# Web-analytiikka verkkoapteekkien sivuilla ja henkilötietojen välittyminen kolmansille osapuolille

Turun yliopisto
Tietotekniikan laitos
TkK-tutkielma
Tietotekniikka
Toukokuu 2025
Robin Carlsson

TURUN YLIOPISTO
Tietotekniikan laitos

ROBIN CARLSSON: Web-analytiikka verkkoapteekkien sivuilla ja henkilötietojen välittyminen kolmansille osapuolille

TkK-tutkielma, 27 s.
Tietotekniikka
Toukokuu 2025

---

Tämä tutkielma käsittelee kolmannen osapuolen analytiikkaa suomalaisten verkkoapteekkien sivuilla. Tutkielmassa selvitetään, mitä mahdollisia asiakkaan henkilö- ja terveystietoja analytiikassa kerätään ja miten asiakkaan suostumusvalinnat vaikuttavat kerättyyn tietoon. Verkkoapteekien yksityisyys on aiheena sekä mielenkiintoinen että yhteiskunnallisesti merkittävä. Tutkielmaa motivoi myös tekijän oma mielenkiinto tietoturvaan, tietosuojaan ja yksityisyyteen yleisemmin.
Tutkielmassa määritellään yksityisyyteen, analytiikkaan ja suostumukseen liittyviä käsitteitä, ja analytiikan toimintaa selvitetään 24 suomalaisen verkkoapteekkin koekäytöllä. Kolmannen osapuolen analytiikassa kerättyjä tietoja tutkitaan nauhoittamalla koekäytön aikaista verkkoliikennettä. Jotta suostumusvalintojen vaikutusta tiedonkeruuseen voidaan verrata, kokeet toistetaan antamalla jokaiselle verkkoapteekille sekä suurin että pienin mahdollinen suostumus evästeisiin ja tiedonkeruuseen.
Eri verkkoapteekkien välillä on suuria eroja apteekkien käyttämässä analytiikassa, käyttäjältä kerätyissä tiedoissa, käyttäjältä kysytyssä suostumuksessa sekä suostumuksen vaikutuksessa kerättyyn tietoon. Kuudesosa verkkoapteekeista lähettää kolmannen osapuolen analytiikkapalveluntarjoajan palvelimille tietoja, joilla voidaan yksilöidä käyttäjä ja yhdistää tämä käyttäjän ostamaan reseptilääkkeeseen.

Asiasanat: verkkoapteekki, analytiikka, yksityisyys, suostumus, terveystieto

# Sisällys

# 1 Johdanto

## 1.1 Tutkielman tarkoitus

Verkkokauppa on jo pitkään ollut osa nykyihmisen arkea, ja yhä useampia tuotteita ja palveluita saa tilattua kätevästi verkosta. Samalla kasvaa myös se joukko ihmisiä, joiden henkilötietoja käsittelevät säännöllisesti sekä verkkokaupat että kauppojen hyödyntämät analytiikkapalvelut. Vaatteiden, pikaruoan ja vessapaperin lisäksi verkosta voi ostaa myös reseptilääkkeitä, joita myyvät suomalaisten apteekkien lukuisat verkkopalvelut.

Yhdysvalloissa tehdyn tutkimuksen mukaan myös verkkoapteekit hyödyntävät analytiikkaa runsaasti [1], ja koronaviruspandemian myötä on ajankohtaista tutkia tilannetta Suomessa. Tekijän kiinnostus tietosuojaan ja kyberturvallisuuteen motivoi aiheen valinnassa, mutta mahdollisilla tietovuodoilla olisi myös yhteiskunnallista merkitystä. Esimerkiksi liikuntarajoitteiselle asiakkaalle verkkoapteekki voi olla välttämättömyys, ja tieto arkaluontoiseen sairauteen liittyvästä lääkeostoksesta voi asettaa asiakkaan erityisen haavoittuvaan asemaan. Psykoterapiakeskus Vastaamon tietomurto johti uhrien omaisten mukaan pahimillaan itsemurhiin [2]. Siksi tutkielmassa selvitetään, mitä kolmansien osapuolien analytiikkaa suomalaisten apteekkien verkkopalvelut käyttävät ja mitä asiakkaan terveys- ja henkilötietoja näille analytiikkapalveluille välittyy.

Koska sivustojen rakenne vaihtelee ja eri käyttäjien selailutottumuksissa on eroja, tut-

kielmassa tarkastellaan myös eri suostumusvalintojen vaikutusta analytiikassa kulkeviin tietoihin. Tutkielman luonteen ja rajoitusten vuoksi tätä yksinkertaistetaan niin, että verkkoapteekeille annetaan sekä suurin että pienin mahdollinen suostumus evästeisiin ja markkinointiin. Verkkoapteekkien koekäyttö ja verkkoliikenteen analyysi tehdään Google Chrome -selaimella ja selaimen DevTools-kehittäjätyökaluilla.

## 1.2 Tutkimuskysymykset

Suomalaisten apteekkien verkkopalveluja etsittiin, seulottiin ja testattiin järjestelmällisesti, jotta saataisiin vastaus kahteen pääkysymykseen:

1. Mitä terveys- tai muita henkilötietoja välittyy kolmannen osapuolen palveluntarjoajille verkkoapteekkien analytiikassa?
2. Miten ääripäiden suostumusvalinnat vaikuttavat kolmannen osapuolen palveluntarjoajille lähetettäviin tietoihin?

Lisäksi huomiota kiinnitettiin siihen, mitä analytiikkapalveluja verkkoapteekeissa käytettiin ja mitä analytiikassa kulkevien tietojen maantieteellisestä päämäärästä saatiin selville. Tutkielma on tehty kahden suomalaisia verkkoapteekkeja käsittelevän tutkimuksen ohella: ”How not to design an online pharmacy: A case study” [3] (Rauti ym. 2024) ja ”Several Online Pharmacies Leak Sensitive Health Data to Third Parties” [4] (Carlsson ym. 2024). Mahdollisista tietovuodoista ja yksityisyyteen liittyvistä ongelmista on kerrottu tietosuojavaltuutetun toimistolle, ja suurin osa ongelmista on jo korjattu.

## 1.3 Tiedonhaku

Tutkielmaa varten etsittiin jo olemassa olevaa tutkimustietoa, joka liittyy joko suoraan verkkoapteekkien hyödyntämään analytiikkaan tai johonkin aiheen osa-alueeseen. Aiemmin julkaistuja tutkimuksia etsittiin internetistä eri kustatantajien tietokannoista, ja tiedonhaussa käytettiin mukana myös Google Scholar -hakupalvelua, jonka avulla tietty hakulauseke voidaan kohdistaa useaan hakukoneen indeksöimään tietokantaan samanaikaisesti. Tietoa etsittiin sekä suomeksi että englanniksi.

Hakulausekkeina käytettiin sekä yksittäisiä sanoja että monimutkaisempia, moniosaisia hakulausekkeita, ja hauissa hyödynnettiin myös loogisia operaattoreita, sanojen katkaisua ja fraasihakua. Esimerkiksi etsiessä verkkoapteekkeja koskevia englanninkielisiä tutkimuksia käytettiin hakulauseketta *”web | internet | digital | online | E-” pharmac**. Näin haku kattaa sekä *apteekki*-sanan eri taivutukset että *verkko*-sanan eri synonyymit. Google Scholarin lisäksi hakuja tehtiin Web of Science Core Collection-, IEEE Xplore- ja ACM digital library -tietokantoihin.

Aiheeseen liittyvien tutkimusten luotettavuutta arvioitiin tarkastelemalla julkaisujen JUFO-luokitusta, tutkimusten käyttämiä lähteitä ja sitä, onko tutkimukset vertaisarvioitu. Tutkielmaan valituista lähteistä suurin osa on englanniksi, sillä suomenkielisillä hauilla löytyi lähinnä opinnäytetöitä, joita ei otettu mukaan. Tietosuojavaltuutetun toimiston suomenkielisiä julkaisuja käytetään kuitenkin lähteinä.

## 1.4 Tutkielman rakenne

Jotta tutkimuksessa käytetyt termit ja koeasetelmaa koskevat valinnat olisivat ymmärrettäviä, luvussa 2 käydään läpi henkilötietojen määrittelyä sekä aiempaa evästeitä ja analytiik-

kaa koskevaa tutkimusta. Luvussa 2 esitellään myös evästeistä riippumattomien yksilöintimenetelmien hyödyntämiä tietoja, pimeitä käytäntöjä suunnittelussa sekä eri käyttäjien asenteita yksityisyyteen ja tietosuojaan.

Luvussa 3 esitellään suomalaisten verkkoapteekkien järjestelmällinen haku ja valikointi hyödyntäen *web scraping* -ohjelmointia. Lisäksi luvussa 3 kuvaillaan myös apteekkien testauksen kulku, mukaan lukien suostumusvalinnat, sivupolku, verkkoliikenteen analysointi ja tulosten kirjaus.

Tuloksia käydään läpi luvussa 4, missä käsitellään myös kokeiden aikana ilmenneitä ongelmia ja haasteita. Tuloksia esitellään yleisesti taulukoilla ja kaavioilla sekä myös tiettyjen yksittäistapausten kautta. Lopulta esitellään vielä johtopäätöksiä luvussa 5, ja samalla pohditaan myös mahdollisia parannuksia tutkimusasetelmaan.

# 2 Henkilötiedot, analytiikka ja suostumus

## 2.1 Terveys- ja muut henkilötiedot

Tietosuojavaltuutetun toimiston mukaan henkilötietoja ovat Suomessa kaikki ne tiedot, ”—joiden perusteella henkilö voidaan tunnistaa suoraan tai välillisesti esimerkiksi yhdistämällä yksittäinen tieto johonkin toiseen tietoon, joka mahdollistaa tunnistamisen.” Muun muassa jo käyttäjän IP-osoite lasketaan henkilötiedoksi. Lisäksi mainitaan, että esimerkiksi ”—terveyttä koskevien tietojen käsittely on lähtökohtaisesti kiellettyä.” [5][6]

Jos henkilötiedot anonymisoidaan, jolloin henkilön tunnistaminen niistä on peruuttamattomasti estetty, niitä ei enää katsota henkilötiedoiksi. Anonymisoituihin tietoihin ei siten sovelleta tietosuojasäännöksiä. [7] Anonymisoinnin on toki oltava riittävän tehokasta. Pseudonymisoidut henkilötiedot – jotka lisätiedoilla voidaan yhdistää jälleen henkilöön – ovat yhä henkilötietoja, ja ne kuuluvat yhä tietosuojasäännösten piiriin. [7]

## 2.2 Web-analytiikka

Verkkosivujen analytiikka ja seuranta (*web analytics, tracking*) on laaja ja mutkikas käsite, jonka piiriin kuuluva teknologia on ajan myötä kehittynyt ja yleistynyt. Analytiikkaa voidaan toteuttaa esimerkiksi verkkosivuille lisätyllä koodilla (kuva 2.1), jota suoritetaan käyttäjän laitteella. [8]

```
<!-- Start MTM -->
<link type="text/css" href="/WebRoot/StoreTypes/7.103
<script type="text/javascript">function _mtmSetConser
<!-- Matomo Tag Manager -->
<script>var _mtm=window._mtm=window._mtm||[];_mtm.pus
<!-- End Matomo Tag Manager -->
<!-- End MTM -->
```

Kuva 2.1: Esimerkki Matomon analytiikkakoodista, joka on liitetty sivun lähdekoodiin.

Analytiikka tarjoaa sivuston omistajille ja ylläpitäjille mahdollisuuden seurata käyttäjien toimintaa sivuilla jopa reaaliajassa, esimerkiksi millä sivuilla käyttäjä vierailee ja mistä käyttäjä on sinne tullut. Käyttäjien liikkeitä voidaan visualisoida muun muassa lämpökartoilla (*heat map*). [9][8]

Analytiikkaratkaisut ovat monesti kuitenkin kolmannen osapuolen tarjoamia palveluja, ja suosituimpien palvelujen takana ovat suuryritykset kuten Google ja Meta (entinen Facebook). Analytiikka voikin kytkeytyä esimerkiksi sosiaalisessa mediassa toteuttuun mainoskampanjaan. Kolmannen osapuolen evästeillä käyttäjää voi seurata sivustolta toiselle. [9][8][10]

### 2.2.1 Analytiikkapalvelujen yleistyminen

Zheutlin ja muut (2022) löysivät tutkimistaan verkkoapteekeista keskimäärin neljä kolmannen osapuolen evästettä tai muuta seurantamekanismia apteekkia kohden. Yli kolme neljäsosaa verkkoapteekeista käytti jonkinlaista seurantaa. Yleisiä olivat varsinkin

Googlen ja Facebookin seurantapalvelut. [1] Sen lisäksi että kolmannen osapuolen analytiikka on yleistynyt roimasti, suosituimmat analytiikkapalvelut kattavat nyt entistä suuremman osan analytiikkaa hyödyntävistä sivustoista. Wambachin ja Bräunlichin (2017) tekemä tutkimus osoitti, että vuonna 2005 sen ajan kolme suosituinta seurantamekanismia kattoivat vain 10 % tutkituista sivuista, ja Googlen analytiikka yhteensä vain 5 %. Vuonna 2015 kolme suosituinta mekanismia kattoivat 73 % sivuista, ja Googlen osuus oli nyt jo 82 %. [9]

### 2.2.2 Analytiikkapalvelujen keräämät henkilötiedot

Vuonna 2015 julkaistussa tutkimuksessa Liu ja muut tutkivat henkilötietojen (esimerkiksi nimen, sukupuolen, kaupungin ja puhelinnumeron) esiintymistä kolmentoista tuhannen ihmisen tavanomaisessa verkkoliikenteessä. Keskivertoa enemmän henkilötietoja keräsivät muun muassa isot, mainospalveluja tarjoavat yritykset kuten Google ja Yahoo. Esimerkiksi Googlen analytiikassaan käyttämä osoite *google-analytics.com* vastaanotti keskimäärin neljää eri henkilötietotyyppiä käyttäjää kohti. [11]

Saman Googlen osoitteen oltiin jo aiemmin todettu vastaanottavan esimerkiksi käyttäjän terveyteen, uskontoon ja seksuaalisuuteen liittyviä tietoja, sekä käyttäjän sijainti kaupungin ja postinumeron tarkkuudella. [12]

Koska esimerkiksi hakukenttään kirjoitettu hakutermi voi näkyä osana tulossivun osoitetta – joka puolestaan voi välittyä analytiikkapalvelulle muun muassa HTTP-pyynnön *Referer*-otsakkeen kautta – voi analytiikkapalvelulle päätyä oikeastaan mitä tahansa henkilötietoa. [13]

### 2.2.3 Käyttäjän yksilöinti

Eräs menetelmä käyttäjien yksilöimiseen verkossa on niin kutsuttu *fingerprinting*, missä yhdistellään käyttäjän selainta ja päätelaitetta koskevia teknisiä tietoja, kuten fonttityylit, näytön koko, käytetyt selainlaajennokset ja HTTP-pyynnön User-Agent-otsake. [10][14][15][16]. Täten arkiselta vaikuttavat tiedot, jotka eivät itsessään yksilöi käyttäjää, voivat silti muodostaa huomionarvoisen uhan käyttäjän yksityisyydelle.

Vaikkei *fingerprinting*-menetelmä olekaan uusi ilmiö, sen perustuminen eri tietojen yhdistelyyn tekee siitä sopeutuvan, ja koska yhdisteltäviä tietoja ei usein tarvitse tallentaa evästeisiin, ei evästeiden estäminen sivustolla takaa käyttäjän yksityisyyttä [14][15][16]. Näin on tärkeää ottaa huomioon verkkoliikennettä tarkastellessa myös ne tiedot, jotka eivät suoraan vaikuta henkilötiedoilta. Pyyntöjen hyötykuorman (*payload*) lisäksi tulee tutkia myös esimerkiksi otsaketietoja.

## 2.3 Suostumus

### 2.3.1 Käyttäjien asenne tietosuojaan

Vuosien varrella monet internetin ja verkkosivujen käyttäjät ovat ilmaisseet tutkimuksissa huolta omasta yksityisyydestään ja kolmansien osapuolien suorittamasta seurannasta [17][18][19][20]. Varsinkin kun käyttäjille esitellään heistä kerättyä tietoa, suurin osa vastaa olevansa huolissaan tiedonkeruun tasosta ja esimerkiksi iän tai sukupuolen pääteltävyydestä [18].

Asenteet yksityisyyteen verkossa vaihtelevat kuitenkin eri ihmisryhmien ja käyttäjätyyppien välillä. Tietyn väestön osan suurempi huoli seurannasta ei myöskään välttämättä johda enempiin toimiin oman yksityisyyden suojelemiseksi. On muutenkin yleistä, että verk-

kopalvelujen käyttäjät eivät huolistaan huolimatta hyödynnä esimerkiksi somepalvelujen yksityisyysasetuksia tai selainten evästeenhallintamekanismeja. [18][19][20]

Tätä näennäistä ristiriitaa käyttäjien asenteiden ja päätoksenteon välillä on kuvattu termillä ”yksityisyyden paradoksi” (*privacy paradox*), ja sen syiksi on esitetty tunteellisuutta, huolimattomuutta, ylikuormitusta tai yksinkertaisesti hyötyjen ja haittojen laskelmointia [19][20][21][22].

Vuosina 2014–2015 erilaisia käyttäjiä opiskelijoista vanhuksiin haastateltiin heidän yksityisyyteensä liittyvästä käyttäytymisestä verkossa ja tiedon jakamiseen vaikuttavista tekijöistä. Vastauksista nousi esille useita uskomuksia, joita vastaajat itsekään eivät pitäneet kovin perusteltuina. Esimerkiksi lukon kuva selaimessa saattoi luoda turvallisuuden tunnetta, vaikka lukon varsinaisesta merkityksestä ei käyttäjällä ollut tarkkaa tietoa. Toisaalta verkkokaupan käyttö tuntui erään haastateltavan mielestä turvallisemmalta kotikoneella kuin mobiililaitteella. Myös erilaiset ilmoitukset, missä tietojen jakamiseen pyydettiin lupaa, murensivat käyttäjän luottamusta kyseiseen sovellukseen. [22]

Valintoihin voivat siis vaikuttaa hetkessä koetut tuntemukset ja saadut vaikutelmat, eikä käyttäytyminen välttämättä seuraa mistään säännönmukaisesta strategiasta. Päätöksiä ohjaavissa heuristiikoissa näkyy kuitenkin yhteyksiä juuri tiettyihin verkkopalvelujen käyttöliittymien osiin [22].

Yksityisyyden paradoksi käsitteenä on saanut myös kritiikkiä, esimerkiksi professori Daniel Soloven artikkelissa *”The Myth of the Privacy Paradox”* (George Washington Law Review, 2021), suomeksi ”Myytti yksityisyyden paradoksista”. Solove esittää, että ihmisten valinnoista yksittäisissä, erityisissä tilanteissa ei ole perusteltua vetää yleistettyjä johtopäätöksiä siitä, miten kyseiset ihmiset arvottavat henkilötietojaan ja yksityisyyttään. Yksityisyys merkitsee ihmisille montaa eri asiaa, käytös ja asenteet ovat kumpikin alttiita

vaikutteille, ja siinä missä asenteet kuvaavat ihmisen arvoja yleisemmin, jokainen valinta kytkeytyy juuri senhetkiseen tilanteeseen ja sen riskeihin. [23]

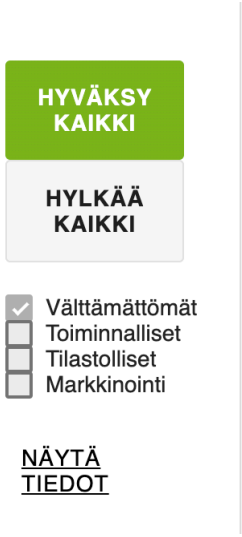
Näin eroa käyttäjien yksityisyyteen liittyvien asenteiden ja tekojen välillä ei ole syytä käsitellä ristiriitana, eikä yksityisyyden paradoksia ole Soloven mukaan olemassa. Sen sijaan yksityisyyttä suojaavista toimenpiteistä luopuminen voi olla jopa perusteltua, kun käyttäjä arvioi sitä valtavaa määrää työtä, minkä yksityisyyden suojeleminen verkossa vaatii. [23]

### 2.3.2 Pimeät käytännöt

Suostumuksen antoon evästeisiin ei siis vaikuta pelkästään käyttäjän oma tietämys ja asenne, ja käyttäjän valintoihin pyritäänkin usein vaikuttamaan suostumusvalikkojen suunnittelulla. Näitä käyttöliittymäsuunnittelun keinoja – missä käyttäjää pyritään harhaanjohtamaan, väsyttämään tai muuten painostamaan – kutsutaan ”pimeiksi käytännöiksi” (*dark patterns*), ja näiden menetelmien on osoitettu toimivan. [24][25][26] Lähestymistapoja pimeiden käytäntöjen määrittelyyn on monia, ja eri tutkimuksissa on käytetty tunnusmerkkeinä käyttöliittymien piirteitä, näiden piirteiden vaikutusta käyttäjään sekä suunnittelijoiden tarkoitusperiä ja tietoisuutta [25].

Euroopan tietosuojaneuvosto on käsitellyt juuri evästeikkunoista löytyvää harhaanjohtavaa suunnittelua raportissaan, jossa ongelmallisia piirteitä ja käytäntöjä on listattu useassa eri kategoriassa. Kategorioihin sisältyy esimerkiksi kieltämistoiminnon piilottaminen vaikeaselkoisten linkkien tai valikkokerrosten taakse, ennalta täytetyt valinnat rastitettavissa ruuduissa sekä erot painikkeiden värissä ja kontrastissa. [27] Kuvan 2.2 esimerkissä näkyy huomattava ero painikkeiden värissä ja kontrastissa, mutta esimerkiksi rastitettavat ruudut on jätetty tyhjäksi.

Raportin laatinut työryhmä päätyykin useamman kategorian kohdalla siihen johtopäätökseen, että kyseiset suunnittelumallit eivät johda GDPR:n tai ePrivacy-direktiivin mukai-




Kuva 2.2: Esimerkki suostumusikkunan painikkeista

seen validiin suostumukseen [27]. GDPR:n mukaan validin suostumuksen on oltava muun muassa ”tietoinen ja yksiselitteinen”, eikä suostumusta pitäisi voida antaa ”valmiiksi rastitetuilla ruuduilla tai jättämällä jokin toimi toteuttamatta” [28]. Siis myös pimeät käytännöt on syytä ottaa huomioon, kun arvioidaan käyttäjien todennäköisiä valintoja verkkosivuille saapuessa.

# 3 Menetelmät

## 3.1 Verkkoapteekkien haku ja valikointi

Listaa selainpohjaisista reseptilääkkeitä myyvistä verkkoapteekeista ei sellaisenaan löytynyt. Esimerkiksi Suomen Apteekkariliiton ylläpitämä apteekki.fi-sivusto vaikuttaa listaavan ainoastaan Apteekkariliittoon kuuluvia apteekkeja. Lääkealan turvallisuus- ja kehittämiskeskus Fimealla on kuitenkin sivuillaan lista kaikista laillisista apteekin verkkopalveluista, sisältäen esimerkiksi myös mobiilisovelluksiin pohjautuvat[1]. Kaikki listan verkkoapteekit eivät Fimean mukaan ole välttämättä toiminnassa.

Fimean listassa oli vuoden 2022 keväällä 253 verkkoapteekkia, eikä sivu tarjonnut mahdollisuutta apteekkien seulontaan toiminnallisuuden tai reseptimyynnin mukaan. Siksi listaa päätettiin karsia *web scraping* -menetelmillä, missä verkkosivun lähdekoodia ja sisältöä käsitellään koneellisesti[29]. Esimerkiksi sivulla olevat linkit voidaan etsiä poimimalla sivun HTML-koodista kaikki a-tunnisteet, jotka ovat yleensä muotoa
`<a href="https://www.esimerkki.fi">Esimerkki</a>`.

Käyttämällä Python-ohjelmointikielen BeautifulSoup-ohjelmointikirjastoa[2] Fimean verk-

[1] https://www.fimea.fi/apteekit/verkkopalvelutoiminta/lailliset_apteekin_verkkopalvelut

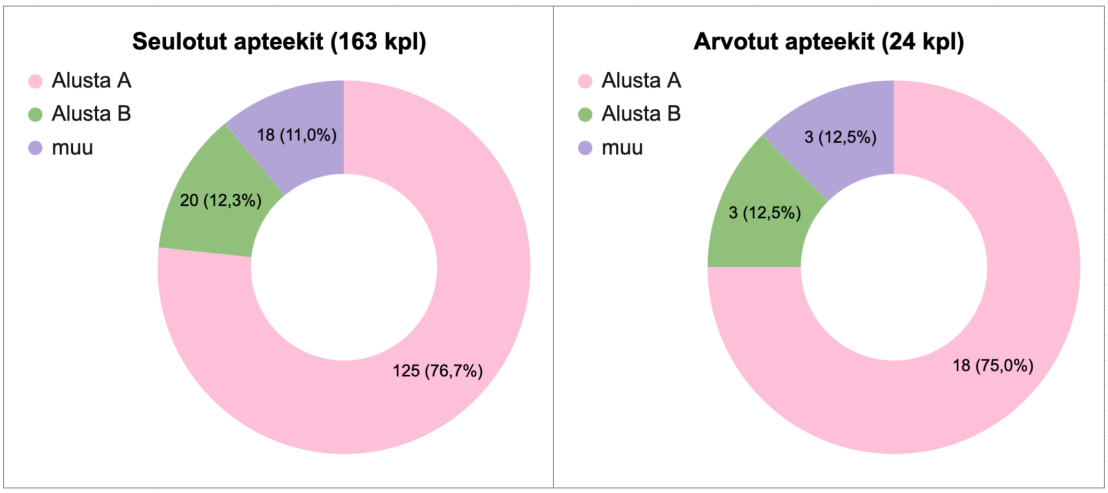
[2] https://pypi.org/project/beautifulsoup4/

koapteekkilistasta seulottiin aluksi ne apteekit, joiden kohdalla oli linkki verkkosivuille. Näin saatiin uusi 183 apteekin lista.

Uutta listaa tutkimalla huomattiin, että monet verkkoapteekit rakentuivat yhteen kahdesta verkkokauppapohjasta, joista toinen oli kotimainen apteekeille suunniteltu pohja ja toinen ulkomainen, yleisempi verkkokauppa-alusta. Käyttämällä samaa ohjelmontikirjastoa listasta eroteltiin näihin pohjiin rakennetut apteekit. Loput käytiin läpi käsin, ja listasta poistettiin rikkinäiset sivustot ja apteekit, jotka eivät myyneet sivuillaan lääkkeitä. Tämän jälkeen seulottiin pois sekä käsin että ohjelmallisin menetelmin vielä ne apteekit, jotka eivät myyneet reseptilääkkeitä.

Jäljelle jääneestä 163 selainpohjaisesta verkkoapteekista 125 käytti samaa ulkomaista verkkokauppapohjaa (Alusta A), 20 käytti samaa kotimaista verkkoapteekkipohjaa (Alusta B) ja 18 ei käyttänyt kumpaakaan.

Seulonnan jälkeen jäljelle jääneestä 163 verkkoapteekista valittiin satunnaisesti 24 apteekkia niin, että eri alustojen määrät pysyisivät suhteessa mahdollisimman samoina (kuva 3.1). Alustaa A käyttäviä apteekkeja valittiin 18, Alustaa B käyttäviä apteekkeja 3 ja muita alustoja käyttäviä apteekkeja myös 3.




Kuva 3.1: Eri alustojen osuudet seulotuista ja testatuista verkkoapteekeista

## 3.2 Verkkoapteekkien testaus

Kun tutkittavat verkkoapteekit oli valittu, niiden käytöstä aiheutuvaa verkkoliikennettä nauhoitettiin kannettavalla tietokoneella Google Chrome -selaimella. Verkkoliikenteen tarkkailuun käytettiin Chromen DevTools -kehittäjätyökaluja, jotka mahdollistivat myös verkkoliikenteen tallentamisen HAR-tiedostoihin. HAR-tiedostot sisältävät nauhoitetut tietoverkkopyynnöt tekstimuotoisine sisältöineen JSON-muodossa, ja tiedostot mahdollistavat verkkoliikenteen tarkastelun myös nauhoittamisen jälkeen.

### 3.2.1 Suostumukset

Neljä verkkoapteekkia ei sivustolle saavuttaessa pyytänyt minkäänlaista suostumusta evästeisiin tai markkinointiin. Monet apteekeista antoivat kuitenkin mahdollisuuden valita esimerkiksi sivuston käyttämät evästeryhmät erilaisilla valikoilla. Kysytty suostumus kirjattiin muistiin jokaisen apteekin osalta.

Jokaisen verkkoapteekin käytöstä aiheutuvaa verkkoliikennettä nauhoitettiin sekä enimmällä että vähimmällä mahdollisella suostumuksella. Jos sivusto tarjosi esimerkiksi mahdollisuuden valita useammasta eri evästekategoriasta (välttämättömien lisäksi), ensimmäisessä testissä valittiin kaikki ja toisessa testissä ei yhtäkään. Jos vaihtoehtoja ei ollut ollenkaan, tehtiin koe vain kerran.

Jotta testattavan verkkoapteekin suostumusvalikkoja voitaisiin aina käyttää kuten ensimmäisellä vierailulla sivustolle, jokaisen kokeen alussa selaimen välimuisti tyhjennettiin ja verkkoapteekin aiemmin asettamat evästeet poistettiin. Tämän jälkeen verkkoliikenteen nauhoitus käynnistettiin ja sivu ladattiin uudestaan. Näissä toimenpiteissä käytettiin kehittäjätyökalujen lisäksi vain selaimen omia valikkoja ja painikkeita.

### 3.2.2 Testauksen sivupolku

Verkkoapteekkien testauksen pohjana oli erään reseptillä saatavan psyykenlääkkeen tarkastelu mahdollisella tuotesivulla sekä kyseisen lääkkeen tilausprosessin aloittaminen. Näin apteekin hyödyntämille analytiikkapalveluille voisi mahdollisesti välittyä tieto käyttäjän aikomuksesta tilata kyseinen reseptilääke, mikä puolestaan paljastaisi arkaluontoista tietoa käyttäjän terveydentilasta. Alustaa B käyttävillä apteekeilla ei kuitenkaan ole erillisiä tuotesivuja reseptilääkkeille, minkä vuoksi kyseisten verkkoapteekkien kohdalla testeissä siirryttin suoraan tilausprosessiin.

Myös tilausprosessin etenemisessä on eroja eri verkkoapteekkien välillä. Koska koekäytössä haluttiin välttää perusteettoman reseptitilauksen esittämistä farmaseutille, tavoitteena oli päästä yhdelle verkkoapteekin sivuilla sijaitsevalle tilausprosessiin kuuluvalle sivulle, poislukien esimerkiksi vahvassa tunnistautumissa käytetyt kolmannen osapuolen verkkosivut. Näin tilausprosessi ei etenisi liian pitkälle, mutta prosessin verkkoliikennettä verkkoapteekin sivuilla kyettäisiin silti taltioimaan.

Jos tilausprosessin aloittaminen vaati kirjautumista verkkoapteekkiin tai vahvaa tunnistautumista, tämä tehtiin lähtökohtaisesti etukäteen, ennen lääkkeen sivulle siirtymistä. Joissain tapauksissa tunnistautuminen oli kuitenkin sidottu tilausprosessiin.

## 3.3 Verkkoliikenteen analysointi ja taulukointi

Kun verkkoapteekin testaus saatiin valmiiksi, verkkoliikenteen nauhoitus pysäytettiin. Nauhoitettu verkkoliikenne tallennettiin heti HAR-tiedostoon, jotta sen sisältöjä voitaisiin tutkia myös myöhemmin.

Sivuston käytöstä aiheutuneita tietoverkkopyyntöjä eroteltiin toisistaan esimerkiksi pyy-

detyn tiedostomuodon, pyynnön ajankohdan sekä kolmannen osapuolen osotteiden perusteella. Pyyntöjen läpikäynti pienemmissä osissa oli tarpeen, sillä pyyntöjä kertyi joka testissä satoja.

Joukosta etsittiin niitä pyyntöjä, joiden sisältö tai vastaanottaja viittasi analytiikkaan. Mahdollisia sivuston hyödyntämiä analytiikka-alustoja voitiin päättellä esimerkiksi kehittäjätyökalujen *Sources*-valikosta (suomeksi ”Lähteet”), joka listaa sivun käyttämiä resursseja. Yksittäisten tietoverkkopyyntöjen kohdalla tarkasteltiin sekä pyynnön otsikkotietoja että hyötykuormaa. Monesti analytiikassa kerätyt tiedot lähetettiin osana vastaanottajan osoitetta, *query string* -parametreina. Erityisesti etsittiin niitä tietoja, joista käy ilmi joko aikomus tilata reseptilääke tai mikä lääke on kyseessä.

Analytiikassa kulkeneista henkilö-, tuote- ja tilaustiedoista tehtiin taulukko, jossa tiedot eroteltiin verkkoapteekin, kokeen, kokeen vaiheen ja tietojen vastaanottajan perusteella. Lisäksi taulukkoon merkittiin apteekin hyödyntämä verkkokauppa-alusta sekä sivuston tarjoamat suostumusvalinnat. Taulukko mahdollisti esimerkiksi analytiikkapalvelujen keräämien tietojen vertailun eri apteekkien ja suostumusvalintojen välillä.

# 4 Tulokset

## 4.1 Käyttäjältä kysytty suostumus

Käyttäjältä kysytty suostumus verkkoapteekkien sivuille saavuttaessa vaihteli. Kaikissa 18 apteekissa, joiden verkkopalvelut rakentuivat alustaan A, käyttäjältä pyydettiin lupaa evästeisiin. Suostumusikkunassa oli mahdollisuus hyväksyä kaikki evästeet, hyväksyä vain pakolliset evästeet tai säätää evästeitä käyttötarkoituksen mukaan. Eivälttämättöminä käyttötarkoituksina listattiin tilastot ja analyysit sekä markkinointitarkoitukset. Alustaa B käyttäneestä kolmesta apteekista yksikään ei pyytänyt käyttäjältä suostumusta evästeisiin, analytiikkaan tai muuhun vastaavaan.

Loput kolme apteekkia olivat keskenään erilaisia: Yksi pyysi lupaa ”evästeiden ja muiden tekniikoiden käyttöön” eri käyttötarkoituksiin. Käyttäjällä oli mahdollisuus hyväksyä kaikki, hylätä kaikki tai valita haluamansa käyttötarkoitukset, missä vaihtoehtoina oli ”välttämättömät”, ”toiminnalliset”, ”tilastolliset” sekä ”markkinointi”. Toinen apteekeista käytti lausetta ”Sallin tietojeni käyttämisen markkinointiin”, johon vastausvaihtoehtona oli ”Ei” tai ”Kyllä”. Kolmas apteekki ei puolestaan kysynyt minkäänlaista suostumusta.

Kaikkiaan neljässä apteekin verkkopalvelussa ei siis ollut mahdollisuutta antaa kahta eriasteista suostumusta evästeisiin, sillä suostumusta ei kysytty.

## 4.2 Apteekkien käyttämät analytiikkapalveluntarjoajat

Apteekkien hyödyntämiä analytiikkapalveluntarjoajia on taulukoitu kuvassa 4.1 Yleisin palveluntarjoaja kolmannen osapuolen analytiikalle oli Google. Apteekkien koekäytöstä aiheutuvassa verkkoliikenteessä oli analytiikkatietoa välittäviä tai esimerkiksi Google Tag Managerin JavaScript-koodia lataavia HTTP-pyyntöjä 16 apteekin kohdalla.

Yhdysvaltalaisen SolarWindsin omistama analytiikkapalvelu Pingdom ilmeni kolmen alustaa B käyttävän apteekin kohdalla.

Yksi apteekeisesta, joka ei rakennu alustaan A tai B, hyödynsi Googlen lisäksi myös Facebookin sekä Microsoftin omistaman Bingin analytiikkaa.

Myös Giosg-nimiselle palvelulle välittyi analytiikkatietoa, kuten tieto käyttäjän liikkumisesta sivuilla. Giosg tarjoaa sivuilta löytyvän livechat-toiminnon, jonka voi katsoa osaksi sivuston välttämätöntä toiminnallisuutta, sillä esimerkiksi reseptilääkkeitä ostaessa lääkelaki edellyttää keskusteluyhteyden farmaseuttiin [30]. On vaikeaa vetää rajaa välttämättömän ja ylimääräisen tiedon välille Giosgille lähtevissä pyynnöissä, eikä Giosgia siksi käsitellä tässä tutkielmassa kolmannen osapuolen analytiikkapalveluntarjoajana.

Apteekkeja, jotka eivät hyödyntäneet yhtäkään kolmannen osapuolen analytiikkapalvelua, oli kuusi. Nämä kaikki oli rakennettu alustaan A.

| | Google | Pingdom | Facebook | Bing |
|---|---|---|---|---|
| **Alusta A (max 18)** | 12 | 0 | 0 | 0 |
| **Alusta B (max 3)** | 1 | 3 | 0 | 0 |
| **Muut (max 3)** | 3 | 0 | 1 | 1 |
| **Yhteensä (max 24)** | 16 | 3 | 1 | 1 |

Kuva 4.1: Apteekkien analytiikkapalveluntarjoajat alustoittain

## 4.3 Analytiikassa kerätyt käyttäjän tiedot

Vaikka eri analytiikkapalveluille lähtevien HTTP-pyyntöjen varsinainen hyötykuorma vaihteli, jokaisen pyynnön otsakkeissa kulki vähintäänkin tieto käyttäjän IP-osoitteesta, päätelaitteen käyttöjärjestelmästä ja selaimen versiosta. Pyyntöjen hyötykuormista, *query string* -parametreina, löytyi yhden apteekin kohdalla jopa valitun reseptilääkkeen nimi, hinta ja tuotekategoria. Valittu lääke saattoi välittyä suoraan tai välillisesti myös tuotesivun osoitteen kautta. Aikomus tilata jokin tuote paljastui monesti tilausprosessiin liittyvän sivun osoitteen kautta, mutta tieto tuotteen lisäämisestä ostoskoriin välittyi suoraan sen apteekin kohdalla, joka välitti myös reseptilääkkeen tuotetiedot.

Analytiikkapalveluille lähtevien pyyntöjen hyötykuormissa oli myös erilaisia teknisiä ja laitetietoja, jotka auttaisivat käyttäjän yksilöinnissä: Pyynnöissä listattiin esimerkiksi käyttäjän päätelaitteen näytön koko, värisyvyys ja mahdollisesti suoritinarkkitehtuuri, esimerkiksi ”x86”. Lisäksi pyynnöissä listattiin myös selainikkunan koko, selaimen kieli sekä tieto siitä, onko käyttäjä kirjautunut verkkoapteekkiin.

Pyynnöt sisälsivät myös monia erilaisia tunnisteita, joiden käyttötarkoituksesta ei aina saatu selvyyttä. Esimerkiksi monet Googlen analytiikkaan liittyvät pyynnöt sisälsivät kuitenkin cid-nimisen tunnisteen, jolla voidaan yksilöidä tietty käyttäjä, laite tai selaininstanssi [31].

## 4.4 Euroopan ulkopuolelle lähetetyt tiedot

Verkkoapteekkien käyttämistä kolmannen osapuolen analytiikkapalveluntarjoajista yleisin oli Google, jonka palvelimien sijainteja oli monesti vaikea varmistaa. Esimerkiksi

iplocation.net-sivuston[1] listaamat tietokannat ja hakupalvelut esittivät useita eriäviä sijainteja Google Analyticsin palvelimille, monesti sekä Euroopan unionissa että Yhdysvalloissa. Facebookin ja Pingdomin käyttämät palvelimet paikantuivat iplocation.net:in lähteiden perusteella Ruotsiin ja Irlantiin.

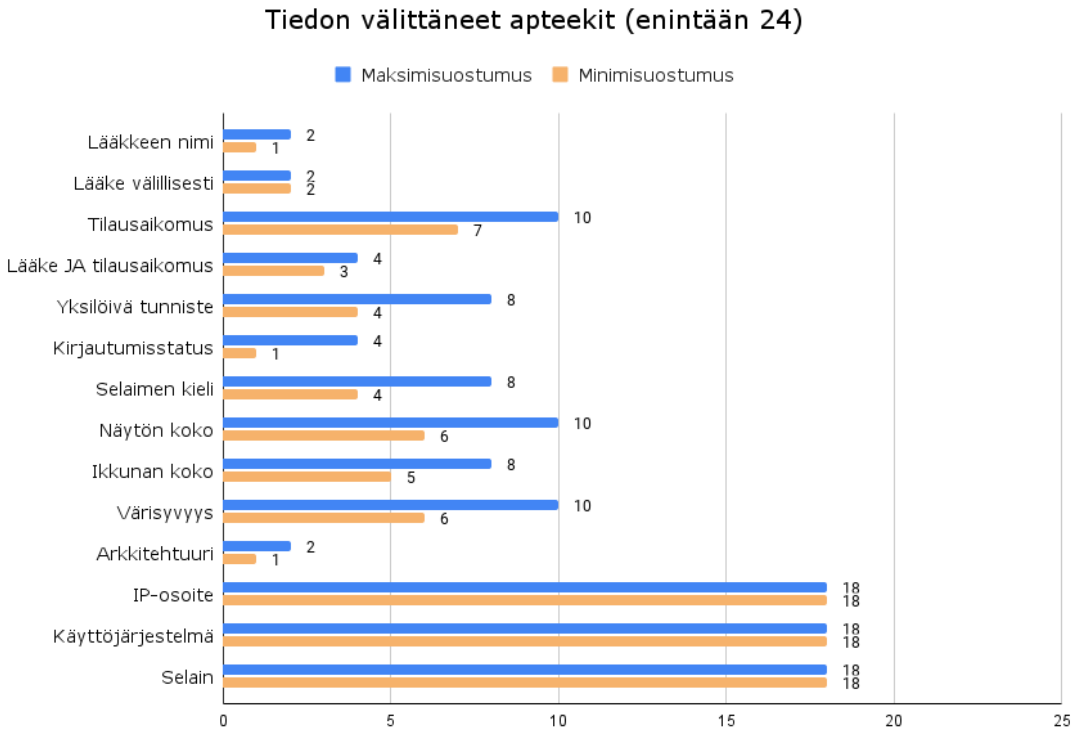
Oli kuitenkin tapauksia, joissa verkkoapteekkien analytiikassa kerättyjä tietoja lähetettiin Pohjois-Amerikassa sijaitseville palvelimille, Euroopan unionin ja Euroopan talousalueen ulkopuolelle. Bingiä käyttäneen apteekin kohdalla analytiikassa kerätyt tiedot lähettiin palvelimelle, joka iplocation.net:in lähteiden mukaan sijaitsi joko Yhdysvalloissa tai Kanadassa. Tietoihin sisältyi käyttäjän liikkeet sivustolla, päätelaitteen ja selaimen tietoja (esimerkiksi näytön värisyvyys, käyttöjärjestelmä ja selaimen versio) sekä erilaisia tunnisteita, joiden toiminta jäi epäselväksi.

Google Analyticsia käyttäneestä kuudestatoista apteekista kuusi lähetti tietoja Googlen palvelimille, jotka sijaitsivat iplocation.net:in lähteiden mukaan Yhdysvalloissa. Pyynnöissä meni aina tieto päätelaitteen käyttöjärjestelmästä ja selaimen versiosta, sekä cid-tunniste. Erään apteekin kohdalla Yhdysvaltoihin menevissä pyynnöissä kulki myös tarkempia laitetietoja, kuten näytön värisyvyys ja mahdollisesti suoritinarkkitehtuuriin viittaava ”x86”.

## 4.5 Suostumusvalintojen väliset erot

Kun kolmannen osapuolen analytiikkapalveluille välittyneitä tietoja vertailtiin maksimi- ja minimisuostumuksella tehtyjen testien välillä, eroja ilmeni vain joidenkin tietojen ja apteekkien kohdalla. Kuvassa 4.2 näkyy suostumusten välinen ero tietoja välittäneiden apteekkien lukumäärissä.

[1] https://www.iplocation.net/




Kuva 4.2: Suostumusvalinnan vaikutus tietojen välittymiseen kolmannen osapuolen analytiikkapalveluntarjoajille

Neljän apteekin verkkoliikenteessä näkyi eroja maksimi- ja minimisuostumuksen välillä. Kolme rakentui alustaan A ja käytti sivuillaan Googlen analytiikkaa. Neljäs hyödynsi sekä Googlen että Facebookin analytiikkaa, eikä käyttänyt alustaa A tai B. Näiden neljän apteekin verkkoliikenteestä hävisivät minimisuostumuksella ne HTTP-pyynnöt, joissa kulki esimerkiksi valitun lääkkeen tuotiedot, tieto sivujen välillä liikkumisesta sekä tieto näytön ja selainikkunan koosta. Jokainen näistä neljästä latasi kuitenkin Googlen analytiikkapalvelujen JavaScript-koodia suostumusvalinnoista riippumatta, joten otsakkeissa kulkeva IP-osoite, käyttöjärjestelmä ja selain ilmenivät 18 kertaa sekä minimi- että maksimisuostumuksella.

# 5 Pohdinta

## 5.1 Kehittäjän näkökulma

Ei ole syytä olettaa, että verkkoapteekin kehittäjä olisi välttämättä sama yritys, joka verkkoapteekin omistaa. Luonteva ratkaisu varsinkin pienemmille apteekeille olisi tilata verkkoapteekkipalvelu ohjelmistokehitykseen ja verkkosuunnitteluun erikoistuneelta yritykseltä. Apteekin on syytä keskustella kehittäjän kanssa yksityisyyteen liittyvistä vaatimuksista, mutta tärkeää on myös kehittäjän oma-aloitteisuus verkkoapteekin tietoturvan ja käyttäjän yksityisyyden huomioinnissa.

Verkkokaupan kehittäjällä on mahdollisuus päättää, miten kauppa on rakennettu ja mitä tietoja se kerää. Vaikka maksuliikenteeseen liittyviä standardiratkaisuja voi olla vaikea välttää – ja ostosten yhteydessä asiakkaalta on pakko kerätä joitain tietoja – kolmannen osapuolen analytiikka tuskin on välttämätöntä. On vaikea uskoa suomalaisen reseptilääkkeitä myyvän verkkoapteekin tarvitsevan analytiikkaratkaisua, jossa tietoja käyttäjästä ja ostotapahtumasta kerätään esimerkiksi Yhdysvaltalaisille palvelimille, varsinkin jos asiakas on kieltäytynyt tiedonkeruusta.

Lähtökohtana voisi olla verkkoapteekki, jonka käyttäjästä ei kerätä mitään muuta tietoa paitsi tunnistautumisessa ja maksuliikenteessä kerätyt välttämättömät tiedot. Verkkoap-

teekkiin lisättäisiin esimerkiksi lokitusta virheenjäljitykseen tai CAPTCHA-testejä[1] ihmiskäyttäjien tunnistamiseen vain, jos niistä saadaan merkittävä hyöty eivätkä ne vaaranna käyttäjän yksityisyyttä.

Jos muuta tiedonkeruuta ei voida kohtuudella rajoittaa, ja apteekki haluaa käyttää esimerkiksi tuotesuunnitteluun ja markkinointiin liittyvää analytiikkaa, analytiikassa kerättävän tiedon tulisi säilyä apteekilla ja sen ylläpitäjällä, tai vähintään jollain luotettavaksi katsotulla taholla Euroopan unionin sisällä. Esimerkiksi Matomo-alustaa (entinen Piwik) on tutkittu vaihtoehtona Google Analyticsille, sillä Matomon analytiikkaa voi käyttää ilman tietojen lähettämistä kolmannen osapuolen palvelimille [32][33].

## 5.2 Käyttäjän näkökulma

Jos verkkoapteekkien käyttäjien tietoja käsittelee useampi taho, ja tietoja on säilötty useampaan paikkaan, tietovuodon riski kasvaa. Vaikka analytiikkapalveluntarjoaja säilöisi käyttäjän tiedot tietoturvallisesti ja laillisesti, eikä tietomurtoa tapahdu, voi käyttäjä silti kokea kolmannen osapuolen analytiikan sopimattomaksi. Jos käyttäjä on suomalaisen verkkoapteekin asiakas, miksi käyttäjän mahdollisesti yksilöiviä laitetietoja tulisi säilyttää yhdysvaltalaisen mainosyhtiön palvelimilla, varsinkin jos tietoihin on liitetty käyttäjän ostama reseptilääke?

Henkilön terveystietojen vuotamisella voi olla merkittäviä seurauksia. Suomalaisen psykoterapiakeskus Vastaamon vuonna 2020 paljastuneessa tietomurrossa vuosi kymmenien tuhansien suomalaisten tietoja, mukaan lukien arkaluontoisia potilastietoja [34][2]. Uhrien omaisten mukaan tietovuoto johti pahimmillaan jopa itsemurhiin [2]. Kuten luvussa

[1] https://fi.wikipedia.org/wiki/CAPTCHA

2 todettiin, moni verkkopalvelujen käyttäjä onkin huolissaan yksityisyydestään, eikä käyttäjillä ole välttämättä valmiuksia tutkia analytiikassa tapahtuvaa tiedonkeruuta itse.

# 6 Yhteenveto

## 6.1 Vastaukset tutkimuskysymyksiin

Tutkielman alussa määriteltiin kaksi tutkimuskysymystä:

1. Mitä terveys- tai muita henkilötietoja välittyy kolmannen osapuolen palveluntarjoajille verkkoapteekkien analytiikassa?
2. Miten ääripäiden suostumusvalinnat vaikuttavat kolmannen osapuolen palveluntarjoajille lähetettäviin tietoihin?

Vastaukset ovat hyvin apteekkikohtaisia, ja jo tutkielmaan arvotussa 24 verkkoapteekkissa oli eroja käytetyn analytiikan lisäksi suostumusikkunoissa, käyttäjältä kerätyissä tiedoissa sekä näiden yhteydessä. Samaan alustaan rakentuvista ja samaa suostumusta kysyvistä apteekeista osa hyödyntää Googlen analytiikkaa tuote- ja tilaussivuilla, kun taas osa ei käytä Googlen analytiikkaa ollenkaan. Joissain analytiikkaa hyödyntävissä apteekeissa tietoa välittyy kolmansille osapuolille vähemmän, kun evästeiden käyttöön annetaan pienin mahdollinen suostumus. Toisissa apteekeissa mitään eroa ei näy. Niistä kolmesta apteekista, jotka eivät käyttäneet alustaa A tai B, kaksi kysyi käyttäjältä suostumusta eri tavoilla, ja yksi ei kysynyt suostumusta ollenkaan. Yleisin kolmannen osapuolen pal-

veluntarjoaja analytiikalle oli Google, mutta verkkoapteekeista löytyi myös Facebookin, Pingdomin ja Bingin analytiikkaa.

Yksi verkkoapteekkeista keräsi Googlen analytiikassa jopa käyttäjän valitseman reseptilääkkeen tuotetiedot. Minimisuostumuksella tuotetietoja ei kuitenkaan kerätty. Lisäksi aikomus ostaa tietty reseptilääke paljastui kolmen muun verkkoapteekin kohdalla välillisesti, kun analytiikassa kerättiin tieto sivuista, joilla käyttäjä on liikkunut. Näiden kolmen apteekin kohdalla käyttäjän antamalla suostumuksella ei havaittu olevan vaikutusta kerättyyn tietoon. Tutkielmaan arvotuista apteekeista 17% oli sellaisia, joissa käyttäjän valitsema lääke oli pääteltävissä analytiikassa kerätyistä tiedoista, jos käyttäjä antoi täyden suotumuksen. Valittu lääke oli pääteltävissä 12,5% apteekeista, jos käyttäjä antoi pienimmän mahdollisen suostumuksen. Jos aikomus ostaa tietty reseptilääke lasketaan asiakkaan terveyttä koskevaksi tiedoksi, sen käsittely on tietosuojavaltuutetun toimiston mukaan ”lähtökohtaisesti kiellettyä” [6].

## 6.2 Rajoitukset ja jatkotutkimus

Apteekkien testaukseen arvottiin useita samoihin alustoihin rakennettuja verkkoapteekkeja. Valittujen apteekkien lukumäärä ei kuitenkaan riitä johtopäätöksiin alustakohtaisista eroista analytiikan käytössä tai asiakkaan yksityisyyden suojelemisessa. Tutkimalla alustoja A ja B voisi esimerkiksi selvittää, onko suostumusikkuna (tai sen puute) osa itse alustaa. Samalla voisi tutkia, miten eri alustat tukevat kolmannen osapuolen analytiikkaratkaisuja ja mitä muita mahdollisia ratkaisuja verkkoapteekit voisivat hyödyntää käyttäjän tietosuojan varmistamiseksi. Parhaimman kuvan analyytikan käytöstä alustoittain saisi testaamalla kaikki 163 reseptilääkkeitä myyvää, selaimessa toimivaa suomalaista verkkoapteekkia. Kenties suurempaan tutkimukseen voisi ottaa mukaan myös mobiilisovellukset. Suomalaisia verkkoapteekkeja voisi myös verrata muiden maiden verkkoapteek-

keihin. Vertailu olisi erityisen mielenkiintoinen, jos muista maista löytyisi samoja alustoja ja ratkaisuja käyttäviä apteekkeja.

Ei voida myöskään olettaa, että apteekkien käyttämät kolmannen osapuolen analytiikkapalvelut tai analytiikassa kerätyt tiedot pysyisivät vuosien kuluessa samana. Tutkimukset, joiden ohella tutkielma tehtiin [3][4], johtivat ilmoitukseen tietosuojavaltuutetun toimistolle, ja suurin osa löydetyistä ongelmista on jo korjattu. Uudet verkkoapteekit, verkkokauppa-alustat ja analytiikkaratkaisut voivat kuitenkin luoda uusia uhkia käyttäjien yksityisyydelle, ja suomalaisten verkkoapteekkien yksityisyyttä olisi hyvä tutkia myös tulevaisuudessa.

# Lähdeluettelo


[1] A. R. Zheutlin, J. D. Niforatos ja J. B. Sussman, ”Data-Tracking Among Digital Pharmacies”, 8, vol. 56, SAGE, s. 958–962.

[2] P. Nykänen, *Vastaamo-uhrien juristi A-studiossa: Osa pystyy jatkamaan elämäänsä kohtalaisen normaalisti, osa ei ole enää keskuudessamme*, 2024. url: `https://yle.fi/a/74-20086539`.

[3] S. Rauti, E. Vuorinen ja R. Carlsson, ”How not to design an online pharmacy: A case study”, teoksessa *Proceedings of the 2023 8th International Conference on Information Systems Engineering*, sarja ICISE ’23, Association for Computing Machinery, 2024, s. 90–94.

[4] R. Carlsson, S. Rauti, S. Mickelsson et al., ”Several online pharmacies leak sensitive health data to third parties”, teoksessa *World Conference on Information Systems and Technologies*, Springer, 2023, s. 164–175.

[5] Tietosuojavaltuutetun toimisto, ”Mikä on henkilötieto?”. url: `https://tietosuoja.fi/mika-on-henkilotieto`.

[6] Tietosuojavaltuutetun toimisto, ”Henkilötietojen käsittely”. url: `https://tietosuoja.fi/henkilotietojen-kasittely`.

[7] Tietosuojavaltuutetun toimisto, ”Pseudonymisoidut ja anonymisoidut tiedot”. url: `https://tietosuoja.fi/pseudonymisointi-anonymisointi`.

[8] I. Bekavac ja D. G. Praničevic´, ”Web analytics tools and web metrics tools: An overview and comparative analysis”, teoksessa *Croatian Operational Research Review*, Trg JF Kennedyja 6, Zagreb: Croatian Society for Operational Research, 2015, s. 373–386. DOI: 10.17535/crorr.2015.0029.

[9] T. Wambach ja K. Bräunlich, ”The Evolution of Third-Party Web Tracking”, teoksessa *Information Systems Security and Privacy*, O. Camp, S. Furnell ja P. Mori, toim., Springer International Publishing, 2017, s. 130–147.

[10] I. Sanchez-Rola, X. Ugarte-Pedrero, I. Santos ja P. G. Bringas, ”The web is watching you: A comprehensive review of web-tracking techniques and countermeasures”, *Logic Journal of the IGPL*, vol. 25, nro 1, s. 18–29, elokuu 2016, ISSN: 1367-0751. DOI: 10.1093/jigpal/jzw041. eprint: https://academic.oup.com/jigpal/article-pdf/25/1/18/9133424/jzw041.pdf.

[11] Y. Liu, H. H. Song, I. Bermudez, A. Mislove, M. Baldi ja A. Tongaonkar, ”Identifying Personal Information in Internet Traffic”, teoksessa *Proceedings of the 2015 ACM on Conference on Online Social Networks*, sarja COSN ’15, Palo Alto, California, USA: Association for Computing Machinery, 2015, s. 59–70. DOI: 10.1145/2817946.2817947.

[12] D. Malandrino, A. Petta, V. Scarano, L. Serra, R. Spinelli ja B. Krishnamurthy, ”Privacy Awareness about Information Leakage: Who Knows What about Me?”, teoksessa *Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society*, sarja WPES ’13, Berlin, Germany: Association for Computing Machinery, 2013, s. 279–284. DOI: 10.1145/2517840.2517868.

[13] Tietosuojavaltuutetun toimisto, ”Apulaistietosuojavaltuutetun päätös käsittelyn lainmukaisuutta, käsittelyn turvallisuutta, sisäänrakennettua ja oletusarvoista tietosuojaa, rekisteröityjen informointia ja henkilötietojen siirtoa kolmansiin maihin koskevassa asiassa”. url: https://tietosuoja.fi/documents/6927448/

146469002/ATSV+p%C3%A4%C3%A4t%C3%B6s+4672.161.22.pdf/dfe9578a-59ec-26f5-1cf7-0651cbc8b298/ATSV+p%C3%A4%C3%A4t%C3%B6s+4672.161.22.pdf?t=1673939099389.

[14] K. Boda, Á. M. Földes, G. G. Gulyás ja S. Imre, ”User Tracking on the Web via Cross-Browser Fingerprinting”, teoksessa *Information Security Technology for Applications*, P. Laud, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, s. 31–46.

[15] P. Eckersley, ”How Unique Is Your Web Browser?”, teoksessa *Privacy Enhancing Technologies*, M. J. Atallah ja N. J. Hopper, toim., Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, s. 1–18.

[16] N. Kaur, S. Azam, K. Kannoorpatti, K. C. Yeo ja B. Shanmugam, ”Browser Fingerprinting as user tracking technology”, teoksessa *2017 11th International Conference on Intelligent Systems and Control (ISCO)*, 2017, s. 103–111. DOI: 10.1109/ISCO.2017.7855963.

[17] European Commission, Directorate-General for Communication. ”Special Eurobarometer 431: Data protection (v1.00)”. (2015), url: http://data.europa.eu/88u/dataset/S2075_83_1_431_ENG.

[18] C. E. Wills ja M. Zeljkovic, ”A personalized approach to web privacy: awareness, attitudes and actions”, 1, vol. 19, Emerald Group Publishing Limited, 2011, s. 53–73. DOI: 10.1108/09685221111115863.

[19] E.-M. Schomakers, C. Lidynia ja M. Ziefle, ”A Typology of Online Privacy Personalities”, 4, vol. 17, 2019, s. 727–747. DOI: 10.1007/s10723-019-09500-3.

[20] N. Gerber, P. Gerber ja M. Volkamer, ”Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior”, *Computers & Security*, vol. 77, s. 226–261, 2018, ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2018.04.002.

[21] S. S. Sundar, H. Kang, M. Wu, E. Go ja B. Zhang, ”Unlocking the Privacy Paradox: Do Cognitive Heuristics Hold the Key?”, teoksessa *CHI ’13 Extended Abstracts on Human Factors in Computing Systems*, sarja CHI EA ’13, Paris, France: Association for Computing Machinery, 2013, s. 811–816. DOI: 10.1145/2468356.2468501.

[22] A. Gambino, J. Kim, S. S. Sundar, J. Ge ja M. B. Rosson, ”User Disbelief in Privacy Paradox: Heuristics That Determine Disclosure”, teoksessa *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, sarja CHI EA ’16, San Jose, California, USA: Association for Computing Machinery, 2016, s. 2837–2843. DOI: 10.1145/2851581.2892413.

[23] D. J. Solove, ”The Myth of the Privacy Paradox”, 1, vol. 89, 2021, s. 1–51. url: https://heinonline.org/HOL/Page?collection=journals&handle=hein.journals/gwlr89&id=16&men_tab=srchresults.

[24] Kilpailu- ja kuluttajavirasto, ”Pimeät käytännöt”. url: https://www.kkv.fi/kuluttaja-asiat/huijaukset/pimeat-kaytannot/.

[25] A. Mathur, M. Kshirsagar ja J. Mayer, ”What Makes a Dark Pattern... Dark? Design Attributes, Normative Considerations, and Measurement Methods”, teoksessa *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, sarja CHI ’21, Yokohama, Japan: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445610.

[26] M. Nouwens, I. Liccardi, M. Veale, D. Karger ja L. Kagal, ”Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence”, teoksessa *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, sarja CHI ’20, Honolulu, HI, USA: Association for Computing Machinery, 2020, s. 1–13.

[27] European Data Protection Board. ”Report of the work undertaken by the Cookie Banner Taskforce”. (2023), url: https://edpb.europa.eu/our-work-tools/our-documents/other/report-work-undertaken-cookie-banner-taskforce_en.

[28] Euroopan parlamentti ja neuvosto ja Euroopan unionin julkaisutoimisto, ”Asetus - 2016/679 - FI - GDPR - EUR-Lex”. url: https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX:32016R0679.

[29] R. Mitchell, *Web Scraping with Python, 2nd Edition*. O’Reilly Media, 2018.

[30] finlex.fi. ”Lääkelaki”. (2024), url: https://finlex.fi/fi/lainsaadanto/1987/395#chp_6__sec_57v20221233__subsec_1v20240934.

[31] Google, ”Measurement Protocol Parameter Reference”. url: https://web.archive.org/web/20230329135642/https://developers.google.com/analytics/devguides/collection/protocol/v1/parameters#cid.

[32] A. Chandler ja M. W. and, ”Using Piwik Instead of Google Analytics at the Cornell University Library”, *The Serials Librarian*, vol. 71, nro 3-4, s. 173–179, 2016. DOI: 10.1080/0361526X.2016.1245645.

[33] D. Quintel ja R. Wilson, ”Analytics and Privacy: Using Matomo in EBSCO’s Discovery Service”, *Information Technology and Libraries*, vol. 39, nro 3, syyskuu 2020. DOI: 10.6017/ital.v39i3.12219.

[34] J. Mantsinen, *Poliisi on selvittänyt kaikkien Vastaamon tietomurron uhrien henkilötiedot – Asianomistajia yli 33 000*, 2023. url: https://www.aamulehti.fi/rikos/art-2000009757940.html.