

Henkilötiedustelu sosiaalisen median kanavissa

Markus Hoffrén

Opinnäytetyö

Toukokuu 2021

Tekniikan ja liikenteen ala

Insinööri (AMK), Tieto- ja viestintätekniikan koulutusohjelma

Tietoverkkotekniikka

Tekijä(t) Hoffrén, Markus	Julkaisun laji Opinnäytetyö, AMK	Päivämäärä Toukokuu 2021
	Sivumäärä 51	Julkaisun kieli Suomi
		Verkojulkaisulupa myönnetty: x
Työn nimi Henkilötiedustelu sosiaalisen median kanavissa		
Tutkinto-ohjelma Tieto- ja viestintätekniikka		
Työn ohjaaja(t) Häkkinen Antti, Kotikoski Sampo		
Toimeksiantaja(t) Jyväskylän Ammattikorkeakoulu / IT-instituutti, CYBERDI-projekti		
<p>Tiivistelmä</p> <p>Ihmiset eivät välttämättä tiedosta, mitä kaikkea he ovat julkaisseet eri sosiaalisen median kanavilla ja kuinka helposti tämä tieto on muiden saatavilla. Sosiaalisen median kanavien ja muiden avoimien lähteiden merkitys on suuri henkilötiedustelussa juuri tästä syystä, että ihmiset julkaisevat itse tietoja, joita yhdistelemällä saadaan rakennettua kokonaisuus, jonka avulla pahimmassa tapauksessa saadaan tehtyä esimerkiksi identiteettivarkauksia.</p> <p>Opinnäytetyö toteutettiin käyttämällä kvalitatiivista tutkimusmenetelmää. Työn tavoitteena oli tuoda ilmi, mitä tietoja sosiaalisen median käyttäjistä saa haettua erilaisten työkalujen avulla ja pystyykö omia henkilötietoja piilottamaan eri sosiaalisen median alustoilla. Opinnäytetyö tehtiin CYBERDI-projektille, jonka tarkoituksena on kasvattaa yhteiskunnassamme tietoisuutta muun muassa digitaaliseen maailmaan liittyvistä uhkista.</p> <p>Tutkimus aloitettiin tekemällä käyttäjätunnukset eri sosiaalisen median alustoille. Tämän jälkeen perehdyttiin sosiaalisen median kanavia hyväksikäyttäviin työkaluihin ja selvitettiin näiden piirteitä ja ominaisuuksia. Käytännön työtä varten valittiin yhteensä seitsemän erilaista työkalua eri sosiaalisen median alustaa varten ja suoritettiin testit käyttämällä näitä työkaluja. Työkalut valittiin sen perusteella, että ne olivat suosittuja, helppokäyttöisiä ja ilmaisia.</p> <p>Työkalujen avulla saatiin rakennettua erilaisia raportteja kohdehenkilön käyttäytymisvoista eri alustoilla ja tietoa muun muassa tämän mielenkiinnon kohteista. Julkisista julkaisuista saatiin hyvin tiedot raportteihin, mutta piilotettuja, henkilökohtaisia tietoja ei työssä käytetyillä työkaluilla saatu selville.</p>		
Avainsanat (asiasanat) OSINT, SOCMINT, henkilötiedustelu, sosiaalinen media		
Muut tiedot (Salassa pidettävät liitteet)		

Author(s) Hoffrén, Markus	Type of publication Bachelor's thesis	Date May 2021 Language of publication: Finnish
	Number of pages 51	Permission for web publication: x
Title of publication Personal data gathering on social media platforms		
Degree programme Information and Communication Technology		
Supervisor(s) Häkkinen Antti, Kotikoski Sampo		
Assigned by JAMK University of Applied Sciences / IT-institute, CYBERDI-project		
Abstract <p>People may not realize what kind of information they publish on different social media channels and how easily this information is available to others. The importance of social media platforms and other open sources in personal data gathering is great because people themselves publish information, which can be combined to build an entity that in the worst-case scenario can lead to identity theft, for example.</p> <p>The thesis was carried out using a qualitative research method. The purpose of the thesis was to reveal what kind of information of social media users can be retrieved with the help of different tools and whether one can hide one's own personal information on different social media platforms. The thesis was done for the CYBERDI project, which is a project whose main purpose is to raise awareness in our society about threats related to the digital world and to combat cybercrime.</p> <p>The study began by making profiles for different social media platforms. After this, familiarized with the tools utilizing social media platforms and figured out their features and characteristics. Practical work was done with a total of seven different tools that were selected for different social media platforms and tests were performed using these tools. The tools were chosen on the basis that they were popular, easy to use and free.</p> <p>The tools were used to build various reports on the target person's behavior on different platforms and information on the objects of this interest, for example. Information of the target person's public posts were covered in reports, but hidden, personal information was not found using the tools that were chosen for the thesis.</p>		
Keywords/tags (subjects) OSINT, SOCMINT, personal data gathering, social media		
Miscellaneous (Confidential information)		

Sisältö

1	Johdanto	4
1.1	Tavoitteet.....	4
1.2	Toimeksiantaja	4
2	Tutkimuksen lähtökohdat	5
2.1	Tutkimusongelma ja -kysymykset	5
2.2	Tutkimusmenetelmät.....	5
2.3	Tutkimusasetelma.....	6
3	Teoreettinen viitekehys	7
3.1	Social media intelligence gathering (SOCMINT)	7
3.2	Sosiaalisen median alustat	8
3.2.1	Facebook	9
3.2.2	Twitter	9
3.2.3	Instagram.....	10
3.3	OSINT-työkalut	10
3.3.1	Twitonomy	10
3.3.2	Tinfoleak.....	11
3.3.3	Sleeping time	11
3.3.4	Social bearing	12
3.3.5	Searchusers	13
3.3.6	Downloadgram	13
3.3.7	Osintgram.....	13
4	Toiminnallinen viitekehys	15
4.1	Profiilin luonti.....	15
4.2	Työkalujen käyttö.....	16
4.2.1	Twitonomy	16
4.2.2	Tinfoleak.....	22
4.2.3	Sleeping time	28
4.2.4	Social bearing	30
4.2.5	Searchusers	35

	2
4.2.6 Downloadgram	36
4.2.7 Osintgram.....	36
5 Tutkimustulokset	43
5.1 Tulokset.....	43
5.2 Reliaabelius ja validius.....	44
6 Johtopäätökset	45
7 Pohdinta	46
Lähteet	47

Kuviot

Kuvio 1. Suosituimmat sosiaalisen median alustat (Digital 2021: Global overview report, 2021.)	8
Kuvio 2. Social Bearing - etusivu	12
Kuvio 3. Osintgram - requirements.txt	14
Kuvio 4. Twitonomy - yhteenveto profiilista.....	16
Kuvio 5. Twitonomy – tweetit	17
Kuvio 6. Twitonomy - Tweets Analytics	18
Kuvio 7. Twitonomy - Tweet history.....	19
Kuvio 8. Twitonomy – Twiittausaikatauluja.....	20
Kuvio 9. Twitonomy – Alustat, joista twiitattu eniten	20
Kuvio 10. Twitonomy - Seuraajat ja seuratut profiilit	21
Kuvio 11. Twitonomy – Tykkäykset	22
Kuvio 12. Twitonomy – listat	22
Kuvio 13. Tinfoleak – etusivu.....	23
Kuvio 14. Tinfoleak – Yleistiedot	24
Kuvio 15. Tinfoleak - Applikaatiot ja sosiaalisen median alustat	24
Kuvio 16. Tinfoleak – Hashtagit	25

Kuvio 17. Tinfoleak - Merkinnät käyttäjistä	26
Kuvio 18. Tinfoleak - Tykkäykset ja tweetit	26
Kuvio 19. Tinfoleak - Useimmiten käytetyt sanat	27
Kuvio 20. Tinfoleak – Metadata.....	27
Kuvio 21. Tinfoleak - Maantieteelliset tiedot.....	28
Kuvio 22. Sleeping time – Yhteenveto	29
Kuvio 23. Social bearing – Yhteenveto	30
Kuvio 24. Social bearing – Yleistiedot	31
Kuvio 25. Social bearing - Twiittien analytiikka.....	32
Kuvio 26. Social bearing - Mistä twiitattu ja millä kielellä.....	32
Kuvio 27. Social bearing - Jaetut domainit.....	33
Kuvio 28. Social bearing - Käytetyt hashtagit.....	34
Kuvio 29. Social bearing – Word cloud, twiiteissä käytetyt sanat.....	34
Kuvio 30. Searchusers – hakutulosten etusivu	35
Kuvio 31. Searchusers – uusin postaus, tykätyn kuva ja kommentoiduin kuva .	36
Kuvio 32. Osintgram – hakuvaihtoehdot	37
Kuvio 33. Osintgram - kuvien osoitetiedot	38
Kuvio 34. Osintgram – kuvatekstit.....	38
Kuvio 35. Osintgram – kommenttidata.....	39
Kuvio 36. Osintgram – kommenttimäärä.....	39
Kuvio 37. Osintgram - seuraajat ja seurattut profiilit.....	39
Kuvio 38. Osintgram - seurattujen profiilien sähköpostiosoitteet	40
Kuvio 39. Osintgram - seurattujen profiilien puhelinnumerot.....	40
Kuvio 40. Osintgram – aihetunnisteet	41
Kuvio 41. Osintgram – info	41
Kuvio 42. Osintgram – tykkäykset	41
Kuvio 43. Osintgram – mediatyypit	42
Kuvio 44. Osintgram – kommentoijat.....	42

1 Johdanto

1.1 Tavoitteet

Opinnäytetyön tavoitteena oli lisätä tietoisuutta henkilötiedustelun mahdollisuuksista kaikille sosiaalisen median alustojen käyttäjille ja selvittää, kuinka helposti sosiaalisen median kanavan käyttäjästä saa tietoa, millaista tietoa löytyy sekä millaisia työkaluja tähän tarkoitukseen on saatavilla. Työ oli tarkoitus rajata kolmeen suosittuun ja sisällöltään erilaiseen sosiaalisen median kanavaan ja näistä kanavista tiedon hankkimiseen. Rajaukseen valitut kolme sosiaalisen median kanavaa ovat Facebook, Twitter ja Instagram. Opinnäytetyötä varten luotiin edellä mainittuihin palveluihin tunnukset, joihin tutkimuksen perusteella valittujen työkalujen käyttö kohdistui. Käytännön työ toteutettiin lopulta vain kahteen edellä mainituista sosiaalisen median alustoista, sillä Facebookin turvallisuuspolitiikan myötä käyttäjätunnusta ei saatu siellä ylläpidettyä. Tästä on kerrottu tarkemmin luvussa 3.2.1.

Internetin ”pimeän puolen” eli dark webin työkaluihin ei tässä työssä perehdytä, vaikka sielläkin olisi tarjolla ilmaisia, henkilötiedusteluun käytettäviä työkaluja. Nämä työkalut eivät ole tälle työlle relevantteja, koska koko pimeän verkon kompleksi on hyvin laaja ja vaatisi tarkempaa ymmärrystä siellä olevista työkaluista ja niiden laillisuudesta, mikäli niitä lähdetäisiin tutkimaan tarkemmin.

1.2 Toimeksiantaja

Työn toimeksiantaja oli Jyväskylän ammattikorkeakoulun ja Poliisiammattikorkeakoulun (POLAMK) yhteistyössä toteuttama CYBERDI-projekti. CYBERDI-projektin tavoitteina on kasvattaa yhteiskunnassamme tietoisuutta digitaaliseen maailmaan liittyvistä uhkista, torjua kyberrikollisuutta kehittämällä ja hyödyntämällä moderneja teknologioita, sekä vahvistaa samalla yhteistyökumppanuuksia tekemällä aktiivista tutkimus- ja kehitysyhteistyötä erilaisissa ryhmissä ja verkostoissa. (CYBERDI – Kansallista & kansainvälistä kyberosaamista kasvattamassa. N.d.)

2 Tutkimuksen lähtökohdat

2.1 Tutkimusongelma ja -kysymykset

Tämän työn tutkimusongelma keskittyy henkilötiedustelun uhkien ympärille sosiaalisen median alustoilla. Sosiaalisissa medioissa, kuten Facebookissa, ihmiset luovat alustalle käyttäjätunnuksen. Käyttäjätunnuksen taakse sisältyy valtavia määriä henkilötietoja, jotka ovat helposti saatavilla kenen tahansa toimesta, varsinkin, jos henkilöt ovat asettaneet tiedot julkisiksi. Sosiaalisen median käyttäjät eivät välttämättä ymmärrä, kuinka paljon informaatiota heistä saadaan kerättyä, koska he ovat tehneet käyttäjätunnukset eri sosiaalisen median alustoille ja julkaisseet siellä tietoja. Työssä selvitettiin millaisen vaaran sosiaalisen median alustoilta tapahtuva tiedustelu muodostaa sosiaalisen median alustojen käyttäjille.

Opinnäytetyön tutkimuskysymykset ovat seuraavanlaiset:

- Kuinka helppo sosiaalisen median käyttäjän profiilista on onkia tietoa?
- Mitä työkaluja tiedon hankkimiseen sosiaalisen median alustoista löytyy?
- Vaikuttaako profiilitietojen piilottaminen työkaluilla saataviin tietoihin?

2.2 Tutkimusmenetelmät

Tutkimusmenetelmät jaetaan yleisesti kvantitatiivisiin eli määrällisiin tutkimuksiin ja kvalitatiivisiin eli laadullisiin tutkimuksiin. Vaikkakaan näiden välisiä eroja ei voida yksiselitteisesti avata, niin lähtökohtaisesti voidaan kvalitatiivista tutkimusta pitää ymmärtävänä ja kvantitatiivista selittävänä tutkimuksena. Laadullisessa tutkimuksessa tutkitaan kohdetta kokonaisvaltaisesti ja sillä pyritään ymmärtämään kohteen ominaisuuksia ja laatua. (Tuomi & Sarajärvi, 2018, 72-73.)

Työ toteutettiin tutkimuksellisena opinnäytetyönä ja tutkimustyyppiltään se on laadullinen eli kvalitatiivinen tutkimus. Työn tutkimusmenetelmänä käytettiin sisällönanalyysiä. Sisällönanalyysi on perusanalyysimenetelmä, jonka tarkoituksena on luoda selkeä ja sanallinen kuvaus tutkittavasta ilmiöstä. Työssä tutkittiin henkilötiedustelua

sosiaalisen median avulla ilmiönä. Aineistoa kerättiin internetistä erilaisia työkaluja hyödyntäen, työkaluista on kerrottu tarkemmin luvussa 3.3. (Tuomi & Sarajärvi, 2018, 91-92, 108.)

2.3 Tutkimusasetelma

Työtä varten luotiin eri sosiaalisen median kanaviin profiilit, joita käytettiin tiedonhankinnan kohteena. Eri sosiaalisen median kanavissa julkaistiin erilaisia julkaisuja tämän profiilin avulla, mutta profiilin perustiedot (ikä, sukupuoli, nimi, sähköpostiosoite) ovat jokaisella käytetyllä alustalla samat. Tavoitteena on tehdä mahdollisimman todentuntuinen profiili, jotta pystytään osoittamaan työssä käytettyjen työkalujen toimivuus tositilanteessa. Kaikki muu tieto asetettiin julkiseksi, mutta henkilökohtaiset tiedot (ikä, sukupuoli ja sähköpostiosoite) jätettiin kaikilla alustoilla piiloon.

Opinnäytetyössä hankittiin tietoa näistä profiileista sosiaalisen median tiedustelun avulla. Tähän tarkoitukseen etsittiin erilaisia työkaluja, joita hyödynnettiin tiedonhankinnassa. Työkalujen etsintään käytettiin hyödyksi Googlen hakukonetta ja myöhemmin työkaluja etsittiin suoraan Github.com -verkkosivuston kautta, kun havaittiin sieltä löytyvän suuri määrä erilaisia tarkoitukseen sopivia työkaluja. Työkalut valittiin sillä perusteella, että ne olivat ilmaisia, avoimen lähteen tiedusteluun tarkoitettuja työkaluja, joilla pystyi kohdentamaan haun tiettyyn sosiaalisen median alustan käyttäjään. Työkalut on esitelty luvussa 3.3 ja työkalujen hyödyntämisestä ja käyttötaivoista on kerrottu tarkemmin luvussa 4.2.

3 Teoreettinen viitekehys

3.1 Social media intelligence gathering (SOCMINT)

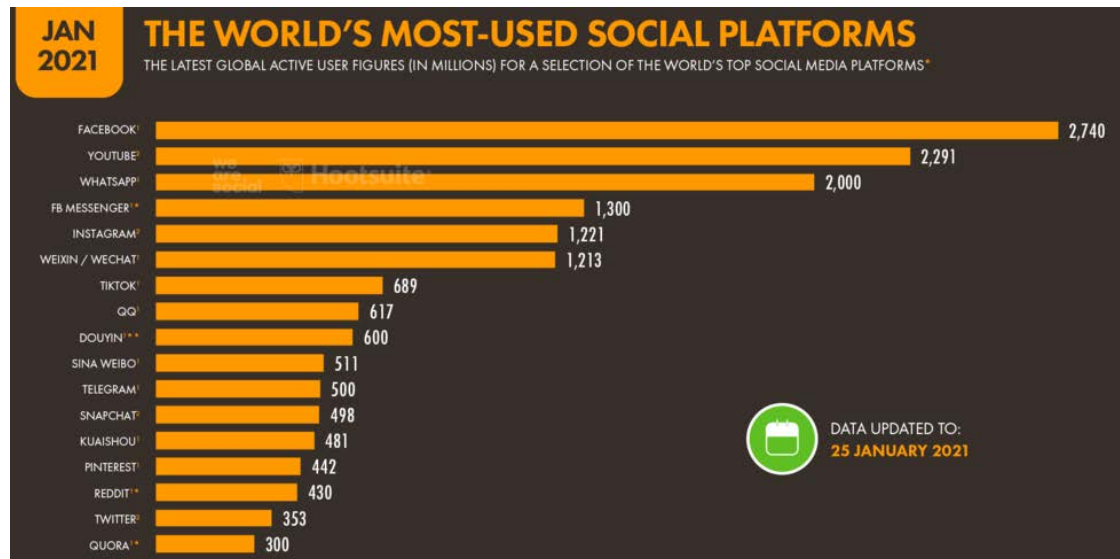
Avointen lähteiden tiedustelua on tehty jo satoja vuosia eri metodeja ja kanavia hyväksikäyttäen. Sota-aikoina tiedustelua tehtiin esimerkiksi hankkimalla vihollisten strategiatietoja seuraamalla heidän uutislähteitensä, mutta nykyään tiedustelua tehdään suurimmaksi osaksi hyväksikäyttäen internetiä ja siellä olevia avoimia lähteitä. (Schaurer & Störger, 2013.)

Avointen lähteiden tiedustelu (Open Source Intelligence, OSINT) tarkoittaa käytännössä avoimista lähteistä kerättyä tietoa ja näiden tietojen yhdistämistä ja analysointia. Avoimiksi lähteiksi luetaan kaikki julkinen tieto, joka on kaikille saatavilla avoimesti ja laillisesti, esimerkiksi verkossa eri sosiaalisen median kanaviin jaettu informaatio, blogipostaukset sekä sanomalehdet ja esitteet. Näistä kanavista saatuja tietoja voidaan hyödyntää yhdistämällä eri lähteistä löytyneet tiedot ja rakentamalla näistä yksi kokonaisuus – esimerkiksi henkilöprofilointia tai yrityksen tietoturvakartoitusta varten. (Hassan, 2020.)

SOCMINT (Social Media Intelligence Gathering) on yksi OSINTin alakategorioista. SOCMINT tarkoittaa yksinomaan sosiaalisen median kanavista saadun tiedon perusteella tapahtuvaa tiedustelua. Tällaisia kanavia ovat muun muassa Facebook, Twitter, Instagram, LinkedIn, Reddit, Pinterest ja Youtube. SOCMINTia sovelletaan nykyään yhä useampaan eri tarkoitukseen, esimerkiksi tiedustelupalvelut hankkivat ja analysoivat tietoja sosiaalisen median kanavista seuratakseen tietynlaisia uhkia, tai yritykset saattavat etsiä tietoa henkilöistä ennen heidän palkkaamistansa rakentaakseen tarkemman henkilökuvan heistä. (Hassan, 2020.)

3.2 Sosiaalisen median alustat

Työtä varten valittiin sosiaalisen median alustat sen perusteella, että ne olivat suosituimpien alustojen listalla ja että niihin löytyi lyhyen alkuselvityksen aikana myös ilmaisia avoimen lähdekoodin työkaluja henkilötiedustelua varten.



Kuvio 1. Suosituimmat sosiaalisen median alustat (Digital 2021: Global overview report, 2021.)

Kuviossa 1 nähdään tarkemmin suosituimpien sosiaalisen median alustojen käyttäjämäärät vuoden 2021 alussa, luvut kuviossa miljoonina.

Jokaisen työssä läpikäydyn sosiaalisen median kanavan asetuksista löytyy vähintään asetus, jolla oman profiilin tai tietyn postauksen saa poistettua. Tämä ei kuitenkaan tarkoita, etteikö aiemmin julkisena olleet tiedot olisi enää saatavilla, sillä nykyään myös tähän on rakennettu erilaisia työkaluja. Tämän vuoksi voidaan olettaa, että kaikki tieto, joka on julkaistu internetiin, on tallennettuna johonkin ja täten sitä ei voida täysin poistaa. Jotkin sosiaalisen median alustat tarjoavat mahdollisuuden piilottaa tietyt julkaisut, mutta mikäli nämä ovat aiemmin olleet kaikkien nähtävillä eli julkisia julkaisuja, näiden piilottaminen tai poistaminen vain mahdollisesti hankaloittaa tämän tiedon etsimistä jälkikäteen.

3.2.1 Facebook

Facebook on lähes kaksikymmentä vuotta vanha sosiaalisen median alusta, joka sai perustamisvuonnaan 2004 heti yli miljoona käyttäjää alustallensa. Tämän työn kirjoitushetkellä Facebookilla on lähes 2,8 miljardia aktiivista kuukausittaista käyttäjää ja sen käyttöliittymä toimii yli sadalla eri kielellä. Facebook on vuosien varrella ostanut itselleen myös muun muassa Instagramin ja Whatsappin, mutta nämä ovat kuitenkin sisällöltään hieman erilaisia alustoja. (Chen, 2021.)

Facebookiin oli tarkoitus tehdä profiili tätä työtä varten, mutta Facebookin turvallisuusominaisuuksien vuoksi tämä osoittautui hieman haasteelliseksi. Profiili saatiin luotua, mutta heti seuraavan kirjautumisyrittäksen jälkeen alkoi Facebook kyselemään puhelinnumeroa profiilin varmentamiseksi. Tätä varten ei ollut antaa puhelinnumeroa tilin varmentamiseksi, joten tili lukkiutui ennen kuin se ehdittiin kunnolla ottaa edes käyttöön. Tästä syystä työssä ei käsitellä työkaluja, joilla saataisiin mahdollisesti hankittua nimenomaan Facebook-tileistä tietoa.

3.2.2 Twitter

Twitter on vuonna 2006 julkistettu mikroblogi- ja sosiaalisen median kanava, johon käyttäjät voivat syöttää maksimissaan 280 merkkiä pitkän päivityksen lähes mitä tahansa älylaitetta käyttäen. Twitterin sisältöä pystyy lukemaan ilman rekisteröitymistä, mutta halutessaan itse julkaista viestejä (twiittejä), täytyy palveluun tehdä ilmainen henkilökohtainen tunnus. Twitterillä on työn kirjoitushetkellä yli 350 miljoonaa aktiivista kuukausittaista käyttäjää (kuvio 1). (Zeevi, 2013.)

Twitteriä käytetään nykyään yksityiselämästä postauksen lisäksi myös ammatillisessa tarkoituksessa. Esimerkiksi journalistit ja poliitikot ovat yleisesti hyvin aktiivisia kyseisen palvelun käyttäjiä, jotka yrittävät saada jaettua agendaansa muille Twitterin käyttäjille tai saada heiltä informaatiota asioista Twitterin välityksellä. Twitterin avulla voi helposti myös parantaa oman yrityksensä imagoa mainostamalla tätä palvelussa, sillä siellä viestit tavoittavat helposti suuren yleisön. (Zeevi, 2013.)

3.2.3 Instagram

Instagram on vuonna 2010 Kevin Systromin ja Mike Kriegerin julkaisema sosiaalisen median alusta, jonka käyttäjämäärä kasvoi räjähdysmäisesti heti julkaisupäivästä alkaen – Instagramin latauskertoja kertyi yli miljoona kappaletta vain muutamassa kuukaudessa. Sovelluksen perinpohjainen tarkoitus on toimia alustana, jossa käyttäjät voivat jakaa kuvia ja videoita sekä kommentoida ja tykätä näistä. (Blystone, 2020.)

Instagramin perustajat möivät alustansa Facebookille miljardilla dollarilla vuonna 2012. Tämän jälkeen Instagramin suosio on vain kasvanut ja nykyään kyseisellä alustalla on yli miljardi aktiivista käyttäjää kuukaudessa (kuvio 1) ja sen markkina-arvon lasketaan olevan yli 100 miljardia dollaria. (Simon, 2021.)

3.3 OSINT-työkalut

OSINT-työkaluja on saatavilla nykypäivänä jo todella paljon ja näitä ilmestyy verkkoon koko ajan vain enemmän, kun ihmisiä alkaa kiinnostaa mitä tietoa heistä löytyy verkosta ja kuinka helposti tämä tieto on saatavilla. Nämä työkalut, joita tässä työssä käytettiin, ovat kaikki ilmaisia, avoimen lähdekoodin työkaluja. Näiden lisäksi löytyy myös useita maksullisia työkaluja, jotka yleensä pystyvät antamaan enemmän ja vielä tarkempaa tietoa tiedustelun kohteesta ja tähän liittyvästä datasta, mutta näihin työkaluihin ei tarkemmin tutustuttu niiden maksullisuuden vuoksi. Työkalut valittiin aiempien kriteereiden lisäksi myös sen perusteella, että ne olivat suosittuja, helppokäyttöisiä ja nämä pystyvät hankkimaan paljon erilaista tietoa profiileista.

3.3.1 Twitonomy

Twitonomy on ilmainen Diginomy Pty Ltd:n kehittämä työkalu julkisen datan hakemiseen Twitteristä käyttäen Twitterin ohjelmointirajapintaa (API). Työkalun käyttöä varten täytyy Twitonomyllle luovuttaa oman Twitter-tilin tunnukset, sillä Twitterin sisältöä on hankala lukea, mikäli sinne ei ole kirjautunut jollain käyttäjätunnuksella sisään – tässä tapauksessa käytettiin opinnäytetyön tekijän omia henkilökohtaisia tunnuk-siani. (Diginomy Pty Ltd, 2021.)

Twitonomy tarjoaa käyttäjilleen myös maksullista Twitonomy Premium-palvelua, joka sisältää muun muassa tarkemmat hakuehdot ja mahdollistaa tietojen lataamisen Excel- tai PDF-tiedostoina haluamalleen laitteelle. Premium-palvelun ominaisuuksiin ei perehdytty tämän tarkemmin, vaan työssä keskityttiin työkalun ilmaisiin ominaisuuksiin, koska työn tarkoituksena oli saada henkilötietoja käyttäen avuksi vain ilmaisia työkaluja. (Diginomy Pty Ltd, 2021.)

3.3.2 Tinfoleak

Tinfoleak on Vicente Aguilera Diazin kehittämä OSINT-työkalu Twitteristä saatavan datan analysointiin. Työkalun kehittäjä kertoo Tinfoleakin Github-sivulla tämän olevan kehittynein OSINT-työkalu Twitter-analysointiin, sillä Tinfoleakilla saa kaivettua hyvin paljon erilaista tietoa kohdehenkilön Twitter-tilistä.

Tinfoleak on hyvin laajasti käytetty työkalu ja se sisältyy nykyään oletuksena useassa eri Linux-jakelupaketeissa, esimerkiksi Kali, Caine ja BlackArch-jakelupaketeissa. (Diaz, 2018.)

Tinfoleakin pystyy itse asentamaan ja määrittämään haluamalleen koneelle tai vaihtoehtoisesti voi käyttää tätä varten kehitettyä verkkosivustoa – tämä toimii samalla periaatteella, käyttäen samaa python-skriptiä taustalla.

Tässä työssä Tinfoleakia käytettiin kyseisen verkkosivuston kautta, sillä se oli helposti saatavilla, helppokäyttöinen sekä selkeästi toteutettu. Työkalua varten tarvitsi kertoa sivustolle vain kohdehenkilön Twitter-tunnus sekä sähköposti, johon lähetettiin linkki raporttia varten. Sähköpostiin tuli hetken päästä linkki raporttiin, jossa näkyi löytynyt data – linkin avaamiseen ei tarvinnut tunnuksia eikä siinä vaikuttanut olevan muitakaan käyttämistä rajoittavia tekijöitä. (Diaz, 2018.)

3.3.3 Sleeping time

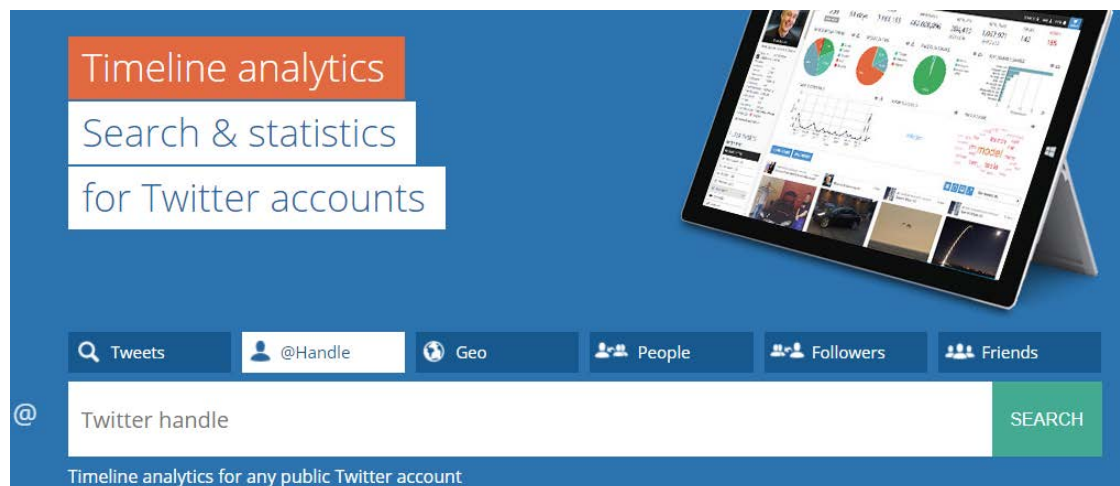
Sleeping Time on Amit Agarwalin kehittämä, hänen Digital Inspiration -projektiin liittyvä Twitter-työkalu, jolla voi hakea haluamansa Twitter-käyttäjän keskimääräisen twiittaus- ja nukkuma-aikataulun. Työkalu toimii sillä periaatteella, että se hakee kohdehenkilön 1000 viimeisintä twiittiä, tarkistaa niistä julkaisukohdan ja kertoo hakutuloksissa, millä aikavälillä käyttäjä saadun datan perusteella nukkuu – eli ei lähetä

twiittejä. Kyseinen työkalu näyttää myös kohdehenkilön nimen sekä profiiliinsa määrittämän kuvaustekstin. (Sleeping time, N.d.)

Työkalu on ilmainen, se toimii verkkosivuston *sleepingtime.org* -kautta ja se vaatii vain sisäänkirjautumisen jollain Twitter-tunnuksella, jolloin työkalulla voi hakea kenen tahansa julkisen Twitter-käyttäjän nukkuma-aikataulun.

3.3.4 Social bearing

Social Bearing on Tom Elliottin kehittämä ilmainen, verkkosivuston kautta toimiva ja Twitterin ohjelmointirajapintaa hyödyntävä OSINT-työkalu. Työkalulla voi hakea hyvin paljon erilaista statistiikkaa ja tietoa julkisesta Twitter-profiilista, esimerkiksi viimeisimmät twiitit, käytetyimmät sanat twiiteissä tai kieli, jota on eniten käytetty kohdehenkilön twiiteissä. Kuviossa 2 näkyy Social bearingin helppokäyttöiseksi ja selkeäksi tehty etusivu ja hakuvaihtoehdot. (Social Bearing, N.d.)



Kuvio 2. Social Bearing - etusivu

Social Bearing -työkalulla voi hakea myös esimerkiksi tietyllä sanalla ja maantieteellisellä lokaatiolla twiittejä kaikilta käyttäjiltä, mutta työssä käytettiin vain @Handle-ominaisuutta. Hakukenttään syötettiin kohdehenkilön Twitter-tunnus ja työkalu hoiti kaiken muun. Näin saatiin paljon erilaista dataa liittyen yhteen profiiliin.

3.3.5 Searchusers

Searchusers on työkalu datan hankkimiseen Instagram-profiileista. Työkalu ei vaadi rekisteröitymistä ja se on täysin ilmainen käyttää. Työkalun käyttö vaatii vain kohdeprofiilin Instagram-käyttäjätunnuksen, joka syötetään <https://searchusers.com> -verkkosivuston etusivulla olevaan hakulaatikkoon. Työkalu tekee tämän jälkeen kaiken työn ja koostaa raportin kohdeprofiilista löytyneistä tiedoista. Tarkemmat hakutulokset löytyvät luvusta 4.2.5. (Searchusers, N.d.)

3.3.6 Downloadgram

Downloadgram on yksi monesta verkkosivustosta, jonka avulla voi ladata kohdeprofiilin Instagram-tililtä kuvat ja videot talteen haluamalleen laitteelle. Käytännössä saman voi tehdä myös itse, mikäli osaa katsoa sivuston lähdekoodia, josta pääsee hakemaan suoraan linkin parempilaatuiseen kuvaan ja lataamaan tätä kautta kuvan laitteelle – Downloadgram tekee tämän kuitenkin helpommaksi. Downloadgram-työkalulle syötettiin vain kohdeprofiilin kuvan tai videon URL, klikattiin ”Download”, minkä jälkeen kuvan sai ladattua laitteellensa. (Downloadgram, N.d.)

3.3.7 Osintgram

Osintgram on Githubista löytyvä Giuseppe Criscionen kehittämä OSINT-työkalu, jolla pystyy hakemaan kaikenlaista tietoa Instagram-profiileista. Työkalun käyttö vaatii ensin tämän ohjelman asennuksen, joten tähän tarkoitukseen käytettiin virtuaalikoneetta, johon on asennettu helppokäyttöinen Ubuntu-käyttäjärjestelmä. Osintgram käyttää hyödykseen Instagramin ohjelmointirajapintaa (API). (Osintgram, N.d.)

Työkalun toimintaa varten asennettiin virtuaalikoneelle git, python3 sekä python3-pip, avaamalla terminaali ja syöttämällä komennot: *sudo apt install git*, *sudo apt install python3*, sekä *sudo apt install python3-pip*. Näiden asennusten jälkeen itse työkalu kloonattiin Githubista komennolla: *git clone https://github.com/Datalux/Osintgram.git*. Työkalu ei suoraan näillä komennoilla vielä lähtenyt toimimaan, sillä python-skriptiä ajettaessa tuli virhesanoma, mutta pienen selvittelyn jälkeen ongelmat saatiin ratkottua asentamalla sovelluskehittäjien kirjasto komennolla *sudo apt install*

libreadline-dev. Näiden jälkeen asennettiin vielä loput tarvittavat sovellukset komennolla `pip3 install -r requirements.txt`. Requirements.txt -tiedoston sisältö löytyy kuvio 3. (Osintgram, N.d.)

```
1 requests==2.24.0
2 requests-toolbelt==0.9.1
3 geopy>=2.0.0
4 prettytable==0.7.2
5 instagram-private-api==1.6.0
6 gnureadline>=8.0.0; platform_system != "Windows"
7 pyreadline==2.1; platform_system == "Windows"
```

Kuvio 3. Osintgram - requirements.txt

Asennusten jälkeen asetettiin vielä Instagram-käyttäjätunnus ja -salasana *config*-kansion sisällä olevaan *credentials.ini*-tiedostoon, sillä Instagramista ei pääse lukemaan dataa, mikäli ei ole kirjautunut sinne jollakin käyttäjällä sisään. Tähän tarkoitukseen käytettiin testikäyttäjämme tunnuksia.

Työkalua pääsi käyttämään asennusten jälkeen ajamalla komento, jossa määritetään, mistä käyttäjätunnuksesta lähdetään dataa etsimään: `python3 main.py opparitepi`. Hakutermeistä ja löytyneestä datasta on mainittu tarkemmin luvussa 4.2.7.

4 Toiminnallinen viitekehys

4.1 Profiilin luonti

Työhön valikoiduilla sosiaalisen median alustoilla on jokaisella hieman erilainen käyttötarkoitus. Facebook on hyvin laaja alusta sisällöltään, sinne pystyy jakamaan kuvia ja videoita lähestulkoon missä muodossa tahansa ja jakaa ne valitulle ryhmälle, Twitter keskittyy lähinnä mikroblogimaiseen ympäristöön merkkirajansa kanssa, kun taas Instagramin pääasiallinen tarkoitus on jakaa kuvia ja videoita, ja kommentoida näitä. Tästä syystä profiilin luontikin täytyy tehdä jokaiseen alustaan hieman eri tavalla, vaikka testikäyttäjämme profiilien pääpiirteet ja lähtökohdat ovat samat kaikilla alustoilla. Kuten aiemmin jo mainittiin, Facebookiinkin ehdimme tehdä tunnuksen, mutta testikäyttäjän tili lukkiutui ennen aikaisesti, joten tämän alustan profiilin tiedot eivät ole relevantteja työlle.

Twitteriin luotiin 22.10.2020 tunnus *@OppariTepi*, jonka käyttäjä on 13.6.1971 syntynyt miespuolinen henkilö nimeltä Tepi Oppari. Syntymäaikaa ei julkaistu julkisena tietona profiiliin, eli tämä pitäisi olla vain käyttäjän omana tietona eikä näkyä muille käyttäjille. Twitteriin julkaistiin muutaman päivän välein twiittejä (yhteensä kahdeksan kappaletta), joista yksi poistettiin myöhemmin, jotta voidaan tarkkailla, pystyvätkö työkalut näyttämään kertaalleen poistettuja tietoja.

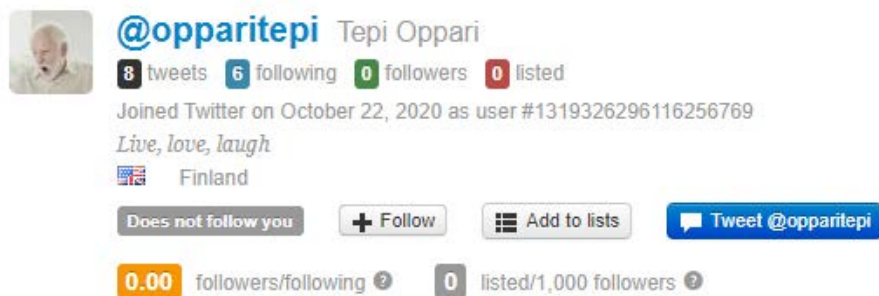
Instagramiin tehtiin tunnus *@tepioppari*, jonka käyttäjä on miespuolinen henkilö nimeltä Tepi Oppari. Sukupuoli ja sähköpostiosoite olivat henkilökohtaisia tietoja, eli niitä ei jaettu julkisesti muille. Instagramiin ei saatu suoraan verkkoselaimen kautta lisättyä kuvia, joten tässä hyödynnettiin Bluestacks-nimistä Android-emulaattoria. Käytännössä siis kyseinen sovellus ladattiin koneelle, ladattiin Play-kaupasta Instagram, kirjauduttiin testitunnuksilla ja ladattiin tilille tekijänoikeusvapaa kuva, joka hankittiin shutterstock.com -sivustolta ilmaiseksi. Myös testikäyttäjän profiilikuva on haettu samalta sivustolta.

4.2 Työkalujen käyttö

4.2.1 Twitonomy

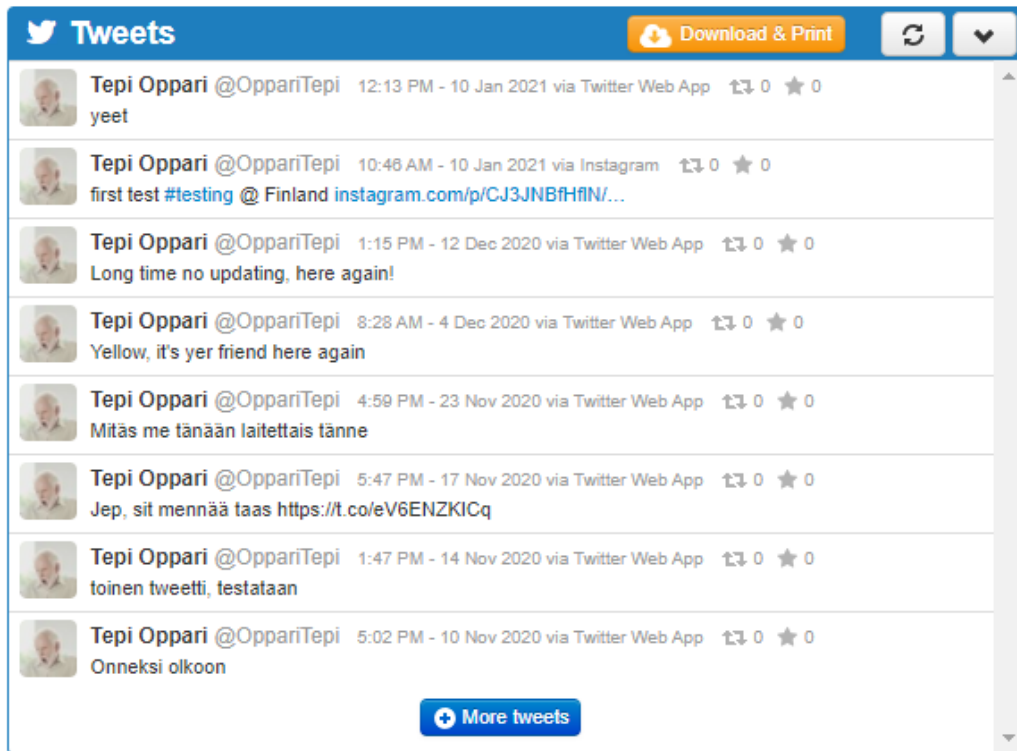
Twitonomyyn kirjautumisen jälkeen työkalua pystyi heti käyttämään ja dataa pystyi etsimään mistä tahansa julkisesta Twitter-profiilista käyttämällä ”Analyze a Twitter profile” -hakua. Hakuun laitettiin käyttäjätunnus, jolta haluttiin dataa hakea, tässä tapauksessa ”@opparitepi”.

Kuviossa 4 näemme yhteenvedon profiilista, johon haku kohdistui. Tässä näemme listattuna muun muassa, kuinka monta kertaa käyttäjä on twiitannut, kuinka montaa eri käyttäjää hän seuraa Twitterissä tai kuinka monta seuraajaa hänellä on. Yhteenvedosta löytyy myös profiilikuva, käyttäjätunnus, nimi sekä päivämäärä, jolloin käyttäjä on Twitteriin liittynyt.



Kuvio 4. Twitonomy - yhteenvedo profiilista

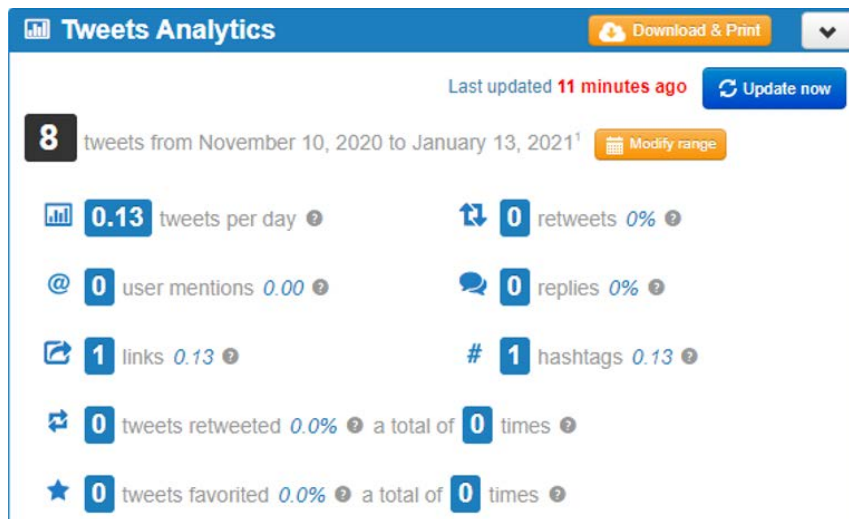
Kuviosta 5 näemme kaikki twiitit, jotka kohdekäyttäjällä on postannut Twitteriin. Tässä tulee huomata, että Twitonomy hakee kaikki twiitit, mitä profiilissa on tällä hetkellä näkyvissä, eli käyttäjän poistamia twiittejä ei tähän listaan tule näkyviin.



Kuvio 5. Twitonomy – tweetit

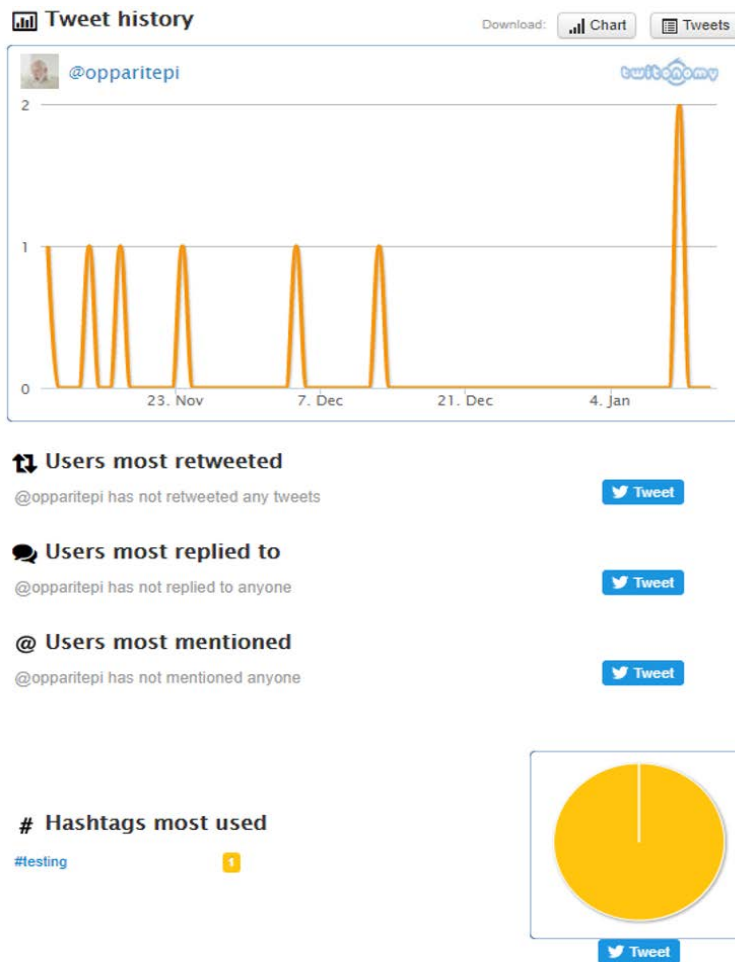
Analytiikkakuvassa (kuvio 6) nähdään selkeässä muodossa tarkempaa tietoa käyttäjän profiilista ja twiiteistä. Kuvioista nähdään mm. kahden desimaalin tarkkuudella, kuinka monta twiittiä käyttäjä postaa keskimäärin päivässä (kohdehenkilöllä kuvanottohetkellä twiittimäärä on 0,13 twiittiä päivässä), kuinka monta kertaa hän on uudelleentwiitannut sekä kuinka moneen twiittiin hän on vastannut. Twitonomy mainitsee tästä kohdasta sivuillaan tarkennuksen, että työkalu hakee tähän analysointikuvaan maksimissaan 3200 uusinta twiittiä käyttäjältä.

Tässä tulee huomioida, että testaushetkellä twiittejä oli yhteensä kahdeksan kappaletta, joista yksi poistettiin myöhempiä testejä varten. Twitonomy ei siis tehnyt raportissaan virhettä, vaan näytti tiedot oikein. Yhden twiitin poiston jälkeen ajoin Twitonomyn raportin uudestaan ja tällöin twiittien yhteismääräksi tuli seitsemän.



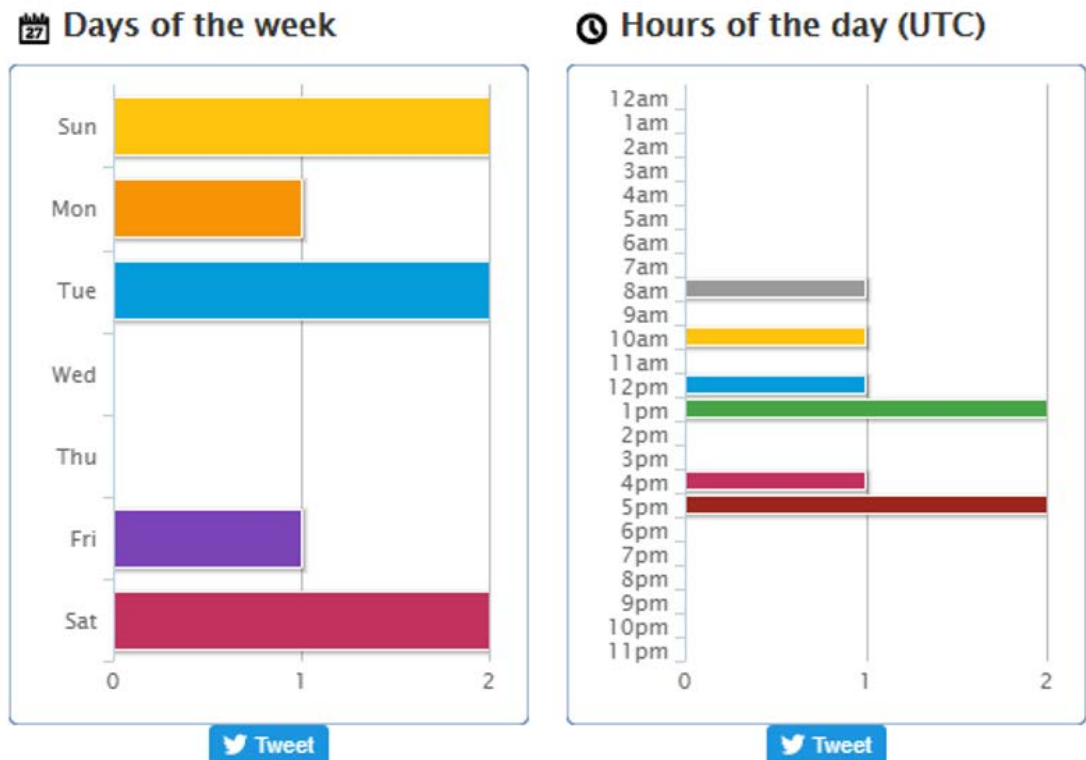
Kuvio 6. Twitonomy - Tweets Analytics

Twitonomy kertoi myös kaaviomuodossa profiilin twiittihistoriasta, kuviossa 7 näkyy tarkemmin nämä tulokset. Twitonomyn antamat tulokset tässä kaaviossa twiittien päivämäärästä eivät ole niin tarkat kuin esimerkiksi kuviossa 5 nähdään, mutta tämä kuitenkin näyttää tiedot visuaalisemmassa muodossa tällä otsakkeella. Kyseisessä kuviossa näkyy myös uudelleentwiittaukset, vastaukset twiitteihin, eniten kohdeprofiilin twiiteissä mainitut käyttäjät sekä käytetyimmät aihetunnisteet.



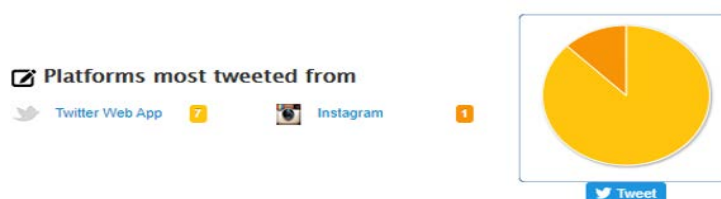
Kuvio 7. Twitonomy - Tweet history

Kuvion 8 aikatauluista nähdään helposti mihin aikaan ja minä päivänä käyttäjä on useimmiten lähettänyt twiittejä. Värit ovat tässä mukana vain erottaakseen päivät ja ajat toisistaan. Kohteena ollut käyttäjä (eli tässä tapauksessa testikäyttäjämme) on lähettänyt twiittejä eniten tiistaisin, lauantaisin ja sunnuntaisin maksimissaan kaksi twiittiä päivässä. Kellonajallisesti eniten twiittejä on lähetetty kello 13 ja 17, muina ajankohtina on lähtenyt vain maksimissaan yksi twiitti tunnissa. Nämä tiedot vaikuttavat olevan oikeassa ja tämä on myös helppo tarkistaa, sillä Twitter näyttää kaikissa julkisissa julkaisuissa minuutin tarkkuudella postauksen ajankohdan.



Kuvio 8. Twitonomy – Twiittausaikatauluja

Testiprofiililla julkaistiin twiittejä Twitterin verkkosivun kautta selaimella sekä kerran Instagramin jakoa käyttäen. Kuviossa 9 näkyy Twitonomyn löytämät sovellukset, joista tälle Twitter-profiilille on julkaistu materiaalia ja tässä näkyy, että seitsemän twiittiä on tehty suoraan Twitter-sivuston kautta selaimella sekä yksi twiitti on julkaistu Instagramia hyväksikäyttäen.



Kuvio 9. Twitonomy – Alustat, joista twiitattu eniten

Testikäyttäjälle ei saatu yhtään seuraajaa, mutta seurantaan laitettiin muutama tunnettu henkilö tai yritys, jotka sattuiivat olemaan Twitterin etusivulla. Nämä Twitter-profiilit näkyvät kuviossa 10. Samaisessa kuviossa näkyy myös muuta tietoa näistä

seuratuista profiileista, esimerkiksi twiittien, seurattujen profiilien ja seuraajien määrä sekä viimeisin twiitti.

Followers Download & Print

No follower yet

Following Download & Print

Tesla @Tesla
 7,620 tweets 74 following 6,954,399 followers 15,904 listed February 2008
 electric cars, giant batteries and solar

Neil deGrasse Tyson @neiltyson
 7,139 tweets 39 following 14,410,227 followers 44,887 listed January 2009
 Astrophysicist

Sauli Niinistö @niinisto
 438 tweets 833 following 248,987 followers 664 listed May 2009
 Tasavallan presidentti. Republikens president. President of the Republic of Finland | Viestit tasavallan presidentin omia. Presidentin kanslia: @TPKanslia

Elon Musk @elonmusk
 13,320 tweets 103 following 42,209,135 followers 58,921 listed June 2009

Sanna Marin @MarinSanna
 11,970 tweets 693 following 200,743 followers 568 listed January 2013
 Suomen pääministeri. Finlands statsminister. Prime Minister of Finland. Kansanedustaja. SDP:n puheenjohtaja.

NASA @NASA
 63,630 tweets 209 following 43,031,437 followers 93,588 listed December 2007
 We're going to the Moon with @NASAArtemis. This time, to stay.

Open in Twitter

Kuvio 10. Twitonomy - Seuraajat ja seuratut profiilit

Twitonomy listaa myös erilliseen näkymään kaikki postaukset, mistä kohdeprofiili on tykännyt. Twitter on aiemmin käyttänyt näistä nimitystä "Favorites / suosikit", mutta korvannut nämä nykyään tykkäyksillä. Twitonomyssa näkee edelleen käytettävän tätä vanhaa nimitystä, mutta käytännössä kyseisen termin tarkoitus on sama ja työkalu toimii oletetulla tavalla. Tämä lista nähdään kuviossa 11.



Kuvio 11. Twitonomy – Tykkäykset

Twitonomy näyttää myös kohdeprofiilin omat listat, sekä kaikki listat mitkä sisältävät kohdeprofiilin. Kuten kuviossa 12 näkyy, testiprofiililla ei ole omia listoja eikä tämä profiili sisälly kenenkään muun listalle.



Kuvio 12. Twitonomy – listat

4.2.2 Tinfoleak

Kuten aiemmin mainittiin luvussa 3.3.2, Tinfoleakia käytettiin tässä työssä heidän verkkosivujensa kautta (tinfoleak.com) syöttämällä kuvion 13 näkymään kohteen Twitter-käyttäjätunnus ja sähköpostiosoite, johon lähetettiin linkki raporttia varten. Tätä varten käytettiin niin sanottua kertakäyttöistä sähköpostipalvelua, joita löytyy

useita eri vaihtoehtoja. Tähän valikoitui tällä kertaa käytettäväksi sivusto *dropmail.me*.


SEARCH FOR LEAKS

Get the report in your inbox.

Note: e-mail address is **exclusively** for the purpose of sending you an e-mail with the URL to the dossier requested. No spam. No third parties.
Note 2: you report may take a while to arrive to you. It requires processing and there are more requests enqueued. Be patient. Resending your request several times won't accelerate it.

@ Twitter username

Your e-mail address

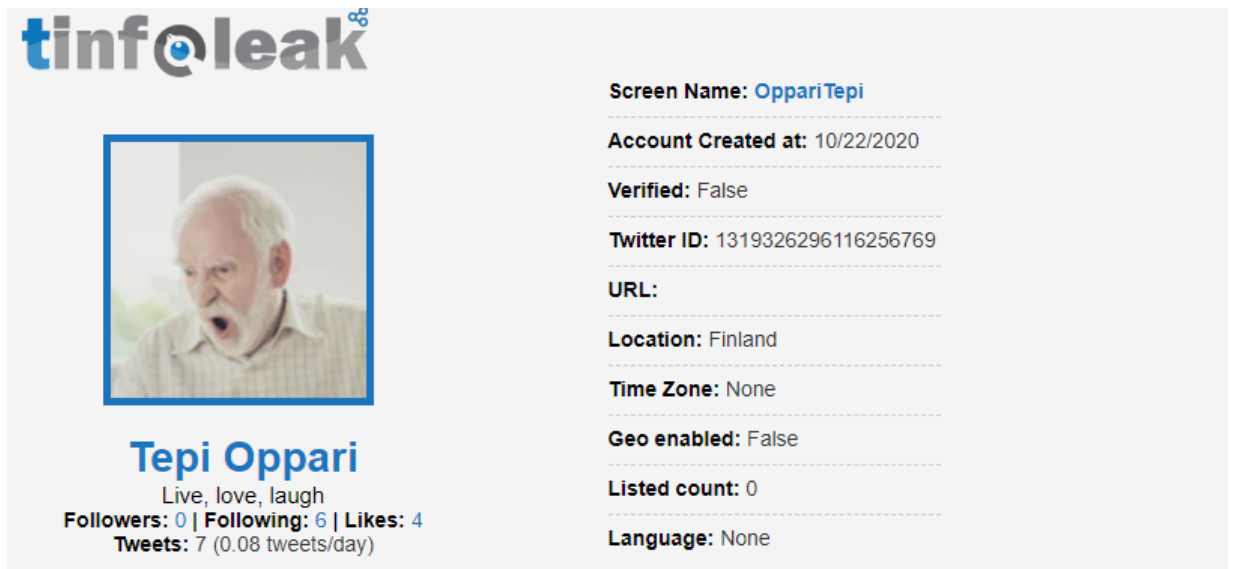
I'm not a robot  reCAPTCHA
Privacy - Terms

Send

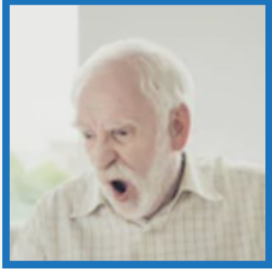
Kuvio 13. Tinfoleak – etusivu

Tinfoleakin etusivulle (kuvio 13) syötettiin testikäyttäjän Twitter-tunnus *OppariTepi* sekä *dropmail.me*-sivustolta saatu väliaikainen sähköpostiosoite. Muutaman minuutin kuluttua tästä sähköpostiin tuli Tinfoleakilta viestiä, joka sisälsi linkin raporttiin, jossa oli kyseisen työkalun hakema informaatio tästä käyttäjästä.

Tinfoleakin raportissa ensimmäiseksi huomio kiinnittyi kuvion 14 yleistietonäkymään. Tämä on yhteenveto kohdeprofiilin tiedoista, sisältäen muun muassa käyttäjän biografian, seuraajamäärän ja twiittimäärän. Jostain syystä Tinfoleakin työkalu ei osannut hankkia tietoa URLista, joka johtaa kyseisen profiilin sivulle. Tinfoleak näytti kyseisessä kuviossa myös käyttäjätilin tekopäivän väärin. Kaikki muut tiedot näyttivät olevan oikein. Testiprofiilissa ei ole kieltä missään eritoten mainittu, eikä Tinfoleakin työkalu sitä osannut "arvata" esimerkiksi twiittien kielestä, joten tästä syystä kielen kohdalla näkyy *Language: None*.



tinfoleak



Tepi Oppari
Live, love, laugh
Followers: 0 | Following: 6 | Likes: 4
Tweets: 7 (0.08 tweets/day)

Screen Name: OppariTepi

Account Created at: 10/22/2020

Verified: False

Twitter ID: 1319326296116256769

URL:

Location: Finland

Time Zone: None

Geo enabled: False

Listed count: 0

Language: None

Kuvio 14. Tinfoleak – Yleistiedot


Kuviossa 15 nähdään mitä eri sovelluksia Twitteriin postauksiin on käytetty kyseiseltä tililtä ja kuinka monta kertaa sovelluksista on postattu. Instagramista on postattu yhden kerran, minkä Tinfoleak osaakin näyttää oikein, mutta tämä ei osaa kertoa minä päivänä Instagramin kautta on ensimmäistä kertaa postattu jotain Twitteriin. Saattaa olla, että tämä ei näy vain siksi, koska Instagramista postattiin vain kerran Twitteriin ja tämä julkaisupäivämäärä näkyy kohdassa *Last use (01/10/2021)*.

CLIENT APPLICATIONS

SOURCE	USES	PERCENTAGE	FIRST USE	FIRST TWEET	LAST USE	LAST TWEET
Twitter Web App	6	85.7 %	11/10/2020	view	01/10/2021	view
Instagram	1	14.3 %	-	-	01/10/2021	view

Total: 2 results.

SOCIAL NETWORKS

SOCIAL NETWORK	USERNAME	PICTURE	NAME	ADDITIONAL INFO
Twitter	opparitepi		Tepi Oppari	Finland

Total: 1 results.

Kuvio 15. Tinfoleak - Applikaatiot ja sosiaalisen median alustat

Testikäyttäjämme twiiteissä ei käytetty kuin yhtä hashtagia kerran, *#testing*. Tämänkin tiedon Tinfoleakin algoritmi osaa hakea oikein ja se osaa kertoa myös kyseisestä

postauksesta muitakin tietoja, kuten uudelleentwiittausmäärän ja tykkäysten määrän, kuten kuviosta 16 nähdään.

HASHTAGS

HASHTAGS IN TWEETS

DATE	TIME	RT	LIKE	TWEET	USER	PROFILE IMG	LOCATION	#HASHTAG
01/10/2021	10:46:42	0	0	view	@OppariTepi		Finland	#testing

Total: 1 results.

HASHTAG DETAIL

DATE (SINCE)	DATE (UNTIL)	RT	LIKE	COUNT	#HASHTAG
01/10/2021	01/10/2021	0	0	1	#testing

Total: 1 results.

TOP HASHTAGS

DATE (SINCE)	DATE (UNTIL)	RT	LIKE	COUNT	#HASHTAG
01/10/2021	01/10/2021	0	0	1	#testing

Total: 1 results.

Kuvio 16. Tinfoleak – Hashtagit

Twitterissä ei testikäyttäjällä käytetty missään postauksissa merkintöjä, eli ei merkattu muita käyttäjiä twiiteissä, eikä kukaan tätä käyttäjääkään merkannut twiitteihin. Tästä syystä kuvion 17 mukainen Tinfoleakin hakema lista merkinnöistä on tyhjä.

USER MENTIONS								
USER MENTIONS IN TWEETS								
DATE	TIME	RT	LIKE	TWEET	USER	PROFILE IMG	LOCATION	@MENTION
Total: 0 results.								
USER MENTION DETAIL								
DATE (SINCE)	DATE (UNTIL)	RT	LIKE	COUNT	NAME	@MENTION		
Total: 0 results.								
TOP MENTIONS								
DATE (SINCE)	DATE (UNTIL)	RT	LIKE	COUNT	NAME	@MENTION		
Total: 0 results.								

Kuvio 17. Tinfoleak - Merkinnät käyttäjistä

Tinfoleak näyttää myös kohdeprofiiliin tykkäykset ja twiitit (kuvio 18), mutta nämäkään eivät jostain syystä näkyneet testikäyttäjän raportissa oikein. Tätäkin testasin muilla julkisilla profiileilla, eikä Tinfoleak näyttänyt kyseistä kohtaa muidenkaan raporteilla oikein.

LIKES							
Unsolicited Function							
TWEETS							
DATE	TIME	USER	PROFILE IMG	NAME	LOCATION	MEDIA	TWEET (FILTER: '[ok'])
Total: 0 results.							




Kuvio 18. Tinfoleak - Tykkäykset ja tweetit

Useimmiten käytetyt sanat -listalle Tinfoleak ilmeisesti hakee satunnaisesti sanat, mikäli näitä on käytetty twiiteissä saman verran. Mitään sanaa ei testikäyttäjä ole käyttänyt kahta kertaa, eli kaikkia käytettyjä sanoja on käytetty saman verran, mutta tälle listalle ei kaikki sanat silti päätyneet, vaan vain kymmenen satunnaista, twiiteissä käytettyä sanaa (kuvio 19).

WORDS MOST USED					
WORD	OCCURRENCES	PERCENTAGE	FIRST OCCURRENCE	LAST OCCURRENCE	
toinen	1	10.0%	2020-11-14 13:47:42	2020-11-14 13:47:42	
Jep,	1	10.0%	2020-11-17 17:47:16	2020-11-17 17:47:16	
tänne	1	10.0%	2020-11-23 16:59:42	2020-11-23 16:59:42	
Mitäs	1	10.0%	2020-11-23 16:59:42	2020-11-23 16:59:42	
Finland	1	10.0%	2021-01-10 10:46:42	2021-01-10 10:46:42	
laitettais	1	10.0%	2020-11-23 16:59:42	2020-11-23 16:59:42	
testataan	1	10.0%	2020-11-14 13:47:42	2020-11-14 13:47:42	
friend	1	10.0%	2020-12-04 08:28:01	2020-12-04 08:28:01	
tänään	1	10.0%	2020-11-23 16:59:42	2020-11-23 16:59:42	
test	1	10.0%	2021-01-10 10:46:42	2021-01-10 10:46:42	

Kuvio 19. Tinfoleak - Useimmiten käytetyt sanat

Kuvion 20 profiilikuvan ja twiiteissä käytettyjen kuvien metadata-tiedoista ei voida täydellä varmuudella sanoa, että kyseiset tiedot olisivat oikein, sillä profiilikuva on tekijänoikeusvapaa arkistokuva ja ladattu ilmaiseksi *shutterstock*-sivustolta.

METADATA							
PROFILE IMAGE							
PICTURE				DESCRIPTION			
							
COPYRIGHT		ORIGINAL DATETIME		MAKE		MODEL	
				Apple Computer Inc.		2000/01/01 00:00:00	
SOFTWARE		SUBJECT DISTANCE		PLATFORM		ICC DATE	
THUMBNAIL				COORDINATES			
USER IMAGES AND VIDEOS							
MEDIA	APP	REPLY TO	RT	LIKE	SOURCE USER	RT USER	TWEET
	Twitter Web App		0	0	 @OppariTepi 11/17/2020 17:47:16		view
Total: 1 results.							

Kuvio 20. Tinfoleak – Metadata

Kuviossa 20 nähdään myös *User images and videos* -otsakkeen alla Twitter Web Appia käyttäen lähetetty media, joka on .gif-muotoinen liikkuva kuva. Tinfoleak näyttää tästä vain tiedon, kuinka monta kertaa kyseinen postaus on uudelleentwiitattu ja kehen toimesta, sekä kyseisen postauksen tykkäysmäärä.

Tinfoleak -työkalu sai hyvin huonosti profiilista maantieteellisiä tietoja irti. Kuviossa 21 nähdään tiedot, mistäpäin maapalloa käyttäjä on twiitannut ja tämän suosituimmat twiittauspaikat maantieteellisesti. Testikäyttäjältämme tähän ei saatu mitään dataa.

GEOLOCATION INFORMATION ▲

TWEETS WITH GEOLOCATION ENABLED

DATE	TIME	COORDINATES	LOCATION	MEDIA	APPLICATION	TWEET
Total: 0 results.						

USER ROUTE

TWEETS	GLOBAL	DATE (SINCE)	TIME (SINCE)	DATE (UNTIL)	TIME (UNTIL)	DAYS	LOCATION	COORDINATES
Total: 0 results.								

TOP LOCATIONS

TWEETS	DATE TIME (SINCE)	DATE TIME (UNTIL)	MO	TU	WE	TH	FR	SA	SU	LOCATION	COORDINATES
Total: 0 results.											

ADVANCED SEARCH ▲

Unsolicted Function

USER CONVERSATIONS ▲

Unsolicted Function

Kuvio 21. Tinfoleak - Maantieteelliset tiedot

Samaisessa kuviossa näkyy myös ”Advanced search ja User conversations” -otsakkeet, mutta kohdeprofiilin raportissa nämä kohdat näkyvät tyhjinä, kommentilla ”Unsolicted Function” – ikään kuin työkalulla ei olisi lupaa hakea näitä tietoja profiilista. Testasin tätä työkalua myös muihin julkisiin profiileihin, eikä Tinfoleak näyttänyt näidenkään maantieteellistä dataa oikein.

4.2.3 Sleeping time

Sleeping time -työkalu toimii niin, että kirjautuu työkalun verkkosivulle (<https://sleepingtime.org>) jollain Twitter-tunnuksella sisään ja sallii applikaation hakea tietoa Twitteristä. Kirjautumisen jälkeen päästään kirjoittamaan hakupalkkiin Twitter-käyttäjätunnus, josta halutaan tietoa hakea. Hakutulokset näyttävät lyhyesti

tiedot profiilista ja milloin tämä ei useimmiten lähetä Twitteriin viestejä. Tästä saadusta datasta sovellus päättelee, että (varsinkin aktiivinen) käyttäjä nukkuu sinä aikana, kun käyttäjä ei Twitteriin mitään postaa.

Kellotaulun alla näkyy myös kuvaus, jonka käyttäjä on omaan profiiliinsa määrittänyt. Kuviossa 22 nähdään tämän työkalun yhteenveto, eli periaatteessa kaikki data, jota kyseinen työkalu antoi.

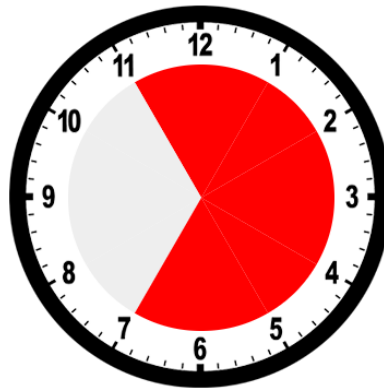
SLEEPING TIME

Find the sleeping schedule of anyone on Twitter

TECH BLOG TALL TWEETS HUNDRED ZEROS PODCAST GALLERY

Tepi Oppari is likely to sleep between 11 p.m. and 7 a.m.

[Tweet](#)



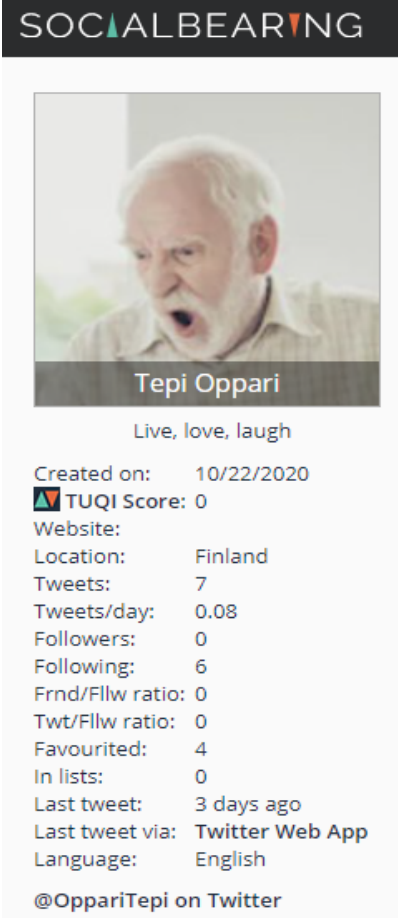
“ *Tepi Oppari - Live, love, laugh*

Kuvio 22. Sleeping time – Yhteenveto

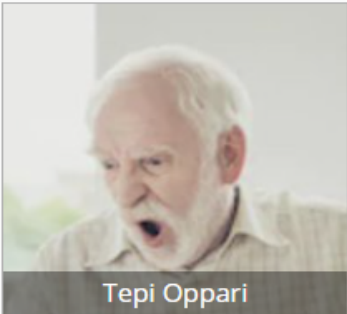
4.2.4 Social bearing

Social Bearing on myös yksi verkkosivuston kautta toimivista OSINT-työkaluista, jotka hyödyntävät Twitterin ohjelmointirajapintaa. Social Bearing toimii hyvin samaan tapaan kuin esimerkiksi aiemmin työssä mainitut Twitonomy ja Tinfoleak, myös Social Bearing kerää Twitter-profiilista hyvin paljon erilaista dataa, jota käyttäjä on sinne julkaissut.

Social bearing -työkalun raportilla testiprofiilistamme näkyy myös eräänlainen yhteenveto tilistä (kuvio 23). Tässä näkyy, kuten aiempienkin työkalujen yhteenvedossa, muun muassa twiittien määrä, seuraajien määrä ja seurattujen profiilien määrä. Muut tällä sarakeella näkyvät tiedot ovat riippuvaisia seuraajien ja listojen määrästä, joten tällä tilillä kyseiset kohdat näyttävät melko tyhjiltä.



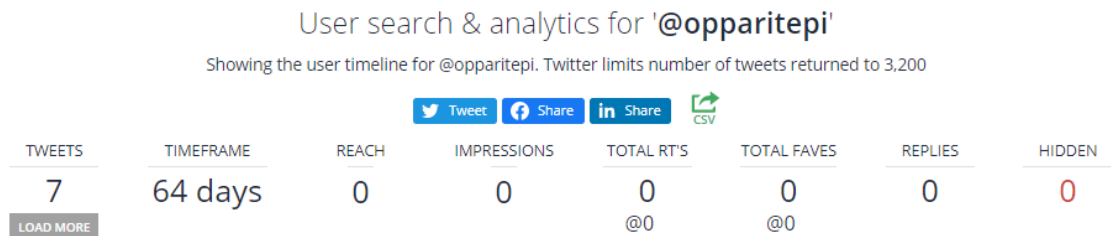
SOCIAL BEARING


Tepi Oppari
 Live, love, laugh

Created on: 10/22/2020
 TUQI Score: 0
 Website:
 Location: Finland
 Tweets: 7
 Tweets/day: 0.08
 Followers: 0
 Following: 6
 Frnd/Fllw ratio: 0
 Twt/Fllw ratio: 0
 Favourited: 4
 In lists: 0
 Last tweet: 3 days ago
 Last tweet via: Twitter Web App
 Language: English
 @OppariTepi on Twitter

Kuvio 23. Social bearing – Yhteenveto

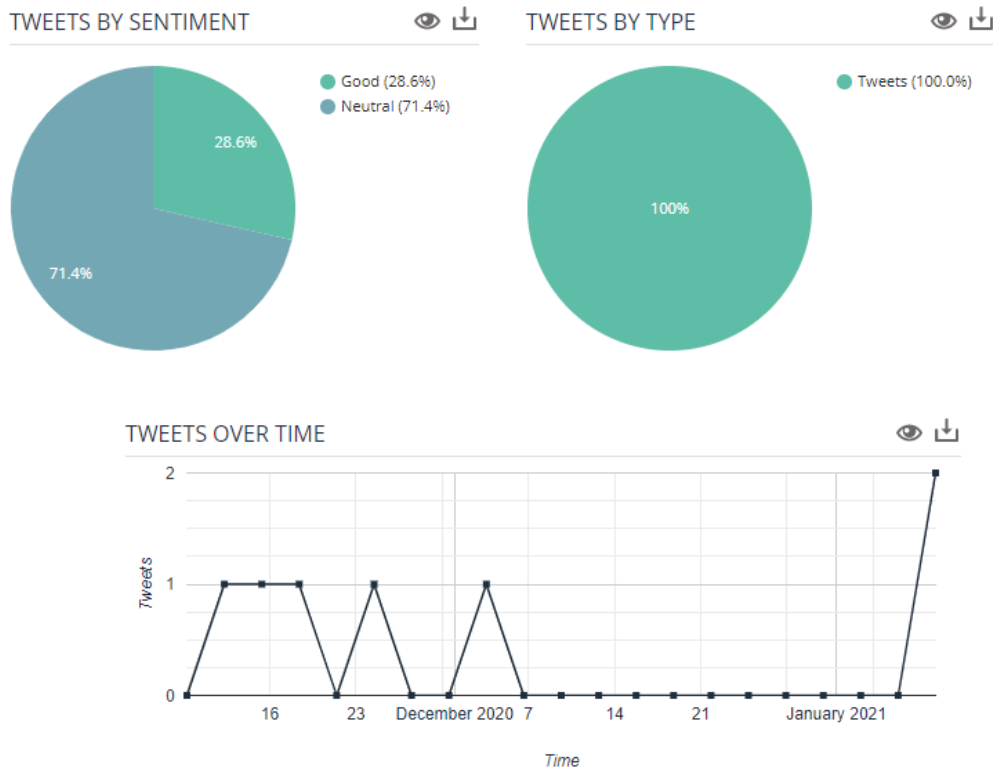
Kuviossa 24 nähdään tarkempia tietoja käyttäjän twiiteistä. Tämä sisältää muun muassa twiittien kokonaismäärän, aikavälin, jonka välillä twiitit on Twitteriin postattu, sekä kuinka monta henkilöä käyttäjän twiitit on tavoittanut ja näiden uudelleentwiitausmäärät.



Kuvio 24. Social bearing – Yleistiedot

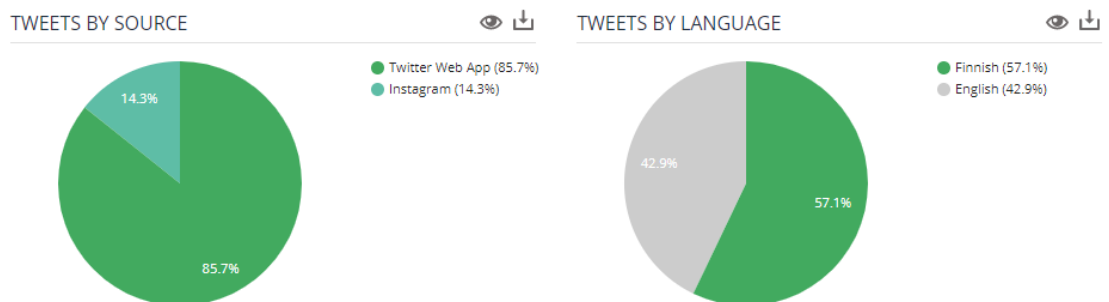
Social bearing keskittyy raportissaan näyttämään analytiikkaa käyttäen hyväksi erilaisia kaavioita sekä prosentuaalisia osuuksia. Tämä on osittain selkeämpi lähestymistapa ja raportin käyttäjän kannalta on mukavaa, että eri työkalut antavat visuaalisesti erilaisia raportteja.

Kuviossa 25 nähdään Social bearingin työkalun mittaama ”Tweets by sentiment”, eli twiittien tunnetaso tai ajatustapa. Kyseisen työkalun mukaan 28,6 % kohdeprofiilin twiiteistä on ollut hyvätahtoisia ja 71,4 % neutraaleja. Näistä tunnetasoprosenteista on hankala sanoa varmuudella, onko nämä tiedot oikein, sillä testikäyttäjän julkaisut ovat esimerkiksi ”testataan” tai ”first test”. Suurin prosentuaalinen osuus on kuitenkin neutraaleja julkaisuja, kuten pitäisikin, joten oletetaan tämän analyysin olevan kohdillaan. Samassa kuviossa nähdään myös eri twiittien tyylien prosentuaalinen osuus – kohdeprofiili ei uudelleentwiitannut tai vastannut muiden twiitteihin, joten kuvio näyttää datan oikein: 100 % kohdeprofiilin twiittien tyyleistä on vain twiittejä.



Kuvio 25. Social bearing - Twiittien analytiikkaa

Social bearing -työkalun kielentunnistus vaikuttaa erinomaiselta. Kuvion 26 kaaviossa nähdään prosentuaalinen jako twiittien kielten mukaan – 57,1 % twiiteistä on suomenkielisiä ja 42,9 % englanninkielisiä. Kaavio vaikuttaa todenmukaiselta, sillä testi-käyttäjältä postatut twiitit sisältävät nimenomaan sekä suomen- että englanninkielistä tekstiä.

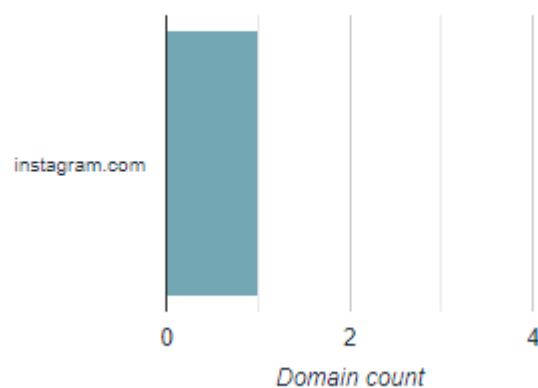


Kuvio 26. Social bearing - Mistä twiitattu ja millä kielellä

Kuviossa 26 on sisällettyä myös kaavio ”Tweets by source”, eli mitä sovelluksia käyttäen käyttäjä on twiitannut ja näiden prosentuaalinen osuus. Kohdeprofiili on postannut yhteensä seitsemän twiittiä, joista yksi on Instagramin kautta ja loput suoraan Twitterin verkkosivuja käyttäen. Kuvion mukainen yhden desimaalin tarkkuudelle pyöristetty prosentuaalinen osuus on siis todenmukainen (Twitter Web App 85,7 %, Instagram 14,3 %).

Social bearing -työkalu hakee julkaistuista twiiteistä myös tarkemmin sisältöä. Kuviossa 27 nähdään, mitä eri domaineja twiiteissä on mainittu ja kuinka paljon. Kuten aiemmin mainittu, testikäyttäjä on postannut kerran Instagramin kautta Twitteriin, jossa postaukseen sisältyy automaattisesti linkki kyseiseen Instagram-postaukseen. Tämän vuoksi Social bearing hakee tämän domainin kyseiseen kaavioon.

TOP DOMAINS SHARED



Kuvio 27. Social bearing - Jaetut domainit

Kuviossa 28 nähdään käyttäjän twiiteissä käyttämät hashtagit, eli avainsanat tai aihe-tunnisteet. Kyseinen kuvio näyttää datan oikein, sillä käytimme vain kerran yhtä aihe-tunnistetta testikäyttäjän tililtä twiitattaessa.



#testing

Kuvio 28. Social bearing - Käytetyt hashtagit

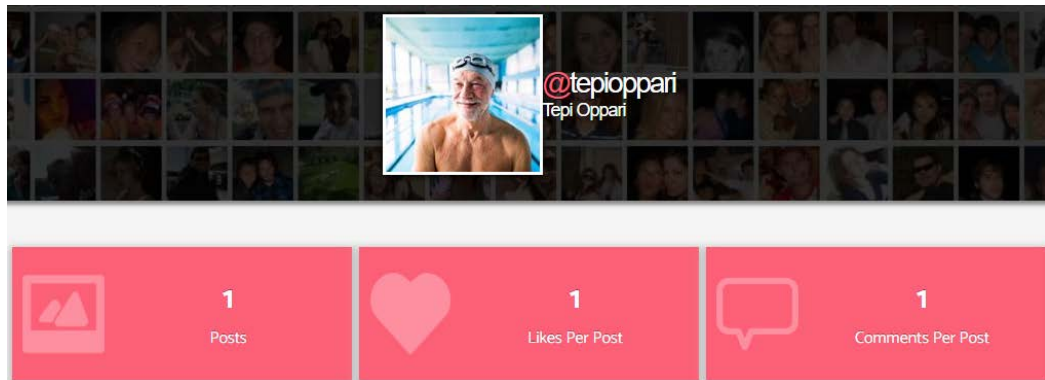
Social bearing -työkalu listaa myös twiiteissä käytetyimmät sanat visuaalisesti hieman erilaiseen muotoon, he käyttävät tästä termiä "word cloud" suom. sanapilvi. Tässä sanapilvessä näkyvät kohdeprofiilin eniten käyttämät sanat satunnaisessa järjestyksessä. Sanapilvi nähdään kuviossa 29.



Kuvio 29. Social bearing – Word cloud, twiiteissä käytetyt sanat

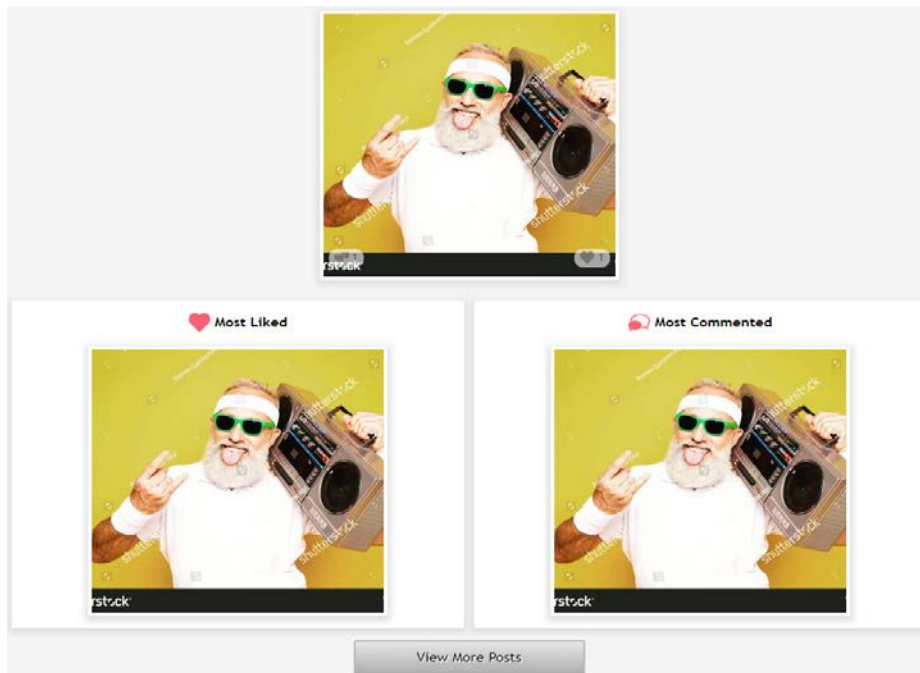
4.2.5 Searchusers

Searchusers-työkalulla saatiin testiprofiilistamme irti profiilikuva sekä postausten, tykkäysten ja kommenttien määrä (kuvio 30). Näitä ei pystytä avaamaan tätä enempää, eli työkalu osaa kertoa vain edellä mainitut määrät, eikä näytä keneltä tykkäykset tai kommentit ovat tulleet.



Kuvio 30. Searchusers – hakutulosten etusivu

Työkalun etusivulla näkyy myös uusimmat julkaisut, tykätyn julkaisu ja kommentoiduin julkaisu. Hakutuloksissa näkyi sama kuva kaikissa edellä mainituissa kohdissa siksi, koska tilille ladattiin vain yksi kuva. Hakutulokset nähdään kuviossa 31.



Kuvio 31. Searchusers – uusin postaus, tykätyin kuva ja kommentoiduin kuva

4.2.6 Downloadgram

Downloadgram on Instagramista kuvien ja videoiden lataamiseen tarkoitettu työkalu. Kaikessa lyhykäisyydessään työkalu toimii niin, että syöttää Downloadgramin etusivulla hakupalkkiin julkisesta Instagram-profiilista löytyvän kuvan tai videon linkin ja valitsee "Download now". Työkalu toimii heidän mukaansa lähestulkoon millä tahansa käyttöjärjestelmällä ja verkkoselaimella – ainakin testeissä käytetyillä Opera- ja Google Chrome -selaimilla Windows 10 -käyttöjärjestelmällä työkalu vaikutti toimivan hyvin.

4.2.7 Osintgram

Osintgram poikkeaa muista työssä käytetyistä OSINT-työkaluista sillä tavalla, että tätä työkalua ei käytetty verkkoselaimen kautta, vaan se asennettiin virtuaalikoneelle (käyttöjärjestelmänä oli Ubuntu) ja työkalua käytettiin ajamalla eri komentoja komentoriviltä.

Tässä tulee myös huomioida, että Osintgram ei pysty hakemaan dataa profiileista, jotka eivät ole julkisia – työkalu antaa heti virhesanoman, mikäli yrittää etsiä tietoa tililtä, jonka omistaja on asettanut sen yksityiseksi Instagramin asetuksista.

Python-skriptillä ajettiin ensin komento *list*, joka listasi meille kaikki käytössä olevat komennot ja lyhyen kuvauksen siitä, mitä nämä tekevät. Nämä komennot näkyvät listattuna kuviossa 32. Lähestulkoon kaikki tämän työkalun eri komennot ajettiin ja tulokset näkyvät myöhemmissä kuvioissa, poislukien komennot, jotka kohdistuivat muualle kuin kohdeprofiiliin tai ensimmäiset kaksi komentoa *FILE=y/n*, *JSON=y/n*, sillä näillä komennoilla voidaan määrätä, minne tulokset raportoidaan, eikä ole siinänsä työlle relevantteja komentoja. Komennoilla *photos*, *propic* ja *stories* pystyi lataamaan halutessaan kuvat ja käyttäjän tarinat, mutta näitäkään komentoja ei tarvittu tai käytetty.

Komento *photodes* olisi antanut jotain tietoa kohdeprofiilin kuvista, mutta tämä komento ei jostain syystä toiminut, vaan antoi vain virhesanoman ja kirjasi ulos työkalusta.

```
Run a command: list
FILE=y/n      Enable/disable output in a '<target username>_<command>.txt' file'
JSON=y/n     Enable/disable export in a '<target username>_<command>.json' file'
addr         Get all registered addressed by target photos
captions     Get target's photos captions
commentdata  Get a list of all the comments on the target's posts
comments     Get total comments of target's posts
followers    Get target followers
followings   Get users followed by target
fwersemail   Get email of target followers
fwingsemail  Get email of users followed by target
fwersnumber  Get phone number of target followers
fwingsnumber Get phone number of users followed by target
hashtags     Get hashtags used by target
info         Get target info
likes        Get total likes of target's posts
mediatype    Get target's posts type (photo or video)
photodes     Get description of target's photos
photos       Download target's photos in output folder
propic       Download target's profile picture
stories      Download target's stories
tagged       Get list of users tagged by target
target       Set new target
wcommented   Get a list of user who commented target's photos
wtagged      Get a list of user who tagged target
```

Kuvio 32. Osintgram – hakuvaihtoehdot

Ensimmäisenä haettiin kuvien osoitetiedot komennolla *addrs*. Kuten kuvioista 33 näkyy, työkalu löysi osoitteen, josta ainoa Instagramiin lataamamme kuva on ladattu. Kyseinen tieto ei kuitenkaan ole täysin oikein. Vaikka käytössä ei ollut VPN:ää tai muuta keinoa, jolla sijaintia olisi voitu huijata, kertoo Osintgram, että kuva olisi ladattu Instagramiin Helsingistä, vaikkakin lataushetkellä testausta suoritettiin Jyväskylässä. Maa on kuitenkin oikea ja päivämäärä ja kellonaika vaikuttaisivat täsmäävän täysin, joten voidaan olettaa, että kyseisellä komennolla saa suuntaa antavaa tietoa sijainnista vähintään maan osalta.

```
Run a command: addrs
Searching for target localizations...

Woohoo! We found 1 addresses
+-----+-----+
| Post | Address
|
|         time         |
+-----+-----+
| 1     | 17, Eerikinkatu, Kamppi, Eteläinen suurpiiri, Helsinki, Helsingin seut
ukunta, Uusimaa, Etelä-Suomen aluehallintovirasto, Manner-Suomi, 00100, Suomi /
Finland | 2021-01-10 12:46:40 |
+-----+-----+
```

Kuvio 33. Osintgram - kuvien osoitetiedot

Captions -komennolla haettiin kaikkien kuvien kuvatekstit. Kuviossa 34 nähdään tulokset: yksi kuvateksti löytyi ja tämä voidaan vahvistaa oikeaksi.

```
Run a command: captions
Searching for target captions...
Found 1
Woohoo! We found 1 captions
first test #testing
```

Kuvio 34. Osintgram – kuvatekstit

Kuviossa 35 nähdään kaikki kommentit, joita kohdeprofiilin kuviin on julkaistu sisältäen julkaisijan käyttäjätunnuksen ja ID:n, sekä kuvan ID:n, johon kommentti on julkaistu.

```

Run a command: commentdata
Retrieving all comments, this may take a moment...
+-----+-----+-----+-----+
| POST ID | ID | Username | Comment |
+-----+-----+-----+-----+
| 2483494196885715277_44513238664 | 44513238664 | tepioppari | comment |
+-----+-----+-----+-----+

```

Kuvio 35. Osintgram – kommenttidata

Komennolla *comments* pystyi hakemaan yhteismäärän kaikista kommenteista, joita kohdeprofiilin kuviin on julkaistu (kuvio 36).

```

Run a command: comments
Searching for target total comments...
1 comments in 1 posts

```

Kuvio 36. Osintgram – kommenttimäärä

Kuviosta 37 nähdään yksiselitteisesti kohdehenkilön seuraajat sekä seuratut profiilit, sekä näiden tilien ID, käyttäjätunnus ja koko nimi. Hakukomennot, joilla nämä tiedot saatiin, olivat *followers* ja *followings*.

```

Run a command: followers
Searching for target followers...

+-----+-----+-----+
| ID | Username | Full Name |
+-----+-----+-----+
Run a command: followings
Searching for target followings...

+-----+-----+-----+
| ID | Username | Full Name |
+-----+-----+-----+
| 528817151 | nasa | NASA |
| 25025320 | instagram | Instagram |
+-----+-----+-----+

```

Kuvio 37. Osintgram - seuraajat ja seuratut profiilit

Kuviosta 38 nähdään kohdeprofiilin seurattujen profiilien sähköpostiosoitteet. Testikäyttäjä seurasi profiilia *NASA* ja vaikka kyseisen profiilin Instagram-sivustolla ei ollut

mainintaa heidän sähköpostiosoitteestaan, löysi Osintgram sähköpostiosoitteen, jota käyttäen kyseinen profiili on rekisteröity.

```
Run a command: fwingsemail
Searching for emails of users followed by target... this can take a few minutes
Caught 1 followings email

+-----+-----+-----+-----+
| ID      | Username | Full Name | Email                               |
+-----+-----+-----+-----+
| 528817151 | nasa     | NASA     | public-inquiries@hq.nasa.gov      |
+-----+-----+-----+-----+
```

Kuvio 38. Osintgram - seurattujen profiilien sähköpostiosoitteet

Komennolla *fwingnumber* pystyi hakemaan seurattujen profiilien puhelinnumerot, mikäli heillä tällaiset olivat julkisina käyttäjätiedoissaan. Testikäyttäjämme seurasi vain kahta sattumalta valittua tunnusta, *Instagramia* sekä *NASAa*. Kummallakaan näistä tunnuksista ei ollut puhelinnumeroita profiileissaan, joten Osintgram-työkaluun ei niitä löytänyt, kuten kuviossa 39 nähdään.

Työkalua testattiin tältä osin myös kohdistamalla haun omaan henkilökohtaiseen profiiliini Instagramissa ja tuloksien perusteella voidaan vahvistaa, että haku löysi puhelinnumeroita kirjoittajan seurattujen henkilöiden profiileista.

```
Run a command: fwingnumber
Searching for phone numbers of users followed by target... this can take a few minutes
Caught 0 followings phone numbers

Sorry! No results found :-(
```

Kuvio 39. Osintgram - seurattujen profiilien puhelinnumerot

Testikäyttäjämme yhdessä kuvassa käytettiin kerran aihetunnistetta *#testing*, jonka myös Osintgram osasi kertoa (kuvio 40).

```
Run a command: hashtags
Searching for target hashtags...
1. #testing
```

Kuvio 40. Osintgram – aihetunnisteet

Kuviossa 41 nähdään kaikki kohdeprofiilin yleistiedot, jotka löytyivät komennolla *info*. Tällä komennolla löytyneissä tiedoissa näkyy muun muassa profiilikohtainen ID, seuraajamäärät sekä linkki, jota kautta pääsee katsomaan ja halutessaan lataamaan käyttäjän profiilikuvan olettaen samanlaatuisena kuin se on ladattukin Instagramin palvelimille. Tämä työkalu osaa myös etsiä tiedon, onko kohdeprofiili yritystili tai vahvistettu tili. Vahvistettu tili tarkoittaa, että tili on vahvistettu olevan esimerkiksi julkisuuden henkilö tai muu tunnetun brändin oikea tili.

```
Run a command: info
[ID] 44513238664
[FULL NAME] Tepi Oppari
[BIOGRAPHY] this is my Bio
[FOLLOWED] 0
[FOLLOW] 2
[BUSINESS ACCOUNT] False
[VERIFIED ACCOUNT] False
[HD PROFILE PIC] https://instagram.fq1f1-1.fna.fbcdn.net/v/t51.2885-19/126086518_136486358252435_3631984275772030567_n.jpg?_nc_ht=instagram.fq1f1-1.fna.fbcdn.net&_nc_ohc=qW99nDWx2QsAX979IYe&edm=AIRHW0AAAA&ccb=7-4&oh=c4a65a2345b208da7bd287c9c59765b5&oe=60907A6D&_nc_sid=e3b034
```

Kuvio 41. Osintgram – info

Kuvio 42 näyttää komennon *likes* tulokset: kuvissa on yhteensä vain yksi tykkäys ja sekin on testikäyttäjän oma tykkäys omasta kuvastaan.

```
Run a command: likes
Searching for target total likes...
1 likes in 1 posts
```

Kuvio 42. Osintgram – tykkäykset

Komennolla *mediatype* pystyi hakemaan, mitä sisältöä ja kuinka paljon kohdeprofiili on julkaissut Instagramiin julkisesti. Työkalu osasi kertoa, että testikäyttäjämme on julkaissut vain yhden kuvan (kuvio 43).

```
Run a command: mediatype
Searching for target captions...
Checked 1 posts
Woohoo! We found 1 photos and 0 video posted by target
```

Kuvio 43. Osintgram – mediatyypit

Kuviossa 44 vielä nähdään erilainen lista kohdeprofiilin kuvien kommentoijista. Komennolla wcommented voidaan hakea tämä lista, jossa näkyy jokaisen kommentoijan kohdalla erikseen käyttäjätunnuksen ja ID:n lisäksi vielä kommenttien määrä.

```
Run a command: wcommented
Searching for users who commented...
+-----+-----+-----+-----+
| Comments | ID           | Username  | Full Name |
+-----+-----+-----+-----+
| 1         | 44513238664 | tepioppari | Tepi Oppari |
+-----+-----+-----+-----+
```

Kuvio 44. Osintgram – kommentoijat

5 Tutkimustulokset

5.1 Tulokset

Osa työkaluista olivat hyvin samankaltaisia ja osasivat etsiä samanlaista tietoa kohdeprofiilista, mutta näiden ulosanti on kuitenkin hieman erilainen. Osa työkaluista näytti esimerkiksi suoraan kaiken mahdollisen datan sellaisenaan käyttäjän analysoitavaksi, kun taas osa työkaluista muunsi datan graafiseksi ja näytti tiedon helppolukuisempana, mutta osittain epätarkempana.

Kaiken kaikkiaan työkalut osasivat etsiä hyvin paljon erilaista tietoa kohdeprofiilista ja osasivat myös yhdistää tietoja kootakseen suuremman kokonaisuuden, esimerkiksi luvussa 2.4.3 läpikäyty *Sleeping time* -työkalu keräsi twiittiaikataulun perusteella kohdeprofiilin nukkuma-aikataulun. Lisäksi erikoishuomion arvoinen maininta *Social bearing* -työkalulle (luku 4.2.4), joka muista työkaluista poiketen analysoi Twitter-profiilin dataa ja näytti kohdeprofiilin twiiteistä tietoa muun muassa, onko twiiteillä tarkoitettu hyvää vai pahaa, sekä kyseinen työkalu analysoi myös twiittien kielet erinomaisesti.

Työkalut löysivät hyvin tietoa, mutta nämä eivät saaneet kuitenkaan profiileista irti henkilökohtaista dataa, eli sähköpostia, ikää tai sukupuolta. Nämä tiedot jätettiin tietoisesti piiloon, joten työkalut toimivat sinänsä oletetulla tavalla eivätkä pystyneet näitä tietoja Twitter- tai Instagram-profiileista hakemaan.

5.2 Reliaabelius ja validius

OSINT-työkaluja löytyi eri sosiaalisen median kanaviin internetistä melko vaivattomasti ja suurin osa näistä olivat myös todella helppokäyttöisiä suoraan verkkoselaimen kautta. Tässä tulee huomioida se, että työkalut toimivat opinnäytetyössä raportoidulla tavalla kirjoitushetkellä ja näiden toimivuudesta jatkossa ei ole varmuutta. Työkaluja tehdään jatkuvasti lisää ja sosiaalisen median kanavat muuttavat sivustojaan niin, että sieltä saadaan todennäköisesti tulevaisuudessa hankittua erilaista tietoa erilaisin menetelmin.

Työkalut käyttivät tiedon hankkimiseen hyödyksi sosiaalisen median alustojen ohjelmointirajapintoja, joita muutetaan koko ajan, joten työkaluja täytyisi päivittää samaan tahtiin, jotta nämä pystyisivät käyttämään ohjelmointirajapintaa samaan tapaan hyödyksi myös tulevaisuudessa.

6 Johtopäätökset

Nykyään löytyy hyvin paljon erilaisia työkaluja sosiaalisen median kanavista profiilien tietojen hankkimiseen ja näistä erilaisten raporttien automaattiseen kasaamiseen, mutta mikään yksittäinen työkalu ei tulosten perusteella osannut hakea kaikkea saatavilla olevaa tietoa. Parhaaseen lopputulokseen pääsee, kun käyttää hyväkseen useita erilaisia työkaluja ja kasaa niistä itselleen raportin työkalujen antamasta datasta.

Periaatteessa kaikki työkalut toimivat niin, että ne hankkivat tietoa muutenkin näytillä olevasta datasta ja kasaavat näistä helppolukuisemman raportin, jota analysoimalla voidaan päätellä tiettyjä asioita ja piirteitä kohdeprofiilista. Piilotettuja henkilökohtaisia tietoja ei saatu työssä käytetyillä työkaluilla irti kohdeprofiilista.

Työkaluja löytyi melko helposti internetistä, kunhan osasi etsiä oikeilla hakusanoilla (osint-tools). Työkalut käyttivät hyväkseen sosiaalisen median alustojen ohjelmointirajapintoja ja täten tarjolla oli vain tiettyjä hakuvaihtoehtoja, joita osa työkaluista osasi hyödyntää paremmin kuin toiset.

Työkalut olivat suurimmilta osin helppokäyttöisiä ja varsinkin verkkosivustojen kautta toimivat työkalut olivat tarkoitettu niin sanotusti peruskäyttäjille. Github.com -verkkosivuston kautta löytyneet työkalut vaativat kuitenkin vähintään perustiedot Linux-käyttöjärjestelmän komentorivin hallinnasta käyttääkseen niitä tarkoitettulla tavalla, vaikka useimmiten työkalun github-sivustolla on tämän asennus- ja käyttöoppaat saatavilla.

Profiilitietojen piilottaminen vaikutti tiedon hankintaan, sillä tavalla, että näitä tietoja ei työssä käytetyillä työkaluilla saatu hankittua. Näitä tietoja olivat sähköpostiosoite, sukupuoli ja syntymäaika. Mikäli näitä tietoja ei olisi piilotettu heti profiilin tekovaiheessa, olisi tällöin tiedot saattaneet tallentua jo johonkin ja täten ne olisivat mahdollisesti, ainakin osittain, saatavilla.

7 Pohdinta

Opinnäytetyön tavoitteena oli selvittää kuinka helposti ja millaista tietoa eri sosiaalisen median alustojen käyttäjistä saa irti ja miten tätä tietoa voidaan mahdollisesti hyödyntää. Tähän tarkoitukseen löytyi melko vaivattomasti useita erilaisia työkaluja, joita kokeilemalla saatiin testikäyttäjistä periaatteessa kaikki julkinen tieto ulos erilaisten raporttien muodossa.

Opinnäytetyö ei kuitenkaan kattanut kaikkia suosituimpia sosiaalisen median alustoja, eikä useita muita työkaluja, joista olisi saattanut saada vielä lisää tai erilaisempaa tietoa kohdeprofiilista. Tämä oli kuitenkin tiedossa jo työn alussa, eikä kaikkia alustoja tai työkaluja olisi ollut edes mahdollista käydä läpi, sillä sekä erilaisia sosiaalisen median alustoja että näihin tehtyjä työkaluja löytyy todella paljon nykyään internetistä. Työssä olisi ollut myös hyödyllistä esimerkiksi Facebookin osalta, mikäli olisi ollut puhelinnumero, mitä käyttää alustalle rekisteröitymistä ja todennusta varten.

Henkilötiedustelusta opinnäytetyö kuitenkin avasi lukijalle hyvin mistä on kyse ja kuinka helposti julkiseen dataan päästään kiinni ja sitä voidaan analysoida eri menetelmillä. Vaikka opinnäytetyö painottui vain kahdelle eri sosiaalisen median alustalle, saa tästä silti hyvän yleisymmärryksen periaatteessa kaikkien sosiaalisen median alustojen toiminnasta ja niiden käytöstä henkilötiedusteluun.

Lähteet

Blystone, D. 2020. The Story of Instagram. Artikkele Investopedia.com -verkkosivustolla. Viitattu 15.5.2021. <https://www.investopedia.com/articles/investing/102615/story-instagram-rise-1-photo0sharing-app.asp>

Bule, G. N.d. A Guide To Open Source Intelligence (OSINT). Artikkele itsec.group -verkkosivustolla. Viitattu 21.4.2021. <https://itsec.group/blog-post-osint-guide-part-1.html>

Chen, J. 2021. 20 Facebook stats to guide your 2021 Facebook strategy. Artikkele sproutsocial.com -verkkosivustolla. Viitattu 21.4.2021. <https://sproutsocial.com/insights/facebook-stats-for-marketers/>

CYBERDI – Kansallista & kansainvälistä kyberosaamista kasvattamassa. N.d. Artikkele jamk.fi -verkkosivustolla. Viitattu 14.5.2021. <https://www.jamk.fi/fi/Tutkimus-ja-kehitys/projektit/CYBERDI/Projektiesittely/>

CYBERDI – Projektiesittely. N.d. Artikkele polamk.fi -verkkosivustolla. Viitattu 14.5.2021. <https://polamk.fi/cyberdi>

Diaz, V. 2018. Tinfoleak-työkalu github.com -verkkosivustolla. Viitattu 21.4.2021. <https://github.com/vaguileradiaz/tinfoleak>

Diginomy Pty Ltd, 2021. Twitonomy-työkalun verkkosivut. Viitattu 8.5.2021. <https://www.twitonomy.com/content/privacy.php>

Downloadgram. N.d. Downloadgram-työkalun verkkosivusto. Viitattu 21.4.2021. <https://downloadgram.com/>

Digital 2021: Global overview report. 27.1.2021. Artikkele datareportal.com -verkkosivustolla. Viitattu 4.4.2021. <https://datareportal.com/reports/digital-2021-global-overview-report>

Hassan, N. 2020. A Guide To Social Media Intelligence Gathering (SOCMINT). Artikkele secjuice.com -verkkosivustolla. Viitattu 21.4.2021. <https://www.secjuice.com/social-media-intelligence-socmint/>

Osintgram. N.d. Osintgram-työkalu github.com -verkkosivustolla. Viitattu 5.4.2021 - <https://github.com/Datalux/Osintgram>

Schaurer, F. & Störger, J. 2013. Artikkele afio.com -verkkosivustolla. Viitattu 21.4.2021. https://www.afio.com/publications/Schauer_Storger_Evo_of_OSINT_WINTERSPRING2013.pdf

Searchusers. N.d. Searchusers-työkalun verkkosivusto. Viitattu 21.4.2021. <https://searchusers.com/>

Simon, E. 2021. How Instagram Makes Money. Artikkele Investopedia.com -verkkosivustolla. Viitattu 15.5.2021. <https://www.investopedia.com/articles/personal-finance/030915/how-instagram-makes-money.asp>

Sleeping time. N.d. Sleeping time-työkalun verkkosivusto. Viitattu. 21.4.2021. <https://sleepingtime.org/>

Social Bearing. N.d. Social bearing-työkalun verkkosivusto. Viitattu 21.4.2021 - <https://socialbearing.com/>

Tinfoleak web interface. N.d. Tinfoleak-työkalun verkkosivusto. <https://tinfoleak.com/>

Tuomi, J. & Sarajärvi, A. 2018. Laadullinen tutkimus ja sisällönanalyysi. Uudistettu painos. Helsinki: Tammi.

Zeevi, D. 2013. Twitter 101: What is Twitter Really About? Artikkele socialmediatoday.com -verkkosivustolla. Viitattu 14.5.2021. <https://www.socialmediatoday.com/content/twitter-101-what-twitter-really-about>