & Sarajärvi,

2018, s. 82.) Tämä oli tutkimuksen kannalta tärkeää, sillä haastatteluja toteutettaessa oli tarve kysyä sekä tarkentaa tiettyjä vastauksia. Tämä ei olisi samalla tavalla ollut mahdollista esimerkiksi kyselyiden avulla. Haastattelut toteutettiin teema- eli puolistrukturoituina haastatteluina. Puolistrukturoiduissa haastatteluissa haastateltavat saavat vastata samoihin kysymyksiin vapaasti, omin sanoin, valmiiden vastausvaihtoehtojen sijaan sekä tutkijan on mahdollista esittää syventäviä ja tarkentavia kysymyksiä (Eskola & Suoranta, 1998, s. 63). Näin varmistettiin se, että asiantuntijoilla oli mahdollisuus lisätä huomioitaan aiheeseen liittyen, mikäli tutkijalla oli jäänyt jokin merkittävä asia huomioimatta. Tuomi ja Sarajärvi (2018) korostavat sitä, että teemahaastattelu rakentuu tutkijan valitsemien teemojen ja niiden alle muodostettujen kysymysten varaan. Teemat muodostetaan tutkimusongelma ja aiempi tieto huomioiden (Tuomi & Sarajärvi, 2018, s. 84–85.)

Toisena empiirisenä aineistonkeruumenetelmänä testattiin tutkijan konstruktiivista mallia empiirisesti PoC (engl. Proof of Consept) hengessä. Testien avulla pyrittiin testaamaan teoreettisen mallin eri kokonaisuuksia käytännön näkökulmasta ja luomaan esimerkkejä lukijalle, miksi tietyt kokonaisuudet tutkijan mallissa tulee huomioida peitettäessä ja suojatessa identiteettiä verkossa. Testaus toteutettiin määrittämällä testattavat kokonaisuudet ja kuhunkin testiin keskenään vertailtavat asiat. Testien tuloksia verrattiin toisiinsa, tehtiin havaintoja sekä selvitettiin mallin toimivuutta käytännössä. Käytännön toteutuksen testaaminen kohdistettiin tutkijan teoreettisen mallin eri kokonaisuuksiin, eikä täydellistä anonymiteettiä testattu. Tämä johtui siitä, että täydellisen anonymiteetin testaamisen voidaan katsoa olevan haastavaa käytännön tasolla. Tämä johtuu siitä, että myös aktiivisilla digitaalisilla jalanjäljillä on merkittävä osa pyrittäessä anonymiteettiin verkossa (Shavers ja Bair, 2016). Käytännön toteutuksen tekeminen voidaankin nähdä osaksi käyttäjän aktiivista toimintaa, jonka seurauksena on mahdollista tehdä virheitä, jättäen jälkeensä digitaalisia jalanjälkiä. Laajempien testien toteuttamisen voitaisiin myös nähdä paisuttavan työtä merkittävästi, työn laajuus huomioiden.

4.2.3 Aineiston analyysimenetelmä

Tutkimuksen aineiston analyysimenetelmänä käytetään teorialähtöistä analyysiä. Analyysissä nojataan tutkijan konstruktiivisen tutkimuksen avulla kehitettyyn teoreettiseen malliin käyttäjän identiteetin peittämiseksi ja suojaamiseksi. Tuomi ja Sarajärvi (2018) jakavat laadullisen tutkimuksen analyysin karkeasti kahteen ryhmään. Ensimmäisessä ryhmässä analyysiä ohjaa teoreettinen tai epistemologinen asemointi, kuten esimerkiksi fenomenologinen analyysi tai grounded theory. Ne analyysimuodot, joita ei ohjaa tietty teoria tai epistemologia kuuluvat toiseen ryhmään. Teoreettisia ja epistemologiaisia lähtökohtia voidaan kuitenkin soveltaa toiseen ryhmään, johon kuuluu esimerkiksi sisällönanalyysi (Tuomi & Sarajärvi, 2018, s. 100.) Tuomi ja Sarajärvi (2018) jakavat laadullisen analyysin muodot aineistolähtöiseen analyysiin, teoriaohjaavaan analyysiin ja teorialähtöiseen analyysiin. Suurin ero näiden välillä muodostuu siitä, kuinka taustalla

vaikuttava teoria ohjaa aineiston hankintaa, analyysia ja raportointia (Tuomi & Sarajärvi, 2018, s. 104–109.)

Tuomen ja Sarajärven (2018) mukaan aineistolähtöisen analyysiin päättelyn logiikka on enemmän induktiivista, kuin deduktiivista. Heidän mukaansa aineistolähtöisen analyysin avulla pyritään luomaan tutkittavasta aiheesta ja aineistosta teoreettinen kokonaisuus. Tehtävänasettelu ja tutkimuksen tarkoitus ohjaavat analyysiyksiköiden valintaa eivätkä ne ole etukäteen valittuja. Tutkittavaan ilmiöön liittyvien aikaisempien tietojen, havaintojen ja teorioiden ei pitäisi vaikuttaa analyysin lopputulokseen. Aineistolähtöisen analyysin ongelmaksi nähdään objektiivisuuden puute. Esimerkiksi tutkijan käyttämien käsitteiden, menetelmien, tutkimusasetelmien voidaan nähdä vaikuttavan tutkimuksen tuloksiin (Tuomi & Sarajärvi, 2018, s. 104–109.)

Teoriaohjaavalla analyysilla voidaan Tuomen ja Sarajärven (2018) mukaan ratkaista aineistolähtöiselle analyysille tyypillisiä ongelmia, ottamalla teoriaa enemmän mukaan. Teoriaohjaavassa analyysissa analyysi ei pohjaudu kuitenkaan suoraan teoriaan ja analyysiyksiköt valitaan edelleen aineistosta. Aikaisempi tieto kuitenkin ohjaa ja auttaa analyysia. Aikaisemman tiedon ei ole tarkoitus testata teoriaa, vaan auttaa analyysia. Teoriaohjaavan analyysin päättelyn logiikka voidaan nähdä abduktiivisena päättelynä, jossa ajatteluprosessissa vaihtelevat valmiit mallit ja aineistolähtöisyys (Tuomi & Sarajärvi, 2018, s. 106–107.)

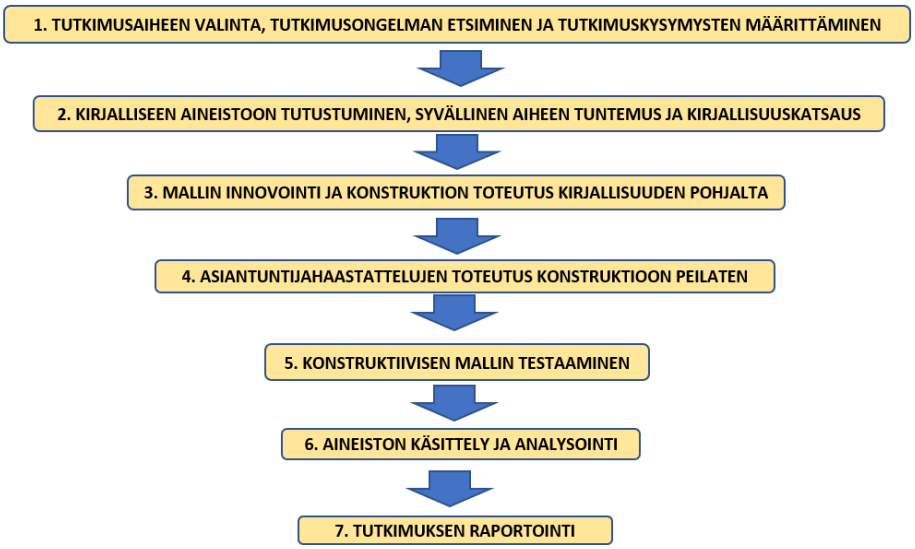
Tuomen ja Sarajärven (2018) mukaan teorialähtöinen analyysi nojaa tiettyyn teoriaan tai malliin. Tutkimuksessa kuvaillaan tämä tunnettu teoria tai malli ja se määrittää tutkimuksen kohteena olevat käsitteet eli tutkittavan ilmiön. Analyysia ohjaakin aiemman tiedon perusteella luotu teoria, malli tai kehys. Tällä pyritäänkin useimmiten aikaisemman tiedon testaamiseen uudessa kontekstissa. Teorialähtöisen analyysin päättelyn logiikka yhdistetään Tuomen ja Sarajärven (2018) mukaan deduktiiviseen päättelyyn. Tietoa pyritään suhteuttamaan tutkimuksen teoriaosassa hahmoteltuihin kategorioihin ja tutkimuskysymykset määräytyvät tutkimuksen teoriaosuuden pohjalta (Tuomi & Sarajärvi, 2018, s. 107–108.)

Teorialähtöinen analyysimuoto soveltuu parhaiten tähän tutkimukseen, sillä ilmiötä tutkitaan teknisestä näkökulmasta, jolloin aiempaa tietoa on käytettävissä merkittävästi. Lisäksi tutkimus pohjautuu tutkijan kehittämään teoreettiseen malliin, joka on kehitetty aiemman tiedon pohjalta.

4.3 Tutkimusprosessi

Tämä tutkimus toteutettiin yksilötyönä, ilman toimeksiantoa. Työ toteutettiin täysin etätyönä ja yhteydenpito tarvittaviin tahoihin hoidettiin sähköpostin ja Zoom-sovelluksen avulla. Työn lähdekirjallisuus oli julkista ja työ toteutettiin julkisena. Haastateltavilta ei myöskään kerätty henkilötietoja. Näin kyettiin aineistoa käsittelemään joustavasti, työtä tekemään etänä ja soveltamaan tarvittavia ohjelmia työn laatimiseksi. Merkittävimmät työkalut, joita työn tekemiseen käytettiin, olivat sähköposti, Zoom ja Microsoft Office. Työn säilöntä ja

versionhallinta toteutettiin paikallisesti tutkijan laitteella. Lisäksi työstä ja aineistosta säilytettiin jatkuvasti varmuuskopio erillisellä fyysisellä USB-muistitikulla.



KUVIO 6 Tutkimusprosessin vaiheet

Tutkimuksen aiheen valinta ja kartoittaminen aloitettiin syksyllä 2022. Ajatus aiheeseen syntyi Turvallisuus ja strateginen analyysi maisteriohjelmaan kuuluvien tiedustelukurssien aikana, kun opiskelijoiden kysyessä suojautumisesta verkossa, eivät opettajat osanneet yksiselitteisesti vastata. Tästä tutkijalle syntyi mielenkiinto ottaa aiheesta selvää ja halu perehtyä digitaalisiin jalanjälkiin teknisestä näkökulmasta. Tutkija näki myös mahdollisuuden tuottaa uutta tietoa, luomalla kokonaisuudesta teoreettisen mallin, jonka avulla laajaa kokonaisuutta olisi mahdollista hahmottaa. Tutkija ehdotti aihetta ohjaajalle loppuvuodesta 2022 ja työn aikataulutavoitteeksi asetettiin syyskuu 2023. Vuoden 2023 ensimmäisen neljänneksen aikana tutkija esitti ohjaajalle tutkimussuunnitelman ja toteutti syvällisen aiheeseen perehtymisen sekä kirjallisuuskatsauksen. Kirjallisuuskatsauksen avulla työ rajautui nykyiseen muotoonsa, malli innovoitiin sekä haastattelut, että testit suunniteltiin ja valmisteltiin. Vuoden 2023 toisen neljänneksen alussa tutkija esitti ohjaajalle suunnitelman empiirisen aineiston keräämiseksi, jonka jälkeen tutkija aloitti aineiston keräämisen. Aineisto kerättiin ja käsiteltiin vuoden 2023 toisen neljänneksen aikana, jonka päätteeksi kirjoitettiin tutkimuksen raportti.

4.3.1 Tiedonhankinta

Tiedonhankinnassa kirjallisuuskatsaukseen hyödynnettiin alkuvaiheessa internetiä, jonka avulla etsittiin aiheeseen liittyvää kirjallisuutta. Hakutermeinä käytettiin digitaaliseen identiteettiin, digitaalisiin jalanjälkiin ja avointen lähteiden tiedusteluun liittyviä termejä niin englanniksi, kuin suomeksi. Mikäli haluttu kirjallisuus ei ollut saatavilla ilmaiseksi internetin avulla, pyrittiin hyödyntämään

saatavilla olevia kirjastopalveluita tai ostamaan teokset joko fyysisinä painoksina tai digitaalisina versioina. Tiedonhankinta aloitettiin hyödyntämällä potentiaalisia hakukoneita ja tietokantoja. Hyödynnettyjä tietolähteitä olivat hakukoneet, kuten Google ja Google Scholar. Kirjastopalveluissa tukeuduttiin Lappeenrannan kaupunginkirjastoon, Jyväskylän yliopiston kirjaston JYKDOK-palveluun ja JYX-julkaisuarkistoon. Merkittävimpiä tietokantoja, joita käytettiin, olivat ResearchGate, IEEE Xplore ja Academia. Mikäli kirjojen katsottiin olevan tutkimuksen kannalta keskeisiä, eivätkä ne olleet muuten saatavilla, ostettiin kirjat Adlibriksen tai Google Books palveluiden kautta. Kirjallisuuskatsauksen tiedonhankinnassa haasteeksi muodostui aiempi tutkimus aiheeseen liittyen. Digitaalisten jalanjälkiin liittyvää aiempaa tutkimusta on paljon, mutta se painottuu vahvasti sosiaalisiin ilmiöihin, teknisten näkökulmien sijaan. Tämä aiheutti haasteita löytää merkittäviä aiempia tutkimuksia aiheeseen liittyen. Toinen merkittävä huomio aiemmissa tutkimuksissa oli se, että ne keskittyivät aina todella tiukasti rajattuihin ja yksityiskohtaisiin kokonaisuuksiin digitaalisiin jalanjälkiin liittyen.

Etsittäessä aineistoja, perehdyttiin löydettyihin aineistoihin huolella. Ensin selvitettiin potentiaalisten aineistojen kohdalla ensimmäiseksi se, liittyvätkö aineistot digitaalisiin jalanjälkiin teknisestä näkökulmasta. Toinen arvioitava kokonaisuus oli aineistojen luotettavuuden arviointi. Sisällön ja luotettavuuden kannalta parhaat aineistot luettiin läpi ja aineiston hankintaa jatkettiin niin pitkään, että saturaatio kirjallisuudessa saavutettiin, eivätkä uudet aineistot lisänneet tietoa ja luoneet arvoa enää merkittävästi kirjallisuuskatsauksen kannalta. Lähteiden viimeinen arviointi tehtiin kirjallisuuskatsauksen yhteydessä, kun tutkimuksen teoriasisältöä koostettiin ja tutkijan mallia innovoitiin. Lähteitä hallinnoitiin erilaisten digitaalisten jalanjälkien kokonaisuuksien mukaan, luomalla aiheisiin soveltuvat kansiot, joissa aineistoja säilytettiin. Tämä oli välttämättömyys, sillä merkittävä osa lähteistä käsitteli vain tiettyjä digitaalisten jalanjälkien kokonaisuuksia. Merkittävimmät kirjallisuuskatsauksen lähteet olivat tunnustettujen asiantuntijoiden kirjoittamaa kirjallisuutta, aiempaa tutkimusta tai alan virallisten toimijoiden kirjallisia aineistoja. Tutkimuksessa käytettiin myös muutamia internet- ja uutisjulkaisuja, joiden tarkoitus oli parantaa tekstin luettavuutta ja antaa käytännön esimerkkejä lukijalle. Näitä julkaisuja ei kuitenkaan käytetty teoriasisällön ja mallin kokonaisuuksien rakentamisessa. Tutkimuksen empiirinen aineisto kerättiin tutkijan konstruktiosta puolistrukturoitujen asiantuntijahaastatteluiden ja testien avulla. Nämä kokonaisuudet käsitellään tulevissa alaluvuissa.

4.3.2 Haastateltavien kohderyhmä ja haastattelujen toteutus

Tutkimusta varten haastateltiin yhdeksää asiantuntijaa, jotka ovat ammattinsa tai harrastuneisuutensa kautta erikoistuneet tietoverkkoihin, digitaalisiin jalanjälkiin tai avointen lähteiden tiedusteluun teknisestä näkökulmasta. Vaikka tekninen osaaminen olikin vaatimuksena, haastateltavien taustoihin pyrittiin hakemaan kuitenkin vaihtelua, jolla pyrittiin saamaan mahdollisimman monipuolinen ymmärrys tutkittavasta ilmiöstä ja vastauksia hieman erilaisista

näkökulmista. Haastateltavien löytäminen muodostikin haasteen, sillä riittävän perehtyneiden haastateltavien löytäminen sekä tavoittaminen muodostui haasteeksi. Haastateltavia etsittäessä sekä soveltuvia kartoitettaessa selvisi esimerkiksi se, että vaikka kentällä toimiikin paljon tietotekniikan teknisiä ammattilaisia, ei valtaosan heistä ole tarvinnut perehtyä identiteetin tai digitaalisten jalanjälkien suojaamiseen osana työtään tai harrastuksiaan. Haastattelujen toteuttaminen ja järjestäminen vei myös enemmän aikaa, kuin tutkija oli osannut ennakoida. Tämä johtui siitä, että vaatimuksena haastatteluihin oli se, että haastateltavat olivat perehtyneet ennakkoaineistoon. Haastateltavien löytämiseksi tutkija lähetti sähköpostiviestejä useisiin organisaatioihin, kävi etsimässä haastateltavia kyberturvallisuusalan tapahtumista, lähestyi suoraan tiettyjä henkilöitä tai käytti jo löydettyjen henkilöiden kontakteja uusien asiantuntijoiden tavoittamiseen. Haastatteluihin pyydettiin varaamaan aikaa kaksi tuntia. Haastateltavien määrä asettui yhdeksään henkilöön. Tämän määrän tutkija katsoi sopivaksi, sillä vastauksissa alkoi merkittävissä määrin ilmenemään saturaatiota, eikä uusien haastateltavien katsottu lisäävän enää merkittävästi arvoa tutkimukselle. Haastateltavilta ei kerätty henkilötietoja ja heille sekä heidän edustamalleen organisaatiolle luvattiin anonymiteetti tutkimukseen liittyen. Tämä katsottiin haastateltavien näkökulmasta tärkeäksi, aiheen arkaluontoisuuden vuoksi.

Kirjallisuuskatsauksen, tutkijan konstruktion ja tutkimuskysymysten pohjalta muodostettiin haastattelurunko ja teemat. Haastattelu jakautui neljään teemaan. Ensimmäisessä teemassa pyrittiin selvittämään haastateltavien näkemyksiä passiivisista digitaalisista jalanjäljistä sekä selvittämään oliko tutkija jättänyt jonkin merkittävän digitaalisen jalanjäljen käsittelemättä tutkimuksensa teoriaosuudessa. Toisessa teemassa pyrittiin selvittämään haastateltavien näkemyksiä menetelmistä passiivisten digitaalisten jalanjälkien peittämiseksi sekä sitä oliko tutkija jättänyt käsittelemättä joitakin keskeisiä menetelmiä digitaalisten jalanjälkien peittämiseksi, tutkielman teoriaosuudessa. Kolmannessa teemassa tarkasteltiin haastateltavien teknisiä ja teoreettisia näkemyksiä tutkijan konstruktiosta, teoreettisesta mallista identiteetin peittämiseksi ja suojaamiseksi verkossa. Näkemyksiä haettiin muun muassa mallin oikeellisuuteen, käytettävyyteen, sovellettavuuteen sekä ymmärrettävyyteen. Neljäntenä teemana oli vapaasana ja keskustelu, jonka avulla pyrittiin vielä selvittämään, oliko haastateltavien mielestä asioita ja kokonaisuuksia, joita ei ollut osattu huomioida. Nämä teemat muodostivat yhteensä 22 kysymystä, jotka on esitetty tutkielman liitteenä (Liite 1).

Tutkimusta varten toteutetut yhdeksän haastattelua toteutettiin maaliskuun 2023 ja kesäkuun 2023 välisenä aikana. Haastateltaville tutkija lähetti ennakkoon tiedotteen tutkimuksesta, teemahaastattelun haastattelukysymykset ja ennakkomateriaalin, joka muodostui tutkielman luvuista 1–3. Haastateltavat saivatkin tutustua haastattelukysymyksiin ennakkoon ja heidän oli tarkoitus vastata kysymyksiin ennakkomateriaalin sekä oman asiantuntemuksensa pohjalta. Samalla haastateltavien kanssa sovittiin myös ajankohta, jolloin haastattelu toteutettaisiin etänä. Haastateltaville lähetetty tiedote tutkimuksesta löytyy tutkimuksen liitteenä (Liite 2). Tiedotteessa kerrottiin, että tutkimukseen osallistuminen on vapaaehtoista, osallistumisen voi halutessaan keskeyttää, tutkimuksesta

ei aiheudu haastateltaville kuluja ja tutkittaville luvattiin ilmoittaa, miten valmiiseen työhön pääsee tutustumaan. Luottamuksellisuutta ja anonymiteettiä korostettiin tutkimuksen kaikissa vaiheissa. Tutkimuksen aihe aiheuttikin mielenkiintoa haastateltavissa ja sai tutkimusaiheena kiitosta.

Haastattelut toteutettiin yksilöhaastatteluina Zoom-sovelluksen avulla, jonne tutkija oli perustanut oman ”huoneen” haastatteluja varten. Haastattelun alussa kerrattiin tiedotteessa olleet keskeisimmät asiat, esittäydyttiin ja aloitettiin vapaamuotoinen keskustelu, jolla pyrittiin luomaan haastatteluun vapautunut ilmapiiri. Haastattelujen aikana tutkija jakoi haastattelulomakkeen näytönjaon avulla haastateltaville ja kirjasi haastateltavien vastauksia haastattelun aikana ylös. Tämä osoittautui erinomaiseksi menetelmäksi kyseiseen tutkimukseen liittyen, sillä haastateltavilla oli mahdollisuus reaaliajassa korjata tutkijan kirjaamia vastauksia ja tutkija kykeni esittämään reaaliajassa tarkentavia kysymyksiä. Näin teknisesti haastavat kokonaisuudet kyettiin käsittelemään ja väärinkäsityksiltä vältyttiin. Samalla kyettiin haastatteluaineisto jo osittain litteroimaan sekä pelkistämään ja haastattelun lopuksi haastateltavilla oli mahdollisuus tarkastaa heidän antamansa vastaukset kirjallisessa muodossa. Tämä ”yhden luukun periaate” katsottiin haastateltavien toimesta hyväksi asiaksi, sillä merkittävä osa heistä korosti töihin liittyviä kiireitään. Ennen kolmannen teeman käsittelyä tutkija kertasi ja esitteli kehittämänsä ”Maksimaalisen digitaalisen turvallisuuden mallin”.

4.3.3 Mallin käytännön testien toteutus

Tutkijan konstruktion testaamiseksi tutkimuksessa toteutettiin neljä PoC (engl. Proof of Consept) henkistä käytännön testiä. Testien tarkoituksena oli testata tutkijan kehittämän teoreettisen mallin kokonaisuuksien soveltamista käytäntöön ja luoda käytännön soveltamisesta esimerkki. Lisäksi pyrittiin luomaan esimerkit siitä, miksi malliin sisältyvät kokonaisuudet tulisi huomioida passiivisia digitaalisia jalanjälkiä sekä identiteettiä suojattaessa verkossa. Testit pyrittiin toteuttamaan niin, että jokaiseen tutkijan ”Maksimaalisen digitaalisen turvallisuuden mallin” kokonaisuuteen kyettiin kohdentamaan yksi testi.

Ensimmäisessä testissä pyrittiin testaamaan mallin identiteetti turvallisuuden kokonaisuutta. Tämä toteutettiin pyytämällä kolmen palveluntarjoajan keräämät tiedot käyttäjästä käyttäjätunnuksen perusteella ja tarkastelemalla sekä vertaamalla näitä tietoja keskenään. Vuonna 2018 voimaan astunut tietosuojaasetus GDPR (engl. General Data Protection Regulation) mahdollistaa EU maassa asuvien ihmisten pyytää organisaatiolta tieto siitä, mitä henkilötietoja heistä on tallennettu ja mihin tarkoitukseen niitä käsitellään. Sen tarkoituksena onkin antaa yksilölle enemmän keinoja henkilötietojen käsittelyyn (Tietosuoja, 2023.) GDPR tarjoaakin jokaiselle mahdollisuuden pyytää itseään koskevat henkilötiedot erinäisiltä organisaatiolta sekä mahdollisuuden pyytää itsensä unohdetuksi eli tietojen poistamista, mikäli tietojen säilyttämiselle ei ole lainmukaista edellytystä.

Muita testejä varten luotiin kaksi erillistä infrastruktuuria, joita testattiin yhdessä ja tuloksia vertailtiin passiivisten digitaalisten jalanjälkien

näkökulmasta. Ensimmäinen infrastruktuuri oli työasema, joka oli toteutettu tutkijan mallin mukaisesti ja toinen infrastruktuuri oli työasema, jolle ei ollut tehty mitään normaalin käyttäjän toimenpiteistä eroavia ratkaisuja. Ensimmäinen infrastruktuuri eli mallin käytännön toteutus on kuvattu tutkimuksen kappaleessa 3. Toinen testi kohdistettiin tutkijan mallin fyysisen turvallisuuden kokonaisuuteen. Testissä molemmat infrastruktuurit liitettiin samaan verkkoon ja Wireshark-pakettianalysaattoria hyödyntämällä tarkkailtiin mitä liikenteestä oli selvitettävissä. Kolmas testi kohdistettiin mallin laitteisto- ja ohjelmistoturvallisuuden kokonaisuuteen. Testillä pyrittiin selvittämään laitteiden julkinen IP-osoite ja IP-osoitteen avulla selvittämään mitä tietoja käyttäjästä oli saatavissa julkisen IP-osoitteen avulla. Tuloksia vertailtiin jälleen keskenään ja havaintoja pohdittiin. Testien toteuttamiseen käytettiin web-sovellusta, joka kykenee keräämään tietoja IP-osoitteista. Neljäs testi kohdistettiin mallin selain turvallisuuden kokonaisuuteen. Testillä pyrittiin testaamaan, millaisia digitaalisia jalanjälkiä selain kykenee keräämään ja miten selaimen digitaalisten jalanjälkien keräämistä kyetään estämään ja peittämään. Testit toteutettiin hyödyntäen web-sovellusta, joka kykenee esittämään selaimen keräämät tiedot visuaalisesti ja ilmaisee kuinka ainutlaatuinen digitaalinen identiteetti selaimen keräämien tietojen avulla käyttäjälle, on mahdollista luoda. Testit ja niiden toteutus on kuvattu yksityiskohtaisemmin kappaleessa 5.2.

4.3.4 Aineiston käsittely ja analyysi

Haastatteluaineiston analysointi ja käsittely muodostui kuudesta vaiheesta. Ensimmäisessä vaiheessa haastatteluaineisto litteroitiin ja ryhmiteltiin haastattelurungon mukaisesti. Tämä aloitettiin jo haastatteluiden aikana, sillä haastattelut kirjattiin suoraan kirjalliseen muotoon, haastattelurungon mukaisesti. Tämä mahdollisti aineiston hyväksyttämisen haastateltavilla, välittömästi haastattelujen jälkeen. Näin varmistuttiin myös siitä, että aineistoon ei jäänyt haastateltavien henkilötietoja. Litteroitua haastatteluaineistoa muodostui yhteensä 45 sivua, käytettäessä fonttia Calibri, kirjasinkokoa 12 ja riviväliä 1,5. Litteroinnin lopuksi, aineisto koostettiin Excel-taulukkoon kysymyksittäin, jatkokäsittelyn helpottamiseksi. Jokainen haastattelu kesti ajallisesti noin 90–120 minuuttia.

Toisessa vaiheessa tutustuttiin aineistoon ja luotiin analyysirunko Excel-taulukon avulla. Analyysirunko muodostettiin teorialähtöisen sisällönanalyysin pohjalta, joka perustettiin tutkijan ”Maksimaalisen digitaalisen turvallisuuden malliin”. Analyysirungon yläluokat olivat tutkijan kehittämän teoreettisenmallin mukaiset kokonaisuudet 1) identiteetti turvallisuus 2) fyysinen turvallisuus 3) laitteisto- ja ohjelmistoturvallisuus sekä 4) selain turvallisuus. Analyysirungon avulla poimittiin ne asiat, jotka kuuluivat analyysirunkoon ja niistä johdettiin uusia alaluokkia. Rungon ulkopuolelle jääneistä asioista voitiin muodostaa uusia alaluokkia aineistolähtöistä sisällönanalyysia noudattaen. Excel-taulukkoa päädyttiin käyttämään, sillä aineisto oli jo valmiiksi litteroitu Excel-taulukkoon, ohjelma oli tutkijalle tuttu ja sen nähtiin soveltuvan aineiston käsittelyyn.

Kolmannessa vaiheessa Excel-taulukossa oleva litteroitu aineisto redusoitiin eli tiivistettiin ja pelkistettiin. Redusointi tarkoittaa sitä, että aineistosta karsitaan tutkimukselle epäolennainen pois ja näin siitä saadaan helpommin käsiteltävä kokonaisuus (Tuomi ja Sarajärvi, 2018, s. 118–120). Koska tutkimuksessa täytyi ymmärtää haastavia kokonaisuuksia ja tarkoituksena oli selvittää asiantuntijoiden näkemyksiä, valittiin käsiteltäviksi ilmauksiksi kokonaiset puheenvuorot. Puheenvuorot käytiin läpi ja ne tiivistettiin sekä pelkistettiin maksimissaan lauseiden mittaisiksi ilmaisuiksi.

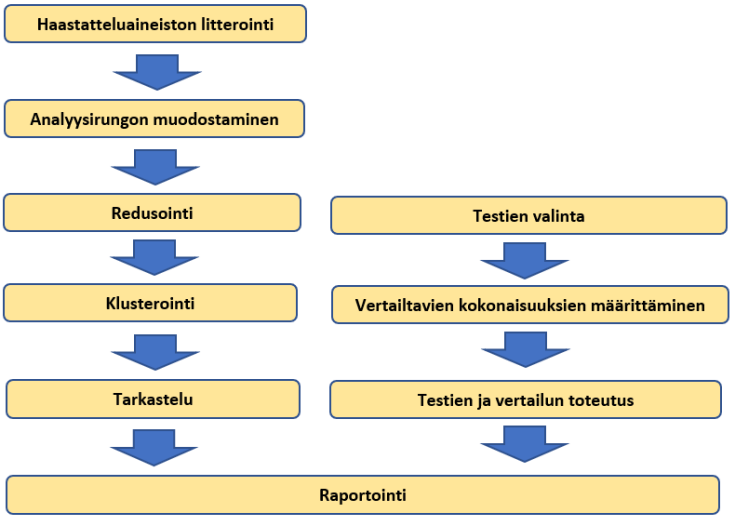
Neljännessä vaiheessa redusoitu aineisto ryhmiteltiin eli klusteroitiin analyysirungon yläluokkien mukaisesti Excel-taulukkoon. Klusteroinnissa eli ryhmittelyssä pelkistetyt ilmaukset ryhmitellään alaluokiksi (Tuomi ja Sarajärvi, 2018, s. 118–120). Ryhmittelyssä käytetyksi analyysiyksiköksi määriteltiin aiemmin redusoidut ilmaukset, jotka olivat tiiviitä ja pelkistettyjä lauseita. Alaluokat pyrittiin muodostamaan mallin mukaisten kokonaisuuksien, alakohtien mukaisesti. Jos tämä ei ollut mahdollista, johdettiin uusia alaluokkia. Esimerkiksi asiantuntijan palomuuriin liittyvä huomio sijoitettiin yläluokkaan ”laitteisto- ja ohjelmistoturvallisuus”, josta johdettiin mallin kokonaisuuden mukainen alaluokka ”palomuuri”. Mikäli vastauksen katsottiin olevan tärkeä, mutta se ei ollut sijoitettavissa mihinkään yläluokkaan, sijoitettiin se kategoriaan ”muut”, josta johdettiin uusia alaluokkia.

Viidennessä vaiheessa taulukkoa sekä analyysirunkoa tarkasteltiin ja arvioitiin kriittisesti. Tämä toteutettiin peilaamalla taulukkoa alkuperäiseen aineistoon. Näin varmistuttiin siitä, että redusoitu materiaali säilytti oikean merkityksensä koko aineiston käsittelyn ajan. Tässä vaiheessa oli myös mahdollisuus korjata ja tarkentaa aineiston redusointia. Samalla pyrittiin varmistumaan siitä, että muodostettu analyysirunko huomioi kaikki tärkeimmät kokonaisuudet. Tämän vaiheen tehtävänä oli antaa palautetta aineiston käsittelyyn liittyen sekä luoda taulukosta entistä selkeämpi ja tiiviimpi kokonaisuus.

Kuudennessa vaiheessa toteutettiin raportointi. Aineistoa tarkasteltiin tutkimuskysymysten ja tutkijan kehittämän mallin näkökulmasta. Tarkoituksena oli löytää aineistosta vastauksia tutkimuskysymyksiin ja tehdä havaintoja sekä tulkintoja vastausten perusteella, siitä miten kattavasti tutkijan kehittämä teoreettinen malli tarjoaa oikein sovellettuna suojaa identiteetille verkossa, teknisestä näkökulmasta. Raportti koostettiin ja ryhmiteltiin tutkimuskysymyksittäin. Näin pyrittiin raportista koostamaan selkeä kokonaisuus, jossa vastaukset eri kokonaisuuksiin olivat helposti eroteltavissa toisistaan.

Konstruktiivisen mallin testaamiseen liittyvän aineiston käsittely muodostui neljästä vaiheesta, jotka olivat 1) testattavien kokonaisuuksien valinta, 2) vertailtavien kokonaisuuksien määrittäminen testeittäin, 3) testien sekä vertailun toteutus ja 4) testien raportointi. Ensimmäisessä vaiheessa valittiin testit, jotka haluttiin toteuttaa. Testeiksi haluttiin valita sellaiset kokonaisuudet, joilla voitiin testata tutkijan mallin jokaista kokonaisuutta minimalistisesti käytännössä. Toisessa vaiheessa testeille määriteltiin vertailtavat kokonaisuudet, joiden avulla voitiin tehdä testeistä havaintoja. Kolmannessa vaiheessa testit toteutettiin käytännössä ja tuloksia vertailtiin sekä luotiin havaintoja. Neljännessä vaiheessa

havainnot raportoitiin lyhyesti ja ne pyrittiin liittämään osaksi tutkijan kehittämää mallia.



KUVIO 7 Aineiston käsittelyn ja analyysin vaiheet

4.3.5 Tietosuoja ja tietoturvallisuus

Tutkimuksen suunnitteluvaiheessa kiinnitettiin erityistä huomiota tietoturvallisuuteen ja tietosuojaan. Tietosuojaan perehdyttiin Jyväskylän yliopiston verkkosivuilla, jotka tarjosivat laajat sekä kattavat kokonaisuudet tietojenkäsittelyyn ja tietosuojaan liittyen. Aiheen arkaluontoisuuden, haastateltavien henkilöllisyyden ja heidän edustamiensa organisaatioiden suojaamiseksi katsottiin parhaaksi, ettei henkilötietoja kerätä. Tämä helpotti myös aineiston käsittelyä. Aineistonkeruusuunnitelma hyväksytettiin myös tutkimuksen ohjaajalla, ennen aineistonkeräämistä.

Tutkimuskysymykset suunniteltiin niin, että haastateltavat kykenivät vastaamaan kysymyksiin yleisellä tasolla, eikä esimerkiksi heidän edustamiensa organisaatioiden käyttämiin yksityiskohtaisiin menetelmiin menty. Tutkimuskysymyksissä huomioitiin myös se, että henkilötietoja tai muita yksilöitävissä olevia tietoja ei kerätty. Haastateltavien henkilöllisyydet ovatkin siis vain tutkijan tiedossa. Aineistossa haastateltavista käytettiin kirjaimia, aakkosjärjestyksen mukaisesti A, B, C ja niin edelleen. Litteroitu materiaali säilytettiin paikallisesti tutkijan työasemalla sekä ulkoisessa muistissa, eikä sitä luovutettu ulkopuolisille.

Rekrytoitavien ja heidän edustamiensa organisaatioiden identiteettiä suojattiin myös rekrytointi vaiheessa. Haastattelukutsut lähetettiin yksityisinä sähköpostiviesteinä henkilöille tai yrityksille. Näin ei yksittäisten henkilöiden tai organisaatioiden ollut mahdollista yrittää päätellä ketkä tutkimukseen olivat mahdollisesti osallistuneet. Rekrytoitaessa haastateltavia, heille lähetettiin ennakkoaineisto, haastattelukysymykset, tiedote tutkimuksesta ja kerrottiin tietojenkäsittelystä sekä tutkimuksesta. Haastattelujen alussa nämä asiat vielä kerrattiin. Näin kyettiin lisäämään tutkimukseen läpinäkyvyyttä ja luomaan rekrytoitaville

selkeä kuva siitä, millaiseen tutkimukseen heitä pyydettiin osallistumaan. Mikäli he suostuivat, sovittiin haastattelulle ajankohta ja heille toimitettiin Zoom-sovellukseen varatun huoneen tiedot, jossa haastattelu sovittuna ajankohtana toteutettiin. Haastateltavat kehuivat haastattelujen järjestelyjä ja kysymysten muotoilua. Heidän mukaansa etähaastattelut ja litteroinnin tarkastaminen haastattelun jälkeen mahdollistivat heille osallistumisen työkiireiden keskellä, sillä jo ennakkoaineistoon perehtyminen vei aikaa. Myös kysymysten muotoilu sai kiitosta, sillä se mahdollisti vastaamisen niin, ettei oman organisaation tekemiin ratkaisuihin tarvinnut ottaa liian yksityiskohtaisesti kantaa.

Tietoturvallisuuden näkökulmasta suurimmaksi riskiksi nähtiin kerätyn aineiston menettäminen. Tämä pyrittiin estämään pitämällä työaseman ohjelmat ajan tasalla, säilyttämällä työn sekä aineiston varmuuskopiot paikallisesti kahdessa eri fyysisessä laitteessa. Työn ja aineiston erillistä salaamista ei katsottu tarpeelliseksi, sillä kerätty aineisto ei sisältänyt arkaluontoisia tietoja ja työn teoriaosuus lähetettiin rekrytointi kutsujen yhteydessä useille ulkopuolisille tahoille. Zoom-sovelluksen huoneelle asetettiin salasana, jossa haastattelut toteutettiin. Näin pyrittiin estämään ulkopuolisten pääseminen kuuntelemaan haastattelua.

5 TUTKIMUSTULOKSET, POHDINTA JA JOHTOPÄÄTÖKSET

Tässä kappaleessa esitellään tutkimuksen kannalta keskeisimmät tulokset. Ensimmäisessä osassa käsitellään asiantuntijoiden näkemyksiä. Asiantuntijoiden haastattelujen tulokset on raportoitu tutkimuskysymysten mukaisesti. Ensimmäisen osan ensimmäisessä alaluvussa käsitellään asiantuntijoiden näkemyksiä passiivisista digitaalisista jalanjäljistä ja keskitytään siihen, mitkä ovat asiantuntijoiden mielestä keskeisimmät passiiviset digitaaliset jalanjäljet, joita käyttäjästä jää verkkoon. Toisessa alaluvussa käsitellään menetelmiä ja tekniikoita passiivisten digitaalisten jalanjälkien peittämiseksi. Kolmannessa alaluvussa käsitellään asiantuntijoiden näkemyksiä tutkijan kehittämästä, teoreettisesta mallista passiivisten digitaalisten jalanjälkien peittämiseksi. Jokaisessa alaluvussa pohditaan esitettyjä tuloksia ja verrataan niitä tutkijan malliin. Kappaleen toisessa osassa käsitellään tutkijan PoC (engl. Proof of Consept) hengessä toteuttamat käytännön testit, joissa tarkastellaan tutkijan malliin kuuluvien kokonaisuuksien toimintaa käytännön toteutuksien avulla, konstruktiivisen tutkimusotteen mukaisesti ja luodaan käyttäjälle käytännön esimerkit aiheesta. Ensimmäisessä alaluvussa vertaillaan ja tarkastellaan palveluntarjoajien käyttäjästä keräämiä tietoja. Toisessa alaluvussa testataan ja tarkastellaan skenaariota, jossa käyttäjän verkkoliikennettä seurataan samassa verkossa. Kolmannessa alaluvussa käsitellään skenaariota, jossa käyttäjän IP-osoitteesta pyritään selvittämään tietoja käyttäjästä. Neljännessä alaluvussa testataan ja tarkastellaan millaisia tietoja selain kykenee käyttäjästä keräämään sekä pyritään luomaan satunnainen sormenjälki selaimelle. Kappaleen viimeisessä osassa esitellään tutkimuksen yhteenveto, pohdinta ja johtopäätökset.

5.1 Asiantuntijoiden näkemykset

Asiantuntijoiden näkemyksiä passiivisista digitaalisista jalanjäljistä, jalanjälkien peittämisestä ja tutkijan kehittämästä mallista selvitettiin teemahaastatteluiden avulla. Haastattelut koostuivat kahdestakymmenestäkahdesta kysymyksestä. Haastatteluilla pyrittiin vastaamaan tutkimuksen tutkimuskysymyksiin, tukemalla tutkimuksessa toteutettua kirjallisuuskatsausta. Kahdessa ensimmäisessä teemassa pyrittiin selvittämään asiantuntijoiden näkemyksiä passiivisista digitaalisista jalanjäljistä ja niiden peittämisestä yleisesti. Kolmannessa teemassa selvitettiin asiantuntijoiden näkemyksiä tutkijan kehittämästä mallista erityisesti käytettävyyden, sovellettavuuden sekä ymmärrettävyyden näkökulmasta.

5.1.1 Passiiviset digitaaliset jalanjäljet

Kysyttäessä haastateltavilta yksityisyyden ja identiteetin suojaamisen näkökulmasta keskeisimpiä passiivisia digitaalisia jalanjälkiä, kaksi kokonaisuutta erottui joukosta. Kaikki vastaajat mainitsivat keskeisimmiksi digitaalisiksi jalanjäljiksi IP-osoitteen ja selaimen keräämät tiedot käyttäjästä. Kolmannes vastaajista korosti myös MAC-osoitteen, käyttäjätilien sekä palveluidentarjoajien tietosuojakäytäntöjen merkitystä.

IP-osoite nähtiin asiantuntijoiden keskuudessa merkittävimmäksi yksittäiseksi passiiviseksi digitaaliseksi jalanjäljeksi, sillä asiantuntijoiden mukaan se näkyy aina verkossa liikennöitäessä ulospäin ja sen peittäminen tai poistaminen kokonaan ei ole mahdollista. Tutkijan mallissa fyysinen IP-osoite sijoittuu fyysisen turvallisuuden kokonaisuuden alle ja IP-osoitteen paljastamia tietoja pyritään peittämään käyttämällä anonyymiä prepaid hotspotia. Merkittävimmiksi fyysisen IP-osoitteen paljastamiksi tiedoiksi nähtiin internet-palveluntarjoaja, geolokaatio eli IP-osoitteen sijainti sekä mahdollisesti käyttäjän edustama organisaatio. IP-osoitteen nähtiin myös liittyvän mallin mukaisen fyysisen turvallisuuden kokonaisuuden lisäksi, mallin laitteisto- ja ohjelmistoturvallisuuteen. Kaksi vastaajista korosti, että mikäli laite kaapataan, on kaappaajan mahdollista selvittää käyttäjän käyttämät langattomat lähiverkot. Tämä on mahdollista, sillä tietokoneet tallentavat usein käyttämänsä langattomat WLAN-verkot, eikä käyttäjä välttämättä muista poistaa käytettyjä verkkoja laitteen muistista. Saatuaan laitteen haltuunsa, voi kaappaaja käyttää www.wigle.net tyyppistä web-palvelua käyttäjän käyttämien verkkojen sijaintien määrittämiseen. Kaksi vastaajista myös korosti, että valittaessa käytettävää WLAN-verkkoa, tulisi perehtyä internet-palveluntarjoajaan. Tämä nähtiin tärkeäksi sen takia, että kaikki internet-palveluntarjoajat eivät tarjoa tai tue geolokaatiota ja tällaisia verkkoyhteyksiä tulisi suosia.

Selaimen keräämät tiedot nähtiin IP-osoitteen jälkeen tärkeimmäksi passiivisten digitaalisten jalanjälkien kokonaisuudeksi. Selaimen keräämistä tiedoista asiantuntijat korostivat IP-osoitteen lisäksi user-agentia, evästeitä, JavaScript- ja Canvas-tekniikkaa sekä käyttöjärjestelmän asetuksia, kuten järjestelmän käyttämää kieltä ja näppäimistö asetuksia. Kaksi asiantuntijaa myös näki IP-osoitteen ja user-agentin olevan keskeisimpiä tietoja, joita eri järjestelmillä, kuten esimerkiksi web-palvelimilla voidaan käyttäjistä lokittaa. User-agentin eli käyttäjäagentin tehtävänä on useimmiten yhdistettäessä verkkosivuun, kertoa verkkosivua ylläpitävälle WEB-palvelimelle HTTP-pyynnön otsikossa tietoja käyttäjän järjestelmästä. Tällaisia tietoja ovat yleensä käytetty selain ja käyttäjän käyttämä käyttöjärjestelmä. Näitä tietoja hyödyntämällä web-palvelin kykenee palvelemaan käyttäjän selainta ja järjestelmää, niiden vaatimusten mukaisesti. Nämä tiedot kuitenkin myös paljastavat tietoja käyttäjän käyttämästä laitteisto- ja ohjelmistokokonaisuudesta. Tutkijan mallissa selaimen keräämiin tietoihin otetaan kantaa selain turvallisuuden kokonaisuudessa. Keskeisimpinä menetelminä turvallisen selaimen valinta, selaimen asetusten määrittäminen ja yksityisyyttä lisäävien selainlaajennosten käyttäminen.

MAC-osoitteen näki tärkeäksi kolmannes haastateltavista. Sen merkityksen nähtiin kuitenkin rajoittuvan lähinnä samassa verkossa toimivien keskuuteen sekä metadataan, johon se eri palveluissa mahdollisesti tallentuu. Tämä johtui siitä, että metadataa lukuun ottamatta, kahden asiantuntijan näkemyksen mukaan MAC-osoite näkyy vain verkon yhdessä segmentissä ja katoaa verkko-operaattorilta nopeasti. He kuitenkin korostivat, että MAC-osoite on laitteen yksilöivä osoite, jonka avulla laite voidaan identifioida. He korostivatkin sitä, että mikäli käyttäjä kantaa useampaa laitetta mukanaan ja yhdistää ne usein esimerkiksi samaan verkkoon, voidaan käyttäjä mahdollisesti tunnistaa liittämällä useamman käytetyn laitteen tunnistetiedot käyttäjään. MAC-osoite on sijoitettu tutkijan mallissa fyysisen turvallisuuden kokonaisuuden alle ja mallissa MAC-osoite muutetaankin, jotta laite ei olisi sen avulla yksilöitävissä. Haastateltavat korostivatkin, että MAC-osoitetta muutettaessa, pelkkä osoitteen muuttaminen ei ratkaise. Heidän mukaansa osoitetta tulisikin muuttaa säännöllisesti, jotta laitteen seuraaminen ja yksilöiminen MAC-osoitteen avulla ei olisi ajan kuluessa edes teoriassa mahdollista.

Kolme asiantuntijaa korosti merkittävänä digitaalisena jalanjälkenä web-palveluihin rekisteröitymisten yhteydessä tarjottuja tietoja, kuten sähköpostiosoitetta, käyttäjätunnusta, kotiosoitetta, salasanaa ja puhelinnumeroa. Asiantuntijoiden mukaan käyttämällä useammissa palveluissa samoja tietoja voidaan käyttäjä identifioida sekä hänen käyttämänsä käyttäjätilit voivat kaikki vaarantua, yhden tilin vaarantuessa. Kolme muuta asiantuntijaa totesivat, että mikäli rekisteröitymisten yhteydessä tarjotut tiedot eivät näy palvelusta ulospäin, eivät nämä tiedot ole digitaalisena jalanjälkenä heidän mielestään yhtä suuri prioriteetti, kuin aiemmin mainitut jalanjäljet. Tiedot näkyvät palveluntarjoajalle, mutta mahdollisen kohteen on tietoja vaikea saada palveluntarjoajalta, mikäli palvelulle ei tapahdu merkittävää tietovuotoa tai palvelu ei sijaitse kohteen kanssa samassa valtiossa. Kolmannes asiantuntijoista korostikin palveluntarjoajien tietosuojakäytäntöihin perehtymistä, palveluita valittaessa. Heidän mukaansa palveluiden välillä on merkittäviä eroja siinä, millaista tietoa käyttäjästä kerätään sekä miten tietoja käsitellään ja tallennetaan. Tutkijan mallissa käyttäjän digitaaliseen identiteettiin otetaankin kantaa identiteetti turvallisuuden kokonaisuudessa, jonka tarkoituksena on luoda käyttäjälle uusi digitaalinen identiteetti ja erottaa se käyttäjän fyysisestä identiteetistä.

Asiantuntijoilta selvitettiin myös heidän näkemyksiään siitä, miten hyvin tutkimuksen teoriaosuudessa digitaaliset jalanjäljet oli käsitelty. Kaikkien yhdeksän haastateltavan näkemyksen mukaan digitaaliset jalanjäljet oli käsitelty tutkimuksessa teknisestä näkökulmasta, tutkimuksen rajaus huomioiden kattavasti ja laadukkaasti. Keskustelua herättivät käyttäjän tekemien ratkaisujen huomiointi eli aktiiviset digitaaliset jalanjäljet ja tutkimuksen päätelaite keskeisyys. Nämä kokonaisuudet liittyivät kuitenkin vahvasti tutkimuksen rajaukseen, jossa aktiivisiin digitaalisiin jalanjälkiin ei haluttu ottaa kantaa ja työ rajattiinkin koskemaan avointen lähteiden tiedustelua suorittavan henkilön päätelaitetta.

Ensimmäisen teeman päätteeksi, asiantuntijoita pyydettiin kertomaan, kuinka he huomioivat arjessaan digitaaliset jalanjäljet toimiessaan verkossa.

Kaikki vastaajat kertoivat kiinnittävänsä huomiota identiteetin ja digitaalisten jalanjälkien peittämiseen sekä suojaamiseen työtehtävissään. He kertoivat erottavansa vapaa-ajan ja työasioiden välisen toiminnan täysin toisistaan. Kysyttäessä digitaalisten jalanjälkien suojaamisesta vapaa-ajalla, oli vastaajien vastauksissa hajontaa. Kaksi asiantuntijoista kertoi, että he ovat vapaa-ajalla laiskoja peittämään digitaalisia jalanjälkiään, koska vastuu on silloin heillä vain itsestään. He eivät myöskään kokeneet tekevänsä vapaa-ajallaan verkossa toimenpiteitä, jotka vaatisivat identiteetin peittämistä tai suojaamista. Kolmannes vastaajista kuvasi itseään myös vapaa-ajalla ”foliohatuksi” ja he kertoivat kiinnittävänsä myös vapaa-ajalla merkittävästi huomiota digitaalisten jalanjälkiensä suojaamiseen. Viimeinen kolmannes kuvasi kiinnittävänsä huomiota digitaalisiin jalanjälkiin valveutuneina käyttäjinä, mutta kertoivat että arjessa korostuu kuitenkin palveluiden käytettävyys ja helppous. Asiantuntijoiden vastausten yleisyyteen perustuen, kiinnitettiin asiantuntijoiden keskuudessa vapaa-ajalla eniten huomiota IP-osoitteen peittämiseen VPN:än avulla. Toiseksi suosituin kokonaisuus, johon kiinnitettiin huomiota, oli selaimen turvallisuus. Kolmanneksi eniten kiinnitettiin huomiota identiteetin turvaamiseen, huomioimalla mitä palveluja käytetään ja mitä tietoja palveluille tarjotaan. Menetelmiä, joita asiantuntijat käyttivät identiteetin turvaamiseen, olivat omien tietojen seuraaminen verkossa, verkkoon laitettavista tiedostoista metadatan poistaminen ennen tiedostojen lataamista palveluun, eri yhteystietojen, salasanojen ja käyttäjätunnusten käyttäminen eri palveluissa sekä sähköpostin välityspalvelut. Yksi haastateltavista totesi avointen lähteiden tiedusteluun tarkoitettujen palveluiden olevan myös erinomaisia työkaluja omien tietojen seuraamiseen verkossa. Hän myös mainitsi, kuinka Googlen hakukoneeseen on mahdollista luoda Google-ilmoitusten avulla toiminto vahtimaan verkkoon julkaistavia tietoja itsestä. Tämä onnistuu luomalla Google-ilmoituksiin halutut hakusanat, jolloin mikäli hakukoneeseen ilmestyy kyseisillä hakusanoilla uutta sisältöä, saa käyttäjä tästä sähköpostiviestin.

Verrattaessa haastateltavien näkemyksiä passiivisista digitaalisista jalanjäljistä, tutkimuksen kirjallisuuskatsaukseen, ei esiin noussut mitään sellaista digitaalisten jalanjälkien kokonaisuutta, jota ei olisi tutkimuksen teoriaosuudessa käsitelty. Kirjallisuuskatsauksessa käsiteltyihin digitaalisiin jalanjälkiin saatiin kuitenkin uusia huomioita yksityisyyden ja suojautumisen näkökulmasta. Tarkasteltaessa haastateltavien vastauksia, voitiin heidän vastauksistaan hahmotella prioriteettijärjestystä keskeisimmille mainituille digitaalisille jalanjäljille. IP-osoite koettiin digitaalisista jalanjäljistä tärkeimmäksi ja toisena tuli selaimen keräämät tiedot, joihin kuuluivat myös käyttöjärjestelmään liittyvät asetukset sekä yleisesti kaikki tiedot, jotka selaimen on mahdollista kerätä käyttäjästä. Kolmantena nähtiin web-palveluiden tileihin liittyvät tiedot ja neljänneksi tärkeimpänä kokonaisuutena verkkokortin MAC-osoite.

Verrattaessa asiantuntijoiden mainitsemia passiivisia digitaalisia jalanjälkiä tutkijan ”Maksimaalisen digitaalisen turvallisuuden malliin”, voidaan todeta, että malli ottaa kantaa jokaiseen asiantuntijoiden mainitsemaan passiiviseen digitaaliseen jalanjälkeen teorian tasolla. Mallissa fyysisen turvallisuuden kokonaisuudessa fyysisen IP-osoitteen paljastamia tietoja pyritään estämään prepaid

hotspotin avulla, jolloin IP-osoitteen paljastuessa, tiedot eivät suoraan johtaisi käyttäjään. Lisäksi fyysinen IP-osoite pyritään peittämään mallin laitteisto- ja ohjelmistokokonaisuuden alla, VPN:än ja Tor:in yhteiskäytöllä. Selaimen tietojen kerääminen ja identiteetin rakentaminen käyttäjästä pyritään estämään laitteisto- ja ohjelmistoturvallisuuden sekä selain turvallisuuden kokonaisuuksien alla. Tämä toteutetaan mallissa kiinnittämällä huomiota käyttöjärjestelmän asetuksiin, käytettävään käyttöjärjestelmään, valitsemalla turvallinen selain, muuttamalla selaimen asetuksia sekä käyttämällä seuraamista estäviä selainlaajennoksia. Käyttäjätilien keräämään dataan otetaan kantaa mallin identiteetti turvallisuuden ja fyysisen turvallisuuden kokonaisuuksien alla, luomalla käyttäjälle uusi digitaalinen identiteetti sekä ottamalla käyttöön laitteet, jotka eivät ole yhdistettävissä käyttäjän oikeaan identiteettiin ja arkisiin toimenpiteisiin. MAC-osoite muutetaan ja peitetään mallissa sekä fyysisen turvallisuuden, että laitteisto- ja ohjelmistoturvallisuuden kokonaisuuksien alla. MAC-osoite muutetaan isäntäkäyttöjärjestelmästä ja sen peittämiseen lisätään myös kerros virtuaalikonetta käyttämällä, jolle on mahdollista määrittää oma MAC-osoite.

5.1.2 Passiivisten digitaalisten jalanjälkien peittäminen

Haastattelujen toisen teeman avulla lähdettiin selvittämään asiantuntijoiden näkemyksiä menetelmistä, joilla aiemmin mainittuja passiivisia digitaalisia jalanjälkiä voidaan peittää ja identiteettiä suojata verkossa. Tutkijan kehittämä malli on kehitetty siitä näkökulmasta, että se tarjoaisi mahdollisimman hyvän anonymiteetin ja suojan identiteetille verkossa, tutkimuksen rajaus huomioiden teknisestä näkökulmasta. Mallin toteuttaminen täydessä laajuudessaan voidaankin nähdä tavalliselle käyttäjälle raskaaksi ja haastavaksi. Haastateltavilta lähdettiin ensin selvittämään, mitkä ovat heidän näkemyksensä mukaan tehokkaimmat yksittäiset menetelmät digitaalisten jalanjälkien peittämiseksi verkossa.

Kaikki haastateltavat listasivat tehokkaimmaksi ja samalla helpoimmaksi yksittäiseksi menetelmäksi VPN:än (engl. Virtual Private Network) eli virtuaalisen erillisverkon käyttämisen. VPN:än eduiksi nähtiin IP-osoitteen peittäminen ja liikenteen salaaminen. Kaksi vastaajista piti VPN:än käyttämistä, jopa itsestäänselvyytenä jokaiselle verkon käyttäjälle. Haastateltavat korostivat kuitenkin, että VPN:än käytössä korostuu palvelun valitseminen ja palveluntarjoajan tietosuojakäytäntöihin tutustuminen. Erityistä huomiota heidän mukaansa tulisi tietosuojakäytännössä kiinnittää siihen, kuinka palveluntarjoaja lokittaa käyttäjän tietoja, kuten fyysistä IP-osoitetta. Kolme haastateltavaa totesi itse luovansa omat VPN-yhteytensä yksityisyyden ja luotettavuuden lisäämiseksi. He kuitenkin totesivat tämän vaativan normaalia enemmän tietoteknistä osaamista ja sen takia palveluntarjoajien tarjoamat VPN-yhteydet soveltuvat paremmin normaaleille käyttäjille.

Haastateltavat korostivat ja muistuttivat kuitenkin fyysisen IP-osoitteen merkityksestä, vaikka käytettäisiinkin VPN:ää. Heidän mukaansa on aina olemassa mahdollisuus, että fyysinen IP-osoite vuotaa, vaikka käytössä olisikin VPN-yhteys ja muistuttivat, että VPN-yhteyden palveluntarjoaja saattaa lokittaa

fyysistä IP-osoitetta. Tämä tarkoittaa sitä, että käyttäjän yksityisyys onkin käytännössä luovutettu VPN-yhteyden ylläpitäjän vastuulle. He näkivätkin tärkeäksi fyysisen IP-osoitteen peittämisen myös VPN-palveluntarjoajalta esimerkiksi vaihtelemalla käytettyä lähiverkkoa säännöllisesti tai hyödyntämällä VPN:ää yhdessä Tor-yhteyden kanssa. Kaikki haastateltavat olivat sitä mieltä, että käyttäjän julkinen IP-osoite tulisi ehdottomasti aina peittää. Kolme asiantuntijoista oli kuitenkin sitä mieltä, että VPN, Tor tai Proxy eivät mikään yksinään ole riittävän tehokas menetelmä IP-osoitteen ja identiteetin peittämiseksi verkossa. He korostivatkin näiden menetelmien yhteiskäyttöä. Yksi haastateltava totesi esimerkiksi, että hänen näkemyksensä mukaan Tor ei automaattisesti ohjaa DNS-kyselyitä Tor-yhteyden läpi ja Tor-verkon solmuissa voi olla mitä tahansa. Hän ei tekisikään mitään sensitiivistä toimintaa, kuten kirjautuisi palveluihin pelkän Tor-verkon avulla. Välityspalvelimien käyttöön liittyen asiantuntijat yhtyivät tutkijan näkemykseen siitä, että ulkopuolisten ja tuntemattomien palvelimien käyttäminen ei ole järkevää. He kuitenkin korostivat proxyja hyvänä menetelmänä, mikäli palvelimet ovat henkilökohtaisessa tai oman organisaation hallinnassa.

Parhaina menetelminä selaimen tietojen keräämiseltä suojautumiseksi pidettiin käyttöjärjestelmän asetusten vaikutusten huomioimista, selaimen valintaa, selaimen asetuksia ja selainlaajennoksia. Haastateltavat suosittelivat käyttäjää perehtymään siihen, millaisia tietoja heidän selaimensa kykenee keräämään heidän käyttöjärjestelmästään sekä työasema kokonaisuudestaan. Esimerkkeinä he mainitsivat käyttöjärjestelmän ja näppäimistön käyttämän kielen näkyvän selaimelle. Yksi asiantuntija mainitsi, että esimerkiksi käyttöjärjestelmän käyttämästä TCP/IP pinosta on mahdollista geneerisesti tunnistaa käytetty käyttöjärjestelmä. Selaimen valinnassa he korostivat selaimen luotettavuutta ja kykyä vaikuttaa selaimen asetuksiin. Selaimen asetusten avulla voidaan vaikuttaa selaimen keräämään dataan. Kaksi asiantuntijaa mainitsi ja suositteli yksityistä Braveselainta kokemattomammille käyttäjille, sillä he totesivat sen olevan suhteellisen yksityinen jo vakioasetuksilla. Kaikki haastateltavat myös mainitsivat selainlaajennosten hyödyntämisen selaimen keräämien tietojen hallinnassa. Tärkeimpinä laajennoksina he pitivät laajennoksia, joilla voidaan vaikuttaa user-agent tietoihin, evästeisiin sekä voidaan estää JavaScript:in käyttäminen. Kaksi asiantuntijoista kertoi, kuinka he toteuttivat selainlaajennosten toteuttamat toiminnot Burp suite-työkalua hyödyntämällä. Haastateltavien mukaan Burp suite-työkalu mahdollistaa whitelist-menetelmän hyödyntämisen, jonka avulla verkkosivujen ja selaimen keräämiä tietoja voidaan suodattaa käyttäjän haluamalla tavalla. Whitelist tarkoittaa menetelmää, jolla luodaan ”säännöt” eli politiikka sille millainen liikenne päästetään läpi ja muu liikenne estetään automaattisesti. Tätä menetelmää käyttävät asiantuntijat korostivat käyttävänsä menetelmää samaan tarkoitukseen kuin mihin aiemmin mainitut selainlaajennokset pyrkivät. He mainitsivat, että Burp suiten etuna on se, että se on varmempi menetelmä kuin selainlaajennokset. He kuitenkin myönsivät kyseisen menetelmän vaativan merkittävästi enemmän osaamista, kuin selainlaajennosten hyödyntäminen.

Muiksi yksittäisiksi ja tehokkaiksi menetelmiksi digitaalisten jalanjälkien ja identiteetin peittämiseksi sekä suojaamiseksi haastateltavat mainitsivat vielä

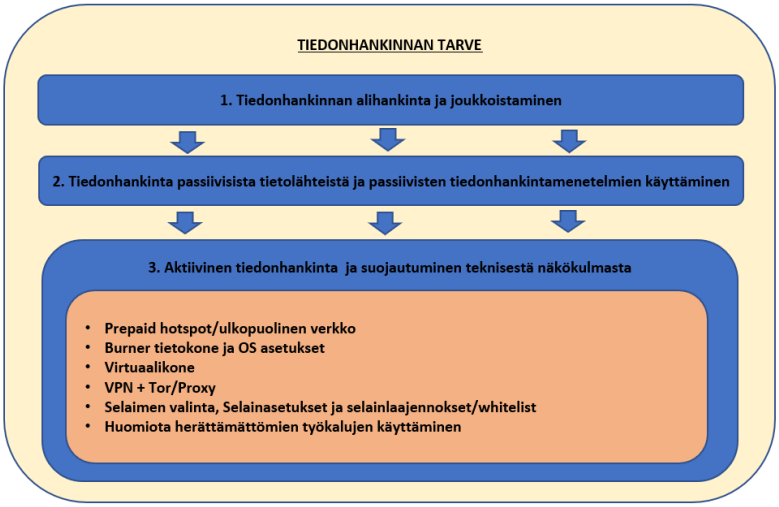
vahvojen salasanojen käyttämisen, ohjelmien ja käyttöjärjestelmän päivittämisen, jälkien pyyhkimisen laitteelta sekä virtualisoinnin, kuten virtuaalikoneiden käyttämisen. Tarkasteltuamme haastateltavien näkemyksiä yksittäisistä, tehokkaista menetelmistä digitaalisten jalanjälkien peittämiseksi, siirryttiin tarkastelemaan kokonaisvaltaisempia ratkaisuja. Tätä selvitettiin kysymällä haastateltavilta millaisilla menetelmillä ja kokonaisuuksilla he peittäisivät digitaalisia jalanjälkiään ja suojautuisivat, mikäli he joutuisivat keräämään tietoja verkossa kyvykkäältä sekä vaaralliselta kohteelta. Vastauksista muodostui kolme merkittävää kokonaisuutta, jotka voidaan nähdä prioriteettijärjestyksenä, jota tulisi tarkastella jokaisella tiedonhankinnan kerralla ja valita kulloinkin parhaiten soveltuva menetelmä.

Turvallisimpana menetelmänä tiedonhankinnalle vaarallisesta kohteesta nähtiin tiedon alihankinta ja joukkoistaminen. Haastateltavat korostivat sitä, kuinka verkko mahdollistaa palveluiden, kuten tiedonhankinnan ostamisen ulkopuoliselta taholta. Asiantuntijoiden mukaan myös joukkoistaminen kehittyy jatkuvasti. Joukkoistaminen (engl. crowdsourcing) tarkoittaa hajautettua ongelmanratkaisumallia, jossa ongelma jaetaan isommalle joukolle, jolloin ongelmaa saadaan ratkaisemaan suurempi joukko ihmisiä. Asiantuntijoiden mukaan he ensisijaisesti käyttäisivätkin tiedonhankintaan jotakin ulkopuolista tahoa tai joukkoistaisivat tiedonhankinnan, jos se vain tiedon sekä tehtävän sensitiivisyys huomioiden olisi mahdollista.

Mikäli tiedonhankintaa ei voida ulkoistaa, käyttäisivät haastateltavat passiivisia tiedonhankinnan menetelmiä ja passiivisia lähteitä. Haastateltavien mukaan verkosta löytyy lukemattomia toimijoita, jotka keräävät tietoa muista verkkosivuista ja lähteistä. Keräämällä tieto näiltä toimijoilta, ei tiedonkerääjän tarvitse olla alkuperäisen lähteen kanssa missään tekemisissä, jolloin kohteelle ei myöskään jää jälkeä tiedonhankinnasta. Esimerkkejä tällaisista palveluista haastateltavien mukaan ovat esimerkiksi Shodan ja Interner Archive. Shodan on palvelu, joka skannaa verkkoon kytkettyjä laitteita ja palvelimia. Se tarjoaakin käyttäjälle hakukoneen, jonka avulla käyttäjä voi etsiä tietoja verkkoon kytketyistä laitteista. Internet Archive on verkkokirjasto, joka tallentaa tietokantaansa verkkosivuja. Palvelusta voikin etsiä esimerkiksi verkkosivujen aiempia versioita.

Vasta kolmantena vaihtoehtona asiantuntijat keskustelisivat suoraan kohdejärjestelmän, kuten web-palvelimen kanssa, mikäli aiemmat menetelmät eivät olisi mahdollisia. Aktiivista tiedonhankintaa varten he kertoivat luovansa kokonaisuuden, jolla voisivat suojautua teknisestä näkökulmasta. Asiantuntijoiden käyttämissä menetelmissä oli havaittavissa pieniä eroja. He kaikki kuitenkin olisivat hyödyntäneet samoja lainalaisuuksia. He olisivat peittäneet fyysisen IP-osoitteensa käyttämällä prepaid hotspotia tai jonkun ulkopuolisen tahon verkkoa. Käyttöön he olisivat määrittäneet burner tietokoneen, vain kyseistä käyttötarkoitusta varten sekä siihen soveltuvan ja kovennetun käyttöjärjestelmän. Koneeseen he olisivat luoneet virtuaalikoneen, jota olisi käytetty ulkoverkossa asiointiin. Yhteyden he olisivat suojanneet VPN, Tor ja Proxy menetelmien yhteiskäytöllä. Selaimen seuraaminen olisi estetty selaimen valinnalla, selainasetuksilla sekä laajennoksilla tai välityspalvelimella, jolla liikennettä olisi suodatettu.

Kaksi asiantuntijaa myös korosti sellaisten työkalujen käyttämistä tiedonhankintaan verkossa, jotka eivät jätä jälkiä. Esimerkiksi mainittiin komentorivipohjainen työkalu Curl. Kolme asiantuntijaa myös pohti sitä, että he saattaisivat rakentaa kokonaisuuden ulkopuoliselle palvelimelle sekä yksi asiantuntija pohti sitä, olisiko näin jäykkään kokonaisuuteen tarvetta, jos käytettäisiin kertakäyttöistä burner tietokonetta, joka hävitettäisiin välittömästi tiedonhankinnan jälkeen. Asiantuntijat tulivat pohdinnassaan kuitenkin siihen johtopäätökseen, että pilvipalveluiden sekä virtuaalikoneiden käyttäminen olisi pidemmän päälle edullisempaa, kuin jatkuva burner tietokoneiden käyttäminen ja korvaaminen.



KUVIO 8 Menetelmät tiedonhankinnan suojaamiseksi verkossa

Toisen teeman päätteeksi selvitettiin, oliko haastateltavien näkemyksen mukaan joitakin keskeisiä menetelmiä digitaalisten jalanjälkien peittämiseksi jäänyt tutkimuksen teoriaosuudessa käsittelemättä. Asiantuntijoiden näkemyksen mukaan mitään merkittäviä kokonaisuuksia ei ollut jäänyt käsittelyn ulkopuolelle. Kaksi asiantuntijaa olisi kaivannut kaksivaiheisen tunnistautumisen ja turvallisten käytäntöjen sekä protokollien käsittelyä työaseman suojautumiseen liittyen. Heidän mielestään esimerkiksi nimipalveluita ja niihin liittyviä turvallisuus näkökulmia, kuten DNS over HTTPS käyttöä sekä DNS nimikyselyiden suodattamista olisi voitu käsitellä tarkemmin. He korostivatkin turvallisten protokollien käyttämistä työasemassa ja antoivat hyvän lisäyksen mallin fyysisen turvallisuuden kokonaisuuteen, mainitsemalla Yubikey avaimet. Yubikey avaimet ovat laitteita, kuten USB-tikkuja, jotka suojaavat pääsyn tietokoneisiin ja verkkoihin, tarjoten kaksivaiheisen tunnistautumisen. Jotta palveluun on mahdollista päästä käsiksi, vaaditaan fyysinen Yubikey avain. Yksi asiantuntija myös korosti käyttäjän työaseman suojaamista sekä lokittamista palomuuria hyödyntämällä. Hänen mukaansa käyttäjän on hyvä myös suojautumisen näkökulmasta lokittaa hänen suuntaansa tulevaa liikennettä, eikä huomioida vain kohteiden lokeja.

Tarkasteltaessa asiantuntijoiden näkemyksiä digitaalisten jalanjälkien peittämisestä ja verratessa niitä Maksimaalisen digitaalisen turvallisuuden malliin, voidaan todeta mallin huomioivan keskeisimmät menetelmät passiivisten digitaalisten jalanjälkien peittämiseksi ja identiteetin suojaamiseksi. Mallissa

fyysinen IP-osoite suojataan käyttämällä prepaid hotspotia ja fyysistä IP-osoitetta peitetään VPN:än ja Tor:in yhteiskäytöllä. Keskeisimmät menetelmät selainten seuraamisen estämiseksi olivat asiantuntijoilla samat, kuin tutkijan mallissa. Keskeisimmät erot mallin ja asiantuntijoiden näkemyksien välillä muodostivat välityspalvelimien ja etäpalvelimien käyttäminen. Tutkija oli rajannut kyseiset menetelmät mallin ulkopuolelle, sillä tutkijan näkemyksen mukaan ne lisäävät malliin kompleksisuutta ja aiheuttavat yksityisyyden sekä anonymiteetin näkökulmasta lisää huomioitavia kokonaisuuksia, kuten palvelimien hankinta. Asiantuntijat näkivät ympäristön rakentamisen etäpalvelimelle sekä välityspalvelimien käyttämisen kuitenkin hyvinä ratkaisuina, mikäli käyttäjän ammattitaito niiden käyttämisen mahdollistaa. He kuitenkin totesivat, ettei tutkijan malli poissulje näiden menetelmien käyttämistä ja niiden käyttäminen sekä soveltaminen mallin pohjalta onnistuu helposti osaavalta toimijalta.

Selvitettäessä asiantuntijoiden näkemyksiä passiivisten digitaalisten jalanjälkien peittämisestä sekä identiteetin suojaamisesta, korostui yksi merkittävä huomio haastattelusta toiseen. Haastateltavat halusivat keskeisenä huomiona tuoda esiin sen, että mikäli käyttäjä tekee itsestään verkossa niin sanotusti ”liian puhtaan”, muodostaa tämä jo itsessään merkittävän anomalian, joka on kohteen havaittavissa. Mikäli muu liikenne ja liikennöitsijät ovat normaaleja ja yksi vaikuttaa liian siistiltä tai puhtaalta, vetää tämä huomiota puoleensa. Eräs haastateltava vertasi tätä turvakieltoon. Turvakielto tarkoittaa tietojenluovutuskieltoa, joka henkilön on mahdollista hakea, jotta hänen tietojaan ei luovuteta ulkopuolisille (Dvv, 2023). Asiantuntijan mukaan usein kuitenkin se, että tiedot eivät ole saatavilla herättää vain lisähuomiota ja kiinnostusta. Haastateltavien mukaan tämä tulisi huomioida rakennettaessa teknistä kokonaisuutta passiivisten digitaalisten jalanjälkien peittämiseksi. Heidän mukaansa olisi ensiarvoisen tärkeää, että verkkoon jäävistä jäljistä saataisiin mahdollisimman geneerisen näköisiä, jolloin ne eivät erottuisi joukosta ja vetäisi huomiota puoleensa. Tutkimuksen rajauksen ulkopuolisina menetelminä, joiden nähtiin liittyvän aktiivisiin digitaalisiin jalanjälkiin, haastateltavat mainitsivat toimintojen volyymin ja jälkien kyllästämisen. Heidän mukaansa esimerkiksi tekemällä verkossa kerran tunnissa toimenpiteen, hukkuu se massaan paremmin, kuin esimerkiksi kymmenen HTTP-pyyntöä minuutissa. Kyllästämisellä haastateltavat tarkoittivat virheellisten jälkien tuottamista, oikeiden jälkien piilottamiseksi massaan. Keskustelua herättivät myös kohteen mahdollisuudet käyttää hunajapurkkia (engl. honeybot), jonka avulla kohde saattaa yrittää hankkia tietoja käyttäjästä. Hunajapurkki saattaa olla esimerkiksi mielenkiintoiselta näyttävä tieto tai kohde, joka on käyttäjän helposti saatavilla. Kuitenkin tietoa tarkastellessa käyttäjästä kyetäänkin kohteen toimesta keräämään tietoa.

5.1.3 Maksimaalisen digitaalisen turvallisuuden malli

Kolmannessa teemassa selvitettiin asiantuntijoiden näkemyksiä Maksimaalisen digitaalisen turvallisuuden mallista. Asiantuntijoilta selvitettiin heidän näkemyksiään mallin toimivuudesta, ymmärrettävyydestä, käytettävyydestä,

käyttötarkoituksesta, sovellettavuudesta sekä kehittämisestä. Mallin tarkastelu aloitettiin toimivuuden ja ymmärrettävyyden näkökulmasta. Haastateltavat totesivat, etteivät olleet havainneet mitään merkittäviä asiavirheitä tutkijan malliin liittyen.

Haastateltavat pitivät mallia loogisena ja selkeänä kokonaisuutena, joka käsitteli sekä otti kantaa kaikkiin tarpeellisiin passiivisiin digitaalisiin jalanjälkiin ylätasolla. Tapa, jolla malli oli jaettu kokonaisuuksiin sekä alaotsikoihin, nähtiin hyväksi. Sen nähtiin olevan yksinkertaisempi tapa jakaa ja jaotella kokonaisuus, kuin esimerkiksi se, että malli olisi suhteutettu OSI-malliin. Tämän nähtiin tukevan paremmin heikomman ATK osaamisen omaavien henkilöiden ymmärrystä kokonaisuudesta. Haastateltavat eivät olisi myöskään muuttaneet mitään kokonaisuuksien alla olevista alaotsikoista. Heidän mukaansa kaikki alaotsikoissa sijaitsevat kokonaisuudet tulee huomioida suojatessa identiteettiä ja digitaalisia jalanjälkiä verkossa. Työaseman suojaaminen aiheutti keskustelua, mutta sen nähtiin kuitenkin liittyvän erottamattomasti identiteetin suojaamiseen toimittaessa verkossa. Yksi haastateltavista totesikin, että jätettäessä yksikin kokonaisuus mallista huomioimatta, menetetään jotakin käyttäjän turvallisuuteen liittyvää. Vaikka identiteetti ei heti paljastuisikaan, saatetaan menettää vaivalla alustettu työasema.

Mallin toimivuuteen liittyen haastateltavilta tiedusteltiin mitä he tekisivät eri tavalla kuin mallissa. Kaikki haastateltavat totesivat, että noudattaisivat samoja teknisiä lainalaisuuksia, kuin mihin mallissakin otetaan kantaa ja mitkä mallissa huomioidaan. Heidän mukaansa he vain kustomoisivat mallin käytännöntoteutuksen omaan käyttötarkoitukseensa sopivaksi. Merkittävimpänä erona kolme haastateltavaa mainitsivat jo aiemmin käsiteltyjen proxy välityspalvelimien käyttämisen ja ympäristön rakentamisen ulkoiselle palvelimelle. Yksi asiantuntija myös estäisi selaimen käyttämiä seuraamistekniikoita Burp suite-työkalun avulla, selainlaajennosten sijaan. Hän myös mainitsi mahdollisten kontti virtualisointi tekniikoiden, kuten Dockerin käyttämisen virtuaalikoneiden tilalla.

Mallin ymmärrettävyydestä keskusteltaessa totesivat kaikki haastateltavat, että malli on heidän taustallaan hyvin ja helposti ymmärrettävissä. Kolme haastateltavaa jopa totesi, että pystyisivät toteuttamaan mallin pohjalta käytännön ratkaisun, vaikka välittömästi. Kaikki haastateltavat näkivät kuitenkin, että mallin ymmärtäminen vaatii ATK taustan tai vähintäänkin ATK perusteiden ymmärtämisen. Puolet haastateltavista jakoi ymmärrettävyyden kahteen eri kohderyhmään. Kohderyhmiksi he nimesivät käyttäjät ja toteuttajat. Heidän näkemyksensä mukaan mallin toteuttaminen käytäntöön vaatii enemmän ymmärrystä ja osaamista, kuin esimerkiksi mallin pohjalta tehdyn käytännön ratkaisun käyttäminen tai kokonaisuuden ymmärtäminen. Mallin ja kokonaisuuden ymmärtäminen nähtiin käyttäjän kannalta kriittiseksi kokonaisuudeksi. Haastateltavat totesivat ensiarvoisen tärkeäksi sen, että käyttäjä on tietoinen mistä kaikesta voi verkkoon jäädä digitaalisia jalanjälkiä ja miten malli tulee toteuttaa käytännössä, jotta jalanjälkiä ei varmasti jätetä toimittaessa verkossa. Kolme haastateltavaa totesikin, että heidän mielestään aihetta ymmärtävälle henkilölle ei passiivisiin digitaalisiin jalanjälkiin liittyvää kokonaisuutta voi tutkimuksen rajaus

huomioiden selkeämmin esittää, sillä kyse on kuitenkin laajasta ja haastavasta kokonaisuudesta. Kaikki haastateltavat korostivatkin mallin kohderyhmän merkitystä ja mallin käyttötarkoitusta. He näkivät käyttäjän tekemät ratkaisut niin kriittiseksi osaksi identiteetin ja digitaalisten jalanjälkien suojaamista, että heidän mukaansa mallia ei kannatakaan kenelle tahansa osaamattomalle tarjota käyttöön.

Mallin käytettävyyttä ja käyttötarkoitusta selvitettiin kysymällä asiantuntijoilta millaisia käyttötarkoituksia he näkevät mallille, onko mallille heidän mielestään tarvetta ja onko malli käyttökelpoinen eli käytettävissä oleva verkossa toteutettavan avointenlähteiden tiedustelun näkökulmasta. Mallin nähtiin soveltuvan parhaiten ylätason viitekehykseksi tai sapluunaksi, siihen mitä asioita tulee huomioida peitettäessä ja suojatessa identiteettiä sekä passiivisia digitaalisia jalanjälkiä verkossa. Yksi haastateltava kuvasi sitä humoristisesti verkon operaatioturvallisuuden ”ruokaympyräksi”. Kolme asiantuntijaa näki, että mallin avulla on todella helppo lähteä hankkimaan aiheesta lisää tietoa ja ymmärtää passiivisten digitaalisten jalanjälkien muodostamaa kokonaisuutta. Kolme haastateltavaa totesi mallin soveltuvan loistavasti koulutuskäyttöön, sillä heidän mukaansa vastaavanlaisia materiaaleja aiheesta löytyy vähän, joissa näin suuria kokonaisuuksia olisi pelkistetty yhdeksi suhteellisen yksinkertaiseksi kokonaisuudeksi. Mallin suurimmiksi käyttötarkoituksiksi nähtiinkin koulutus ja viitekehyksenä toimiminen, jonka pohjalta on mahdollista konseptoida käytännön ratkaisu passiivisten digitaalisten jalanjälkien sekä identiteetin peittämiseksi verkossa. Kolme asiantuntijaa näki myös, että heidän näkemyksensä mukaan jotkin hyökkäyksellistä tietoturvaa toteuttavat yritykset olisivat voineet, jopa maksaa vastaavanlaisen tutkimuksen toteuttamisesta ja mallin kehittämisestä. Heidän mukaansa osalla tietoturvayhtiöistä on omat ratkaisunsa ja mallinsa organisaation ja sen henkilöstön identiteettien sekä digitaalisten jalanjälkien peittämiseksi toimittaessa verkossa. Nämä tiedot, mallit ja menetelmät ovat heidän mukaansa myös usein yrityssalaisuuksia. Mallille nähtiinkin näin olevan käyttöä sekä tarvetta useammasta eri näkökulmasta ja käyttötarkoituksesta.

Asiantuntijoita pyydettiin arvioimaan teoriassa mallin kykyä peittää ja suojata passiiviset digitaaliset jalanjäljet sekä mallin tarjoamaa anonymiteettiä verkossa toteutettavan avointen lähteiden tiedustelun näkökulmasta. Vastauksissaan haastateltavat korostivat käyttäjän ja käytännön toteutuksen merkitystä. Kaikki asiantuntijat totesivat kuitenkin, että teoriassa ja teknisestä näkökulmasta mallia oikein toteuttamalla voidaan passiiviset digitaaliset jalanjäljet ja identiteetti varmasti peittää riittävän tehokkaasti toteutettaessa avointen lähteiden tiedustelua verkossa. Viiden haastateltavan mukaan mallia oikein toteuttamalla kyetään luomaan valtavasti työtä ja haittaa sille taholle, joka pyrkii käyttäjän identiteetin selvittämään. He kuvasivatkin, että mallia käyttämällä ei vastustajalle tarjota ainakaan ”ilmaista lounasta”. Mallin oikean toteuttamisen uskottiin pystyvän peittämään passiiviset digitaaliset jalanjäljet ainakin kaupallisilta toimijoilta. Haastateltavat eivät osanneet ottaa kantaa valtiollisiin toimijoihin. He epäilivät, että valtiolliset tiedustelu toimijat kykenevät selvittämään käyttäjän identiteetin verkossa tarvittaessa ennen pitkään. Kysymykseen tulee vain se,

kuinka paljon resursseja tämä vaatii. Yksi haastateltava totesikin, että usein parhaat tekniikat pyritään pitämään salassa, jolloin tämän hetken parhaita menetelmiä, ratkaisuja tai käytäntöjä digitaalisten jalanjälkien peittämiseksi tai paljastamiseksi eivät massat välttämättä edes tiedä. Asiantuntijat katsoivat mallin kuitenkin tarjoavan hyvän anonymiteetin ja edellytykset käyttäjän tiedonhankinnan suojaamiseksi verkossa teknisestä ja teoreettisestä näkökulmasta, sillä käyttäjän seuraaminen ja identiteetin paljastaminen vaativat mallia oikein toteutettaessa valtavia resursseja. Tärkeänä muistutuksena he pitivät sitä, että tutkimuksessa ei maksuliikennettä ollut täysin käsitelty ja tämä tuleekin tarvittaessa käyttäjän erikseen huomioida.

Mallin soveltaminen ja kehittäminen aiheuttivat haastateltavissa poikkeuksellisen paljon keskustelua. Haastateltavien mukaan malli on nyt kehitetty mahdollisimman turvallisesta näkökulmasta ja on sovellettavissa henkilön toimesta, jolla on korkea ammatti- ja tietotaito. Asiantuntijat kuitenkin näkivät, että näin massiivisiin sekä rankkoihin suojausmenetelmiin ei välttämättä aina ole tarvetta ja myös käytettävyys voi kärsiä osana kontrollien toteuttamista. Asiantuntijat olivatkin kaikki sitä mieltä, että malli tulisi liittää tulevaisuudessa osaksi käyttäjän tekemää riski- ja kohdeanalyysia, joiden pohjalta malli huomioiden kontrollit tulisi valita tapauskohtaisesti digitaalisten jalanjälkien ja identiteetin suojaamiseksi. Lisäksi nähtiin, että kontrollien valintaan olisi hyvä luoda jokin käyttäjää helpottava ratkaisu, jonka avulla kontrolleja kyettäisiin suhteuttamaan sekä valitsemaan oman riski- ja kohdeanalyysin pohjalta. Ratkaisun olisi tärkeä priorisoida mallin mukaisia kohtia käyttäjälle, jolloin käyttäjä ymmärtäisi mitä kokonaisuuksia, missäkin tilanteessa tulisi huomioida ja ylimääräiset kontrollit voitaisiin jättää huomioimatta esimerkiksi käytettävyyden lisäämiseksi sekä ajan säästämiseksi. Kolme asiantuntijaa ehdotti, että tulevaisuudessa mallista kehitettäisiin esimerkiksi kolmiportainen ratkaisu, joka koostuisi kolmesta eri tasoisesta kokonaisuudesta, jotka voitaisiin aina valita riski- ja kohdeanalyysin pohjalta. Tasot menisivät suojauksen kovuuden mukaan esimerkiksi järjestyksessä kevyt-, keski- ja kovataso. Mallin nähtiin myös olevan helposti konseptoitavissa käyttöön. Organisaation käyttöön konseptoitaessa nähtiin, että mallin tulisi jatkossa ottaa kuitenkin enemmän kantaa myös organisaation identiteetin peittämiseen, sillä malli nähtiin tällä hetkellä hyvinkin käyttäjä ja työasema keskeiseksi.

Mallin soveltamisessa ja kehittämisessä nähtiin myös jatkossa tärkeäksi ajan huomioiminen. Aikaan liittyviksi tekijöiksi nähtiin käyttäjätilien kasvattaminen, luotujen ympäristöjen sekä laitteistojen identiteettien elinkaaret sekä resurssien käytön tehokkuus. Avointen lähteiden tiedusteluun käytettävien käyttäjätilien luontiin ja kasvattamiseen nähtiin tarvittavan aikaa, jotta kyetään saamaan uskottavia käyttäjätilejä erityisesti sosiaalisenmedian alustoille ja vältetään näin tiedonhankinnan paljastuminen. Pohdintaa haastateltavissa aiheutti myös se, miten mallilla kyettäisiin ottamaan kantaa siihen, kuinka kauan laitteiden kuten burner tietokoneiden ja mallin mukaisesti luotujen ympäristöjen identiteettejä voidaan pitää ”elossa”, sekä miten ne hävitetään asiaankuuluvalla tavalla. Soveltamiseen liittyen keskustelua herätti myös se, miten kyetään huomioimaan

se taso, jossa digitaaliset jalanjäljet ja identiteetti kyetään suojaamaan tehokkaasti resurssit, kuten esimerkiksi käytettävissä oleva aika huomioiden.

Haastateltavilta kysyttiin lopuksi vielä näkemystä siitä, miten mallia heidän näkemyksensä mukaan kannattaisi testata käytännössä. Haastateltavat näkivät laajojen testien toteuttamisen haasteelliseksi aiheen luonteen ja tutkimuksen rajaus sekä laajuus huomioiden. Parhaiksi ja soveltuvimmiksi käytännön testeiksi asiantuntijat näkivät pienet "demo" testit mallin eri kokonaisuuksista. Asiantuntijoiden näkemysten mukaan testit voisi sitoa tosi elämän esimerkkeihin ja testeillä kyettäisiin myös osoittamaan lukijalle minkä takia joitakin tiettyjä kokonaisuuksia mallissa tulisi huomioida passiivisten digitaalisten jalanjälkien sekä identiteetin näkökulmasta.

Haastattelukysymysten ulkopuolella haastateltavat kehuivat mallia laadukkaasti toteutetuksi. Kolme haastateltavaa totesi, että mallin tarkastelu auttaisi monia perehtymään aiheeseen syvällisemmin, sillä malliin kerätty materiaali on heidän näkemyksensä mukaan hajanaisesti eri kirjallisuudessa sekä vaatisi ilman kyseistä mallia todella laajaan materiaaliin perehtymistä. Haastateltavat totesivat, että malli on heidän mielestään koostettu poikkeuksellisen hyvin, yksinkertaiseksi kokonaisuudeksi valtavasta aihealueesta.

Yhteenvetona mallin nähtiinkin käsittelevän ja ottavan kantaa kaikkiin keskeisimpiin passiivisiin digitaalisiin jalanjälkiin sekä käsittelevän kattavasti tehokkaimmat yleisesti tiedossa olevat menetelmät niiden peittämiseksi. Asiantuntijoiden näkemysten mukaan malli soveltuu käyttötarkoitukseltaan parhaiten koulutuskäyttöön, liittyen operaatioturvallisuuteen tehtäessä avointen lähteiden tiedustelua verkossa sekä malliksi, joka tarjoaa huomioitavat kokonaisuudet peitettäessä passiivisia digitaalisia jalanjälkiä sekä identiteettiä toimittaessa verkossa. Malli tarjoaa teoreettisestä ja teknisestä näkökulmasta riittävät keinot identiteetin suojaamiseksi, toteutettaessa tiedonhankintaa verkon julkisista lähteistä. Suurimmat haasteet malliin liittyen, liittyvät mallin toteuttamiseen käytännössä, joka tutkimuksessa on jätetty käyttäjän omalle vastuulle. Haasteita muodostavat osaamisvaatimukset, riski- ja kohdeanalyysin toteuttaminen sekä kontrollien valitseminen. Mallin hyödyntäminen ja toteuttaminen käytännössä vaativat osaamista ja jo itsessään väärin toteutettu käytännön toteutus mallista voi jättää merkittäviä digitaalisia jalanjälkiä käyttäjästä. Tulevaisuudessa mallin käyttämistä ja soveltamista käytäntöön hyödyntäisivät myös mallia varten kehitetyt työkalut, jotka helpottaisivat riski- ja kohdeanalyysi toteuttamista tapauskohtaisesti sekä ohjaisivat näin sopivien, tapauskohtaisten kontrollien valintaa.

5.2 Mallin käytännön testaaminen esimerkkeinä

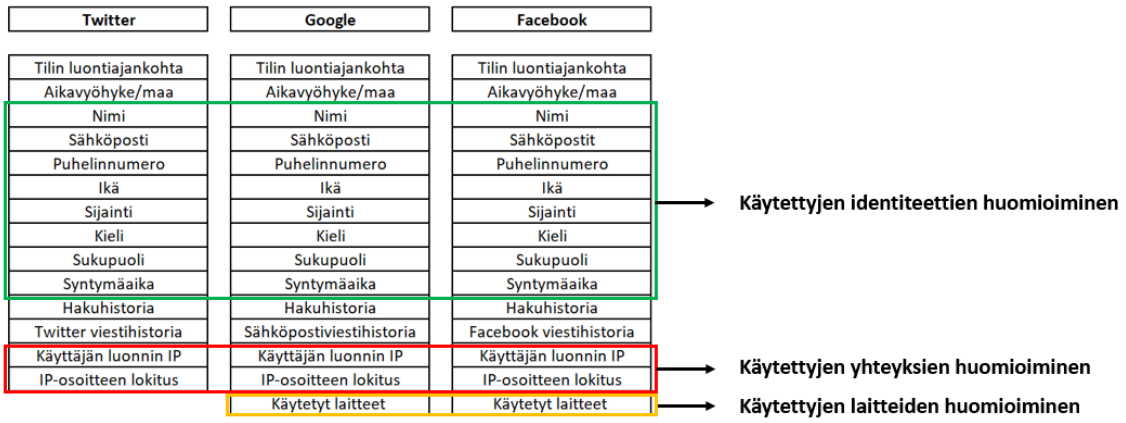
Passiivisia digitaalisia jalanjälkiä ja menetelmiä niiden peittämiseksi testattiin neljän PoC (engl. Proof of Consept) testin avulla. Testien tarkoituksena oli tarkastella asiantuntijoiden mainitsemia keskeisiä passiivisia digitaalisia jalanjälkiä sekä menetelmiä niiden peittämiseksi käytännön tasolla. Testeillä pyrittiin tukemaan kirjallisuuskatsauksessa ja asiantuntijoiden haastatteluissa ilmenneitä

havaintoja. Lisäksi lukijalle haluttiin luoda käytännön esimerkkejä siitä, miksi tietyt passiiviset digitaaliset jalanjäljet tulee huomioida suojatessa identiteettiä verkossa. Ensimmäisessä testissä vertailtiin palveluntarjoajien käyttäjästä keräämiä tietoja. Toisessa testissä tarkasteltiin käyttäjän yksityisyyttä lähiverkon näkökulmasta. Kolmannessa testissä tarkasteltiin IP-osoitteista paljastuvaa dataa sekä neljännessä testissä pyrittiin estämään selainta muodostamasta käyttäjästä yksilöitävissä olevaa sormenjälki kokonaisuutta.

5.2.1 Identiteetti turvallisuus: Käyttäjätilien datan vertailu

Ensimmäisen testin skenaariona oli selvittää millaisia käyttäjätiliin liittyviä tietoja palveluntarjoajat keräävät käyttäjästä ja millaisia tietoja käyttäjästä saattaa paljastua mahdollisten tietovuotojen yhteydessä. Testi toteutettiin pyytämällä kolmelta palveluntarjoajalta käyttäjästä kerätyt tiedot ja vertaamalla tietoja palveluiden välillä. Testin aikana rajattiin tarkasteltavat tiedot kattamaan vain käyttäjätiliin liittyvät tiedot. Tämä johtui siitä, että testiä toteutettaessa paljastui käyttäjästä kerätyn tiedon määrän olevan valtava. Tarkastelussa esitetyt tiedot eivät näin olekaan poissulkevia muiden tietojen keruun osalta. Tiedot pyydettiin Twitteriltä, Googlelta ja Facebookilta. Tietojen pyytäminen on mahdollista aiemmin mainitun GDPR tietosuoja-asetuksen ansiosta.

Ensimmäinen havainto palveluntarjoajien keräämästä datasta oli se, että kaikki palveluntarjoajat lokittivat käyttäjän käyttämiä IP-osoitteita ja palvelut olivatkin tallentaneet muun muassa tilien luonnin yhteydessä käytetyn IP-osoitteen. Toinen merkittävä passiivinen digitaalinen jalanjälki, jota kerättiin Googlen ja Facebookin toimesta oli käyttäjän käyttämät laitteet ja niiden yksilöinti. Kaikki palvelut pyrkivät tallentamaan mahdollisimman paljon yhteystietoja käyttäjästä. Twitterin tiedoista löytyi jopa osio ”personointi”, jossa Twitter pyrki päättelemään käyttäjän ikää sekä kiinnostuksen kohteita, vaikka käyttäjä ei näitä tietoja olisi suoraan palvelulle antanutkaan.



KUVIO 9 Yhteenveto palveluiden käyttäjästä keräämistä tiedoista

Kaikki palvelut tallensivat käyttäjän palveluissaan tekemät haut sekä viestit. Googlen ja Facebookin tiedoista selvisi, että ne pyrkivät tallentamaan tietoja käyttäjästä myös muiden palveluiden avulla. Esimerkiksi Facebook käytti myös Instagramia tietojen keräämiseen. Merkittävimmät havainnot yksityisyyteen liittyen tai tarkemminkin yksityisyyden puuttumiseen liittyen muodostivat Googlen keräämät tiedot. Googlen keräämää dataa tarkastellessa selvisi, että jokainen Googlen verkkoon kytketty tuote, laite tai ohjelma kerää tietoja käyttäjästä. Käyttäjästä kerätty data osoitti, että Google oli tallentanut käyttäjän kaikkiin sen sovelluksiin kirjoitetut tiedot itselleen ja esimerkiksi Chrome-selaimen keräämille tiedoille oli oma kansionsa sekä sähköpostiviestien metadata tallennettiin. Kaksi huolestuttavaa kokonaisuutta liittyivät käyttäjän kontaktien sekä sijainnin tallentamiseen ja keräämiseen. Google oli onnistunut keräämään käyttäjän puhelimeensa tallentamien yhteystietojen puhelinnumerot ja sähköpostiosoitteet. Käyttäjän sijaintia Google pyrki selvittämään jatkuvasti. Sijaintitiedot olikin järjestetty dataan vuosien ja kuukausien mukaan. Sijainnin keräämisen lisäksi Google pyrki selvittämään mistä käyttäjä liikkui ja minne sekä mikä oli käyttäjän liikkumismenetelmä.

KUVIO 10 Googlen keräämää paikkatietoa

Palveluntarjoajien keräämien tietojen tarkasteleminen vahvisti näkemystä siitä, että käyttäjien tulisi yksityisyyden näkökulmasta aina perehtyä palveluiden tietosuojakäytäntöihin ennen palveluiden käyttöönottoa. Tietojen tarkasteleminen osoitti, että palvelut keräävät käyttäjästä merkittävästi enemmän tietoa, kuin mistä käyttäjä itse välttämättä on tietoinen. Tarkasteltujen tulosten perusteella voidaankin todeta, että Google on palveluna huono, käyttäjän yksityisyyden näkökulmasta. Googlen palveluista irtautuminen voi olla tänä päivänä haastavaa, sillä Google alkaa olemaan läsnä lähes kaikkialla. Osa laitteista esimerkiksi saattaa edellyttää Google-tilien sekä -palveluiden käyttöä. Mikäli käyttäjä haluaa

irtautua googlesta, on tarjolla degooglattuja laitteita sekä käyttöjärjestelmiä, kuten esimerkiksi mobiilikäyttöjärjestelmä GrapheneOS. Mikäli käyttäjä haluaa säilyttää mahdollisimman korkean yksityisyyden tai jopa pyrkiä anonymiteettiin palveluita käytettäessä, tulisi käyttäjän kiinnittää huomiota palvelun tietosuojakäytäntöön, käytettyyn verkkoyhteyteen, laitteeseen sekä identiteettiin. Myöskään muiden suurten palveluntarjoajien kuten Microsoftin palveluiden käyttäminen ei ole suositeltavaa yksityisyyden näkökulmasta.

Perlrothin (2021) mukaan FBI sai haltuunsa joukkoampumisen toteuttaneiden terroristien Farookin ja Malikin iphonen vuonna 2015. FBI pyysi Applea avaamaan puhelimen salauksen, mutta Apple kieltäytyi tästä vedoten käyttäjiensä ja brändinsä yksityisyyteen. FBI vei asian oikeuteen, mutta juttu peruttiin. Tämä johtui siitä, että FBI oli saanut pääsyn iphoneen nollapäivähaavoittuvuuden avulla eikä tarvinnut siihen enää Applen apua. Tämä onkin hyvä osoitus siitä miksi itsestä kerättyihin tietoihin tulisikin kiinnittää huomiota. Sillä vaikka palveluntarjoaja käsittelisikin tietoja luotettavasti, voi ulkopuolinen toimija päästä tietoihin käsiksi (Perlroth, 2021, s. 270–277.)

5.2.2 Fyysinen turvallisuus: Liikenteen vertailu samassa verkossa

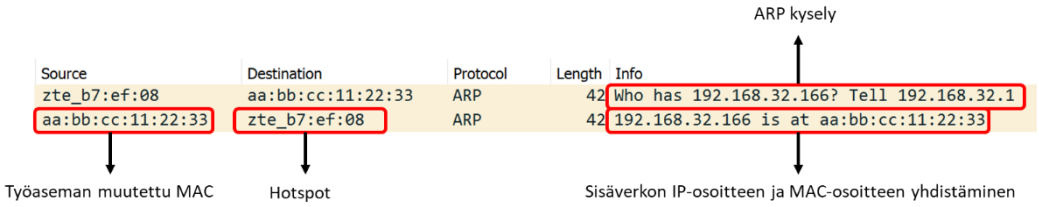
Toisen testin skenaariona oli se, että ulkopuolinen taho on päässyt käyttäjän kanssa samaan verkkoon ja pyrkii seuraamaan käyttäjän liikennettä ja toimintaa verkossa. Testeillä pyrittiin selvittämään verkkoliikenteen luettavuutta ja ARP-protokollaa hyödyntäen yhdistämään käyttäjän MAC-osoite, verkon sisäisen IP-osoitteen kanssa sekä selvittämään, mitä tietoja käyttäjän MAC-osoitteesta on mahdollista selvittää. Testi tehtiin kahdella eri tavalla konfiguroidulla työasemalla ja molemmissa työasemissa yhteys muodostettiin erikseen hankitun prepaid hotspotin avulla, jonka sisäinen IP-osoite oli tutkimuksen toteuttamisen aikana 192.168.32.166.

Toiseen työasemaan ei tehty minkäänlaisia muutoksia, normaaliin työasemaan verrattuna. Ensimmäinen työasema noudatti tutkijan mallia. Työasemassa käytettiin ProtonVpn:än VPN-yhteyttä, joka suojaa liikenteen WireGuard-protokollaa hyödyntämällä. Työaseman MAC-osoite muutettiin Windows rekisterieditorista muotoon ”aa:bb:cc:11:22:33”. Molemmilla työasemilla pyrittiin tarkastelemaan suojattua HTTPS- ja suojaamatonta HTTP-liikennettä. HTTP-liikenteen tarkasteluun ja käyttöön hyödynnettiin verkkosivua http://httpforever.com/, jonka IP-osoite oli tutkimuksen toteuttamisen aikana 146.190.62.39. HTTPS-liikenteen tarkasteluun hyödynnettiin verkkosivua https://example.com/, jonka IP-osoite oli tutkimuksen toteuttamisen aikana 93.184.216.34. MAC-osoitteista pyrittiin etsimään tietoa verkkosivun https://dnschecker.org/ avulla.

Liikenteen luettavuuden kannalta merkittävänä havaintona oli se, että HTTP-liikenne oli luettavissa täysin selväkielisenä. Tarkasteltaessa HTTP-liikennettä samassa verkossa, voitiin kaikki liikenne, kuten esimerkiksi käyttäjän kirjoittamat salasanat lukea täysin selväkielisenä. HTTPS-liikennettä käyttävät verkkosivut salasivat liikenteensä, mutta pakettianalysaattorin avulla oli vielä

mahdollista päätellä mitä protokollaa työasemassa käytettiin liikennöintiin. HTTP- ja HTTPS-protokollat mahdollistivat nähdä vastaanottajan IP-osoitteen, jonka kanssa työasema liikennöi. Testissä käytetty VPN käytti WireGuard-protokollaa liikenteen salaamiseksi. Testien havaintojen perusteella WireGuard esti liikenteen lukemisen selkokielisenä. VPN:än toisena etuna testeissä havaittiin se, että pakettianalysaattorin avulla ei ollut mahdollista selvittää pakettien lopullisen vastaanottajan IP-osoitetta. Tämä johtui siitä, että liikenne kierrätettiin VPN-palvelimen kautta. Havaintojen pohjalta liikenteen salaamiseksi samassa verkossa, parhaaksi ratkaisuksi nähdäänkin VPN:än ja HTTPS-protokollaa käyttävien verkkosivujen yhteiskäyttö.

MAC-osoitteeseen liittyvä havainto testeissä oli, että verkon ARP-pyynnöt näkyivät pakettianalysaattorissa, riippumatta siitä käytettiinkö VPN-yhteyttä vai ei. MAC-osoite oli mahdollista yhdistää sisäverkon IP-osoitteeseen ja MAC-osoitteesta voitiin yrittää hankkia tietoja. Mikäli käytettiin verkkokortin alkuperäistä MAC-osoitetta, kyettiin hakemaan laitteen valmistajan tiedot. Mikäli MAC-osoite oli muutettu, ei MAC-osoitteen avulla tietoja luonnollisesti löytynyt. Testien tulosten perusteella tehtyjen havaintojen mukaan, suurimmat riskit, jotka liittyvät MAC-osoitteeseen muodostuvat mahdollisuudesta yhdistää käyttäjän käyttämä IP-osoite ja MAC-osoite toisiinsa, mikäli laitetta kyetään seuraamaan pidemmän ajanjakson ajan. Toinen riski muodostuu siitä, että alkuperäisen MAC-osoitteen avulla on mahdollista selvittää laitteen valmistaja ja sen kautta mahdollisesti yksilöidä laite. Yksityisyyden kannalta parhaana käytäntönä voidaankin pitää sitä, että MAC-osoite muutetaan säännöllisesti ennen yhteyden muodostamista eri verkkoihin.



KUVIO 11 ARP pyyntö suojatussa työasemassa

KUVIO 12 Hotspotin MAC-osoitteen avulla haetut valmistajan tiedot

MAC-osoitteeseen liittyviä havaintoja tukee myös Kevin Mitnickin (2017) kertomus. Mitnickin mukaan, CIA:n entisen pääjohtaja David Petraeusin ystävät alkoivat saada uhkauksia sähköpostiviesteinä, hänen päätettyään suhteensa rakastajattarensa Paula Broadwellin kanssa. Mitnickin mukaan Broadwell oli lähettänyt uhkaukset sähköpostiviesteinä, käyttämällä eri hotellien julkisia verkkoja ja näin yrittänyt peittää fyysisen IP-osoitteensa piilottaakseen identiteettinsä. IP-osoitteiden geolokaation ansiosta poliisi kykeni kuitenkin keräämään Broadwellin käyttämien hotellien reitittimistä lokitiedot ja yhdistämään Broadwellin uhkauksiin, hänen MAC-osoitteensa avulla. Broadwell ei ollut muuttanut MAC-osoitettaan ja olisikin Mitnickin mukaan voinut estää kiinnijäämisen, muuttamalla MAC-osoitteen uniikiksi jokaisen yhteyden muodostamisen yhteydessä (Mitnick & Vamosi, 2017, s. 117–119.)

5.2.3 Laitteisto- ja ohjelmistoturvallisuus: IP-osoite datan vertailu

Kolmannen testin skenaario oli se, että kohde pyrkii selvittämään käyttäjän sijaintia julkisen IP-osoitteen geolokaation avulla. Testit suoritettiin vertailemalla tutkijan hotspotin fyysistä IP-osoitetta, VPN-yhteyden tarjoamaa IP-osoitetta sekä Tor-verkon tarjoamaa IP-osoitetta keskenään. Tietoja IP-osoitteista selvitettiin verkkosivun https://tools.keycdn.com/geo avulla. VPN-yhteyteen käytettiin VPN-palveluntarjoajan Intian palvelinta.

HOTSPOT		VPN		TOR	
LOCATION		LOCATION		LOCATION	
City	Helsinki	City	New Delhi	City	Sofia
Region	Uusimaa (18)	Region	National Capital Territory of Delhi (DL)	Region	Sofia-Capital (22)
Postal code	00771	Postal code	110001	Postal code	1000
Country	Finland (FI)	Country	India (IN)	Country	Bulgaria (BG)
Continent	Europe (EU)	Continent	Asia (AS)	Continent	Europe (EU)
Coordinates	60.1797 (lat) / 24.9344 (long)	Coordinates	28.6328 (lat) / 77.2204 (long)	Coordinates	42.6951 (lat) / 23.325 (long)
Time	2023-06-05 12:59:59 (Europe/Helsinki)	Time	2023-06-05 14:50:31 (Asia/Kolkata)	Time	2023-06-05 12:17:19 (Europe/Sofia)
NETWORK		NETWORK		NETWORK	
IP address	176.72.65.203	IP address	146.70.142.23	IP address	130.204.161.3
Hostname	mobile-access-b04841-203.dhcp.inet.fi	Hostname	146.70.142.23	Hostname	unknown.interbgc.com
Provider	Telia Finland Oyj	Provider	M247 Europe SRL	Provider	A1 Bulgaria EAD
ASN	1759	ASN	9009	ASN	13124

KUVIO 13 IP-osoitteiden geolokaatio

Havaintojen perusteella IP-osoite paljastaa internet-palveluntarjoajan ja yhteyden maan sekä sijainnin. Vaikka sijainti ei olekaan välttämättä käyttäjän tarkka fyysinen sijainti, voi se antaa käyttäjää seuraavalle taholle käsityksen siitä mistä maasta yhteys on muodostettu. Käyttäjää seuraava taho voi myös pyrkiä hankkimaan lisätietoja palveluntarjoajalta. Testien perusteella oman fyysisen yhteyden käyttäminen onkin yksityisyyden näkökulmasta huono valinta, sillä kyseisen yhteyden IP-osoite voi johtaa käyttäjän identifioimiseen. VPN-yhteyden etuna testeissä havaittiin se, että VPN tarjoaa käyttäjälle mahdollisuuden valita minkä maan palvelinta käytetään. Näin ulospäin näkyykin, että yhteys tulee valitusta maasta. Tor-yhteys mahdollisti myös IP-osoitteen oikean geolokaation peittämisen, mutta Tor ei tarjonnut mahdollisuutta valita uutta sijaintia yhtä joustavasti kuin VPN. VPN osoittautui helpoimmaksi ja joustavimmaksi ratkaisuksi IP-osoitteen geolokaation peittämisessä. Palvelussa voidaan kuitenkin

nähdä merkittävänä tekijänä palveluntarjoajan luotettavuus. Tor-verkon käyttö herättää kysymyksen liikennettä salaavien ja välittävien solmujen luotettavuudesta. Havainnot tukivatkin asiantuntijoiden näkemystä siitä, että paras ratkaisu IP-osoitteen ja geolokaation peittämiseksi onkin useamman kuin yhden menetelmän käyttäminen.

Havaintoa useamman menetelmän yhteiskäytöstä tukee myös tosielämän esimerkki. Shaversin ja Bairin (2016) mukaan FBI on onnistunut ainakin kerran hankkimaan kohteiden aidot IP- ja MAC-osoitteet Tor-selaimen käytöstä huolimatta. Tämän mahdollisti Firefoxin CVE-2013-1690 haavoittuvuus (Shavers & Bair, 2016, s. 26.)

5.2.4 Selain turvallisuus: Selaimen keräämän datan vertailu

Neljännen testin skenaariona oli testata selaimen toteuttamaa seuraamista. Testit toteutettiin vertailemalla työaseman eri konfiguraatioita keskenään. Työaseman konfigurointia säädettiin tulosten ja havaintojen perusteella. Menetelminä seuraamisen estämiseksi olivat selainlaajennokset, käyttöjärjestelmäasetukset, selaimen valinta ja selainasetukset. Selaimen keräämää dataa ja sen perusteella käyttäjästä muodostettavaa sormenjälki (engl. fingerprint) kokonaisuutta tarkasteltiin https://www.amiunique.org/ ja https://coveryourtracks.eff.org/ verkkosivujen avulla. Verkkosivut osoittivat, kuinka uniikki kokonaisuus on mahdollista muodostaa käyttäjän selaimen avulla kerätystä datasta, suhteessa muihin käyttäjiin. Verkkosivu https://www.amiunique.org/ mahdollisti selaimen keräämien tietojen tarkastelun ja tämän lisäksi se eritteli kuinka uniikkeja käyttäjän käyttöjärjestelmän ja selaimen asetukset, sekä kokonaisuudet olivat verrattuna muihin käyttäjiin. Tämä mahdollisti järjestelmän konfiguroinnin niin, että merkittävimmät selaimen keräämät passiiviset digitaaliset jalanjäljet voitiin peittää ja silti kyettiin luomaan selaimen keräämistä tiedoista kokonaisuus, joka ei erotu merkittävästi massasta.

Ensimmäinen havainto testien perusteella oli se, että mitä enemmän selaimen keräämiä tietoja yritettiin peittää tai tietojen keräämistä estää, sitä uniikimpi kokonaisuus oli selaimen avulla käyttäjästä mahdollista muodostaa. Toinen havainto oli se, että mitä enemmän seuraamista estettiin, sitä enemmän myös selaimen ja verkkosivujen käytettävyys kärsi. Nämä havainnot tukivatkin asiantuntijoiden näkemyksiä siitä, että mikäli jälkiä peitetään liikaa, saattaa tämä myös luoda käyttäjästä anomalian verkossa. Tietyissä tilanteissa toivotumpaa voikin olla niin sanotusti piiloutua massaan. Kolmas havainto oli, että eniten selaimen keräämiin tietoihin vaikuttivat käyttäjän käyttöjärjestelmä, laitteisto ja selain kokonaisuudet, asetuksineen.

Testien perusteella tärkeimpiä kokonaisuuksia käyttöjärjestelmässä olivat käytetty käyttöjärjestelmä, käyttöjärjestelmän versio, käyttöjärjestelmän käyttämä kieli sekä aikavyöhyke. Selaimen osalta tärkeimpiä kokonaisuuksia olivat käytetty selain, selaimen versio sekä selaimen yksityisyyteen liittyvät asetukset ja selainlaajennokset. Parhaaseen kokonaisuuteen päästiin jäljittelemällä käyttöjärjestelmän ja selaimen konfigurointi kokonaisuuksia, jotka olivat yleisesti

muidenkin käyttäjien käyttämiä. Paras tulos saavutettiin käyttämällä Windows-käyttöjärjestelmää, Yhdysvaltojen ja Kanadan aikavyöhykettä UTC-05:00, Englannin kieltä käyttöjärjestelmässä sekä Brave-selainta.

Lähelle Brave-selaimen tulosta päästiin myös käyttämällä Firefox-selainta selainlaajennosten sekä yksityiskohtaisten asetusten kanssa. Tulos ei kuitenkaan ollut yhtä geneerinen kuin Brave-selaimen kanssa saavutettu tulos. Tämä johtui siitä, että testien osoittamien havaintojen perusteella, Firefoxin asetukset sekä lisäosat ja laajennokset mahdollistivat selaimelle uniikimman kokonaisuuden muodostamisen käyttäjästä, sillä kyseisten asetusten ja lisäosien käyttäminen ei ollut yleistä muille käyttäjille. Brave-selaimeen ei myöskään tarvinnut itse tehdä muutoksia asetuksiin tai lisätä selainlaajennoksia. Havainnot myös osoittivat, että Brave-selain esti jo itsessään merkittävästi selaimen ja verkkosivujen toteuttamaa seuraamista eikä itsessään vaatinut selainlaajennosten käyttämistä.

Blocking tracking ads?	Yes
Blocking invisible trackers?	Yes
Protecting you from fingerprinting?	your browser has a randomized fingerprint

Asetus	Arvo	Asetuksen suhde muihin käyttäjiin
USER AGENT	Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/114.0.0.0 Safari/537.36	1:97.34
HTTP_ACCEPT HEADERS	text/html, */*; q=0.01 gzip, deflate, br en-GB,en;q=0.6	1:443.33
BROWSER PLUGIN DETAILS	randomized by first party domain	1:6.39
TIME ZONE OFFSET	-240	1:8.14
TIME ZONE	America/New_York	1:10.69
SCREEN SIZE AND COLOR DEPTH	1920x1080x24	1:7.14
SYSTEM FONTS	Arial, Arial Black, Calibri, Cambria, Cambria Math, Comic Sans MS, Consolas, Courier, Courier New, Georgia, Helvetica, Impact, Lucida Console, Lucida Sans Unicode, Microsoft Sans Serif, MS Gothic, MS PGothic, Palatino Linotype, Segoe Print, Segoe Script, Segoe UI, Segoe UI Light, Segoe UI Semibold, Segoe UI Symbol, Tahoma, Times New Roman, Trebuchet MS, Verdana, Wingdings (via javascript)	1:25.98
ARE COOKIES ENABLED?	Yes	1:1.1
LIMITED SUPERCOOKIE TEST	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No, openDatabase: true, indexed db: true	1:2.6
HASH OF CANVAS FINGERPRINT	randomized by first party domain	1:3.65
HASH OF WEBGL FINGERPRINT	randomized by first party domain	1:4.17
WEBGL VENDOR & RENDERER	Google Inc. (NVIDIA)~ANGLE (NVIDIA, NVIDIA GeForce RTX 3070 Laptop GPU Direct3D11 vs_5_0 ps_5_0,D3D11)	1:1607.25
DNT HEADER ENABLED?	FALSE	1:1.81
LANGUAGE	en-US	1:1.74
PLATFORM	Win32	1:2.52
TOUCH SUPPORT	Max touchpoints: 0; TouchEvent supported: false; onTouchStart supported: false	1:1.61
AD BLOCKER USED	no javascript	1:1.0
AUDIOCONTEXT FINGERPRINT	randomized by first party domain	1:4.3
CPU CLASS	N/A	1:1.11
HARDWARE CONCURRENCY	randomized	7.36
DEVICE MEMORY (GB)	4	1:10.77

KUVIO 14 Yhteenveto selaimen sormenjäljen testaamisesta EFF:n testillä

Selaimen muodostamien sormenjälkien testaamiseen käytetyt verkkosivut osoittautuivat testien perusteella erinomaisiksi työkaluiksi määritettäessä oman käyttöjärjestelmän ja selaimen asetuksia sellaisiksi, että selain ei kykenisi uniikkia sormenjälkeä käyttäjästä muodostamaan. Määritettäessä järjestelmän asetuksia ja pyrittäessä estämään selaimen sormenjäljen uniikkius, muodostui internet hyödylliseksi tietolähteeksi. Internetistä löytyi merkittävästi tilastotietoa siitä, mitkä ovat yleisimpiä käyttäjien käyttämiä asetuksia, jotka vaikuttavat selaimen muodostamaan sormenjälkeen. Merkittävä havainto oli myös se, kuinka montaa

erilaista tietoa selain käyttää selaimen sormenjäljen määrittämiseksi. Tietoja voivat selaimen ja käyttöjärjestelmän lisäksi olla esimerkiksi tiedot käytetystä laitteesta sekä sen ominaisuuksista. Näin ei voidakaan havaintojen perusteella antaa geneerisiä asetuksia, joilla kuka vain voisi välttää selaimen muodostaman sormenjäljen uniikkiuden. Tämä johtuu siitä, että jo käyttäjän käyttämän laitteen voidaan nähdä vaikuttavan merkittävästi selaimen muodostamaan sormenjälkeen. Jokainen, joka haluaa tehdä selaimen sormenjäljestä satunnaisen voi kuitenkin käyttää internetistä löytyvää dataa sekä sormenjäljen testaamiseen tarkoitettuja verkkosivuja määrittäessään järjestelmäänsä ja käyttämiään ohjelmia.

5.3 Yhteenveto, pohdinta ja johtopäätökset

Kirjallisuuskatsauksen, asiantuntijoiden näkemysten sekä testien perusteella merkittävimmiksi passiivisiksi digitaalisiksi jalanjäljiksi voidaan nähdä IP-osoite, selaimen ja web-palveluiden keräämät tiedot sekä MAC-osoite. IP-osoite nähtiin merkittävimmäksi jalanjäljeksi, sillä se näkyy liikennöitäessä verkossa ulospäin ja IP-osoitteen poistaminen kokonaan käytöstä ei ole mahdollista. IP-osoitteesta voidaankin selvittää käyttäjän internet-palveluntarjoaja sekä geolokaatio. Selaimen keräämät tiedot kohdistuvat tietoihin käyttäjän laitteistosta, käyttöjärjestelmästä ja selaimesta. web-palvelut kykenevät keräämään käyttäjästä monipuolisesti tietoa tietosuojakäytäntönsä mukaisesti ja käyttäjä ei usein välttämättä olekaan tietoinen kaikista hänestä kerätyistä tiedoista. MAC-osoitteen merkitys rajautuu laitteen yksilöitävyyteen samassa verkossa, sillä osoite on mahdollista yhdistää samassa verkossa sisäisen IP-osoitteen kanssa. MAC-osoitteen avulla voidaan myös selvittää laitteen valmistaja ja pyrkiä tätä kautta hankkimaan lisätietoa käyttäjän järjestelmästä.

Tutkimuksen mukaan yksittäisenä menetelmänä normaalin tietokoneen käyttäjän käytettävyyden sekä vaikuttavuuden näkökulmasta VPN on paras vaihtoehto yksityisyyden ja identiteetin suojaamiseksi verkossa. Hyvä VPN-yhteys kykenee salaamaan käyttäjän liikenteen sekä peittämään käyttäjän fyysisen IP-osoitteen. Käytettäessä palveluntarjoajan VPN-yhteyttä tulee kuitenkin huomioida, että oma yksityisyys luovutetaan palveluntarjoajan vastuulle. VPN:ää käytettäessä tulisikin kiinnittää huomiota käytettyyn fyysiseen IP-osoitteeseen sekä palveluntarjoajalle luovutettuihin tietoihin. Mikäli IP-osoite halutaan peittää varmasti, tulisi käyttää useampaa menetelmää, menetelmistä VPN, Tor ja proxy. Selainten ja web-palveluiden tietojen keräämisessä korostui käyttäjän tietoisuus hänestä kerätyistä tiedoista. Jotta tietoja kyetään selaimelta ja web-palveluilta peittämään, tulee käyttäjällä olla käsitys siitä, mitä tietoja hänestä kerätään. web-palveluissa tieto kerätyistä tiedoista on saatavilla tietosuojakäytännöistä tai pyytämällä omat tiedot palveluntarjoajalta. Selaimen keräämät tiedot voi tarkastaa selaimen asetuksista tai testaamiseen tarkoitetun web-palvelun avulla. MAC-osoitteen muuttaminen onnistuu helpoiten käyttöjärjestelmän asetusten avulla

tai käyttämällä ulkoista verkkoadapteria. Asiantuntijoiden näkemyksien mukaan hankittaessa tietoja verkossa vaaralliselta kohteelta, tulisi tiedonhankintaa lähestyä siitä näkökulmasta, että pyritään itse välttämään teknisten ratkaisujen käyttämistä tiedonhankinnassa tai hankitaan tieto mahdollisuuksien mukaan passiivisia menetelmiä sekä lähteitä hyödyntämällä. Mikäli ei itse käytetä teknisiä järjestelmiä tiedonhankintaan, ei myöskään jätetä digitaalisia jalanjälkiä. Mikäli tiedonhankinta tulisi kuitenkin verkossa teknisesti toteuttaa, huomioisivat asiantuntijat tutkijan mallin mukaiset kokonaisuudet. Erityistä huomiota he kiinnittäisivät käytettyyn yhteyteen, laitteisiin, käyttöjärjestelmään ja ohjelmiin. Asiantuntijoiden näkemykset sekä selaimen sormenjälkeen liittyvä testaus osoittivat, että peitettäessä digitaalisia jalanjälkiä, tulisi kiinnittää huomiota siihen, ettei erotuttaisi muusta massasta.

Tutkimustulokset osoittavat, että tutkijan kehittämä Maksimaalisen digitaalisen turvallisuuden malli ottaa huomioon merkittävimmät passiiviset digitaaliset jalanjäljet sekä menetelmät niiden peittämiseksi, verkossa toteutettavan julkisen tiedonhankinnan näkökulmasta. Asiantuntijat totesivatkin mallin olevan laadukkaasti toteutettu ja muodostuvan loogisista kokonaisuuksista, jotka pitävät sisällään oleellisimmat asiat passiivisten digitaalisten jalanjälkien peittämiseksi. Mallin nähtiinkin soveltuvan koulutuskäyttöön sekä viitekehykseksi tai sapluunaksi, siihen mitä asioita tulee ottaa huomioon suojatessa identiteettiä verkossa teknisestä ja teoreettisesta näkökulmasta. Mallin toteuttamisen motiivina ja tavoitteena olikin luoda teoreettinen malli, joka antaisi kenelle tahansa verkossa avointen lähteiden tiedustelua toteuttavalle taholle ymmärryksen niistä kokonaisuuksista, jotka tulisi huomioida peitettäessä ja suojatessa identiteettiä teknisestä näkökulmasta. Asiantuntijoiden näkemysten mukaan mallin ymmärtäminen sekä soveltaminen vaativat kuitenkin merkittävää ATK osaamista käyttäjältä, jotta käyttäjä kykenee mallin mukaiset asiat ymmärtämään ja soveltamaan mallia käytäntöön. Mallin nähtiin tarjoavan kuitenkin myös kokemattomille käyttäjille sapluunan, jonka avulla voi perehtyä aiheeseen syvällisemmin. Asiantuntijat arvioivat mallin tarjoavan oikein toteutettuna käytännönratkaisuna hyvän ja riittävän anonymiteetin sekä suojan käyttäjän identiteetille tehtäessä avointen lähteiden tiedustelua verkossa, teknisestä näkökulmasta. Mallin sovellettavuuden kannalta nähtiin, että malli tulisi tulevaisuudessa sitoa tapauskohtaisesti riski- ja kohdeanalyysiin sekä näiden perusteella valita mallin mukaiset, tarvittavat kontrollit passiivisten digitaalisten jalanjälkien peittämiseksi. Näin mallin perusteella luotu käytännöntoteutus kyettäisiin suhteuttamaan kohteen vaatimuksiin. Mallin käytännöntoteutuksen ja kontrollien valinnan helpottamiseksi tulevaisuudessa voitaisiin mallista myös kehittää esimerkiksi kolmiportainen ratkaisu, joista voitaisiin valita käytännönratkaisu aina kohteen vaatimusten mukaisesti.

Tutkimuksentulosten perusteella tutkimuksen merkittävimmiksi saavutuksiksi voidaankin nähdä passiivisten digitaalisten jalanjälkien sekä niiden peittämiseen soveltuvien menetelmien laaja käsitteleminen, tutkimuksen avulla kehitetty Maksimaalisen digitaalisen turvallisuuden malli sekä asiantuntijoiden vastauksien avulla turvallisuuden näkökulmasta muodostettu kolmiportainen

lähestymistapa tiedonhankintaan verkosta. Tutkimus rajattiin käsittelemään passiivisia digitaalisia jalanjälkiä. Tutkimuksen toteuttamisen aikana korostui kuitenkin merkittävästi myös aktiivisten digitaalisten jalanjälkien merkitys identiteetin suojaamisen näkökulmasta ja näiden kahden kokonaisuuden erottaminen toisistaan koettiinkin välillä tutkimusta toteutettaessa jopa poikkeuksellisen haastavaksi. Koska käyttäjän tekemät ratkaisut eli aktiiviset digitaaliset jalanjäljet ovat niin keskeisessä roolissa suojatessa identiteettiä verkossa, ei voidakaan käyttäjälle uskottavasti tarjota valmista käytännönratkaisua täydellisen anonymiteetin saavuttamiseksi verkossa. Voidaan vain tarjota teoreettinen malli, jonka avulla käyttäjä voi itse perehtyä yksityisyyteen vaikuttaviin teknisiin kokonaisuuksiin ja pyrkiä huomioimaan ne toiminnassaan. Tutkimusta toteutettaessa korostui myös käsitys siitä, että käyttäjä voidaan ainakin teorian tasolla selvittää ja identifioida verkossa aina, mikäli käytettävissä on vain riittävästi resursseja, kuten aikaa, ammattitaitoa ja rahaa. Tutkimustulosten perusteella voidaan kuitenkin todeta, että tutkimuksen avulla kehitetty Maksimaalisen digitaalisen turvallisuuden malli tarjoaa teoreettisesta ja teknisestä näkökulmasta riittävän määrän esteitä, joiden avulla käyttäjän identiteetin selvittäminen voidaan tehdä kohteelle merkittävän työlääksi ja haastavaksi toteutettaessa avointen lähteiden tiedustelua. Tiedustelun kohdistuessa avoimiin lähteisiin, voidaan mallin myös katsoa nostavan käyttäjän identifioimiseen vaaditut resurssit niin korkeiksi, että kohde päättää luopua käyttäjän identiteetin selvittämisestä, sillä käyttäjän toteuttama tiedustelu kohdistuu jo valmiiksi julkisesti saatavilla oleviin tietoihin.

6 TUTKIMUKSEN ARVIOINTI JA JATKOTUTKIMUS

Tässä kappaleessa tarkastellaan ja arvioidaan tutkimuksen luotettavuuteen ja eettisyyteen liittyviä tekijöitä sekä tutkimustulosten ja tutkimuskysymysten välistä suhdetta. Lisäksi esitetään mahdollisia jatkotutkimusaiheita, joita tutkimuksen tekemisen aikana on ilmennyt. Tutkimuksen ensimmäinen alakysymys oli millaisia passiivisia digitaalisia jalanjälkiä verkkoon jää tehtäessä avointen lähteiden tiedustelua? Kysymykseen vastattiin alaluvussa 5.1.1. Tulosten perusteella voidaan todeta, että ensimmäiseen alakysymykseen kyettiin vastaamaan. Tutkimuksen toinen apukysymys oli, miten passiivisia digitaalisia jalanjälkiä voidaan peittää tehokkaasti verkossa? Alakysymykseen vastattiin alaluvussa 5.1.2 sekä merkittävimpien passiivisten digitaalisten jalanjälkien peittämistä testattiin luvussa 5.2. Tulosten perusteella voidaan todeta, että kysymykseen onnistuttiin vastaamaan. Tutkimuksen päätutkimuskysymys oli, kykeneekö tutkijan kehittämää teoreettista mallia hyödyntämällä peittämään ja suojaamaan käyttäjän identiteetin tehokkaasti teknisestä näkökulmasta, tehtäessä avointen lähteiden tiedustelua verkossa? Kysymykseen vastattiin alaluvussa 5.1.3 ja malliin vaikuttavia kokonaisuuksia havainnollistettiin luvussa 5.2. Tulosten perusteella kysymykseen onnistuttiin vastaamaan monipuolisesti.

6.1.1 Tutkimuksen luotettavuuden tarkastelu

Tuomen ja Sarajärven (2018) mukaan validiteetti ja reliabiliteetti soveltuvat parhaiten luotettavuuden arviointiin määrällisessä tutkimuksessa (Tuomi & Sarajärvi, 2018, s. 157). Eskolan ja Suojärven (1998) mukaan laadullisen tutkimuksen arvioinnissa korostuukin tutkimuksen luotettavuuden arviointi. Tutkimuksen luotettavuutta tarkastellaankin Eskolan ja Suojärven esittämien käsitteiden uskottavuus, siirrettävyys, varmuus ja vahvistuvuus avulla. Lisäksi tarkastellaan tutkijan roolia, tutkimuksen luotettavuuden kannalta (Eskola & Suojärvi, 1998, s. 152–153.)

Eskolan ja Suojärven (1998) mukaan uskottavuus tarkoittaa sitä, vastaavatko tutkijan tekemät käsitteellistykset ja tulkinnat, tutkittavien käsityksiä (Eskola & Suojärvi, 1998, s. 153). Tutkimuksen uskottavuutta sekä tulkintojen tarkkuutta lisäävänä tekijänä voidaan nähdä tutkijan ja haastateltavien samankaltainen tausta sekä perehtyneisyys tutkittavaan aiheeseen. Haastateltavat valikoituivat tutkimukseen aiheeseen perehtyneisyytensä perusteella. Tämän voidaan nähdä vähentäneen virheellisten tulkintojen tekemistä aineiston keräämisen ja käsittelyn aikana. Haastattelut toteutettiin puolistrukturoituina haastatteluina, joiden aikana vastaukset kirjattiin ylös reaaliajassa, jolloin haastateltavien oli mahdollisuus korjata tai tarkentaa vastauksiaan. Haastattelun lopuksi vastaukset myös käytiin läpi ja hyväksytettiin haastateltavilla. Puolistrukturoitu haastattelu antoi myös mahdollisuuden haastattelun aikana kysyä ja tarkentaa sekä kysymyksiä että vastauksia, virheellisten tulkintojen välttämiseksi. Haastattelukysymykset ovat nähtävissä tutkimuksen liitteenä (Liite 1). Kysymykset pyrittiin

rakentamaan niin, että ne eivät johdattelisi haastateltavia ja jättäisivät tilaa myös haastateltavien omille näkemyksille. Uskottavuutta pyrittiin lisäämään myös kuvaamalla tutkimusprosessi ja sen vaiheet yksityiskohtaisesti sekä perehtymällä aiheeseen liittyvään kirjallisuuteen kattavasti. Haastateltaville lähetetty ennakkoaineisto voidaan myös nähdä uskottavuutta lisäävänä, sillä merkittävä osa haastattelusta perustui haastateltaville lähetettyyn ennakkoaineistoon ja sen käsitteistöön. Ennakkoaineiston voidaan nähdä olleen myös uskottavuutta heikentävä tekijä. Osa haastateltavista on saattanut perustaa vastauksiaan myös liikaa ennakkoaineistoon ja aineiston takia kiinnittää huomiotaan myös tutkimuksen kannalta vääriin käsitteisiin, omien huomioidensa sijaan. Uskottavuutta heikentävänä tekijänä voidaan nähdä myös tutkijan tekemät käännökset englanninkielisistä käsitteistä, joille ei kaikille ole välttämättä suoraa suomenkielistä vastinetta. Näiden käsitteiden osalta on mahdollisuus siihen, että käännettäessä käsitteitä on tutkijalla tapahtunut virhe tai haastateltavat ovat ymmärtäneet käsitteen väärin.

Siirrettävyys tarkoittaa Eskolan ja Suojärven (1998) mukaan sitä, kuinka hyvin tutkimustulokset ovat yleistettävissä eli siirrettävissä muihin tutkimuskohteisiin ja tilanteisiin (Eskola & Suojärvi, 1998, s. 153). Tämän tutkimuksen tuloksia ei voidakaan siirtää sellaisenaan koskemaan muita verkkoon kytkettäviä älylaitteita tai verkossa toteutettavia laittomia toimenpiteitä. Erityyppiset laitteet jättävät niille tyypillisiä passiivisia digitaalisia jalanjälkiä ja sisältävätkin näin omia ominaispiirteitään. Koska tutkimuksen rajauksessa on jätetty käsittelemättä maksuliikenteen peittäminen yksityiskohtaisesti ja tutkimus on rajattu käsittelemään avointen lähteiden tiedonhankintaa eivät tulokset ole suoraan sovellettavissa verkossa mahdollisesti toteutettavaan laittomaan toimintaan. Tutkimus ei myöskään käsittele kattavasti aktiivisia digitaalisia jalanjälkiä, jolloin tutkimustulokset eivät ole siirrettävissä sosiaaliseen kontekstiin. Tutkimustulosten siirrettävyyteen ulkomaille voi vaikuttaa myös valtion oma lainsäädäntö, sillä kaikki tutkimuksessa esitetyt menetelmät passiivisten digitaalisten jalanjälkien peittämiseksi eivät esimerkiksi ole laillisia kaikissa valtioissa. Tehtäessä tutkimusta tekniikasta, voidaan myös tutkimuksen tietyllä tavalla nähdä olevan oman aikansa tuote, sillä tekniikat vanhenevat ja uusia syntyy tilalle. Tutkimuksen siirrettävyyttä ajallisesti pyrittiin parantamaan luomalla viitekehys, tarttumatta liikaa käytettyihin ohjelmiin identiteetin peittämiseksi. Tällä pyrittiin lisäämään tutkimuksessa kehitetyn mallin elinikää.

Varmuutta Eskolan ja Suojärven (1998) mukaan on mahdollista lisätä tutkimukseen, ottamalla huomioon tutkimukseen ennustamattomasti vaikuttavia tekijöitä (Eskola & Suojärvi, 1998, s. 153). Tutkimuksen varmuutta pyrittiin lisäämään toteuttamalla laadukas tutkimus- ja aineistonkeruusuunnitelma. Näin kyettiin huomioimaan parhaalla mahdollisella tavalla aineistonkeruuseen sekä tutkimukseen liittyvät ennustamattomat tekijät. Aineisto kerättiin laajan kirjallisuuden, haastattelujen sekä testien avulla. Näin pyrittiin välttämään yllätyksiä sekä lisäämään tutkimukseen varmuutta. Puolistrukturoidut haastattelut antoivat joustavuutta aineiston keräämiselle asiantuntijoilta ja mahdollistivat tarkentavat kysymykset. Haastattelukutsut sekä ennakkomateriaali lähetettiin tahoille,

joiden tiedettiin omaavan tietoa tutkittavasta aiheesta. Ennakkomateriaalina haastateltaville toimitettiin haastattelukysymykset sekä neljäkymmentä sivuinen materiaali luettavaksi haastattelua varten. Näin haastateltavat pystyivät vielä itse pohtimaan halukkuuttaan sekä soveltuvuuttaan tutkimukseen osallistumiseen. Tutkijan ja haastateltavien välille rakennettiin myös luottamusta lupaamalla haastateltaville sekä heidän organisaatioilleen anonymiteetti. Haastateltavien perehtyneisyyttä ennakkoaineistoon pyrittiin lisäämään tarjoamalla heille riittävästi aikaa aineistoon perehtymiseen.

Vahvistuvuus tarkoittaa Eskolan ja Suojärven (1998) mukaan sitä, että tutkimuksessa tehdyt tulkinnat saavat tukea, muista samaa ilmiötä tarkastelleista tutkimuksista (Eskola & Suojärvi, 1998, s. 153). Tutkija ei tutkimusta toteuttaessaan löytänyt täysin samankaltaista aiempaa tutkimusta ja tutkimus tuottaakin tästä näkökulmasta uutta tietoa. Vahvistuvuutta pyrittiin lisäämään perehtymällä kattavasti ulkomaiseen asiantuntija kirjallisuuteen sekä keräämällä aineisto monipuolisesti tutkimusta varten. Tutkimuksessa asiantuntija kirjallisuus, haastattelut sekä tutkijan toteuttamat testit tukivat havaintoina toisiaan. Tiettyjä tutkimukseen liittyviä, yksityiskohtaisia kokonaisuuksia kuten, käyttäjätietojen keräämistä verkossa oli tutkittu. Kyseisiä tutkimuksia hyödynnettiinkin tutkimuksen kirjallisuuskatsauksessa vahvistuvuuden lisäämiseksi.

Eskolan ja Suojärven (1998) mukaan laadullisessa tutkimuksessa merkittävin luotettavuuden kriteeri on tutkija itse ja luotettavuuden arvioinnin tuleekin koskea koko tutkimusprosessia (Eskola & Suojärvi, 1998, s. 152). Tutkijan objektiivisuutta lisäsi se, että hän ei ollut aiemmin työskennellyt tutkittavan aiheen parissa, tutkimusta ei tehty millekään tietylle organisaatiolle ja tutkittava aihe valikoitui tutkijan halusta oppia täysin uutta. Tutkijan riittävästä pätevyydestä ja asiantuntemuksesta pyrittiin varmistumaan laajalla ja monipuolisella kirjallisuuteen perehtymisellä. Suurimmat tutkijan ammattitaitoon liittyvät riskit liittyvätkin mahdollisiin virhetulkintoihin kirjallisuuteen liittyen. Tutkijan tekemien tulkintojen ja havaintojen luotettavuutta pyrittiin lisäämään myös useammalla aineistonkeruumenetelmällä. Luotettavuutta voidaan nähdä lisänneen tutkimuksen rajaaminen tekniseen näkökulmaan, sosiaalisen näkökulman sijaan. Tekniset lainalaisuudet toimivat usein loogisesti, kun taas sosiaaliset ilmiöt voidaan nähdä monisyisempinä ilmiöinä. Tutkija pyrki myös noudattamaan tutkimuksen eettisiä periaatteita. Tutkimusmenetelmä ja prosessi kuvattiin yksityiskohtaisesti. Kaikki haastattelut toteutettiin etänä, jolloin tutkijan olemuksen vaikutukset haastatteluihin voitiin minimoida. Tutkimuksessa käsiteltiin ja arvosteltiin tutkijan kirjallisuuden avulla kehittämää mallia. Tutkija pyrkikin tietoisesti tunnistamaan sekä estämään tähän liittyviä kognitiivisia harhoja ja vinoumia.

Tutkimuksen luotettavuus sekä uskottavuus herättivät myös ajatuksia aineistonkeruun toteutumisen jälkeen. Erityisesti asiantuntijahaastattelut loivat erityisen haasteen luotettavuuden arvioinnille. Tutkijan mallin ymmärrettävyyttä tavanomaisen käyttäjän näkökulmasta olisi uskottavammin voitu testata selvittämällä näkemyksiä normaaleilta käyttäjiltä. Nyt kaikki haastateltavat

olivat käytännössä ammattilaisia ATK taidoiltaan. Merkittävimmät ja haastavimmat kysymykset, jotka herättivät pohdintaa, olivat:

- Kuinka paljon ennakkoon lähetetty aineisto ohjasi vastaajien vastauksia ja näkemyksiä?

- Kuinka voidaan varmistua siitä, että asiantuntijat olivat perehtyneet ennakkoaineistoon?

- Jättivätkö asiantuntijat jotakin oleellista vastaamatta ammattinsa takia?

Edellä mainittuihin kysymyksiin vaikuttaminen ja niiden arviointi on haastavaa. Voidaan kuitenkin nähdä, että kyseisiin tekijöihin on kyetty vaikuttamaan. Ennakkoon lähetetyn aineiston vaikutuksia pyrittiin vähentämään sillä, että haastateltaviksi valittiin ammattilaisia, joilla oli vahva oma kokemus sekä näkemys aiheesta. Asiantuntijoille annettiin heidän tarvitsemansa aika ennakkoaineistoon perehtymiseen ja heidän annettiin esittää, heille sopivaa ajankohtaa haastattelulle. Haastattelujen perusteella voidaan todeta, että aineistoon oli haastateltavien osalta perehdytty. Tutkija näkee, että parhaat tämän hetken tekniikat ja menetelmät anonymiteetin saavuttamiseksi eivät välttämättä ole julkisessa tiedossa. Haastateltavia onnistuttiinkin saamaan monipuolisesti erilaisista taustoista sekä vastauksiin saatiin selkeää saturaatiota. Tällä pyrittiin varmistumaan siitä, että olennaisimmat asiat saataisiin esille haastatteluissa.

6.1.2 Tutkimusetiikan tarkastelu

Tutkimuseettinen neuvottelukunta (2019) jakaa ihmisiin kohdistuvat tutkimuksen eettiset periaatteet seitsemään osa-alueeseen, jotka ovat 1) yleiset eettiset periaatteet, 2) tutkittavan kohtelu ja oikeudet, 3) alaikäinen tutkittavana, 4) vajaakykyinen tutkittavana, 5) henkilötietojen käsittely tutkimuksessa, 6) yksityisyyden suoja tutkimusjulkaisuissa sekä 7) tutkimusaineiston avoimuus (Kohonen, Kuula-Luumi & Spoof, 2019, s. 7–13). Koska tutkimukseen ei osallistunut alaikäisiä tai vajaakykyisiä henkilöitä, jätetään kyseiset kohdat tarkastelun ulkopuolelle.

Yleisiä eettisiä periaatteita on kolme. Ensimmäinen periaate määrittelee, että tutkijan tulee kunnioittaa tutkittavien henkilöiden ihmisarvoa ja itsemääräämisoikeutta. Toinen periaate velvoittaa tutkijaa kunnioittamaan aineellista ja aineetonta kulttuuriperintöä sekä luonnonperintöä ja luonnon monimuotoisuutta. Kolmas periaate vaatii tutkijaa toteuttamaan tutkimuksensa siten, että tutkimuksesta ei aiheudu tutkittavina oleville ihmisille, yhteisöille tai muille tutkimuskohteille merkittäviä riskejä, vahinkoja tai haittoja (Kohonen, Kuula-Luumi & Spoof, 2019, s 7.) Voidaankin todeta, että tutkimuksessa näitä periaatteita on noudatettu onnistuneesti.

Toisessa osa-alueessa määritellään tutkittavan kohtelu ja oikeudet. Tutkittavalla on oikeus osallistua tutkimukseen vapaaehtoisesti mutta myös kieltäytyä osallistumasta. Tutkittava voi myös peruuttaa tai keskeyttää osallistumisensa

milloin tahansa sekä saada tietoa tutkimuksen sisällöstä, henkilötietojen käsittelystä ja tutkimuksen käytännön toteutuksesta (Kohonen, Kuula-Luumi & Spoof, 2019, s. 8–9.) Tutkimuksessa tutkittavia kohdeltiin asianmukaisesti ja heidän oikeutensa esitettiin heille haastattelukutsujen mukana toimitetussa tutkimustiedotteessa, joka löytyy tutkimuksen liitteenä (Liite 2). Lisäksi tutkittaville kerrottiin tutkimuksen käytännön toteutuksesta ja henkilötietojen käsittelystä sekä lähetettiin haastattelukysymykset ennakkoon.

Viides osa-alue määrittelee henkilötietojen käsittelyn tutkimuksessa. Henkilötietojen käsittelyssä tulee noudattaa suunnitelmallisuutta, vastuullisuutta ja lainmukaisuutta. Kuudes osa-alue määrittelee yksityisyyden suojaa tutkimusjulkaisuissa, jonka mukaan tutkimukseen osallistuneiden henkilöiden yksityisyyttä tulee suojella (Kohonen, Kuula-Luumi & Spoof, 2019, s. 11–13.) Haastateltaville ja heidän edustamilleen organisaatioille luvattiin anonymiteetti, eikä haastateltavien henkilötietoja kerätty. Aineistonkeruuta varten laadittiin aineistonkeruusuunnitelma, joka myös hyväksytettiin tutkimuksen ohjaajalla. Lisäksi aineiston käsittelyssä, haastattelukutsuissa ja raportoinnissa kiinnitettiin huomiota siihen, että henkilötietoja ei kirjata ylös. Tutkimuksen kaikissa vaiheissa noudatettiin henkilötietojen käsittelyyn liittyen yliopiston tarjoamia ohjeistuksia sekä lainsäädäntöä. Haastateltavien henkilöllisyydet ovatkin vain tutkijan tiedossa.

Seitsemäs osa-alue määrittelee tutkimusaineiston avoimuuden, joka on edellytys tieteellisen tutkimuksen kriittiselle arvioinnille sekä kehitykselle (Kohonen, Kuula-Luumi & Spoof, 2019, s. 13). Haastattelukutsuja, haastatteluaineistoa ja aineistosta tehtyä litterointia lukuun ottamatta kaikki muu aineisto on määritelty julkiseksi. Näin pyritään varmistumaan siitä, että haastateltavien henkilöllisyydet ja heidän organisaationsa eivät ole pääteltävissä aineiston avulla. Tutkimus toteutettiin julkisena tutkimuksena, johon käytetty kirjallisuus löytyy tutkimuksen lähdeluettelosta. Merkittävä osa tutkimukseen käytetyistä julkaisuista on saatavilla ilmaiseksi verkosta. Tutkimusta varten toteutetut käytännön testit ovat kenen tahansa toistettavissa, tutkimuksen mukaisesti.

Tutkimuseettinen neuvottelukunta (2023) määrittelee myös hyvät tieteelliset käytännöt, joiden perusperiaatteita ovat luotettavuus, rehellisyys, arvostus ja vastuunkanto. Lisäksi määritellään kahdeksan hyvää tieteellistä menettelytapaa, jotka käsittelevät 1) toimintaympäristöä, 2) koulutusta, ohjausta ja mentorointia, 3) tieteellisen työn tekemistä, 4) eettisyyttä ja ennakointia, 5) tutkimusaineistojen käsittelyä ja hallintaa, 6) yhteistyötä, 7) tekijyyttä, julkaisemista ja viestintää sekä 8) asiantuntija- ja arviointitehtäviä (Keiski ym., 2023, s. 11–15.) Tutkimuksessa onkin noudatettu hyvien tieteellisten käytänteiden perusperiaatteita sekä menettelytapoja.

6.1.3 Jatkotutkimus

Tässä tutkimuksessa toteutettu teoreettinen malli toteutettiin mahdollisimman tiukasta yksityisyyden ja suojautumisen näkökulmasta. Mallin soveltamista käytäntöön voitaisiinkin tutkia enemmän. Lisäksi jatkotutkimusaiheita nousi esiin

tutkimuksen toteuttamisen aikana verkossa toteutettavaan avointen lähteiden tiedusteluun sekä yksityisyyteen liittyen.

Tutkimuksessa toteutettua mallia voitaisiin myös tulevaisuudessa kehittää ja tutkia lisää. Tutkimustulosten perusteella malliin voitaisiin kehittää riskienarviointi ja kontrollien valinta työkalut, joiden pohjalta mallia voitaisiin soveltaa käytäntöön. Mallin soveltamista käytäntöön voitaisiin myös tutkia konseptoimalla malli oikean organisaation käyttöön. Mallin toimivuutta voitaisiin myös testata laajemmin, kuin miten tässä tutkimuksessa testaaminen on toteutettu. Tämä voisi olla mahdollista esimerkiksi JyvSecTec:in internet-simulaattorin avulla. Mallia voitaisiin myös testata tiedonhankinnan näkökulmasta. Testauksessa voitaisiin tarkastella käytettävyyden ja suojauksen suhdetta.

Suojautumisen toteutumista verkossa voitaisiin tutkia myös lisää. Tämä onnistuisi esimerkiksi toteuttamalla kokonaisvaltainen tutkimus siitä, miten tällä hetkellä avointen lähteiden tiedustelua suorittavat tahot peittävät identiteettinsä ja digitaaliset jalanjälkensä verkossa. Toinen mielenkiintoinen tutkimus olisi sellainen, jolla mitattaisiin avointen lähteiden tiedustelua verkossa toteuttavien tahojen ATK osaamista erinäisistä näkökulmista.

Yksityisyyteen liittyviä tutkimusaiheita voisivat olla tutkimus liittyen aktiivisiin digitaalisiin jalanjälkiin. Tällainen tutkimus voisi tutkia esimerkiksi sitä, kuinka verkossa toimineet tahot ovat paljastuneet oikeassa elämässä. GDPR mahdollistaa myös tutkimuksen toteuttamisen palveluiden tiedonkeruuseen liittyen. Tutkimuksen avulla olisi mahdollista vertailla palveluiden tietosuojakäytänteitä sekä palveluiden tosiasiallisesti keräämiä tietoja. Tämä olisi mahdollista pyytämällä GDPR:än mukaisesti palveluntarjoajien keräämät tiedot ja vertaamalla niitä palvelun tietosuojakäytänteen lauserakenteisiin. Tämän avulla voitaisiin osoittaa mitä käytännössä mitkäkin sanamuodot ja lauserakenteet tietosuojakäytännössä tosiasiallisesti tarkoittavat.

LÄHTEET

33MAIL. (2023, 3. maaliskuuta). UNLIMITED PRIVATE EMAIL ALIASES. Haettu 3.3.2023 osoitteesta https://33mail.com

Abramson, M. & Aha, D, W. (2013). User Authentication from Web Browsing Behavior. Naval research lab Washington DC. Haettu 4.3.2023 osoitteesta https://apps.dtic.mil/sti/pdfs/ADA599778.pdf

Ahlgren, K. (2019). Behavioral Tracking – Käyttäjätietojen kerääminen ja käyttäjän jälki verkossa. Jyväskylän ammattikorkeakoulu JAMK & JYVSECTEC, 17-29. Haettu 11.5.2023 osoitteesta: https://www.theseus.fi/bitstream/handle/10024/266784/Ahlgren_Katri.pdf?sequence=2&isAllowed=y

Ayed, G, B. (2011). Digital Identity Metadata Scheme. IEEE, 607-608. Haettu 25.02.2023 osoitteesta: https://ieeexplore-ieee-org.ezproxy.jyu.fi/document/5763568/

Bazzell, M. (2016a). Hiding from the Internet: Eliminating Personal Online Information. (Third Edition), 93-94, 156-190, 274, 329, 352-353.

Bazzell, M. (2016b). Personal Digital Security: Protecting yourself from online crime. Library of Congress Cataloging-in-Publication Data, 13-14, 16-18, 42-66, 80-81, 135.

Bazzell, M. (2018). Open source intelligence techniques: Resources for searching and analyzing online information. Library of Congress Cataloging-in-Publication Data: Application submitted. Sixth Edition. IV, 3-8, 89, 382.

Bellingcat & Global Legal Action Network. (2022). Methodology for online open source investigations into incidents taking place in Ukraine since 24 february 2022, 10-13. Haettu 03.03.2023 osoitteesta https://www.glanlaw.org/online-open-source-methodology

Bernardos, C, J., Zuniga, J, C. & O'Hanlon, P. (2015). Wi-Fi Internet connectivity and privacy: hiding your tracks on the wireless Internet. IEEE. Haettu 4.3.2023 osoitteesta https://ieeexplore-ieee-org.ezproxy.jyu.fi/stamp/stamp.jsp?tp=&arnumber=7390443

BleachBit. (2023, 3. maaliskuuta). Clean Your System and Free Disk Space. Haettu 3.3.2023 osoitteesta https://www.bleachbit.org

Bleedingcomputer. (2020, 29. heinäkuuta). Microsoft now detects CCleaner as a Potentially Unwanted Application. Haettu 3.3.2023 osoitteesta https://www.bleepingcomputer.com/news/microsoft/microsoft-now-detect.

Brave. (2023). Brave Browser Privacy Policy. Haettu 11.5.2023 osoitteesta: https://brave.com/privacy/browser/

CCleaner. (2023, 02. maaliskuuta). Speed up and optimize your pc with CCleaner. Haettu 02.03.2023 osoitteesta https://www.ccleaner.com

Christensson, P. (2014, 26. Toukokuuta). Digital footprint definition. Haettu 25.02.2023 osoitteesta https://techterms.com/definition/digital_footprint

Dvv. (2023, 20. Heinäkuuta). Turvakielto. Haettu 20.07.2023 osoitteesta https://dvv.fi/turvakielto

Eskola, J. & Suoranta, J. (1998). Johdatus laadulliseen tutkimukseen. Tampere: Vastapaino, 63, 152-153. Haettu 25.02.2023 osoitteesta https://www.ellibslibrary.com/book/978-951-768-035-6

Fan, X, X., Chow, K, P. & Xu, F. (2014). Web user profiling based on browsing behavior analysis. Berlin: Springer. Haettu 4.3.2023 osoitteesta https://link.springer.com/content/pdf/10.1007/978-3-662-44952-3_5.pdf

Flow, S. (2021). How to hack like a ghost: Breaching the cloud. San Francisco: No starch press, 4-6.

Hassan, N, A. & Hijazi, R. (2018). Open Source Intelligence Methods and Tools: A Practical Guide to Online Intelligence. Library of Congress, 2, 22, 31, 33-38, 46-47, 52-56, 59, 65-67, 86-87, 138, 140.

Hickey, M. & Arcuri, J. (2020). Hands on hacking: Become an expert at next gen penetration testing and purple teaming. Indiana: John Wiley & Sons, 29-31.

Human Rights Center & United Nations Human Rights. (2022). Berkeley Protocol on Digital Open Source Investigations, 31-34. Haettu 25.02.2023 osoitteesta https://humanrights.berkeley.edu/resources/berkeley-protocol-digital-open-source-investigations/

Huuskonen, V. (2020). Palvelimen suojaaminen DDoS-hyökkäyksiltä. Jyväskylä: Jyväskylän yliopisto. Haettu 27.02.2023 osoitteesta https://jyx.jyu.fi/bitstream/handle/123456789/70023/1/URN%3ANBN%3Afi%3Ajyu-202006174240.pdf

Iltalehti. (2022a, 10. lokakuuta). Sota-asiantuntija Emil Kastehelmi Iltalehden erikoistoimittajaksi. Haettu 25.02.2023 osoitteesta https://www.iltalehti.fi/ulkomaat/a/6e4ab757-d061-49e4-ad61-7e8e93520988

Iltalehti. (2022b, 28. lokakuuta). Läpimurto Vastaamo-vyyhdissä – poliisi kertoo, miten siinä onnistuttiin. Haettu 27.02.2023 osoitteesta https://www.iltalehti.fi/digiuutiset/a/90ce5a8e-3375-4dd5-8cc4-d5a2ef64a092

Iltasanomat. (2021, 4. helmikuuta). Veli-Pekka Kivimäki auttoi Bellingcatin läpimurtoon MH17:ssä – ja sai osansa trolleista ja hakkereista. Haettu 25.02.2023 osoitteesta https://www.is.fi/digitoday/art-2000007780858.html

Ironvest. (2023, 3. maaliskuuta). Secure your accounts. Haettu 3.3.2023 osoitteesta https://ironvest.com

iStorage. (2023, 3. maaliskuuta). Compact design: Ultimate security. Haettu 3.3.2023 osoitteesta https://istorage-uk.com/product/datashur/

Jamal, N. & Zain, J, M. (2022). A Review on Nature, Cybercrime and Best Practices of Digital Footprints. IEEE. Haettu 04.03.2023 osoitteesta https://ieeexplore-ieee-org.ezproxy.jyu.fi/stamp/stamp.jsp?tp=&arnumber=9995834

Juan, W., Shimin, C., Jun, Z., Bin, H. & Lei, S. Identification of Tor anonymous network traffic based on machine learning. IEEE. Haettu 4.3.2023 osoitteesta https://ieeexplore-ieee-org.ezproxy.jyu.fi/document/9674056

Kali Linux. (2023, 3. maaliskuuta). The most advanced Penetration testing Distribution. Haettu 3.3.2023 osoitteesta https://www.kali.org

KeePassXC. (2023, 02. maaliskuuta). KeePassXC – Cross-Platform Password Manager. Haettu 02.03.2023 osoitteesta https://keepassxc.org

Keiski, R., Hämäläinen, K., Karhunen, M., Löfström, E., Näreaho, S., Varantola, K., Spoof, S-K., Tarkiainen, T., Kaila, E. & Aittasalo, M. (2023). Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa. Tutkimuseettinen neuvottelukunta, 11-15. Haettu 12.6.2023 osoitteesta: https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf

Kmille. (2022). PoC: Implementing evil maid attack on encrypted /boot. Chaos Computer Club. Haettu 11.5.2023 osoitteesta: https://media.ccc.de/v/gpn20-32-poc-implementing-evil-maid-attack-on-encrypted-boot#t=116

Knowledge Base. (2014, 22. lokakuuta). Welcome to the Mozilla Knowledge Base. Haettu 3.3.2023 osoitteesta https://kb.mozillazine.org/Knowledge_Base

Kohonen, I., Kuula-Luumi, A. & Spoof, S-K. (2019). Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakkoarviointi Suomessa. Tutkimuseettinen neuvottelukunta, 7-13. Haettu 12.6.2023 osoitteesta: https://tenk.fi/sites/default/files/2021-01/Ihmistieteiden_eettisen_ennakkoarvioinnin_ohje_2020.pdf

Li, V. (2022). Bug Bounty Bootcamp: The Guide to Finding and Reporting Web Vulnerabilities. San Francisco: No Starch Press, 61-62.

Lukka. (2001). Konstruktiivinen tutkimusote. Metodix. Haettu 12.5.2023 osoitteesta: https://metodix.fi/2014/05/19/lukka-konstruktiivinen-tutkimusote/

Malwarebytes. (2023, 02. maaliskuuta). FIX TODAY. PROTECT FOREVER. Haettu 02.03.2023 osoitteesta https://www.malwarebytes.com

Microsoft. (2023, 02. maaliskuuta). Malware has met its match. Haettu 02.03.2023 osoitteesta https://microsoft.com/en-us/windows/comprehensive-security

Mitnick, K, D. & Vamosi, R. (2017). The Art of Invisibility: The World's Most Famous Hacker Teaches You How to Be Safe in the Age of Big Brother and Big Data. New York: Little, Brown and Company, 21, 32-48, 64, 78-80, 101-102, 114-121, 181.

Mozilla. (2023, 3. maaliskuuta). Pysyvästi nopea. Haettu 3.3.2023 osoitteesta https://www.mozilla.org

Mozilla. (2023, 3. maaliskuuta). Random User-Agent. Haettu 3.3.2023 osoitteesta https://addons.mozilla.org/en-US/firefox/addon/random_user_agent/

Mozilla. (2023, 3. maaliskuuta). Smart HTTPS. Haettu 3.3.2023 osoitteesta https://addons.mozilla.org/en-US/firefox/addon/smart-https-revived/

OpenPGP. (2023, 02. maaliskuuta). OpenPGP: Email encryption. For all operating systems. Standing the test of time. Haettu 02.03.2023 osoitteesta https://www.openpgp.org

PatchMyPC. (2023, 02. maaliskuuta). Simplify third-party patching on your PC. Haettu 02.03.2023 osoitteesta https://patchmypc.com/home-updater

Pearson. (2022, 5. helmikuuta). Basic Data Transmission in Networks: Mac Tables and ARP tables. Haettu 27.02.2023 osoitteesta https://www.pearsonitcertification.com/articles/article.aspx?p=2339639&seqNum=3

Perlroth, N. (2021). This is how they tell me the world ends: The cyber weapons arms race. Boolmsbury publishing, 270-277.

Privacy Badger. (2023, 3. maaliskuuta). Privacy Badger. Haettu 3.3.2023 osoitteesta https://privacybadger.org

ProtonVPN. (2023, 02. maaliskuuta). Proton VPN Service Privacy Policy. Haettu 02.03.2023 osoitteesta https://protonvpn.com

Puolustusvoimat. (2019). Sotilaan käsikirja 2020. Helsinki, 48-49. Haettu 11.5.2023 osoitteesta:
https://puolustusvoimat.fi/documents/1948673/2258487/Sotilaan+k%C3%A4sikirja+2020/50d5f534-adfd-8f14-340b-9a340fb5b6b6/Sotilaan+k%C3%A4sikirja+2020.pdf

Qubes. (2023, 28. helmikuuta). What is Qubes OS? Haettu 28.02.2023 osoitteesta https://www.qubes-os.org/intro/

Ranakoti, P., Yadav, S., Apurva, A., Tomer, S. & Roy, N, R. (2017). Deep web & online anonymity. IEEE. Haettu 4.3.2023 osoitteesta https://ieeexplore-ieee-org.ezproxy.jyu.fi/document/8284479

Salminen, A. (2011). Mikä kirjallisuuskatsaus? Vaasa: Vaasan yliopiston julkaisuja, 7-8.

Shavers, B & Bair, J. (2016). Hiding behind the keyboard: Uncovering covert communication methods with forensic analysis. Cambridge: Syngress is an imprint of Elsevier, vi, 11-19, 26, 29-35, 115-124, 163-168, 170, 187-188.

Statista. (2023, 24. Helmikuuta). Number of internet and social media users worldwide as of january. Haettu 25.02.2023 osoitteesta https://www.statista.com/statistics/617136/digital-population-worldwide/

Tails. (2023, 3. maaliskuuta). Tails: is a portable operating system that protects against surveillance and censorship. Haettu 3.3.2023 osoitteesta https://tails.boum.org

Tietosuoja. (2023). Usein kysyttyä EU:n tietosuoja-asetuksesta. Haettu 16.5.2023 osoitteesta: https://tietosuoja.fi/gdpr

Tor. (2023, 3. maaliskuuta). Browse Privately: Explore Freely. Haettu 3.3.2023 osoitteesta https://www.torproject.com

Tuomi, J & Sarajärvi, A. (2018). Laadullinen tutkimus ja sisällönanalyysi. Helsinki: Kustannusosakeyhtiö Tammi, 19-23, 82, 84-85, 100, 104-109, 118-120, 157. Haettu 25.02.2023 osoitteesta https://www.ellibslibrary.com/book/9789520400118 s.21, 69-70

Tutanota. (2023, 3. maaliskuuta). Best email service: end-to-end encryption and no ads. Haettu 3.3.2023 osoitteesta https://www.tutanota.com

uBlock Origin. (2023, 3. maaliskuuta). uBlock Origin: Free, open-source ad content blocker. Haettu 3.3.2023 osoitteesta https://ublockorigin.com

VeraCrypt. (2023, 02. maaliskuuta). VeraCrypt. Haettu 02.03.2023 osoitteesta https://veracrypt.fr

VirtualBox. (2023, 3. maaliskuuta). VirtualBox: Welcome to VirtualBox.org. Haettu 3.3.2023 osoitteesta https://www.virtualbox.org

Virustotal. (2023, 02. maaliskuuta). How it works. Haettu 02.03.2023 osoitteesta https://support.virustotal.com/hc/en-us/articles/115002126889-How-it-works

Vladij, N., Madani, P. & Nguyen, E. (2017). Anonymity of TOR Users Demystified. IEEE, 111. Haettu 4.3.2023 osoitteesta https://ieeexplore-ieee-org.ezproxy.jyu.fi/stamp/stamp.jsp?tp=&arnumber=8560770

Whonix. (2023, 28. helmikuuta). Superior Internet Privacy. Haettu 28.02.2023 osoitteesta https://www.whonix.org

Yle. (2022, 26. joulukuuta). Venäjä etsintäkuulutti Bellingcatin toimittajan. Haettu 25.02.2023 osoitteesta https://yle.fi/a/74-20008814/64-3-129941

ZeroTier. (2023). Securely Connect Any Device, Anywhere. Haettu 25.02.2023 osoitteesta https://www.zerotier.com/

LIITE 1 HAASTATTELUKYSYMYKSET

Teema 1: Digitaaliset jalanjäljet

1. Mitkä ovat keskeisimmät digitaaliset jalanjäljet, jotka tulisi huomioida suojatessa yksityisyyttä ja identiteettiä verkossa?
2. Miten kuvailisitte digitaalisten jalanjälkien käsittelyä tutkimuksessa?
3. Millaisia digitaalisia jalanjälkiä tutkimuksessa ei mielestänne ole riittävästi käsitelty?
4. Huomioitteko te digitaaliset jalanjäljet toimiessanne verkossa?

Teema 2: Digitaalisten jalanjälkien peittäminen

1. Minkä digitaalisten jalanjälkien peittämisen koette tärkeimmäksi peitettäessä ja suojattaessa identiteettiä verkossa?
2. Mitkä ovat tehokkaimmat menetelmät digitaalisten jalanjälkien peittämiseksi verkossa?
3. Millaisia keskeisiä menetelmiä digitaalisten jalanjälkien peittämiseksi ei mielestänne tutkimuksessa ole käsitelty?
4. Minkälaisilla ratkaisuilla itse peittäisitte digitaalisia jalanjälkiänne kerättäessä tietoa verkossa potentiaalisesti vaarallisesta lähteestä/kohteesta?

Teema 3: Maksimaalisen digitaalisen turvallisuuden malli

1. Miten arvioisitte mallin ymmärrettävyyttä?
2. Onko malli jaettu mielestänne loogisiin kokonaisuuksiin?
3. Millaisia muita kokonaisuuksia mallissa tulisi huomioida?
4. Onko mallin kokonaisuuksien sisällä mielestänne jotakin ylimääräistä?
5. Puuttuuko mallin kokonaisuuksien sisältä mielestänne jotakin, mitä ei ole otettu huomioon?
6. Miten arvioisitte mallin käytettävyyttä?
7. Miten arvioisitte mallin sovellettavuutta?
8. Mitä tekisitte itse vaihtoehtoisella tavalla?
9. Kykeneekö mallin mukaisella tavalla peittämään digitaaliset jalanjäljet riittävän tehokkaasti verkon avointen lähteiden tiedustelussa?
10. Miten arvioisitte mallin teoriassa tarjoamaa anonymiteettiä ja suojaa verkossa?

Teema 4: Vapaasana/keskustelu

1. Oletteko huomanneet merkittäviä asiavirheitä tutkimuksessa?
2. Mitä asioita tutkija ei ole osannut mielestänne ottaa tutkimuksessa huomioon rajaus huomioiden?
3. Onko teillä joitakin muita keskeisiä huomioita tutkimukseen liittyen?
4. Miten testaisitte mallia käytännössä?

LIITE 2 TIEDOTE TUTKIMUKSESTA

JYVÄSKYLÄN YLIOPISTO

INFORMAATIOTEKNOLOGIAN TIEDEKUNTA

26.03.2023

TIEDOTE TUTKIMUKSESTA

Tutkimuksen nimi ja rekisterinpitäjä

Identiteetin peittäminen ja suojaaminen tehtäessä avointen lähteiden tiedustelua verkossa

Pyyntö osallistua tutkimukseen

Sinua pyydetään mukaan tutkimukseen, jossa pyritään selvittämään millaisia digitaalisia jalanjälkiä jää suoritettaessa avointen lähteiden tiedustelua verkossa ja kuinka näitä digitaalisia jalanjälkiä voidaan peittää. Lisäksi tarkoituksena on selvittää, soveltuuko tutkijan kehittämä malli identiteetin peittämiseen verkossa. Tämä tiedote kuvaa tutkimusta ja siihen osallistumista. Liitteessä on kerrottu henkilötietojen käsittelystä.

Tutkimukseen pyydetään yhteensä n. 5–10 haastateltavaa, jotka ovat ammattinsa toimesta perehtyneet tietoverkkoihin ja digitaalisiin jälkiin.

Vapaaehtoisuus

Tähän tutkimukseen osallistuminen on vapaaehtoista. Voit kieltäytyä osallistumasta tutkimukseen tai-keskeyttää osallistumisen milloin tahansa tutkimuksen aikana.

Tutkimuksen kulku

Kukin haastateltava haastatellaan yhden (1) kerran yksilöhaastatteluna. Ennen haastattelua haastateltavalle on lähetetty tutkimuksen rajaus, teoria ja malli kappaleet ennakkoon luettaviksi sekä haastattelun kysymykset ja tiedote tutkimuksesta. Ennen haastattelua haastateltavien on myös mahdollista pyytää lisätietoja tutkimukseen liittyen sekä perehtyä tietosuojailmoitukseen. Haastattelun jälkeen haastateltaville tarjotaan mahdollisuus tutustua litteroituun haastattelumateriaaliin ja esittää siihen tarvittaessa muutosehdotuksia.

Tutkimuksesta mahdollisesti aiheutuvat haitat ja epämukavuudet

Tutkimuksesta ei aiheudu haastateltaville kuluja.

Tutkimuksen kustannukset

Tutkimukseen osallistumisesta ei makseta palkkiota

Tutkimustuloksista tiedottaminen ja tutkimustulokset

Tutkittaville ilmoitetaan, mistä valmiiseen työhön pääsee tutustumaan. Tutkimuksessa valmistuu tutkimuksen tekijän Pro Gradu-tutkielma.

Tutkittavien vakuutusturva

Jyväskylän yliopiston toiminta ja tutkittavat on vakuutettu.

Jyväskylän yliopiston vakuutukset korvaavat etänä suoritettavissa tutkimuksissa ainoastaan sellaiset vahingot, jotka liittyvät suoraan annettuun tutkimustehtävään ja jotka ovat sattuneet varsinaisen ohjeistetun tutkimustehtävän aikana. Vakuutus ei korvaa taukojen aikana sattuneita vahinkoja.

Jyväskylän yliopiston vakuutukset eivät ole voimassa etänä suoritettavissa tutkimuksissa, jos tutkittavan kotikunta ei ole Suomessa.

Vakuutus sisältää potilasvakuutuksen, toiminnanvastuuvakuutuksen ja vapaaehtoisen tapaturmavakuutuksen. Tutkimuksissa tutkittavat (koehenkilöt) on vakuutettu tutkimuksen ajan ulkoisen syyn aiheuttamien tapaturmien, vahinkojen ja vammojen varalta. Tapaturmavakuutus on voimassa mittauksissa ja niihin välittömästi liittyvillä matkoilla.