

Justus Uurtimo

Suositusjärjestelmien vaikutus disinformaation leviämiseen

Tietotekniikan Kirjallisuuskatsaus

7. toukokuuta 2021

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Justus Uurtimo

Yhteystiedot: uurtimo.justus@gmail.com

Ohjaaja: Timo Tiihonen

Työn nimi: Suositusjärjestelmien vaikutus disinformaation leviämiseen

Title in English: Recommendation systems impact on spread of disinformation

Työ: Kirjallisuuskatsaus

Opintosuunta: Tietotekniikka

Sivumäärä: 33+0

Tiivistelmä: Suositelujärjestelmät ovat erottamaton osa nykyaikaa ja informaation lisääntyminen verkossa tarkoittaa, että niitä tarvitaan tulevaisuudessa yhä enemmän ja enemmän. Suositelujärjestelmissä on kuitenkin ongelmia, jotka heijastuvat yhteiskuntaamme. Suositelujärjestelmät nopeuttavat tiedon vastaanottamista ja rajaavat epäolennaista tietoa käyttäjiltä, mutta ne myös saattavat tällöin rajata tietoa liian yksipuolisesti aiheuttaen kaikukammioita. Tutkielmassa selvisi, että näitä kaikukammioita voidaan hyödyntää osana disinformaation levitystä ja niitä on käytetty kohdentamaan disinformaatiota halutulle yleisölle.

Avainsanat: Disinformaatio, Kaikukammio, Suositusjärjestelmät, Sosiaalinen botti, Trollitehdas, Informaatiovaikuttaminen

Abstract:

In this information age recommender systems are inseparable part of our day to day lives and the ever growing amount of information in the internet makes sure we will need them even more in the future. The outcome of the research for this thesis was that recommender systems are integral part of how the customers consume media. Recommender systems can also cause problems which can reflect even on our society. The most notable of these problems are the echo chambers. Echo chambers occur when recommender systems recommend information from too one-sided point of view. We also find that these echo chambers have been used to target disinformation to the group most susceptible to it.

Keywords: Disinformation, Recommendation systems, echo chambers, Social bot, Troll farm, Information operations

Jyväskylässä 7. toukokuuta 2021

Kuviot

Kuvio 1. (Vaarennetty vaalimainos)	14
Kuvio 2. (Vaarennetty Applen mainos)	15

Sisällys

1	JOHDANTO	1
2	SUOSITTELUJÄRJESTELMÄT JA KAIKUKAMMIOT.....	4
	2.1 Suosittelevjärjestelmien tarve	4
	2.2 Suosittelevjärjestelmät	5
	2.2.1 Sisältöpohjainen suosittelujärjestelmä	6
	2.2.2 Yhteisöpohjainen suosittelujärjestelmä.....	6
	2.2.3 Hybridi suosittelujärjestelmät	7
	2.3 Kaikukammiot.....	7
	2.4 Kaikukammioiden vaikutukset	8
3	DISINFORMAATION LEVITTÄMINEN	10
	3.1 Disinformaatio.....	10
	3.2 Disinformaation levittämisen motiivit.....	12
	3.2.1 Taloudelliset intressit	12
	3.2.2 Poliittiset intressit	13
	3.2.3 Hedonistiset syyt.....	14
4	DISINFORMAATION LEVITTÄJIEN TOIMINTAMALLIT	16
	4.1 Klikkiotsikot.....	16
	4.2 Sosiaaliset botit.....	16
	4.3 Trollitehtaat.....	17
5	YHTEENVETO	19
	LÄHTEET	23

1 Johdanto

Suosittelujärjestelmät ovat erottamaton osa jokapäiväistä elämäämme. Järjestelmät suosittelvat meille elokuvia ja sarjoja suoratoistopalveluissa, ehdottavat ajankohtaisia ja juuri meitä kiinnostavia uutisia, sekä mainostavat tuotteita, joista todennäköisemmin kiinnostumme. (Ricci, Rokach ja Shapira 2015) Hyvin toimivat suosittelujärjestelmät tunnistavat käyttäjän persoonalliset mieltymykset ja luovat käyttäjälle henkilökohtaisen näkymän palveluun. Tällöin käyttäjä näkee palvelussa heti asioita, jotka häntä eniten kiinnostavat ja tämä taas johtaa parempaan käyttäjäkokemukseen. (Knijnenburg ym. 2012) Suosittelujärjestelmät ovatkin nykypäivänä niin olennainen osa käyttäjäkokemusta, ettemme välttämättä edes huomaa tai tiedosta niiden olemassaoloa ja vaikutusta tekemiimme valintoihin.

Tämä tutkielma käsittelee sosiaalisen median hallinnassa käytettäviä suosittelujärjestelmiä, sekä niiden vaikutusta disinformaation leviämiseen. Sosiaalisella mediallyä tarkoitetaan tässä työssä palveluja kuten Facebook, Twitter, Instagram ja Reddit, joissa ihmiset voivat tehdä päivityksiä, reagoida sisältöön positiivisesti tai negatiivisesti, ja joissa ihmiset voivat lukea ja jakaa ystäviensä, sekä muiden ihmisten tuottamaa sisältöä. Positiivisella ja negatiivisella reagoinnilla tarkoitetaan kommentointia ja muita tapoja osoittaa julkisesti mielipide tietystä sisällöstä.

Tutkielmassa suosittelujärjestelmillä tarkoitetaan digitaalisia algoritmisia järjestelmiä, joita käytetään sisällön henkilökohtaiseen suodatukseen. Näiden algoritmisten järjestelmien avulla voidaan automaattisesti suositella käyttäjille sisältöä, joka kutakin käyttäjää todennäköisimmin kiinnostaa. Sisällön suosittelu voi perustua esimerkiksi käyttäjien aikaisempiin valintoihin, tai samankaltaisten käyttäjien tekemiin valintoihin.

Suosittelujärjestelmät voivat aiheuttaa myös ongelmia, jotka pahimmillaan saattavat heijastua jopa yhteiskunnalliseen käyttäytymiseen. Jatkuvasti samankaltaisen sisällön tarjoaminen käyttäjälle on hyvin ongelmallista. Käyttäjän saadessa näkyviinsä sisältöä, josta hän on aikaisemminkin ollut kiinnostunut, johtaa hyvin todennäköisesti siihen, että hän myös jatkossa valitsee vain samankaltaista sisältöä (Jiang ym. 2019). Tällöin suosittelujärjestelmien käyttö voi johtaa esimerkiksi siihen, että ihmiset näkevät sosiaalisessa mediassa sisältöä ainoastaan aiheista, joista he ovat aikaisemmin olleet kiinnostuneita. Lisäksi tarjottavat uutiset on voitu kirjoittaa nimenomaan käyttäjän omia mielipiteitä vastaavista näkökulmista, jolloin ihmiset todennäköisemmin suosivat niitä uutisia, joiden sisällön kanssa he ovat jo valmiiksi samaa mieltä. (Vike-Freiberga Vaira ja Pessoa 2013)

Tässä kirjallisuuskatsauksessa tutkitaan erityisesti sitä, voimistavatko sosiaalisen median suosittelujärjestelmät disinformaation leviämistä. Mahdollisia disinformaation levittämismekanismeja tarkastellaan erityisesti suosittelujärjestelmien aiheuttamien kaikukammioiden, sekä järjestelmiä hyväksikäyttävien disinformaation levittäjien toimintatapoja analysoimalla. Sosiaalisen median alustoista tarkastellaan erityisesti niitä, jotka hyödyntävät suosittelujärjestelmiä tarjotessaan käyttäjilleen sisältöä, ja joita on todettu käytettävän osana informaatiovaikuttamista. Tällaisia sosiaalisen median alustoja ovat esimerkiksi: *Twitter, Facebook, Reddit ja YouTube* Lisäksi työssä käydään läpi, miten suosittelujärjestelmien aikaansaama sisällön vahva personalisointi voi vaikuttaa informaatiokuplien syntymiseen palveluiden sisällä. Työssä käsitellään myös kaikukammioiden muodostumista ja sitä, miten disinformaation levittäjät käyttävät niitä hyödykseen.

Keskeisin tutkimuskysymys, johon työssä pyritään vastaamaan, on: ”*Miten suosittelujärjestelmät vaikuttavat disinformaation leviämiseen sosiaalisessa mediassa?*” Kun on selvitetty, miten suosittelujärjestelmät vaikuttavat disinformaation leviämiseen, voidaan järjestelmiä kehittää ehkäisemään informaatiokuplien syntymistä ja systemaattisesti vaikeuttamaan disinformaation levittäjien toimintaa.

Tutkielma on suoritettu kirjallisuuskatsauksena. Työssä aineistona käytetään vertaisarvioituja tieteellisiä artikkeleita, sekä muita tutkimuksia. Keskeisiä työssä käytettyjä lähteitä ovat olleet esimerkiksi: (*Ricci, Rokach ja Shapira 2015*), (*Woolley 2016*) ja (*Fallis 2015*) Työssä on pyritty käyttämään aineistoa, joka on mahdollisimman uutta. Varsinkin sosiaalisen me-

dian käyttämiin suosittelujärjestelmiin liittyvissä aineistoissa on pyritty käyttämään maksimissaan 8 vuotta vanhoja tutkimuksia.

Suosittelujärjestelmiä on kuitenkin tutkittu jo yli 20 vuoden ajan, joten osa tutkielmassa käytetyistä lähteistä ovat pyrittyä vanhempia. Tutkimuksessa käytetyt aineistot ovat valtaosin englanninkielisiä, mutta osa lähteistä on suomenkielisiä. Työn lähteiden hakemisessa tärkeimpiä hakusanoja ovat olleet esimerkiksi *recommender systems*, *echo chamber*, *information bubble*, *social media*, *misinformation*, *disinformation*, *fake news*, *socialbot*, sekä erilaiset kombinaatiot näistä. Lähteitä haettiin ensisijaisesti *Google scholar*, *JYKDOK*, sekä *Scopus* hakupalveluista. Osa hauista on tehty myös inkognito-tilassa, jotta palveluiden mahdolliset suosittelujärjestelmät eivät vääristäisi hakutuloksia.

Tutkielma on rakenteeltaan jaettu kolmeen varsinaiseen käsittelyosioon, joista ensimmäisessä määritellään suosittelujärjestelmät, käydään läpi niiden käyttötarkoituksia, sekä merkitystä nykyajan digitaalisissa palveluissa. Toisessa käsittelyluvussa tarkastellaan disinformaatiota. Luvussa määritellään disinformaatio tutkielman kontekstissa, jonka jälkeen käydään läpi disinformaation aktiiviseen levittämiseen liittyviä motiiveja. Kolmannessa käsittelyluvussa avataan yleisimpiä keinoja, joita disinformaation levittäjät käyttävät pyrkiessään levittämään väärää tietoa. Lopuksi käydään läpi tutkielman tulokset ja vastataan kootusti esitettyihin tutkimuskysymyksiin, pohditaan, mitä disinformaation leviämisen ehkäisemiseksi on jo tehty, ja mitä sen eteen voitaisiin tehdä tulevaisuudessa, sekä käsitellään mahdollisia uusia tutkimusaiheita.

2 Suosittelujärjestelmät ja kaikukammiot

Tässä luvussa käsitellään mitä suosittelujärjestelmät oikeastaan ovat, mihin niitä tarvitaan ja lisäksi määritellään niistä työn kannalta olennaisimmat suosittelujärjestelmät. Suosittelujärjestelmä käsitteenä määritellään aliluvussa (2.1). Samassa aliluvussa määritellään myös työn kontekstin kannalta tärkeimmät suosittelujärjestelmätyypit. Aliluvussa (2.2) avataan ensin kaikukammio käsitteenä ja tämän jälkeen käsitellään miten suosittelujärjestelmät vaikuttavat niiden syntymiseen. Samassa luvussa käsittelemme myös informaatiokuplan ja sen syntymiseen vaikuttavia tekijöitä.

2.1 Suosittelujärjestelmien tarve

Internet on suosionsa johdosta dramaattisesti vähentänyt uuden tiedon julkaisemisen kustannuksia, sekä mahdollistanut sen että ihmisillä on jatkuvasti saatavilla erittäin suuri määrä tietoa vain muutaman klikkauksen päässä. Verkkolehtien ei tarvitse hankkia useita kallista painolaitteita julkaistakseen informaatiota suoraan verkkoon. Lisäksi sosiaalisen median suuren suosion myötä ihmiset voivat nopeasti ja tehokkaasti jakaa suosikkiartikkeleitaan kaikille ystävilleen. Elämmekin niin sanotusti informaation yltäkylläisyyden aikaa.

Tämän yltäkylläisyyden vuoksi ihmiset tarvitsevat apuvälineitä selvittääkseen informaation tulvasta internetissä. Käyttäjien henkilökohtaisiin tarpeisiin muotoillut informaation suodatuskeinot ovatkin nykyään käytännössä välttämättömyys. (Lops, De Gemmis ja Semeraro 2011) Vuonna 2012 Netflix raportoi, että 75% sisällöstä, jota ihmiset valitsevat tulee suoraan heidän suosittelujärjestelmästänsä. (Amatriain ja Basilico 2012) Vuoteen 2015 mennessä luku oli kasvanut jo 80 % jokaisesta katsotusta tunnista (Gomez-Uribe ja Hunt 2015) ja suosittelujärjestelmien parantuessa tämä luku todennäköisesti jatkaa kasvuaan tulevaisuudessa.

2.2 Suositteijärjestelmät

Suosittelujärjestelmät ovat eräänlaisia tiedonsuodatusjärjestelmiä. Ne ovat ohjelmiston työkaluja, jotka tarjoavat käyttäjilleen suosituksia heitä todennäköisimmin kiinnostavasta sisällöstä. Suositteijärjestelmiä on perinteisesti hyödynnetty esimerkiksi verkkokauppojen käyttäjänäkymissä, jossa käyttäjä ensimmäisenä näkee hänelle todennäköisesti mieluisia tuotteita. Viime vuosikymmenen aikana datan määrä verkossa on kasvanut räjähdysmäisellä nopeudella.

Datan kasvua ilmentää se, että vuodesta 2005 digitaalisen datan määrä on tuplaantunut kahden vuoden välein. (Levin, Wanderer ja Ehrenfeld 2015) Suositteijärjestelmät ovatkin, etenkin käytettävyyden kannalta, välttämättömiä, sillä käyttäjillä ei ole mitenkään aikaa tutustua kaikkeen sivustoilla esillä olevaan sisältöön. (Ricci, Rokach ja Shapira 2015) Suositteijärjestelmien tehokkuus onkin todella tärkeää nykyajan yrityksille, joiden ansaintalogiikka perustuu vahvasti sisällön tarjoamiseen.

Suosittelujärjestelmien kehityksessä suuressa roolissa oli suoratoistopalvelu Netflixin järjestämä kilpailu, jossa se lupasi miljoonan dollarin palkinnon ensimmäiselle joukkueelle, joka pystyisi rakentamaan yhtiön silloista järjestelmää *Cinematch* :iä vähintään 10% tarkemman suosittelujärjestelmän. (Koren, Bell ja Volinsky 2009) Kilpailu toikin merkittävää edistystä suosittelujärjestelmien kehitykselle, kun monet kilpailuun osallistuneet joukkueet perustivat yrityksiä kehittämänsä teknologian pohjalle. Esimerkiksi toiseksi tullut joukkue *Gravity* perusti yrityksen *Gravity RD* (Tamás 2009)

Nykyään suosittelujärjestelmiä hyödynnetään myös todella paljon sosiaalisen median alustoilla. Sosiaalisessa mediassa toimivien suosittelujärjestelmien päätehtävänä on avustaa, informaation muuten ylikuormittamia, ihmisiä löytämään heille mieluisaa sisältöä. (Ricci, Rokach ja Shapira 2015) Sosiaalisia medioita on joskus kuvailtu jopa ”*huomiotalouksiksi*” (eng. ”attention economy”, käänös minun) (Marwick ja Lewis 2017. Tällöin datan määrän jatkuvan kasvun vuoksi sosiaalisten medioiden on ollut pakko ottaa käyttöön suosittelujärjestelmiä, jotta ihmisten mielenkiinto säilyy palveluita käytettäessä.

Suosittelujärjestelmät voidaan jakaa (Ricci, Rokach ja Shapira 2015) mukaisesti kuuteen eri luokkaan. Tässä työssä keskitytään kuitenkin lähinnä henkilökohtaiseen uutisvirtaan perustuviin suosittelujärjestelmiin. Nämä järjestelmät voidaan käytännössä jakaa kolmeen osaan: *Sisältöpohjaiseen järjestelmään, yhteisö toiminnalliseen järjestelmään, sekä hybridi malleihin* näistä kahdesta (Li ym. 2011)

2.2.1 Sisältöpohjainen suosittelujärjestelmä

Sisältöpohjaisessa suosittelujärjestelmässä palvelujen sisällön suodattaminen perustuu käyttäjän profilointiin, käyttäjän tekemien valintojen perusteella. Järjestelmän käyttäjästä luoman profiilin avulla voidaan käyttäjälle tarjota sisältöä, jotka ovat hyvin samanlaisia, tai jotka liittyvät käyttäjää aikaisemmin kiinnostaneeseen sisältöön. (Pazzani 1999) Sisältöpohjaisessa suosittelujärjestelmässä prosessi koostuu käytännössä siis käyttäjäprofiilin ja sisällön ominaisuuksien vertailuun. Tämän lisäksi järjestelmä pyrkii löytämään niiden välisiä yhteyksiä. Näiden yhteyksien, tai niiden puutteiden, avulla järjestelmä kykenee päättämään, mitä sisältöä käyttäjälle kannattaa esittää ja mitä ei. (Lops, De Gemmis ja Semeraro 2011)

2.2.2 Yhteisöpohjainen suosittelujärjestelmä

Aikaisemmat tutkimukset ovat osoittaneet ihmisten arvostavan huomattavasti enemmän ystäviensä suosituksia, kuin suosituksia, joita kone tekee. (Sinha, Swearingen ym. 2001) Tästä syystä varsinkin verkkokaupat ehdottavat usein tuotteita fraasilla: ”*saman kaltaiset ihmiset katsoivat myös tätä*”. Yhteisöpohjaiset suodatusjärjestelmät pyrkivät jäljittelemään juuri tätä ominaisuutta. Käytännössä palvelun käyttäjien samankaltaisuutta lasketaan vertailemalla käyttäjän suosimaa sisältöä muiden käyttäjien suosimaan sisältöön. Tämän jälkeen käyttäjälle voidaan suositella uutta sisältöä, josta muut samankaltaiset käyttäjät ovat pitäneet. (Ricci, Rokach ja Shapira 2011)

2.2.3 Hybridi suosittelujärjestelmät

Hybridi suosittelujärjestelmillä tarkoitetaan suosittelujärjestelmiä, jotka ovat useiden suosittelujärjestelmien yhdistelmiä. Yhdistelemällä erityyppisten suosittelujärjestelmien ominaisuuksia, pyritään paikkaamaan tietyn tyyppisten järjestelmien heikkouksia. Esimerkiksi puhtaasti yhteisö suodatusjärjestelmien heikkous on uuden tuotteen ongelma; eli ”*Miten voidaan suositella sisältöä, jota kukaan ei ole arvioinut vielä?*” Tämä ongelma voidaan korjata esimerkiksi yhdistämällä puhtas yhteisösuodatusjärjestelmä puhtaaseen sisältöpohjaiseen järjestelmään, jossa sisällöllä on valmiiksi omat ominaisuudet, joiden mukaan niitä suositellaan. (Ricci, Rokach ja Shapira 2015)

2.3 Kaikukammiot

Kaikukammiot ovat olleet jo jonkin aikaa vahvasti esillä yhteiskunnallisessa keskustelussa, sekä varsinkin kun tutkitaan sosiaalisen median vaikutusta ihmisten poliittisiin mielipiteisiin ja äänestyskäyttäytymiseen. (Shao ym. 2017) Kaikukammioiden olemassaoloa, sekä niiden vaikutusta poliittiseen päätöksentekoon on kuitenkin kritisoitu tutkimuksissa viime aikoina. (Dubois ja Blank 2018, Bruns 2017) Tässä aliluvussa käydäänkin läpi kaikukammion muodostumiseen vaikuttavia tekijöitä, sekä mitä kaikukammiot voivat aiheuttaa.

Ihmisillä on luontainen pyrkimys suosia tietoa, joka vahvistaa heidän olemassa olevaa maailmankuvaansa. (Nickerson 1998) Tätä ilmiötä kutsutaan *vahvistusharhaksi*. Vahvistusharha on kognitiivinen vinouma, jossa henkilö pyrkii suosimaan omia ennakkokäsityksiään tai hypoteesejaan tukevaa informaatiota. (Nickerson 1998) Vinouman vuoksi henkilöt usein hakevat lisäinformaatiota, joka tukee heidän valmiita käsityksiään aiheesta ja pyrkivät jopa tulkitsemaan tätä tietoa niin että se tukee heidän väitteitään. Kääntäen Ihmiset eivät hae ja jopa aktiivisesti välttelevät tietoa, joka todistaisi heidän olettamuksensa vääräksi. (Koriat, Lichtenstein ja Fischhoff 1980)

Vahvistusharha on tärkeä osa kaikukammioiden syntymistä, sillä vahvistusharhalla on todettu olevan hyvin voimakkaat vaikutukset informaation kulutukseen, etenkin sosiaalisen median alustoilla. (Bessi ym. 2015) Kun suuri osa informaatiosta, mitä ihminen näkee, tulee suoraan suosittelujärjestelmiltä, se voi johtaa kaikukammion syntymiseen.

Määritellään kaikukammio (Cinelli ym. 2020) mukaisesti, niin että se koostuu kahdesta pääominaisuudesta.

- Kaikukammiossa olevilla ihmisillä on ristiriitaiset mielipiteet, johonkin aiheeseen liittyen, verrattuna muihin ihmisiin. Aihe voi olla poliittisesti kiistanalainen ja siihen liittyvät mielipiteet hyvin polarisoituneita. Esimerkiksi samaa sukupuolta olevien avioliitto ja rokotevastaisuus ovat hyvin poliittisesti polarisoituneita aiheita.
- Ihmisten väliset kanssakäymiset, tai ihmisten näkemä informaatio koostuu lähinnä näistä polarisoituneista aiheista, tai niihin liittyvästä keskustelusta. Ihmiset siis pyrkivät keskustelemaan aiheesta vain muiden aiheesta samaa mieltä olevien kanssa.

Kaikukammioiden on todistettu olevan läsnä lähes jokaisessa sosiaalisen median variantissa. Niiden on todistettu olevan läsnä blogeissa (Gilbert, Bergstrom ja Karahalios 2009), keskustelufoorumeilla (Edwards 2013), sekä suurilla sosiaalisen median alustoilla kuten esimerkiksi Facebookissa (Quattrociocchi, Scala ja Sunstein 2016) ja Twitterissä (Barberá ym. 2015)

2.4 Kaikukammioiden vaikutukset

Kaikukammio on vakavasti otettava mekanismi poliittisen polarisaation syntymisessä. Kaikukammioilla on kyky vahvistaa joukon entuudestaan olemassa olevia mielipiteitä, sekä pahimmillaan polarisoida kammiossa olevien ihmisten mielipiteitä entistä enemmän. Ihmisten mielipiteiden polarisoituminen, voi johtaa ääriajattelun lisääntymiseen ja jopa fanaattisuuden syntymiseen aiheen puolesta.

Kaikukammion polarisaatiota lisäävää ilmiötä tukee myös *joukkopolarisaatio teoria*. (Sunstein 1999) määrittelee joukkopolarisaatio teorian näin: *“Members of a deliberating group move toward a more extreme point in whatever direction is indicted by the members’ pre-deliberation tendency.”* Tällöin joukkopolarisaation vuoksi keskusteluun osallistuvat henkilöt ovat vähemmän halukkaita jakamaan tietoa, mikäli he olettavat, että muut keskusteluun tai ryhmään osallistuvat vastustaisivat sitä. Teorian mukaan ryhmien keskustelun myötä ryhmän kokonaismääräinen näkemys asiasta oli lopussa radikaalimpi, kuin mitä keskimääräisen ryhmään osallistuneen henkilön näkemys asiasta alun perin oli. (Sunstein 1999)

Vuonna 2016 noin 86% yhdysvaltalaisista käytti internetiä säännöllisesti. Ihmisten käyttämistä palveluista internetissä ylivoimaisesti suosituimpia olivat sosiaalisen median palvelut ja applikaatiot. Yhdysvalloissa internetiä säännöllisesti käyttävistä aikuisista 79% käyttivät Facebookkia, joista 76 % kertoi käyttävänsä sitä päivittäin. Vuonna 2020 noin 53 % yhdysvaltalaisista aikuisista kertoi saavansa uutisensa sosiaalisen median kautta ja tämän luvun oletetaan tulevaisuudessa nousevan. (Shearer ja Mitchell 2021) Sosiaalisen median käyttäminen uutisten lähteenä on suuri ongelma, sillä suuri osa ihmisistä, jotka sosiaalisen median kautta lukevat uutisensa, eivät kykene erottamaan valeuutisia oikeista uutisista. (Allcott ja Gentzkow 2017).

Kaikukammion suurin haitta onkin juuri se, että ihmiset ovat tällöin lähinnä vuorovaikutuksessa aiheesta samaa mieltä olevien käyttäjien kanssa. Tämä taas voi johtaa vahvistusharhan syvenemiseen. Vahvistusharhan syventyminen, sekä kaikukammion muodostuminen ovat otollista maaperää informaatiokuplan syntymiseen. Termi informaatiokupla on yhdysvaltalaisen kirjailijan ja internetaktivistin Eli Parserin keksimä termi, jolla hän kuvaa potentiaalista internetin personalisointia, joka johtaa ihmisten eristäytymiseen eriäviltä mielipiteiltä ja sisällöltä. (Pariser 2011)

3 Disinformaation levittäminen

Tässä luvussa ensimmäisenä määritellään disinformaation merkitys, sekä rajataan hieman määrittelyä, jotta se sopii paremmin tutkielmaan. Tämän jälkeen luvussa käsitellään motiiveja disinformaation tahalliseen levitykseen, sekä mitä konkreettista sillä voidaan saavuttaa. Lopuksi luvussa käsitellään tapoja, joita yleisimmin käytetään osana disinformaation levitystä ja informaatiovaikuttamista. Disinformaatiota käsitellään erillisenä lukuna, sillä se on hyvin monimuotoinen käsite. Lisäksi sen ymmärtäminen on hyvin olennainen osa työn ymmärtämistä.

3.1 Disinformaatio

Jotta voisimme käsitellä disinformaation toimintatapoja, on meidän ensin määriteltävä se. Disinformaatio on tämän tutkimuksen kannalta ehkäpä tärkein käsite, joten sen tarkka määrittely on ensiarvoisen tärkeää tutkielman ymmärtämisen kannalta. Tässä aliluvussa käsittelemme, mitä disinformaatiolla oikeastaan tarkoitetaan ja tämän jälkeen rajataan sen käsitettä hieman tutkielmaa varten.

Perinteiseltä määritelmältään disinformaatio tarkoittaa väärän tiedon levitystä tahallisesti, tarkoituksenaan johtaa harhaan, pettää, tai hämätä toista lukijaa. (Fetzer 2004) Disinformaatiota ei pidä kuitenkaan sekoittaa misinformaatioon. Misinformaatiolla tarkoitetaan yleisesti pelkästään väärää tai harhaanjohtavaa tietoa, sekä sen levitystä. Misinformaatiolta puuttuu siis tahallisuus, joka erottaa sen disinformaatiosta.

Työtä varten tarkennetaan kuitenkin disinformaation määritelmää hieman, sillä disinformaation määritelmä on perinteisenä määritelmänä hieman liian laaja. Esimerkiksi satiiriset ja parodiset julkaisut pitävät sisällään usein väärää tietoa ja niillä on tavoitteena levittää tätä väärää informaatiota tahallisesti. Alkuperäisen määritelmän mukaisesti ne voitaisiin tällöin laskea disinformaatioksi. Tällaisia julkaisuja ei kuitenkaan tämän työn puitteissa voida laskea disinformaatioksi, sillä julkaisuja ei odoteta otettavaksi tosissaan eikä niillä täten pyritä johtamaan ihmisiä harhaan.

Siispä tarkennetaan disinformaation määritelmää hieman (Fallis 2015) mukaisesti; ”Disinformaatio on harhaan johtavaa informaatiota, jolla on todellinen funktio harhaanjohtamiseen” (käännös minun). Tässä funktiolla tarkoitetaan sitä, että disinformaatiolla on varteen otettava tapa olla uskottava ja sillä on todellinen funktio levitä laajalle muiden ihmisten levittämänä.

Tämän lisäksi työssä lasketaan disinformaation alaisuuteen myös termin ”*valeuutiset (eng. fake news)*” alaisuuteen kuuluvat informaation lähteet. Valeuutiset terminä nousi yleiseen suosioon vuoden 2016 Yhdysvaltojen presidentinvaalien yhteydessä. Termi on ollut viime vuosina suosittu myös akateemikoiden ja journalistien keskuudessa, joissa sitä on käytetty lähinnä kuvaamaan sosiaalisessa mediassa levitettyjä seipitettyjä uutisia. Termi on myös ollut hyvin suosittu, kun pyrkimyksenä on ollut valtavirta median uutisten kritisoiminen, pyrkimyksenä horjuttaa niiden luotettavuutta. (Quandt ym. 2019)

Tässä työssä käsittelemme termiä ominaisuudessa, joka kuvaa seipitettyjä uutisia. Lisäksi yhdistämme termin disinformaation alaisuuteen, sillä termit sisältävät hyvin paljon samoja ominaisuuksia ja monet tieteelliset artikkelit käyttävät termiä valeuutiset, disinformaation sijaan, sen saaman huomion ja poliittisen käytön johdosta.

Tässä työssä luetaan myös disinformaatioksi disinformaatio, jota jaetaan edelleen, vaikka sen uudelleen jakaja rehellisesti uskoo sen sisällön. Määritelmän mukaisesti, mikäli tiedon levittäjä ei tiedä tiedon olevan väärää, se ei ole disinformaatiota, vaan pikemminkin misinformaatiota, sillä tällöin levitys ei pidä sisällään tarkoituksellista vääräntiedon levitystä. Työn luonteen vuoksi on kuitenkin olennaista tutkia myös tämän tiedon leviämistä, sillä sen alkuperäinen lähde on osa disinformaation levityskampanjaa ja sillä on todellinen funktio olla uskottava ja levitä laajalle ihmisten keskuudessa.

3.2 Disinformaation levittämisen motiivit

Edellisessä aliluvussa (3.1) määriteltiin disinformaatio, sekä rajattiin hieman sen käsitettä tutkielmaa varten. Tässä aliluvussa käsitellään disinformaation levityksen motiiveja ja mitä sitä leivittävät toimijat pyrkivät levityksellä saavuttamaan.

Kuten aikaisemmassa aliluvussa määriteltiin, niin disinformaation ja misinformaation suurimpana erona on se, että disinformaation levityksessä levittäjällä on jokin tarkoitus, jonka vuoksi hän levittää väärää tietoa. Vaikka tarkoituksellisen väärintiedon levittäminen voi tuntua mielettömältä, on kuitenkin useita motiiveja, jonka takia ihmiset valitsevat levittää tahallaan valheellista tietoa.

Tässä työssä disinformaation levityksen syyt jaetaan kolmeen pääkategoriaan; *Taloudelliset intressit, poliittiset intressit ja hedonistiset syyt*. Disinformaation levityksen motivaatio kuitenkin usein kulminoituu jonkinlaiseen kombinaatioon näistä kolmesta. (Marwick ja Lewis 2017)

3.2.1 Taloudelliset intressit

Sosiaalisen median saavuttaman suuren suosion myötä useat toimijat ovat käyttäneet etenkin sosiaalista mediaa ja niissä toimivia mainosverkostoja tahallisesti levittääkseen huhuperäistä, tai täysin väärää tietoa tarkoituksenaan ansaita sillä rahaa. Ilmiön lisääntymistä on edesauttanut Yhdysvaltojen silloisen presidentin Donald Trumpin äänekäs perinteisen median mollaaminen, kutsumalla valtamediaa usein puheissaan termillä ”*Valeuutiset*” (eng. ”*Fake news*”). (*Trump Calls Fox 'Fake News' | NowThis* 2020)

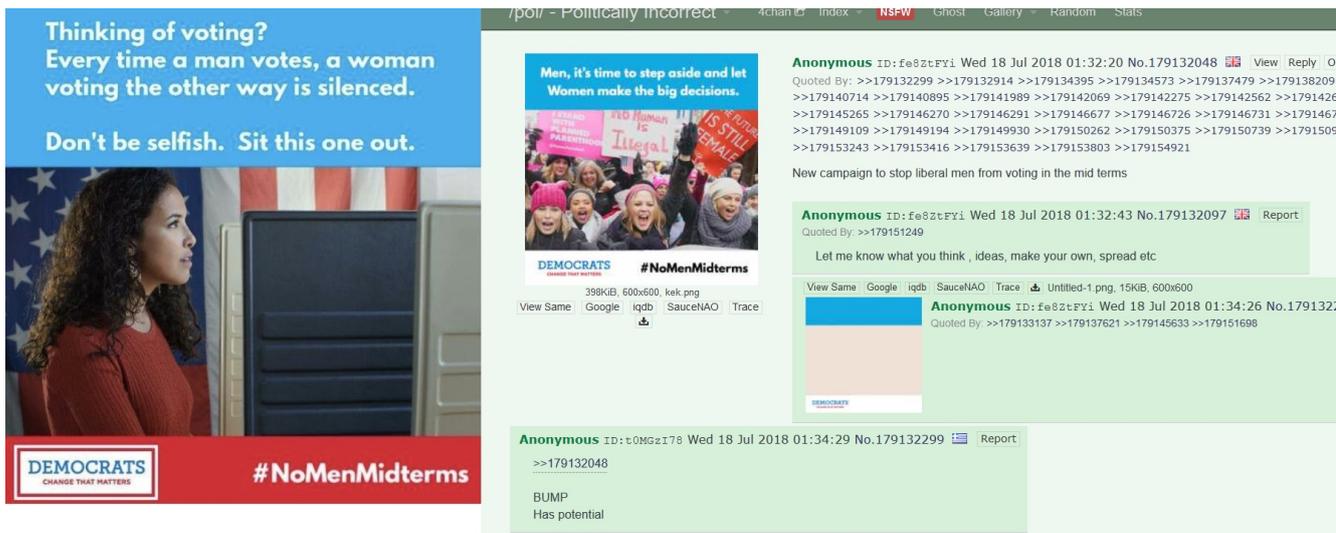
Esimerkiksi Paris Wadein ja Ben Goldmanin luoma valeuutissivusto keräsi vuonna 2016 yli 805 000 seuraajaa sivuston Facebook sivulle. Sivusto itsessään houkutteli lukijoita käyttämällä liioiteltuja ja jopa valheellisia otsikoita, joita he levittivät sosiaalisissa medioissa. Parsi Waden ja Ben Goldman ansaitsivatkin sivustonsa avulla parhaimmillaan jopa 40 000\$ mainostilaa myymällä. (McCoy 2016).

Ilmiön laajuutta kuvaa myös se, miten suuri osa vuoden 2016 Yhdysvaltojen presidentinvaaleihin kohdistuneesta disinformaatiosta oli lähtöisin pienestä makedonialaisesta kaupungista Velesistä, jossa luotiin yli sata valeutisiin tarkoitettua sivustoa. (Hughes ja Waismel-Manor 2021) Veles on suhteellisen köyhä kaupunki, jonka keskitulo on noin 380\$ / kk ja jossa on noin 24% työttömyysaste (Lidija Kostovska 2016). Mirko Ceselkoski perustikin juuri Velesiin hänen ”*the Facebook Marketing University*”:sä, joka kouluttaa nuoria ja aikuisia netti- ja Facebook sivujen ylläpitoon, sekä kuinka luoda hyviä klikkiotsikkoja ja miten ansaita niillä rahaa (Hughes ja Waismel-Manor 2021). Satunnaisesta valeutissivustosta saatava taloudellinen hyöty ei välttämättä houkuttele suomalaisia luomaan niitä massoittain, mutta jopa 200\$ kuukausittainen lisäansio on keskiverto ihmiselle Velesissä merkittävä summa.

3.2.2 Poliittiset intressit

Poliittista disinformaatiota levitetään yleensä propaganda tarkoituksessa, eli sen tarkoituksena on vaikuttaa ihmisten poliittisiin mielipiteisiin. Tällaisia ovat esimerkiksi kampanjat kuten #NoMenMidterms, jonka ideana oli yrittää saada liberaali miehet olemaan äänestämättä Yhdysvaltojen vuoden 2018 välivaaleissa. Tämä disinformaatio kampanja oli lähtöisin sivuston 4chan /pol/ kanavalta.

Ideologiseen disinformaation levitykseen voidaan myös liittää yritykset vaikuttaa ihmisten äänestys käyttäytymiseen, tai muuhun poliittiseen mielipiteeseen. Varsinkin sosiaalisen median käyttäminen osana poliittista vaikuttamista sai suurta huomiota vuoden 2018 Cambridge Analytica – skandaalin johdosta, jossa noin 90 miljoonan Facebook käyttäjän dataa oli kerätty salaa. Tätä dataa käytettiin poliittisten mainosten kohdentamiseen osana Ison-Britannian Brexit kampanjaa, sekä vuoden 2016 Yhdysvaltojen presidentin vaaleissa. (Isaak Jim 2018) Vuonna 2019 julkaistun Muellerin raportin mukaan sosiaalisen median alustoja, varsinkin Facebookia ja Twitteria, oli käytetty venäläisten toimijoiden toimesta tarkoituksena levittää disinformaatiota, sekä yllyttämään ihmisiä väkivaltaisuuksiin Yhdysvaltojen vuoden 2016 presidentin vaalien aikaan. (Mueller III 2019)



Kuvio 1. Esimerkki väärennetystä demokraattien mainoksesta ja 4chanin /pol/:issa sijaitse-
vasta langasta (eng. thread), jossa kannustettiin ihmisiä luomaan ja levittämään disinformaa-
tioita (*NoMenMidTerms* 2018)

3.2.3 Hedonistiset syyt

Levittäessä disinformaatiota hedonististen syiden vuoksi henkilö yleensä tekee sen jossain sosiaalisessa mediassa, kuten Facebookissa, Redditissa, tai Twitterissä. Näitä sosiaalisia medioita yhdistää ominaisuus, jossa arvostusta voidaan mitata tykkäysten, jakojen tai muiden reagoitien avulla. Tällaista huomiota voidaan sitten soveltaa rahan saamiseen, kuten LibertyWritersNews teki omalla sosiaalisen median sivullaan. Disinformaatiota saatetaan levittää myös tarkoituksena aiheuttaa ihmisille mielipahaa. Tällaisesta on useita esimerkkejä, kun 4chan:issä aikaansa viettävät ihmiset ovat aloittaneet laajoja disinformaatio kampanjoita mm. huijatakseen ihmisiä lataamaan iPhone puhelimiaan mikroaaltouunissa. Tällaiset pilailu ja mielipahan aiheuttamistarkoituksessa aloitetut disinformaatio kampanjat ovat kuitenkin lähinnä yksittäistapauksia, eikä niistä tämän takia ole tehty laajempaa tutkimusta.



Kuvio 2. Kuva, jota käytettiin sosiaalisessa mediassa huijaamaan ihmisiä lataamaan puhelimensa mikroaaltouunissa (Paton 2014)

4 Disinformaation levittäjien toimintamallit

Edellisessä aliluvussa (3.2) käsiteltiin, mitä motivaatioita disinformaation levitykseen on. Tässä luvussa käsitellään, millaisia tapoja disinformaation levittäjät perinteisesti käyttävät. Tapojen tunnistaminen on etenkin yhteiskunnallisesti tärkeää, jotta ihmiset kykenevät tunnistamaan milloin he lukevat sisältöä, joka saattaa olla disinformaatiokampanjan tuottamaa. Tunnistaessaan yleisimpiä tapoja voidaan informaatiovaikuttamista ehkäistä esimerkiksi raportoimalla sisältö alustan ylläpidolle.

4.1 Klikkiotsikot

Klikkiotsikoilla tarkoitetaan yleisesti otsikoita, jotka on tarkoituksella kirjoitettu provosoi-vasti. Niiden tavoitteena on saada ihmisten huomio nopeallakin vilkaisulla ja saada ihmiset klikkaamaan itsensä artikkeliin, joka taas yleensä sijaitsee artikkelin kirjoittajan omilla verkkosivuilla. Tavoitellessaan mittavaa taloudellista hyötyä disinformaation levittäjät pyrkivät saamaan ihmisiä omille sivuilleen, joihin he voivat myydä mainoksia. Tätä varten he yleensä hyödyntävät esimerkiksi sosiaalisen median alustoja, luoden klikkiotsikoita, jotka houkuttelevat ihmisiä siirtymään pois sosiaalisen median alustalta ja heidän valeutissivustolleen. Tällainen on juuri esimerkiksi aikaisemmin työssä mainitun Valesista toimivien valeutissivustojen, sekä LibertyWritersNews:in toimintamalli.

4.2 Sosiaaliset botit

Sosiaalisilla boteilla tarkoitetaan tietokoneohjelmia, joiden päätarkoituksena on esiintyä oikeana ihmisenä sosiaalisessa mediassa. Botin toissijaisena tarkoituksena voi olla tiedon kerääminen ja oikeiden ihmisten kanssakäyminen. (Woolley 2016)

Poliittisessa disinformaation levityksessä, hyvin yleinen keino on käyttää edistyneitä sosiaalisia botteja, jotka saattavat muodostaa suuria bottiverkkoja. Sosiaaliset botit ovatkin varsin tehokkaita mis- ja disinformaation levittäjiä, sillä niiden ylläpito on todella halpaa, jonka vuoksi niitä myös on todella paljon. Vuonna 2014 Twitter julkisti raportissaan, että jopa 23

miljoonaa käyttäjää sivulla olivat todellisuudessa sosiaalisia botteja. (Woolley 2016)

Poliittiset toimijat voivat käyttää sosiaalisia botteja esimerkiksi poliitikon seuraajamäärän keinotekoiseen kasvattamiseen, joka antaa vaikutelman siitä, että poliitikko olisi huomattavasti suositumpi, kuin mitä hän oikeasti on. Botteja voidaan lisäksi käyttää tukemaan tiettyä näkökulmaa poliittisesta aiheesta, asettamalla ne kirjoittamaan postauksia, jotka tukevat näkökulmaa, tai sitä ajavaa poliitikkoa. (Woolley 2016)

Tahot, joilla on tarvittavat resurssit, kykenevätkin valjastamaan käyttöönsä jopa tuhansia botteja edistääkseen asiaansa, tai hyökkäämään heidän agendaansa vastustavien tahojen kimppuun. (Bessi ja Ferrara 2016) Esimerkiksi Twitterissä botit voivat käyttää samaa hashtagia, kuin muut asian puolella olevat ihmiset. Tämä antaa vaikutelman, että asialla on paljon suurempi kannatus, kuin mitä sillä todellisuudessa on. Bottien käyttö on myös varsin tehokas vaikuttamisen keino, sillä ihmiset luottavat todella paljon muihin sosiaalisen piirinsä ihmisiin. (Jagatic ym. 2007) Sosiaaliset botit herättävätkin ihmisten luottamuksen käyttäytymällä kuten oikeat ihmiset.

4.3 Trollitehtaat

Trollitehtaat (eng. *Troll farms*) ovat organisaatioita, joiden tehtävä on levittää propagandaa ja vaikuttaa yhteiskunnalliseen keskusteluun. Trollitehtaat maksavat ihmisille, joiden tehtävänä tehtaalla on pyörittää useita sosiaalisen median tilejä, blogeja tai vastaavia sosiaalisen median kanavia, joissa he esiintyvät tuiki tavallisina ihmisinä, mutta käytännössä pyrkivät ohjaamaan poliittista keskustelua heidän palkkaaman tahon haluamaan suuntaan. (Nasrin ym. 2019) Eräs tunnetuimpia trollitehtaita on venäläinen IRA (*Internet Research Agency*), jota myös joskus sanotaan RU-IRA:ksi, jotta sitä ei sekoitettaisi Irlannin tasavaltalaisarmeijaan. RU-IRA on vuonna 2013 perustettu informaatiovaikuttamisen keskus, jossa työskentelee hieman yli tuhat ihmistä. IRA on hyvin merkittävä trollitehdas, koska sillä on hyvin läheiset suhteet Kremliin (Kenworthy 2019)

Joillakin trollitehtailla työntekijöille annettiin päivittäin teemat, joiden pohjalta heidän tuli kirjoitella viestejä, useilla eri käyttäjillä, sosiaaliseen mediaan. (Aro Jessika 2020) Käytännössä keskustelussa käytetään kiusaamistaktiikoita, joissa esimerkiksi erimieltä olevia nimittellään ja uhkaillaan. Toisin sanoen tavoitteena on estää asiasta erimieltä olevien osallistuminen keskusteluun ja kaapata täten narratiivi. Trollit saattavat satunnaisesti levittää hyvin samanlaisia viestejä. Käytännössä kuitenkin tehtaot ovat hyvin ankaria työpaikkoja, joissa tehtaiden työntekijöitä rangaistaan sakoilla, mikäli he eivät pääse julkaisutavoitteisiin, tai he jäävät kiinni siitä, että he ovat kopioineet viestejään suoraan. (Walker 2018)

5 Yhteenveto

Edellisissä kolmessa luvussa käsiteltiin suosittelujärjestelmiä, niiden merkitystä, toimintaa ja sitä miten ne luovat kaikukammioita. Lisäksi käsiteltiin disinformaatiota, sekä sen levittämisen syitä ja tapoja. Vaikka tässä kirjallisuuskatsauksessa on levittämiseen käytetyt tavat eroteltu kolmeen yleisimpään, on kuitenkin muistettava, että levitys tyyliä kulminoituvat hyvin usein jonkinlaiseen kombinaatioon näistä kolmesta. Esimerkiksi Venäjällä toimiva informaatio vaikuttamiseen keskittynyt keskus RU-IRA käyttää toiminnassaan sosiaalisten bottien verkostoja, sekä trollitehtaita, jotka tukevat toisiaan. (Kenworthy 2019).

Suosittelujärjestelmien tehoa ja toimivuutta on perinteisesti tutkittu paljon etenkin algoritmien tarkkuuden, sekä onnistuneiden suositteluiden perusteella. Näissä tutkimuksissa tarkastelun kohteena on usein ollut lähinnä tulevien valintojen ennustus suosittelujärjestelmien avulla. (Pu, Chen ja Hu 2011) Tässä työssä pyrittiin kokoamaan kattavaa kirjallisuuskatsausta, miten suosittelujärjestelmät vaikuttavat poliittiseen diskurssiin, etenkin dis- ja misinformaation leviämiseen vaikuttavana tekijänä.

Tässä viimeisessä luvussa pyritään yhdistämään miten suosittelujärjestelmät ja disinformaation leviämisen mallit yhdistyvät. Käymme läpi, miten disinformaation levittäjät hyödyntävät suosittelujärjestelmiä levittäessään väärää tietoa. Tämän lisäksi käydään myös läpi mitä asialle on jo tehty ja mitä sille voitaisiin vielä tehdä.

Tämän kirjallisuuskatsauksen perustella on ilmeistä, että suosittelujärjestelmillä on massiivinen rooli siinä mitä mediaa ihmiset kuluttavat ja etenkin, miten he sitä kuluttavat. Tämä voi kuitenkin luoda suuria ongelmia, sillä mikäli järjestelmät suosittelevat käyttäjille vain uutisia, jotka heitä kiinnostavat, tai uutisia, joista järjestelmä ajattelee heidän pitävän, sillä on suuri mahdollisuus kaventaa käyttäjien maailmankuvaa. Tutkimuksissa on todettu, että mikäli käyttäjien valitsema uutispalvelu käytti suosittelujärjestelmää uutisten suodatuksessa, jopa 50% käyttäjien lukemista uutisista tuli suoraan suosittelujärjestelmältä. (Yang 2016)

Aikaisemmassa luvussa käsittelin sosiaalisten bottien toimintaa disinformaation levityksessä. Sosiaaliset botit ovatkin hyvin yleinen ja suosittu keino, sillä ne ovat hyvin kustannus tehokkaita, eivätkä nykyään välttämättä vaadi edes ohjelmointitaitoa niitä käyttäviltä tahoilta. Automatisoitujen bottien tuottaman sisällön määrä onkin niin mittavaa, ettei normaaleilla ihmisillä voi mitenkään olla aikaa tarkastaa jokaisen uutisen lähdettä. Suosittelevjärjestelmien aiheuttamat kaikukammiot pahentavat tilannetta entisestään, sillä tunnistamalla kaikukammion, voidaan disinformaatiota sisältävä narratiivi kohdentaa suoraan halutulle yleisölle. (Shao ym. 2017) Tällöin tämä disinformaatio leviää luonnostaan ihmisten keskuudessa, jotka ovat jo valmiiksi samaa mieltä sen sisältämän narratiivin kanssa.

Varsinkin Twitter käyttää hashtageja ("#") rajaamaan aiheita ja niiden avulla käyttäjät voivat osallistua helpommin eri aiheiden keskusteluihin. ("How to use hashtags" 2020) On kuitenkin muistettava, että vaikka kaikukammiot ja informaatiokuplat määrittyvät niiden erottautumisella muusta sisällöstä, niin hashtagien merkityksistä käydään ajoittain aktiivista taistelua. (Hadgu, Garimella ja Weber 2013) Hyvänä viimeaikaisena esimerkkinä on *ProudBoys*, joka "valloitettiin" äärioikeistolaiselta liikkeeltä, joka kantaa samaa nimeä. (Halminen 2020)

Twitter myös esittää käyttäjilleen sisältöä, joka on "nousussa" (eng. "trends") Tämä sisältö määrittyy lähtökohtaisesti suosittelualgoritmien kautta. Sisältöön vaikuttaa etenkin ketä käyttäjä seuraa, missä käyttäjä on maantieteellisesti ja mitä käyttäjä on ilmoittanut, tai mitä järjestelmä on havainnut, käyttäjää kiinnostavan. Nousussa oleva sisältö myös usein kokoaa samankaltaisia hashtageja samaan näkymään. ("Twitter Trends FAQ" 2020)

Sosiaalisten bottien avulla on mahdollista valloittaa varsinkin vähemmän käytössä olevia hashtageja ja täten vaikuttaa ihmisten "nousussa" näkymään, sillä algoritmit eivät välttämättä ehdi huomaamaan yksittäisen hashtagin tarkoituksen vaihtumista ja liittää sen yhä vanhaan asiasisältöön. Sosiaalisilla boteilla, voidaan myös kasvattaa tietyn näkökulman sosiaalisen median päivityksen saamia jakoja, tykkäyksiä tai muita positiivisia reagoiteja hyvin nopeasti ja todella paljon.

Mikäli päivitys saa suuren määrän positiivisia reagoiteja, se johtaa todennäköisemmin siihen, että myös oikeat ihmiset jakavat sitä ja sen sanoma leviää myös niiden ihmisten keskuuteen, jotka eivät seuraa normaalisti tätä näkökulmaa tukevia päivityksiä. Tämä taas puolestaan saattaa aiheuttaa suljetun kehän, sillä nämä kammion ulkopuoliset ihmiset lisäävät päivityksen suosiota reagoimalla siihen ja jakamalla sitä.

Trollitehtaiden trollit ovat olleet erityisen aktiivisia suomalaisten käyttämällä sivuilla varsinkin Ukrainan sodan alkamisen jälkeen. (Korkki 2015) Trollit käytännössä käyttävät suosittelujärjestelmiä hyödykseen iskeytymällä esimerkiksi Twitterin hashtagihin, jossa ne levittävät omaa narratiiviaan ja kiusaavat erimieltä olevia pois keskustelusta. Lisäksi Facebookissa ja Redditissä näiden trollien on helppoa soluttautua niiden sisältämiin sisäisiin ryhmiin. Nämä ryhmät ovat käytännössä orgaanisia kaikkukammioita, jotka perustuvat jonkin asian ympärille. Tällöin kaikki sisältö, jota trollit ryhmään tuottavat näytetään kaikille ryhmän jäsenille heidän etusivuillaan.

Sosiaalisen median alustoja ei ole alun perin luotu uutisten levittämistä varten, joten niistä puuttuu usein keinot väärän tiedon ilmiantamiseen. Tämä on suuri ongelma, sillä nykyään noin 62 % yhdysvaltalaisista aikuisista saa uutisensa sosiaalisen median kautta. (Greenwood, Perrin ja Duggan 2016). Tähänkin on jo onneksi alettu puuttumaan, mutta vasta viime aikoina esimerkiksi Twitter ja Facebook ovat tehneet konkreettisia toimia disinformaation, sekä misinformaation leviämisen ehkäisemiseen.

Facebook on poistanut paljon misinformaatiota sisältäviä postauksia koronavirukseen liittyen ja kieltänyt kokonaan koronarokotteen vastaiset mainokset, sekä postaukset. (Kang-Xing 2020) Twitter on muun muassa lisännyt merkintöjä postauksiin, jotka saattavat sisältää disinformaatioita, tällaisia ovat esimerkiksi olleet entisen Yhdysvaltojen presidentin Donald Trumpin Twiitit liittyen vaalien varastamiseen. (Roth Yoel 2020) Suosittelujärjestelmien aiheuttaman disinformaation leviämisen tutkiminen on todella tärkeää, sillä disinformaation leviämisen ehkäiseminen on tärkeää länsimaisten demokratioiden olemassa-olon kannalta.

Olemme tottuneet elämään demokraattisissa hyvinvointivaltioissa, jossa on vapaa ja sensuroimaton media ja Suomessa onkin maailman toiseksi vapain lehdistö. (Harjumaa 2020) Lisäksi lehdistömme nauttii yleisesti suurta luottamusta kansalaisten keskuudessa. (Matikainen Janne 2020) Uutislähteisiin sokeasti luottaminen on kuitenkin myös ongelmallista, sillä useat, etenkin englanninkieliset, valeuutislähteet pyrkivät joskus matkimaan tunnettuja valtavirtamedian lähteitä. (“List of fake news websites” 2020) Tämän takia on aina hyvä tarkistaa uutislähteen linkin todenperäisyys, sekä varsinkin lukea otsikkoa pidemmälle.

Palvelut kuitenkin siirtyvät yhä enemmän ja enemmän internettiin ja niiden käyttämät suosittelujärjestelmät rajaavat hyvin paljon sitä informaatiota, jota näemme verkossa. Tämän ongelman tiedostaminen on elintärkeää demokraattisen valtion tulevaisuuden kannalta. Tulevaisuudessa tutkimuksen aiheena voisikin olla: *”Miten paljon jatkuva polarisoituneille mielipiteille altistuminen vaikuttaa yhteiskunnalliseen käyttämiseen”*

Lähteet

Allcott, Hunt, ja Matthew Gentzkow. 2017. “Social media and fake news in the 2016 election”. *Journal of economic perspectives* 31 (2): 211–36.

Amatriain, Xavier, ja Justin Basilico. 2012. “Netflix recommendations: Beyond the 5 stars (part 1)”. *Netflix Tech Blog* 6. <https://netflixtechblog.com/netflix-recommendations-beyond-the-5-stars-part-1-55838468f429>.

Aro Jessika, Mäkeläinen Mika. 2020. “Kioski Venäjän some-propagandan alkulähteillä”. *Yle*, <https://kioski.yle.fi/omat/kioski-pietarin-trollitehtaalla>.

Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua A Tucker ja Richard Bonneau. 2015. “Tweeting from left to right: Is online political communication more than an echo chamber?”. *Psychological science* 26 (10): 1531–1542.

Bessi, Alessandro, Mauro Coletto, George Alexandru Davidescu, Antonio Scala, Guido Caldarelli ja Walter Quattrociocchi. 2015. “Science vs conspiracy: Collective narratives in the age of misinformation”. *PloS one* 10 (2): e0118093.

Bessi, Alessandro, ja Emilio Ferrara. 2016. “Social bots distort the 2016 US Presidential election online discussion”. *First Monday* 21 (11-7).

Bruns, Axel. 2017. “Echo chamber? What echo chamber? Reviewing the evidence”. Teoksessa *6th Biennial Future of Journalism Conference (FOJ17)*.

Cinelli, Matteo, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi ja Michele Starnini. 2020. “Echo chambers on social media: A comparative analysis”. *arXiv preprint arXiv:2004.09603*.

Dubois, Elizabeth, ja Grant Blank. 2018. “The echo chamber is overstated: the moderating effect of political interest and diverse media”. *Information, communication & society* 21 (5): 729–745.

- Edwards, Arthur. 2013. “(How) do participants in online discussion forums create ‘echo chambers’?: The inclusion and exclusion of dissenting voices in an online forum about climate change”. *Journal of Argumentation in Context* 2 (1): 127–150.
- Fallis, Don. 2015. “What is disinformation?” *Library trends* 63 (3): 401–426.
- Fetzer, James H. 2004. “Information: Does it have to be true?” *Minds and Machines* 14 (2): 223–229.
- Gilbert, Eric, Tony Bergstrom ja Karrie Karahalios. 2009. “Blogs are echo chambers: Blogs are echo chambers”. Teoksessa *2009 42nd Hawaii International Conference on System Sciences*, 1–10. IEEE.
- Gomez-Uribe, Carlos A, ja Neil Hunt. 2015. “The netflix recommender system: Algorithms, business value, and innovation”. *ACM Transactions on Management Information Systems (TMIS)* 6 (4): 1–19.
- Greenwood, Shannon, Andrew Perrin ja Maeve Duggan. 2016. “Social media update 2016”. *Pew Research Center* 11 (2): 1–18.
- Hadgu, Asmelash Teka, Kiran Garimella ja Ingmar Weber. 2013. “Political hashtag hijacking in the US”. Teoksessa *Proceedings of the 22nd international conference on World Wide Web*, 55–56.
- Halminen, Laura. 2020. “Proud boys -äärioikeistoryhmän nimi valjastettiin uuteen käyttöön Twitterissä: aihe-tunniste täyttyi onnellisten miespariskuntien kuvilla”. *Helsingin Sanomat*, <https://www.hs.fi/ulkomaat/art-2000006659785.html>.
- Harjumaa, Marika. 2020. “Suomi jatkaa kakkosena lehdistönvapaudessa – pandemia kaventaa sananvapautta, arvioi toimittajajärjestö”. *Yle*, <https://yle.fi/uutiset/3-11315019>.
- “How to use hashtags”. 2020, <https://help.twitter.com/en/using-twitter/how-to-use-hashtags>.
- Hughes, Heather C, ja Israel Waisman-Manor. 2021. “The Macedonian Fake News Industry and the 2016 US Election”. *PS: Political Science & Politics* 54 (1): 19–23.
- Isaak Jim, Hanna Mina J. 2018. “User data privacy: Facebook, Cambridge Analytica, and privacy protection”. *Computer* 51 (8): 56–59.

- Jagatic, Tom N, Nathaniel A Johnson, Markus Jakobsson ja Filippo Menczer. 2007. “Social phishing”. *Communications of the ACM* 50 (10): 94–100.
- Jiang, Ray, Silvia Chiappa, Tor Lattimore, András György ja Pushmeet Kohli. 2019. “Degenerate feedback loops in recommender systems”. Teoksessa *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 383–390.
- Kang-Xing, Jing. 2020. “Keeping people safe and informed about the coronavirus”. *About Facebook*, <https://about.fb.com/news/2020/12/coronavirus/>.
- Kenworthy, Steven. 2019. “Homophily and the Internet Research Agency: How a Russian Troll Farm Exploits American Social Media Behavior”, https://www.researchgate.net/publication/333092246_Homophily_and_the_Internet_Research_Agency_How_a_Russian_Troll_Farm_Exploits_American_Social_Media_Behavior.
- Knijnenburg, Bart P, Martijn C Willemsen, Zeno Gantner, Hakan Soncu ja Chris Newell. 2012. “Explaining the user experience of recommender systems”. *User Modeling and User-Adapted Interaction* 22 (4): 441–504.
- Koren, Yehuda, Robert Bell ja Chris Volinsky. 2009. “Matrix factorization techniques for recommender systems”. *Computer* 42 (8): 30–37.
- Koriat, Asher, Sarah Lichtenstein ja Baruch Fischhoff. 1980. “Reasons for confidence.” *Journal of Experimental Psychology: Human learning and memory* 6 (2): 107.
- Korkki, Jari. 2015. “Suomalaisia herätellään informaatio-sotaan”. *Yle*, <https://yle.fi/uutiset/3-7764063>.
- Levin, Matthew A, Jonathan P Wanderer ja Jesse M Ehrenfeld. 2015. “Data, big data, and metadata in anesthesiology”. *Anesthesia & Analgesia* 121 (6): 1661–1667.
- Li, Lei, Ding-Ding Wang, Shun-Zhi Zhu ja Tao Li. 2011. “Personalized news recommendation: a review and an experimental investigation”. *Journal of computer science and technology* 26 (5): 754–766.
- Lidija Kostovska, Director. 2016. “Statistical album of the Republic of Macedonia, 1991-2011”, https://www.stat.gov.mk/PrikaziPoslednaPublikacija_en.aspx?id=75.

- Lops, Pasquale, Marco De Gemmis ja Giovanni Semeraro. 2011. “Content-based recommender systems: State of the art and trends”. *Recommender systems handbook*, 73–105.
- Marwick, Alice, ja Rebecca Lewis. 2017. “Media manipulation and disinformation online”. *New York: Data & Society Research Institute*.
- Matikainen Janne, Komulainen Pasi. 2020. “Luottamus mediaan korkealla, kriittinen suhtautuminen silti kasvussa”. *Helsingin Yliopisto*, <https://www2.helsinki.fi/fi/uutiset/talous-yhteiskunta/luottamus-mediaan-korkealla-kriittinen-suhtautuminen-silti-kasvussa>.
- McCoy, Terrance. 2016. “Inside a Long Beach Web operation that makes up stories about Trump and Clinton: What they do for clicks and cash”. *Los Angeles Times*, <https://www.latimes.com/business/technology/la-fi-tn-fake-news-20161122-story.html>.
- Mueller III, Robert S. 2019. “Report On The Investigation Into Russian Interference In The 2016 Presidential Election. Volumes I & II.(Redacted version of 4/18/2019)”, <https://digitalcommons.unl.edu/usjusticematls/47/>.
- Nasrin, Nayeema, Kim-Kwang Raymond Choo, Myung Ko ja Anthony Rios. 2019. “How Many Users Are Enough? Exploring Semi-Supervision and Stylometric Features to Uncover a Russian Troll Farm”. Teoksessa *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*, 20–30.
- Nickerson, Raymond S. 1998. “Confirmation bias: A ubiquitous phenomenon in many guises”. *Review of general psychology* 2 (2): 175–220.
- Pariser, Eli. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
- Paton, Callum. 2014. “A screenshot of the iPhone 6 wave hoax currently being circulated on social media. iPhone 'Wave': iOS 8 hoax claims you can charge your iPhone in the microwave - you can't”, <https://www.independent.co.uk/life-style/gadgets-and-tech/news/ios-8-hoax-claims-you-can-charge-your-iphone-microwave-using-new-operating-system-s-wave-technology-you-can-t-9743281.html>.
- Pazzani, Michael J. 1999. “A framework for collaborative, content-based and demographic filtering”. *Artificial intelligence review* 13 (5): 393–408.

- Pu, Pearl, Li Chen ja Rong Hu. 2011. “A user-centric evaluation framework for recommender systems”. Teoksessa *Proceedings of the fifth ACM conference on Recommender systems*, 157–164.
- Quandt, Thorsten, Lena Frischlich, Svenja Boberg ja Tim Schatto-Eckrodt. 2019. “Fake news”. *The international encyclopedia of Journalism Studies*, 1–6.
- Quattrociocchi, Walter, Antonio Scala ja Cass R Sunstein. 2016. “Echo chambers on Facebook”. Available at SSRN 2795110.
- Ricci, Francesco, Lior Rokach ja Bracha Shapira. 2011. “Introduction to recommender systems handbook”. Teoksessa *Recommender systems handbook*, 1–35. Springer.
- . 2015. “Recommender systems: introduction and challenges”. Teoksessa *Recommender systems handbook*, 1–34. Springer.
- Roth Yoel, Pickles Nick. 2020. “Updating our approach to misleading information”. *Twitter Blog*, https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information.html.
- Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Alessandro Flammini ja Filippo Menczer. 2017. “The spread of fake news by social bots”. *arXiv preprint arXiv:1707.07592* 96:104.
- Shearer, Elisa, ja Amy Mitchell. 2021. “News use across social media platforms in 2020”. *Pew Research Center*.
- Sinha, Rashmi R, Kirsten Swearingen ym. 2001. “Comparing recommendations made by online systems and friends.” *DELOS* 106.
- Sunstein, Cass R. 1999. “The law of group polarization”. *University of Chicago Law School, John M. Olin Law & Economics Working Paper*, numero 91.
- Tamás, Bodoky. 2009. “Mátrixfaktorizáció egymillió dollárért”. *index*, https://index.hu/tech/net/2009/08/07/matrixfaktorizacio_egymillio_dollarert/.
- Trump Calls Fox 'Fake News' | NowThis*. 2020. <https://www.youtube.com/watch?v=-5Rd64QXkjI>.

“Twitter Trends FAQ”. 2020, <https://help.twitter.com/en/using-twitter/twitter-trending-faqs>.

Walker, Shaun. 2018. “This article is more than 6 years old The Russian troll factory at the heart of the meddling allegations”, <https://www.theguardian.com/world/2015/apr/02/putin-kremlin-inside-russian-troll-house>.

“List of fake news websites”. 2020. *Wikipedia*, https://en.wikipedia.org/wiki/List_of_fake_news_websites.

Viķe-Freiberga Vaira, Hammersley Ben, Däubler-Gmelin Herta, ja Maduro Luis Miguel Pires Pessoa. 2013. “A free and pluralistic media to sustain European democracy”. *Report, High Level Group on Media Freedom and Media Pluralism*.

Woolley, Samuel C. 2016. “Automating power: Social bot interference in global politics”. *First Monday*, <https://firstmonday.org/ojs/index.php/fm/article/download/6161/5300>.

Yang, JungAe. 2016. “Effects of popularity-based news recommendations (“most-viewed”) on users’ exposure to online news”. *Media Psychology* 19 (2): 243–271.