

Hyvät pahat algoritmit - Asiantuntijoiden näkemyksiä sosiaalisen median algoritmeista ja käyttäjän mahdollisuuksista vaikuttaa sisällön personointiin

Laura Trémouille

Helsingin yliopisto
Valtiotieteellinen tiedekunta
Viestintä
Pro gradu -tutkielma
Marraskuu 2020



Tiedekunta – Fakultet – Faculty Valtiotieteellinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree Programme Politiikan ja viestinnän koulutusohjelma	
Tekijä – Författare – Author Laura Vilhelmiina Trémouille			
Työn nimi – Arbetets titel – Title Hyvät pahat algoritmit - Asiantuntijoiden näkemyksiä sosiaalisen median algoritmeista ja käyttäjän mahdollisuuksista vaikuttaa sisällön personointiin			
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Viestintä			
Työn laji – Arbetets art – Level Pro Gradu -tutkielma 30op		Aika – Datum – Month and year Marraskuu 2020	Sivumäärä – Sidoantal – Number of pages 56
Tiivistelmä – Referat – Abstract			
<p>Tämän tutkimuksen tarkoitus on kartoittaa asiantuntijoiden suhtautumista sosiaalista mediaa ja niiden algoritmeja sekä personointia kohtaan. Tutkimuksen lähtökohtana on sosiaalisen median muuttama tapamme vastaanottaa tietoa maailmasta ja ajankohtaisista aiheista. Sosiaalisen median vaikutus esimerkiksi sosiaalisten kuplien muodostumiseen on puhuttanut julkisuudessa paljon, mutta aihe jakaa tutkijoiden mielipiteitä vahvasti. Vaikka monet tutkijat ovat kyseenalaistaneet kuplien muodostumisen, niistä puhutaan edelleen paljon. Ja vaikka julkisessa keskustelussa on lisääntynyt puhe sosiaalisen median mahdollisista haittapuolista ja sen seurauksena luottamus palveluita kohtaan on käyttäjien keskuudessa horjunut, palveluiden käyttö ei ole suinkaan vähentynyt. Tässä tutkimuksessa pyrim selvittämään asiantuntijoiden näkemyksiä siitä, kuinka sosiaalinen media vaikuttaa tapamme saada tietoa ja mikä rooli käyttäjän omalla toiminnalla on tiedon kulussa.</p> <p>Tutkimus toteutettiin haastattelemalla neljää sosiaalisen median ja algoritmien parissa työskentelevää asiantuntijaa. Kaksi haastateltavaa työskentelee tutkijana ja kaksi sosiaalisen median algoritmien parissa konsultteina. Haastattelut analysoitiin teoriaohjaavaa sisällönanalyysia hyödyntäen.</p> <p>Haastateltujen asiantuntijoiden näkemykset sosiaalisesta mediasta ja algoritmeista sekä suhtautuminen niitä kohtaan mukailivat vahvasti yleistä asiantuntijoiden suhtautumista: he näkivät palveluissa paljon hyvää, mutta suhtautuivat kriittisesti moniin vallitseviin suuntauksiin ja kehityskuluihin. Haastateltavien näkemyksen mukaan palveluiden käyttäjillä on merkittävä vaikutus siihen, millaista sisältöä algoritmit näille personoivat. Käyttäjät eivät kuitenkaan ole tietoisia omasta vaikutuksestaan tai mahdollisuuksistaan vaikuttaa sisältöjen personointiin. Kaikissa haastatteluissa korostui sosiaalisen median kasvanut merkitys ihmisten arjessa. Sosiaalisen median hyödyt yksittäisille käyttäjille ovatkin niin suuret, että mahdollisia haittoja on helppo ylenkatsoa. Sen sijaan, että käyttäjien tietoisuutta potentiaalisista ongelmista pyrittäisiin lisäämään, ehkä tärkeämpää olisikin saada käyttäjät välittämään.</p> <p>Vaikka sosiaalisen median vaikutusta sosiaalisten kuplien muodostumiselle on monissa tutkimuksissa kyseenalaistettu, ei ajatusta niistä kannata kuitenkaan haudata ja hylätä. Erityisesti, koska sosiaalisen median palvelut ja niihin liittyvä teknologia kehittyvät huomattavasti nopeammin kuin kattava tutkimustieto aiheesta, on sosiaaliseen mediaan liittyvä keskustelu ajankohtaista ja tärkeää.</p>			
Avainsanat – Nyckelord – Keywords Sosiaalinen media, algoritmit, sosiaaliset kuplat, personointi			
Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors Juha Herkman			
Säilytyspaikka – Förvaringställe – Where deposited Helsingin yliopiston kirjasto, Helsingfors universitets bibliotek, Helsinki University Library			
Muita tietoja – Övriga uppgifter – Additional information			

Sisällysluettelo

1 Johdanto	1
2 Teoreettinen tausta	4
2.1 Keskeiset käsitteet	4
2.1.1 Algoritmit	4
2.1.2 Personointi.....	5
2.1.3 Sosiaaliset kuplat.....	7
2.2 Kuinka algoritmit ovat muuttaneet ihmisten tapaa vastaanottaa tietoa?	9
2.3 Datan keruun yksityisyyden suoja sosiaalisen median palveluissa	13
2.4 Kuinka sisältöä personoidaan sosiaalisen median palveluissa?	15
3 Tutkimustehtävä ja tutkimuskysymykset.....	17
4 Tutkimuksen toteutus.....	18
4.1 Tutkimuskohteen kuvaus.....	18
4.2 Haastattelututkimus.....	18
4.3 Haastatteluiden toteutus.....	20
4.4 Analyysi ja luokittelu.....	20
5 Tutkimustulokset ja niiden tulkintaa	22
5.1 "Some on tosi-tv:tä pahimmillaan" – Haastateltavien suhtautuminen sosiaalisen median palveluihin	22
5.2 Käyttäjän toiminnan vaikutus personointiin	31
5.2.1 Käyttäjistä kerättävät tiedot - "Kaikki mitä irti lähtee"	32
5.2.2 Miten sisältöjen personointi tapahtuu?.....	36
5.3 Mitä sosiaalisen median palveluiden käyttäjien tulisi palveluista ymmärtää	40
5.4 Johtopäätökset.....	44
6 Luotettavuus.....	47
7 Pohdintaa.....	49
Lähteet.....	52
Liitteet	56

1 Johdanto

Sosiaalisesta mediasta on lyhyessä ajassa tullut keskeinen osa ihmisten elämää. Käytämme palveluja niin verkostoitumiseen, viihtymiseen kuin tiedon hankintaan ja vietämme palvelujen parissa tai vähintään niiden tavoitettavissa suuren osan päivästä. Samalla sosiaalinen media on muuttanut tapaamme muodostaa käsitystä ympäröivästä maailmasta ja sen tapahtumista. Palveluihin asettuu sulaan sopuun ja samalle viivalle sekä kansainväliset uutiset, läheisten merkkipäivät, mainokset, että entisen naapurin iltapalakuvat ja tätä loputonta sisältövirtaa selaamalla muodostamme kuvaa siitä, mitä maailmassa tapahtuu ja mikä on tärkeää.

Vuonna 2016 Yhdysvalloissa tehdyssä tutkimuksessa havaittiin, että amerikkalaiset saivat suurimman osan uutisistaan sosiaalisesta mediasta (Dunkle, 2016) ja sama kehitys on havaittu myös Suomessa. Digital News Report 2019 -kyselytutkimuksen mukaan puolet suomalaisista ilmoitti verkon pääasialliseksi uutislähteekseen. Erityisesti tämä korostui nuorissa ikäryhmissä. (Media-alan tutkimussäätiö, 2019.) Monet tutkijat ovat kuitenkin huolissaan siitä, että sosiaalisen median kautta saadut uutiset luovat käyttäjille yksipuolisen kuvan maailmasta ja sen tapahtumista, sillä palveluita personoidaan kullekin käyttäjälle sopivaksi. Jos se mitä minä näen palveluiden kautta ja sitä kautta pidän ajankohtaisena ja tärkeänä, on esimerkiksi ristiriidassa sen kanssa mitä puolisoni näkee, kiinnostavien päivälliskeskustelujen käyminen ajankohtaisista aiheista voidaan kuvitella olevan haastavaa. Ja kun tämä tapahtuu yhteiskunnallisella tasolla eri ryhmien välillä, sillä saattaa olla seurauksensa yhteisen konsensuksen ja suunnan löytämiselle.

Sosiaalisen median tuomia muutoksia yhteiskuntaan on tutkittu runsaasti. Monet tutkimukset ovat todenneet sosiaalisen median vaikuttaneen merkittävästi muun muassa populismin nousuun, mielipiteiden polarisaatioon ja luottamukseen instituutioita ja tiedettä kohtaan (ks. esim. Hiilamo, 2018; Hatakka, 2019). Vaikka edelleen monet käyttävät sosiaalista mediaa uutislähteenä, viime aikoina suhtautuminen sosiaalista mediaa kohtaan on muuttunut. Esimerkiksi Sanomalehtien Liiton vuonna 2018 teettämän Yhteisöllistyvä media -tutkimuksen mukaan luottamus sosiaalista mediaa kohtaan on vähentynyt merkittävästi

(Sanomalehtien Liitto, 2018). Luottamusta on osaltaan vähentänyt runsas negatiivinen huomio, jota monet sosiaalisen median yhtiöt ovat saaneet osakseen (Matikainen ym., 2020). Erityisesti Facebookia on kritisoitu muun muassa sen tavasta kerätä ja myydä tietoa käyttäjistä markkinoijille ja erityisesti Yhdysvaltojen vuoden 2016 presidentinvaalien jälkeen sosiaalisen median palveluiden valta ja vaikutus on herättänyt laajaa huolta. Huolimatta siitä, että käyttäjien ja valtaa pitävien päättäjien tietoisuus sosiaalisen median mahdollisista riskeistä on lisääntynyt, palveluiden käyttö ei ole suinkaan vähentynyt.

Suuri osa julkisuudessa näkyvyyttä saaneesta sosiaalisen median kritiikistä on ansaitusti kohdistunut nimenomaan palveluita ja yhtiöitä kohtaan. Vaikka tämä kritiikki on aiheellista ja keskustelu siitä erittäin tärkeää, kaikki sosiaaliseen mediaan liittyvät haasteet eivät liity alustoihin ja yrityksiin. Sosiaalisen median palvelut eivät toimi tyhjiössä, vaan palveluiden sisällöt tulevat käyttäjiltä ja se, miltä palvelu kullekin näyttää ja millaista tietoa käyttäjä levittää ja vastaanottaa, riippuu merkittävästi käyttäjästä itsestään. Lisäksi alustojen markkinointimahdollisuuksia hyödyntävät mainostajat muovaavat osaltaan sitä todellisuutta, mikä sosiaalisen median kautta käyttäjille muodostuu.

Sisältöjen personointia sosiaalisessa mediassa ja sen tuomia vaikutuksia yhteiskunnalle on tutkittu paljon. Algoritmeista kirjoitetaan mediassa jatkuvasti, mutta se miten kokonaisuus toimii ja kuinka algoritmit vaikuttavat tiedon kulkuun sosiaalisessa mediassa ei kaikesta huolimatta tunnu olevan sosiaalisen median palveluiden käyttäjille selvää. Algoritmit kehystetään mediassa usein jonain hallitsemattomana, itsenäisenä toimijana, joka ohjaa käyttäjien tiedon saantia (ks. esim. Sundman, 2019). Algoritmit kuitenkin viime kädessä ovat itsessään koodinpätkiä, joilla sinänsä ei ole toimijuutta, vaan jotka toteuttavat niille asetettua tehtävää. Käyttäjillä itsellään on kuitenkin merkittävä vaikutus algoritmeihin ja siihen millä perusteilla algoritmit suodattavat näille erilaisia sisältöjä. Algoritmien toiminnan parempi ymmärtäminen voi auttaa käyttäjiä suhtautumaan kriittisemmin sosiaalisen median kautta saamaansa tietoon ja sen kattavuuteen, ja siksi käyttäjien ymmärryksen kasvattaminen on tärkeää.

Tässä tutkimuksessa tarkoitukseni on kartoittaa sosiaalisen median ja algoritmien parissa työskentelevien asiantuntijoiden ja tutkijoiden näkemyksiä sosiaalisen median algoritmeista sekä palveluiden käyttäjien mahdollisuuksista vaikuttaa algoritmeihin ja sitä kautta sisältöjen personointiin. Tutkimus on

toteutettu haastattelemalla neljää algoritmien parissa työskentelevää asiantuntijaa. Tutkimuksessani pyrin vastaamaan seuraaviin kysymyksiin:

1. Miten asiantuntijat suhtautuvat sosiaaliseen mediaan ja algoritmeihin?
2. Miten asiantuntijat näkevät käyttäjien toiminnan vaikuttavan sosiaalisen median sisältöjen personointiin?
3. Mitä sosiaalisen median palveluita käyttävän tulisi palveluista ymmärtää välttääkseen kuplautumista?

Luvussa 2 taustoitin tutkimustani esittelemällä aiheeseen liittyvää aikaisempaa tutkimusta ja kirjallisuutta liittyen erityisesti siihen, kuinka algoritmit ovat muuttaneet tapamme vastaanottaa tietoa, kuinka sosiaalisen median palvelut keräävät dataa käyttäjistä ja kuinka datan keruuta säännellään sekä kuinka käyttäjien on mahdollista vaikuttaa sosiaalisen median kautta saamaansa tietoon. Luvuissa 3 ja 4 esittelen tutkimuskohteeni ja tutkimuksen toteutustavan. Luvussa 5 esitän tutkimukseni tulokset ja peilaan niitä luvussa 2 esiteltyyn aikaisempaan tutkimukseen aiheesta. Luvussa 6 pohdin tutkimukseni luotettavuutta ja luvussa 7 esittelen tutkimuksen teon herättämiä pohdintoja ja ajatuksia siitä, millaista tutkimusta sosiaaliseen mediaan liittyen kaivataan lisää sekä mihin suuntaan sosiaalisen median kehitys on, tai pitäisi olla, menossa.

2 Teoreettinen tausta

Tässä luvussa esittelen algoritmeihin ja sosiaaliseen mediaan liittyvää aiempaa tutkimusta ja kirjallisuutta. Avaan ensin tutkimukseni keskeisiä käsitteitä ja sen jälkeen keskustelua siitä, kuinka algoritmit ovat muuttaneet tapamme vastaanottaa tietoa, kuinka sosiaalisen median palvelut keräävät dataa käyttäjistä ja kuinka datan keruuta säännellään sekä siitä, kuinka käyttäjien on mahdollista vaikuttaa sosiaalisen median kautta saamaansa tietoon.

2.1 Keskeiset käsitteet

Tutkimuksessani pyrin selvittämään käyttäjän toiminnan vaikutuksia siihen, millaista tietoa hän sosiaalisesta mediasta saa. Tutkimukseni keskeisiä käsitteitä ovat algoritmit, personointi ja sosiaaliset kuplat.

2.1.1 Algoritmit

Algoritmia kuvataan usein yksityiskohtaisena toimintaohjeena; se kertoo mitä tehdä, missä järjestyksessä ja millä ehdoilla. Monet kuvaavat algoritmia reseptinä, joka kertoo kuinka haluttuun lopputulokseen päästään. Tekoälyalgoritmit käsittelevät niille syötettyä dataa ja muodostavat datasta saadun tiedon perusteella erilaisia ohjeistuksia tai toimintakehotteita algoritmista riippuen. Esimerkiksi Spotifyn algoritmi analysoi käyttäjän kuuntelutottumuksia ja suosittelee käyttäjälle musiikkia tämän analyysin perusteella.

Aiemmin algoritmeja on kehitetty niin, että ihminen on ohjelmoinut algoritmit toimimaan tietyllä tavalla ja suorittamaan tiettyä tehtävää (Hosanagar, 2019). Algoritmi on esimerkiksi ohjelmoitu tulkitsemaan, että jos käyttäjä katsoo ohjelman x, suosittele sille ohjelmaa y. Nykyään tekoälyalgoritmien kehityksessä suurena trendinä on kuitenkin koneoppiminen, jossa algoritmi jatkuvasti oppii sille syötetystä datasta ja adaptoi toimintaansa uudelleen (Hildebrandt, 2016). Tällöin algoritmi itse tulkitsee käyttäjien datasta mitä sen kannattaisi suositella ohjelman x katsoneelle. Tähän analyysiin vaikuttaa esimerkiksi se, mistä muusta ohjelman x katsoneet muut käyttäjät ovat pitäneet. Ohjelmoijalla on edelleen valta vaikuttaa siihen, kuinka algoritmi keräämäänsä dataa käsittelee, mutta vaikutus lopputuotokseen on vähentynyt, kun algoritmi kehittää toimintaansa autonomisesti.

Algoritmien voidaan nähdä toimivan ihmisen ja tiedon välillä sekä tulkkeina että portinvartijoina. Ensimmäisen näkemyksen mukaan algoritmit tuntevat ja ymmärtävät ihmisen toisinaan paremmin kuin tämä itse. Nämä algoritmit ikään kuin *lukevat* ihmisen mieltä ja toimivat sen mukaan. Jos esimerkiksi käyttäjä etsii tietoa Sauli Niinistöstä ja kirjoittaa Googlen hakukoneeseen ”Sauli” ja se ehdottaa automaattisesti ”Sauli Niinistö”, voidaan hakukoneen nähdä ikään kuin tulkinneen käyttäjän ajatuksia. Toisen näkemyksen mukaan internetin algoritmit taas rajoittavat tiedonsaantia, sillä algoritmit suodattavat aina pois osan vaihtoehdoista. Algoritmien puolestapuhujat kehystävät ne yleensä ensimmäisen näkökulman mukaan, kun niitä kyseenalaistavat ja kritisoivat taas kuvaavat niitä jälkimmäisen näkökulman kautta (Roth, 2019, 3).

Sosiaalisessa mediassa kaikki käyttäjän näkemä sisältö on algoritmien suodattamaa. Se, mitä käyttäjä etusivullaan esimerkiksi Facebookissa näkee, perustuu muun muassa käyttäjän ilmaisemaan toiveeseen vastaanottaa tiettyjä sisältöjä (esimerkiksi tykkäämällä tietystä sivusta tai liittymällä tiettyyn ryhmään) ja Facebookin algoritmin tulkintaan siitä, mitä käyttäjä haluaisi nähdä. (Roth, 2019, 4.) Lisäksi näkymään vaikuttaa se, kuka Facebookille on maksanut sisältöjen esittämisestä, eli mainostamisesta (Hern, 2017). Sosiaalisen median algoritmeista puhuttaessa on myös hyvä pitää mielessä, että niiden takana on yritykset, joiden liiketoiminnan tulonlähde perustuu käyttäjien palvelun parissa viettämään aikaan. Sosiaalisen median algoritmeilla pyritään siis myös maksimoimaan palvelun parissa vietetty aika. Tässä tutkimuksessa algoritmeilla tarkoitetaan nimenomaan sosiaalisen median algoritmeja eikä oteta kantaa muunlaisiin algoritmeihin.

2.1.2 Personointi

Thurman ja Schifferes (2012) määrittävät personoinnin tarkoittavan käyttäjän ja ohjelmiston välistä vuorovaikutusta, jossa teknologiaa hyödyntäen sisältöjä ja niiden toimitusta sekä järjestämistä mukautetaan yksittäisen käyttäjän mieltymysten mukaan. Personointiteknologiat ovat alun perin syntyneet vastauksena internetin kautta saatavilla olevalle informaatiotulvalle. Ne ovat siis kehitetty vastaamaan tarpeeseen, joka edelleen on hyvin olennainen: kaivaa kaikesta verkon kautta saatavilla olevasta massasta kullekin käyttäjälle sopivat

sisällöt ja informaatiot. Sinänsä personointiteknologiat ovat siis verkkoselailun kannalta hyvinkin olennaisia ja tarpeellisia. (ks. esim. Smyth, 2008; Montaner ym., 2003.) Personointia on kuitenkin alettu hyödyntämään myös kaupallisessa tarkoituksessa eri julkaisijoiden toimesta, joiden tavoitteena on saada käyttäjä viettämään enemmän aikaa näiden julkaisemien sisältöjen parissa. Siinä missä aiemmin personointiteknologioiden kehittäjät näkivät ne ratkaisuna tiedon löydettävyyteen liittyvään ongelmaan, on niistä tullut yhä enenevässä määrin esimerkiksi organisaatioiden keskeinen strategia käyttäjien sitouttamiseen (Thurman & Schifferes, 2012).

Thurman ja Schifferes (2012) jakavat personoinnin eksplisiittiseen ja implisiittiseen personointiin. Eksplisiittisellä personoinnilla tarkoitetaan palvelua, jossa käyttäjä itse määrittää, millaisia sisältöjä ja suosituksia hän haluaa. Tällöin algoritmi personoi sisältöjä käyttäjän määrittämien valintojen mukaan. Implisiittinen personointi taas viittaa palveluun, jossa algoritmi ehdottaa sisältöä käyttäjille perustuen käyttäjää seuraamalla kerättyyn dataan tämän aiemmasta toiminnasta. (em. 3) Eksplisiittistä personointia hyödynnetään esimerkiksi monilla uutissivustoilla tai näiden sovelluksissa, joissa käyttäjä voi määrittää haluavansa saada ilmoituksia esimerkiksi maailman politiikkaan, Yhdysvaltojen presidentinvaaleihin tai urheiluun liittyvistä uutisista.

Suurin osa sosiaalisen median palveluista perustuu pääosin implisiittiseen personointiin, eli aiemman käyttäytymisen perusteella tapahtuvaan algoritmiseen suositteluun. Kuitenkin monissa palveluissa, kuten esimerkiksi Facebookissa, on mahdollista ”tilata” ilmoituksia esimerkiksi tietyissä ryhmissä julkaistuista tai tiettyjen käyttäjien tekemistä sisällöistä. Tällainen sosiaalisessa mediassa tapahtuva personointi voidaan luokitella Thurmanin ja Schifferesin jaossa kuuluvan eksplisiittiseen personointiin. Mikä tästä tekee kuitenkin sosiaalisessa mediassa jossain määrin epäselvää, on se, että tietyistä sivuista tykkääminen tai tiettyjen käyttäjien lisääminen verkostoon voitaisiin myös luokitella eksplisiittiseksi ilmaisuksi toivotuista sisällöistä. Kuitenkin palvelut suodattavat käyttäjän näkymästä pois ison osan näistä. Eksplisiittinen-implisiittinen jaosta käytetään myös nimiä ”itsevalikoitu personointi” (*self-selected personalization*) ja ”ennalta valikoitu personointi” (*pre-selected personalization*) (Zuiderveen Borgesius ym., 2016).

Winter, Maslowska ja Vos (2020) puolestaan jakavat personoinnin vihjepohjaiseen personointiin (*cue-based personalization*) ja piirrepohjaiseen personointiin (*trait-based personalization*). Vihjepohjaisella personoinnilla tarkoitetaan sisällön kohdentamista käyttäjälle johonkin tähän liittyvän selkeästi ilmaistun tiedon avulla. Vihjepohjaisessa personoinnissa hyödynnetään esimerkiksi käyttäjän nimeä, asuinpaikkaa, ikää, sukupuolta tai työnantajaa, joiden perusteella käyttäjälle voidaan personoida tiettyjä mainoksia tai sisältöjä niin, että niistä käy selkeästi ilmi, että käyttäjä näkee tämän, koska on esimerkiksi merkinnyt asuinpaikakseen tietyn kaupungin. Piirrepohjaisella personoinnilla kirjoittajat kuvaavat personointia, josta personointi ei käy selkeästi ilmi, mutta jossa sisältö on kohdennettu käyttäjälle hyödyntäen tämän aikaisemmasta verkkokäyttäytymisestä kertynyttä dataa. Esimerkiksi tällaisesta personoinnista kirjoittajat antavat juomamainoksen, joka ekstrovertiksi luokitellulle käyttäjälle esitetään varustettuna kuvalla, jossa ystävykset nauttivat juomasta sosiaalisessa kanssakäymisessä, kun taas introvertiksi luokitellulle käyttäjälle vastaava mainos esitetään kuvalla, jossa juoma esitetään pöydällä, jonka ääressä istuu ongelmaa pohdiskeleva henkilö.

Winter, Maslowska ja Vos (2020) huomauttavat, että koska vihjepohjaisesta personoinnista käy selkeämmin ilmi, että sisältö on personoitu käyttäjälle, se saattaa herättää vastareaktioita vastaanottajissa, jotka toisinaan kokevat tällaisen personoinnin häiritsevänä yksityisyyden kannalta. Piirrepohjainen personointi on kirjoittajien mukaan siis hienovaraisempaa ja siksi usein käyttäjille vaikeampaa tunnistaa personoinniksi, minkä vuoksi se herättää vähemmän vastarintaa ja enemmän samaistuttavuutta mainokseen. (Winter, Maslowska & Vos, 2020.) Jaottelu kuvaa kiinnostavasti vastaanottajassa heräävää negatiivista reaktiota havaittuun personointiin, joka puuttuu silloin, kun personointia on tehty niin taitavasti, ettei sitä ymmärretä tapahtuvan.

2.1.3 Sosiaaliset kuplat

Sosiaalisen kuplan termiä ovat kehittäneet Esa Väliverronen ja Janne Seppänen kirjassaan *Mediayhteiskunta* (2014) ja se perustuu internet-aktivisti Eli Pariserin tunnetuksi tekemään *Filter Bubble* termiin, jota Pariser käsittelee kirjassaan *The Filter Bubble* (2012). Sosiaalisella kuplalla tarkoitetaan samanmielisten ryhmittymiä, joita sosiaalisessa mediassa muodostuu. Kuten

aiemmin todettua, nykyään monien saadessa uutisensa sosiaalisen median välityksellä, suodattuu tiedonkulku algoritmien toiminnan mukaan. Kun algoritmit suosittelevat käyttäjille personoitua sisältöä tavoitteenaan saada käyttäjä pysymään palvelun parissa, näytetään tälle sisältöjä, jotka sopivat käyttäjän mieltymyksiin ja kiinnostuksen kohteisiin. Pariserin mukaan näin syntyy kuplia, joiden sisällä olevat jakavat samat asenteet ja arvot. Samalla, kun tämä saattaa vahvistaa ryhmän sisäistä yhteisöllisyyttä ja yhteenkuuluvuuden tunnetta, se ajaa eri ryhmiä kauemmas toisistaan, kun nämä eivät kuplissaan altistu eriäville mielipiteille ja näkemyksille. (Seppänen & Väliverronen, 2014, 212.) Tällainen julkisuuden pirstoutuminen ja samoin ajattelevien yhteisöjen vahvistuminen vaikuttavat Seppäsen ja Väliverronen (2014, 213) mukaan väistämättä yhteiskunnan dynamiikkaan. Yleisesti tutkijoiden pelkona on, että kuplissa ennakkoluulot vahvistuvat, kun uutiset, jotka ovat ristiriidassa käyttäjän mielipiteiden kanssa, on suodatettu pois. Sosiaaliset kuplat ovat herättäneet huolta erityisesti siksi, että niiden on pelätty johtavan yhteiskunnan polarisoitumiseen. Kuplissa, joissa vahvistuu vain yksi kanta ja näkökulma käsiteltävään aiheeseen, on myös monien näkemyksien mukaan suurempi riski dis- ja misinformaation leviämislle (Difranzo & Gloria-Garcia, 2017).

Sosiaalisten kuplien olemassa oloa on kuitenkin myös tutkimuksissa kyseenalaistettu. Joidenkin tutkimusten mukaan sosiaalisen median kautta käyttäjä altistuu useille erilaisille lähteille ja myös vastakkaisille mielipiteille eikä sosiaalisia kuplia näin tosiasialisesti pääse syntymään. (ks. esim. Zuiderveen Borgesius ym., 2016.) Flaxman, Goel ja Rao (2016) keräsivät dataa 50 000 amerikkalaisen verkkouutisselauksesta ja havaitsivat, että sosiaalisen median kautta löydettyjen artikkelien kulutuksessa oli havaittavissa selkeämpää ideologista jaottelua kuin suoraan uutissivustoilla kulutetuissa sisällöissä. Kuitenkin he havaitsivat myös, että sosiaalisen median kautta käyttäjät altistuivat myös vastakkaista näkökulmaa edustaville uutisille. Axel Bruns (2019) kyseenalaistaa myös sosiaalisten kuplien olemassa olon kirjassaan *"Are filter bubbles real?"*. Bruns näkee, että ryhmiin ajautumisesta sosiaalisessa mediassa ei voi syyttää teknologiaa, sillä ihmiset itse päättävät mihin ryhmään haluavat kuulua valitsemalla keitä seuraavat tai keitä lisäävät tai hyväksyvät kavereikseen. Bruns pitää epätodennäköisenä, että käyttäjän verkostoon kuuluisi esimerkiksi vain demokraatteja tai vain republikaaneja, vaan näkee, että verkostonsa kautta

kaikki altistuvat jossain määrin erilaisille näkökulmille ja mielipiteille. Lisäksi Bruns näkee, ettei Pariserin ja muiden sosiaalisia kuplia teoretisoineiden tutkijoiden ennustukset informaation ja sen myötä yhteiskunnan segregoitumisesta ole mahdollisia, sillä ihmisen on mahdotonta tai ainakin hyvin vaikeaa kuulua vain yhteen kuplaan ja saada kaikkea tietoaan vain sieltä. (Bruns, 2019, 52.) Keskeistä sosiaalisen median palvelujen kautta saatavan tiedon laadussa näyttäisikin olevan verkoston rakenne, eli se millaisia henkilöitä kunkin verkostoon kuuluu ja millaisiin erilaisiin kupliin kuulutaan.

Kritiikistä ja kyseenalaistuksesta huolimatta sosiaaliset kuplat nähdään tämän hetkessä viestinnän tutkimuksessa tärkeänä tutkimuskohteena ja käsitteenä, joista on puhuttu julkisessakin keskustelussa erityisen paljon myös 2020 Yhdysvaltojen presidentinvaalien alla.

2.2 Kuinka algoritmit ovat muuttaneet ihmisten tapaa vastaanottaa tietoa?

Tiedonkulku on keskeisessä osassa sekä yhteiskunnan että yhteisöllisyyden tunteen muodostumisessa ja ylläpidossa. Erityisesti demokratian toteutumisen kannalta tiedon vapaa liikkuminen ja julkisuudessa käytävät keskustelut ovat keskeisessä asemassa. Se, mitä julkisuuden alueelle pääsee ja miten asioita julkisuudessa käsitellään vaikuttaa merkittävästi siihen, millainen kuva ihmisille muodostuu ympäröivästä maailmasta. Siinä missä aikaisemmin perinteinen media toimi julkisuuden portinvartijana määritellen mistä asioista keskustellaan, on sosiaalinen media poistanut tämän rajan. (Seppänen & Väliaverronen, 2014.)

Sosiaalinen media on myös muuttanut tapamme vastaanottaa uutisia ja tietoa ympäröivästä maailmasta. Kuten luvussa 1 mainittiin, sekä amerikkalaisissa, että suomalaisissa tutkimuksissa on havaittu, että suuri osa ihmisistä saa uutisensa pääosin sosiaalisen median ja internetin kautta. Sosiaaliseen mediaan tiedon välittäjänä liittyy kuitenkin ongelmia. Palveluiden algoritmit personoivat käyttäjien näkymää ja vaikuttavat siihen millaista sisältöä kukin palveluissa näkee. Näin ollen palvelut vaikuttavat siihen, millainen kuva todellisuudesta, yhteiskunnasta ja ajankohtaisista aiheista muodostuu. Algoritmien vaikutusta maailmankuvan muodostumiseen on testattu erityyppisillä

kokeiluilla. Esimerkiksi Epstein ja Robertson (2015) testasivat hakukonehaun tulosten järjestyksen vaikutusta poliittisen mielipiteen muodostumisessa ja havaitsivat, että se missä järjestyksessä tulokset esitettiin, vaikutti merkittävästi: ensimmäisinä esitettyihin tuloksiin luotettiin ja mielipide haun kohteena olleesta asiasta muodostettiin niiden perusteella. Se, kuinka algoritmit toimivat käytännössä, riippuu kuitenkin monesta asiasta, kuten palveluiden liiketoimintamallista sekä suurilta osin myös käyttäjien omasta toiminnasta palveluiden parissa (ks. esim. Bruns, 2019).

Sosiaalinen media ja algoritmit ovat herättäneet kiinnostusta tutkijoiden keskuudessa viime vuosien aikana merkittävästi. Erityisesti vuoden 2016 Yhdysvaltojen presidentin vaalien jälkeen sosiaalinen media, algoritmit ja niiden vaikutus on herättänyt keskustelua myös julkisuudessa, sillä monien tutkimusten mukaan sosiaalisella medially näyttäisi olleen suuri merkitys vaalien lopputulokselle (Benkler ym., 2018). Donald Trumpin kampanjointiin Facebook-käyttäjien personointidataa kerännyt ja hyödyntänyt konsulttiyhtiö Cambridge Analytica ajautui konkurssiin kampanjoinnin keräämän negatiivisen julkisuuden seurauksena. Myös Britannian eroa Euroopan Unionista ajanut Brexit-kampanja hyödynsi kampanjoinnissaan Cambridge Analytican palveluja. (Benkler ym., 2018.) Vaikka molemmista kampanjoista ja Cambridge Analytican toiminnasta heräsi suuri kohu, konkreettista muutosta esimerkiksi sosiaalisen median palveluiden käytössä ei ole tapahtunut. Käyttäjien luottamus sosiaalista mediaa kohtaan koki kuitenkin merkittävän kolauksen. (Sanomalehtien Liitto, 2018.)

Vaikka algoritmien tekemästä sisältöjen suosittelusta ja rajaamisesta kuohutaan paljon, maailmankuvaa rajaavia toimintamalleja on mediassa ollut olemassa jo pitkään. Jo ennen internetin yleistymistä ihmiset ovat tutkimusten mukaan hakeutuneet sellaisen tiedon pariin, joka vahvistaa näiden ennakkoletuksia sen sijaan, että haastaisi niitä. Tällaiset tutkimukset pohjautuvat usein Leon Festingerin (1957) teoriaan kognitiivisesta dissonanssista, jonka mukaan ihminen pyrkii välttämään sellaisia tilanteita, joissa tiedot ja asenteet ovat ristiriidassa. (Hilden, 2017.) Ihmiset ovat esimerkiksi lukeneet tiettyä sanomalehteä tai katsoneet tietyn kanavan uutisia. Myös tällä tavalla toimiessaan osa tarjolla olevasta informaatiosta on automaattisesti jäänyt saamatta ja uutisia on vastaanotettu vain tietystä näkökulmasta käsin. Algoritmien voidaan osaltaan nähdä toisintavan tätä mallia niiden tarjotessa ihmisille sisältöjä, jotka vahvistavat

näiden olemassa olevia ajatusmalleja. Sosiaalisen median algoritmien tapaan nostaa näkyville erityisesti niitä sisältöjä, joiden tekijöiden kanssa käyttäjä on ollut eniten tekemisissä toisintaa myöskin perinteistä mallia, jossa lähellä olevien ihmisten elämästä ollaan tietoisempia kuin niiden, joiden kanssa ollaan yhteydessä vain satunnaisesti. Ihmisillä on myös perinteisesti ollut taipumus viettää enemmän aikaa saman- kuin erimielisten ihmisten kanssa. (ks. esim. Frey, 1986; Munson & Resnick, 2010.)

Mediayhtiöiden toiminta on jo ennen sosiaalista mediaa perustunut kahteen liiketoimintamalliin. Ensinnäkin yhtiöiden tavoitteena on myydä sisältöjä (lehteä, tv-ohjelmia, radio-ohjelmia) yleisölleen, mutta sen lisäksi tavoitteena on myydä lukijoita mainostajille. (ks. esim. Seppänen & Väliaverron, 2014.) Sinänsä siis palvelujen käyttäjien tuotteistamisessa mainostajille ei ole mitään uutta. Tässä suhteessa keskeisenä erona aikaisempaan on kuitenkin se, että joukkoviestintävälineissä mainosten kohdentaminen perustui joko yleisön kokoon (esimerkiksi televisiomainokset) tai ilmaistuun kiinnostukseen kohteeseen (esimerkiksi kalastuslehdessä mainostaminen). Sosiaalisessa mediassa mainoksia voidaan esittää kuitenkin huomattavasti rajatummin valitulle yleisölle. Esimerkiksi Joseph Turow (2011, 6) huomauttaa, että internetissä ja sosiaalisessa mediassa mainostajille ei enää myydä "yleisöjä" vaan "yksilöitä". Turow näkee tämän ongelmallisena erityisesti siksi, että palveluiden määritettyä esimerkiksi käyttäjän sosioekonomisen statuksen, tälle esitetään sisältöjä, joiden nähdään sopivan tämän elämäntilanteeseen ja budjettiin. Turow pelkää tämän johtavan tilanteeseen, jossa maailma ja sen tarjoamat mahdollisuudet näyttäytyvät hyvin erilaisina eri henkilöille. Siinä missä internetin alkuaikoina sen nähtiin tarjoavan käyttäjille mahdollisuuden luoda itse oma mediamaisemansa, Turow näkee nykyisen mallin tarjoavan tosiasiaissa mainostajille mahdollisuuden luoda jokaiselle käyttäjälle sellainen mediamaisema, joka edistää parhaiten mainostajien etuja. (Em. 12-13.)

Eryyisenä erona vanhaan malliin vastaanottaa tietoa on algoritmisessa personoinnissa sen läpinäkymättömyys ja käyttäjien tiedostamattomuus. Aiemmin ihminen saattoi tietoisesti valita esimerkiksi tietyn lehden sen mukaan, jonka näki edustavan sellaisia arvoja, jotka myös itse jakoi. Algoritmien personoimassa sisältöjen valinnassa tätä valintaa ei tehdä tietoisesti, vaan algoritmit tulkitsevat käyttäjän mieltymyksiä tämän aiemmasta toiminnasta.

Algoritmien vaikutus voidaankin nähdä kyseenalaisena erityisesti silloin, jos käyttäjä ei ymmärrä tätä toimintalogiikkaa, vaan uskoo näkevänsä personoimatonta sisältöä eikä tällöin myöskään osaa kyseenalaistaa näkemäänsä. Tutkimuksissa onkin havaittu, että esimerkiksi eniten disinformaatiota sosiaalisessa mediassa levittää painettuun sanaan luottava vanhempi ikäpolvi, joille sosiaalisen median ja internetin toimintalogiikka ei välttämättä ole yhtä tuttu kuin nuoremmalle ikäpolvelle (Vehkoo, 2019).

Algoritmien tutkimuksesta haastavan tekee se, että algoritmit ovat yritysten tarkoin varjeltuja liikesalaisuuksia. Sosiaalisen median palvelut ovat yrityksiä siinä missä muutkin, ja algoritmit sekä niiden käsittelemä data ovat näiden yritysten yksi tärkeimmistä voimavaroista, eikä palvelut siksi (ainakaan toistaiseksi) ole avanneet algoritmiensa toimintalogiikkaa julkisesti. (ks. esim. Willson, 2016.) Kuitenkin samalla, kun ymmärrys algoritmien vaikuttavuudesta kasvaa, kasvaa myös paine yrityksiä kohtaan avata algoritminsa julkisuuteen tarkasteltavaksi ja arvioitavaksi. Kuten Barack Obama totesi vuoden 2009 tiedonvälitysvapauden muistojuhlan puheessa: ”Demokratia vaatii vastuullisuutta, ja vastuullisuus vaatii läpinäkyvyyttä” (Obama, 2009, vapaa suomennos), ja siksi vaatimusta algoritmien julkistamisesta on perusteltu demokratian toteutumisen kannalta. (Bucher, 2018.) Esimerkiksi Taina Bucher (2018) sekä Michael Koliska ja Nicholas Diakopoulos (2018) puoltavat näkemystä siitä, että algoritmien läpinäkyvyyttä koodin tasolla tulisi avata.

Vaikka algoritmeilla on keskeinen rooli sosiaalisen median kautta tapahtuvalle tiedonkululle, algoritmit eivät toimi tyhjiössä. Tätä näkemystä on tuoneet vahvasti tutkijoista esiin esimerkiksi Mike Ananny ja Kate Crawford (2018), jotka ovat huomauttaneet, ettei läpinäkyvyys itsessään takaa algoritmien vastuullisuutta. Nämä tutkijat ovat korostaneet, että pelkän koodin tarkastelun sijaan tulisi tarkastella sitä kokonaisuutta ja verkostoa, jossa algoritmi toimii, sillä algoritmit eivät itsessään ole vastuullisia tai ei-vastuullisia, vaan niiden toiminta on riippuvainen esimerkiksi kerätystä datasta, joka taas riippuu käyttäjistä. Myös esimerkiksi David Beer (2017) on todennut, ettei algoritmeja tulisi tutkia irrallaan kontekstistaan, sillä niiden toiminta määräytyy viimekädessä sosiaalisissa prosesseissa käyttäjien toiminnan seurauksena. Michele Willson (2016) jatkaa Anannyn ja Crawfordin ajatuksia siitä, että algoritmeja tulisi tutkia osana ympäristöään ja toteaa, ettei niitä tulisi nähdä vain teknologisina ratkaisuina, vaan

prosesseina ympäröivässä monimutkaisessa ekosysteemissä, joka käsittää teknologisten ratkaisujen ja ohjelmoijan suunnittelun lisäksi muun muassa yleisöt ja käyttötavat. Myös Skrubbeltrang Mahnke (2019) huomauttaa, että algoritmeista ajatellaan usein, että ne yksistään valitsevat mitä käyttäjä sosiaalisen median palveluissa näkee. Vuonna 2016 hän tutki Instagramin käyttäjien reaktioita palvelun ilmoitukseen sisältöjen algoritmista personoinnista ja havaitsi, että se herätti vastarintaa erityisesti palvelun aktiivisten käyttäjien parissa. Näistä suurin osa kuitenkin luuli Skrubbeltrang Mahnken mukaan algoritmien päättävän itsenäisesti mitä käyttäjät näkevät ja mitä eivät ymmärtämättä, että sillä keitä käyttäjiä nämä seuraavat on edelleen suurin merkitys sisältöjen esittämisessä. (Em., 2.)

Oman haasteensa algoritmien tutkimukselle asettaa sosiaalisen median yritysten tapa muokata algoritmeja jatkuvasti. Esimerkiksi Facebookilla on yli 2,5 miljardia kuukausittaista käyttäjää (Clement, 2020) mikä mahdollistaa yritykselle sen, että erilaisia algoritmeja voidaan ketterästi testata eri ryhmille ja näin seurata kuinka algoritmien toiminta vaikuttaa käyttäjien toimintaan palvelun parissa. (Bucher, 2018.)

2.3 Datan keruun yksityisyyden suoja sosiaalisen median palveluissa

Toukokuussa 2018 voimaan astuneessa Euroopan unionin yleisessä tietosuojasetuksessa (2016/679) (GDPR) määritetään verkossa tapahtuvalle henkilötietojen keruulle ja käsittelylle tiettyjä vaatimuksia. Hilden (2017) jakaa asetuksen henkilötietojen käsittelyyn ja personointiin liittyvät osiot kolmeen osioon: henkilötietojen käsittelyä koskevat periaatteet ja käsittelyn laillisuus (art. 5 ja 6), käyttäjän oikeudet (III luku) ja palveluntarjoajan velvollisuudet (VI luku).

Hildenin mukaan yleisen tietosuojasetuksen 5. artiklassa määritetään, että henkilötietojen rekisterinpitäjän täytyy noudattaa kuutta periaatetta, joiden mukaan henkilötietoja on käsiteltävä:

1. Tietoja on käsiteltävä lainmukaisesti, asianmukaisesti ja läpinäkyvästi
2. Tietoja on käsiteltävä tiettyä, nimenomaista ja laillista tarkoitusta varten
3. Käsittely on rajoitettava siihen, mikä on tarpeellista

4. Tietojen tulee olla täsmällisiä ja tarvittaessa päivitettyjä
5. Tietoja saa säilyttää vain niin kauan kuin on tarpeen
6. Tietoja on käsiteltävä tavalla, jolla varmistetaan niiden turvallisuus ja luottamuksellisuus

Lisäksi 6. artiklassa määritetään, että tietojen käsittelyä varten rekisteröidyn pitää olla antanut suostumuksensa henkilötietojensa käsittelyyn ja että käsittelyn tulee olla tarpeen rekisterinpitäjän tai kolmannen osapuolen oikeutettujen etujen toteuttamiseksi. Rekisterinpitäjän pitää myös selkeästi ilmoittaa rekisteröidylle, mitkä nämä oikeudet ovat. Artiklassa määritellään myös, että rekisterinpitäjän on pystyttävä osoittamaan, että periaatteita on noudatettu. (Hilden, 2017, 15.)

9. artiklan ensimmäisessä kohdassa säädetään lisäksi erityisiä henkilötietoryhmiä koskeva rajoitus tietojen keruuseen liittyen:

1. Sellaisten henkilötietojen käsittely, joista ilmenee rotu tai etninen alkuperä, poliittisia mielipiteitä, uskonnollinen tai filosofinen vakaumus tai ammattiliiton jäsenyys sekä geneettisten tai biometrinen tietojen käsittely henkilön yksiselitteistä tunnistamista varten tai terveyttä koskevien tietojen taikka luonnollisen henkilön seksuaalista käyttäytymistä ja suuntautumista koskevien tietojen käsittely on kiellettyä. (Hilden, 2017, 17)

Kuitenkin esimerkiksi Facebookissa käyttäjien toimintaan perustuvaan datan avulla esimerkiksi näiden poliittisesta kannasta pystytään tekemään hyvinkin tarkkoja päätelmiä (Merrill, 2016). Käyttäjän suostumuksella myös arkaluontoisten tietojen käsittely on laillista (Hilden, 2017, 17) ja tähän myös sosiaalisessa mediassa tapahtuva tiedon keruu perustuu.

Käyttäjän oikeuksista tietosuoja-asetuksessa määritellään, että rekisteröidylle täytyy toimittaa tietojen käsittelyyn liittyviä tietoja ja että käyttäjällä on oikeus päästä omiin tietoihinsa sekä oikeus niiden oikaisemiseen, poistamiseen, käsittelyn rajoittamiseen, siirtämiseen järjestelmästä toiseen ja oikeus vastustaa henkilötietojen käsittelyä. Oikeus tietojen käsittelyn vastustamiseen koskee erityisesti päätöksentekoa, joka perustuu täysin automatisoituun henkilötietojen käsittelyyn, jolla on oikeusvaikutuksia tai joka merkittävästi vaikuttaa rekisteröityyn henkilöön. (Hilden, 2017, 17-18.) Sosiaalisen median palveluiden personointiin liittyen keskeinen kysymys onkin, vaikuttaako personointi

merkittävästi palvelu käyttäjään? Kuten aiemmin esitettiin, tutkimuksissa on havaittu uutisten ja tiedon saannin siirtyneen enenevässä määrin sosiaalisen median palveluihin, joten tästä näkökulmasta olisi perusteltua olettaa, että sosiaalisen median palveluiden personointi vaikuttaa käyttäjään merkittävästi. Kuitenkin, kuten sosiaaliin kupliin liittyvässä osiossa kuvattiin, sosiaalisen median vaikutus tiedon yksipuolistumiseen jakaa tutkijoiden mielipiteitä.

Lisäksi tietosuoja-asetuksessa määritetään, että personointia tekevän toimijan on annettava käyttäjälle ”merkitykselliset tiedot käsittelyyn liittyvästä logiikasta” (Hilden, 2017, 18, ref. Yleinen tietosuoja-asetus 13(2)(f)), eli palveluiden käyttäjille tulisi avata millä logiikalla personointi tapahtuu. Kuitenkin esimerkiksi Facebookin mainosasetuksissa ja profiloinnin läpinäkyvydessä on Hildenin mukaan havaittavissa selkeitä puutteita.

Yleinen tietosuoja-asetus on kerännyt sekä kehuja, että kritiikkiä. Eniten kritiikkiä on kohdistunut siihen, että asetus pohjautuu vahvasti käyttäjän suostumukseen tiedon keruusta ja käsittelystä. Esimerkiksi Breen, Ouazzane ja Patel (2020) ovat kyseenalaistaneet, onko käyttäjien suostumus tietojen käsittelylle riittävä kriteeri ja haastavat henkilötietoja tallentavien yritysten tapaa pyytää lupaa tietojen käsittelylle. Yleisin tapa suostumuksen antamiselle tietojen käsittelyä on pyytää käyttäjää klikkaamaan raksin ruutuun laatikossa, jossa on linkki erilliselle tietosuojaselosteelle. Breen ja kumppanit näkevät, että tällainen luvan kysymistapa ei noudata tietosuoja-asetuksen vaatimaa selkeyttä ja tarkkuutta siitä, mihin kerättävää tietoa käytetään. Lisäksi he huomauttavat, että tietosuojaselosteet ovat usein pitkiä ja kirjoitettu lakikielellä. Yleisen tietosuoja-asetuksen mukaan vastuu siitä, että käyttäjä ymmärtää mihin tästä kerättävää tietoa tullaan käyttämään on tiedon kerääjällä ja siksi Breen ja kumppanit näkevät, ettei lakikielellä kirjoitettujen tietosuojaselostusten hyväksyminen riitä suostumuksen antamiselle.

2.4 Kuinka sisältöä personoidaan sosiaalisen median palveluissa?

Pullinen (2019) kuvaa sosiaalisen median algoritmista personointia buffetpöytänä, joka järjestyy sen mukaan, kuka pöydän ääreen saapuu. Algoritmi tarjoaa pöydän etualalla ihmisen lempiruokia ja siirtää taaemmaksi ne, joista tämä ei pidä. Järjestely saa ihmisen viettämään pöydän ääressä enemmän aikaa

ja kuluttamaan enemmän herkkuja kuin tälle luultavasti on hyväksi, eikä sen myötä ihminen myöskään törmää uusiin makuelämyksiin. (Pullinen, 2019.) Siihen, millä perusteella algoritmit päättelivät käyttäjän mieltymykset, vaikuttaa kuitenkin moni eri asia. Hilden (2017, 8) erottelee sisältöjen personointimenetelmiä viiden eri tyyppisen käyttäjästä kerätyn datan mukaan:

- Demografisiin tietoihin perustuva suosittelu
- Aiempaan mediakäyttöön ja sisältöjen samankaltaisuuteen perustuva suosittelu
- Samankaltaisten käyttäjien mediakäyttöön perustuva suosittelu
- Käyttäjien palautteeseen perustuva suosittelu
- Sosiaalisiin verkostoihin perustuva suosittelu

Menetelmät vaikuttavat siihen laajennetaanko vai kavennetaanko personoinnilla käyttäjien näkymää. Hilden (2017, 9) huomauttaa, että kaikissa personointimenetelmissä on omat heikkoutensa:

”Demografisiin tietoihin perustuva suosittelu ei ole tarpeeksi tarkka, aiempaan käyttöön perustuvat suosittelut voivat olla liian kapeita, samankaltaisten käyttäjien käyttöön perustuva suosittelu vaatii paljon käyttäjiä ja sisältöjä, ja palautteeseen perustuva suosittelu vaatii käyttäjän aktiivista osallistumista. Lisäksi ongelmana on, että algoritmi ei välttämättä tunnista laatua, ja tämän takia erityisesti mediayhtiöt palkkaavat työntekijöitä kuratoimaan sisältöjä.”

Sosiaalisen median palveluissa suosittelu perustuu harvoin vain yhteen menetelmään vaan on kombinaatio näitä kaikkia. (Hilden, 2017.) Se, millä painotuksilla kukin palvelu käyttäjien sisältöä personoi ei kuitenkaan ole julkista tietoa.

3 Tutkimustehtävä ja tutkimuskysymykset

Tämän tutkimuksen tarkoitus on kartoittaa sosiaalisen median ja algoritmien parissa työskentelevien asiantuntijoiden ja tutkijoiden näkemyksiä sosiaalisen median algoritmeista sekä palveluiden käyttäjien mahdollisuuksista vaikuttaa algoritmeihin ja personointiin. Julkisessa keskustelussa sosiaalisen median algoritmit nähdään usein jonain hallitsemattomana itsenäisenä toimijana, joka ohjaa tai rajoittaa sosiaalisen median palveluissa liikkuvaa tietoa. Palveluiden käyttäjillä itsellään on kuitenkin merkittävä vaikutus algoritmeihin: niiden tehtävä on nimenomaan analysoida käyttäjien toimintaa palveluiden parissa ja päätellä siitä millainen sisältö näitä kiinnostaa tai ei kiinnosta. Siksi se, kuinka käyttäjät palveluiden parissa toimivat, vaikuttaa algoritmeihin automaattisesti. Tämä yhteys ja se mitä kaikkea käyttäjien toimintaa analysoidaan, ei kuitenkaan suurimmalle osalle käyttäjistä tunnu olevan selvää. Tutkimuksessani pyrin algoritmien parissa työskenteleviä asiantuntijoita haastatteleamalla selvittämään miten nämä asiantuntijat näkevät käyttäjien toiminnan vaikuttavan algoritmeihin ja sitä kautta sisältöjen personointiin. Tutkimuksessani pyrin vastaamaan seuraaviin kysymyksiin:

1. Miten asiantuntijat suhtautuvat sosiaaliseen mediaan ja algoritmeihin?
2. Miten asiantuntijat näkevät käyttäjien toiminnan vaikuttavan sosiaalisen median sisältöjen personointiin?
3. Mitä sosiaalisen median palveluita käyttävän tulisi palveluista ymmärtää välttääkseen kuplautumista?

4 Tutkimuksen toteutus

Tässä osiossa avaan tutkimukseni toteutustapaa, tutkimusta varten kerättyä aineistoa sekä aineiston analyysimenetelmää ja tulosten luokittelua. Tutkimukseni on toteutettu puolistrukturoituna haastattelututkimuksena, jossa haastateltavana oli neljä sosiaalisen median parissa työskentelevää asiantuntijaa. Haastatteluiden analyysimenetelmänä on käytetty teoriaohjaavaa sisällönanalyysia.

4.1 Tutkimuskohteen kuvaus

Tutkimukseni tarkoitus on kuvata algoritmeihin ja personointiin liittyvää asiantuntijakeskustelua ja kartoittaa käyttäjien mahdollisuuksia vaikuttaa algoritmeihin ja niiden tekemään personointiin. Algoritmien tutkimisessa on kuitenkin omat haasteensa. Ensinnäkin, algoritmit ovat sosiaalisen median palveluiden taustalla olevien yritysten tarkoin vaalimia liikesalaisuuksia, eikä niiden tarkkaan toimintalogiikkaan tai niille asetettuihin tehtäviin pääse suoraan tutustumaan. Toiseksi, algoritmit kehittyvät ja muokkautuvat jatkuvasti: palveluiden kehittäjät muokkaavat niitä yritysten liiketoiminnan tavoitteiden mukaan, minkä lisäksi algoritmit kehittyvät koneoppimisen kautta muokkaamalla omaa toimintaansa. Palvelujen toimintaa tutkimalla ja analysoimalla on kuitenkin mahdollista tehdä tiettyjä päätelmiä ja yleistyksiä algoritmien toimintalogiikasta, siitä millaista tietoa käyttäjistä kerätään ja miten käyttäjien toimintaa analysoidaan ja sisältöä personoidaan kerätyn datan perusteella. Siksi tutkimuskohteenani on algoritmien ja sosiaalisen median parissa toimivien asiantuntijoiden näkemykset ja mielipiteet algoritmeista ja käyttäjien mahdollisuuksista vaikuttaa personointiin.

4.2 Haastattelututkimus

Tutkimukseni aineisto on kerätty puolistrukturoituina teemahaastatteluina. Haastattelu menetelmänä mahdollistaa tiedonkeruun kattavasti, sillä sen avulla voidaan kerätä faktatiedon lisäksi näkemyksiä, mielipiteitä ja henkilökohtaisia kokemuksia sekä perehtyä haastateltavien taustoihin, mikä mahdollistaa vastausten laittamiseen kontekstiin (Atkins, L., Wallace, S., & British Educational

Research Association, 2012, 86). Puolistrukturoitu teemahaastattelu asettuu haastattelumenetelmän strukturoidun lomakehaastattelun ja avoimen teemahaastattelun väliin ja mahdollistaa keskustelun ennalta määritettyjen kysymysten ympärillä rajoittamatta liikaa sitä, mistä voidaan keskustella. Teemahaastattelu mahdollistaa haastateltavien omien tulkintojen ja näkemysten avaamisen sen jättäessä avoimeksi tarkat suunnat siitä, mihin keskustelu voi edetä. Puolistrukturoitu haastattelurunko kuitenkin varmistaa sen, että kaikkien haastateltavien kanssa on käyty läpi samat kysymykset. (Eskola & Suoranta, 1998, 65.) Koska tutkimukseni perustuu haastateltavien näkemyksiin ja mielipiteisiin, haastattelumenetelmä mahdollistaa tarkempien perustelujen kaivamisen vastausten takaa. Sosiaalisen median algoritmit ja käyttäjien vaikutus niihin on monisyinen aihe ja haastattelumenetelmän tarjoama vastavuoroinen dialogi haastateltavien kanssa mahdollistaa tarkentavien lisäkysymysten esittämisen oikein ymmärtämisen varmistamiseksi.

Tutkimukseen haastateltiin sellaisia asiantuntijoita ja tutkijoita, jotka työnsä puolesta ovat syvällisesti tutustuneet sosiaaliseen mediaan, algoritmeihin ja personointiin. Tutkimukseen haastateltiin yhteensä neljää henkilöä, joista kaksi ovat tutkijoita ja kaksi toimivat sosiaalisen median ja algoritmien parissa työssään muuten. Haastateltavista kolme on miehiä ja yksi nainen ja he ovat iältään noin 30-45 vuotiaita. Kvalitatiivista tutkimusta tehdessä yksi yleisin kysymys liittyy aineiston kokoon (Eskola & Suoranta, 1998, 45). Eskolan ja Suorannan (1998) mukaan kvalitatiivisessa tutkimuksessa aineiston keruu perustuu tarkoituksenmukaisuuteen ja harkinnanvaraisuuteen, sillä tutkimuksen päämääränä ei ole tilastollinen yleistys vaan tietyn ilmiön tulkinta ja kuvaus (em., 45). Myös tässä tutkimuksessa haastateltavien määrän sijaan olennaista on haastateltavien asiantuntemus tutkimuksen kohteena olevasta temasta.

Haastattelututkimuksessa, kuten muissakin tutkimusmenetelmissä, yksi tärkeä seikka huomioida on omat ennakoasenteet ja odotukset tutkimusta ja sen tuloksia kohtaan. Tässä tutkimuksessa ja omalla kohdallani ennakoasenteeni on, että käyttäjät vaikuttavat toiminnallaan algoritmeihin merkittävästi, mutta käyttäjät itse eivät ole tästä tietoisia. Ja siksi voittoa tavoittelevat sosiaalisen median yritykset pystyvät käyttämään palveluiden käyttäjiä hyväksi tavalla, joka ei ole näille itselleen selvää. Sosiaalisen median yritysten vallasta puhutaan paljon ja yrityksiä kritisoidaan yleisesti. Henkilökohtaisesti ajattelen siis, etteivät

algoritmit itsessään ole ongelmallisia, vaan ongelmallista on se, etteivät käyttäjät itse ymmärrä kuinka heitä käytetään. Ja ratkaisuksi tähän näen sen, että käyttäjien ymmärrystä palveluiden toimintalogiikasta ja oman toiminnan vaikutuksista lisätään. Lisäksi oma ennako-oletukseni on, että haastateltavilla tutkijoilla on todennäköisesti myös kriittinen suhtautuminen sosiaalisen median algoritmeihin ja personointiin.

4.3 Haastatteluiden toteutus

Haastattelut toteutettiin etätyösuosituksesta johtuen etänä Zoom-ohjelmaa käyttäen lokakuun 2020 aikana. Haastatteluista tallennettiin ääniraita keskustelun litterointia ja analysointia varten. Verkon yli toteutetusta haastattelusta jäi uupumaan kasvokkaisen keskustelun mahdollistama läsnäolon tuoma parempi yhteys haastateltavaan, mutta vallitsevassa tilanteessa etähaastattelu oli kaikkien kannalta parempi vaihtoehto. Toisaalta uskon etähaastattelun madaltaneen haastateltavien kynnystä osallistua tutkimukseen, sillä se ei vaatinut heiltä ajankäytöllisesti samanlaista panostusta kuin paikan päällä toteutettava haastattelu olisi vaatinut. Haastatteluiden kesto vaihteli 45minuutista 1,5 tuntiin. Haastateltaville lähetettiin ennen haastattelua tutkimuksen taustaa ja tutkimusongelmaa avaava infopaketti, sekä ylätasoa kysymyksiä tutkimusaiheeseen liittyen, joista haastattelussa tulitaisiin keskustelemaan. Haastattelun teemat mukailivat pitkälti edellisessä luvussa esitettyjä tutkimuskysymyksiä, eli keskittyivät haastateltavien suhtautumiseen sosiaalista mediaa ja sen algoritmeja kohtaan sekä käyttäjien mahdollisuuksiin vaikuttaa personointiin. Haastattelurunko löytyy tutkielman lopusta liitteenä.

4.4 Analyysi ja luokittelu

Haastattelut analysoitiin sisällönanalyysin avulla. Sisällönanalyysia pidetään laadullisen tutkimuksen perusanalyysimenetelmänä ja sillä tarkoitetaan erilaisten sisältöjen analyysia väljänä teoreettisena kehyksenä (Tuomi & Sarajärvi, 2018, 78). Tuomin ja Sarajärven (2018, 86) mukaan sisällönanalyysimenetelmällä pyritään tiivistettyyn ja yleiseen kuvaukseen tutkimuksen kohteena olevasta aiheesta tai ilmiöstä. Menetelmän haasteena on pidetty sitä, että se ei itsessään anna johtopäätöksiä tutkittavasta aiheesta tai

ilmiöstä, minkä vuoksi monia sisällönanalyysillä toteutettuja tutkimuksia on kritisoitu keskeneräisyydestä (em. 87). Sisällönanalyysi vaatiikin erityistä huomiota johtopäätösten koostamisen osalta.

Tuomi ja Sarajärvi (2018, 88) korostavat sisällönanalyysin ja sisällön erittelyn eroina sitä, että sisällön erittelyllä pyritään kuvaamaan kvantitatiivisesti aineiston sisältöä ja sisällönanalyysissa taas keskitytään sanalliseen kuvaamiseen. He huomauttavat kuitenkin, että kvantifiointia voidaan käyttää sisällönanalyysin apuna. Omassa tutkimuksessani näen tämän olennaisena, sillä vaikka ajatukseni onkin keskittyä nimenomaan haastatteluissa nouseviin kuvauksiin tutkittavista teemoista, on myös kiinnostavaa se, kuinka usein samat asiat nousevat eri haastatteluissa.

Tuomi ja Sarajärvi (2018) mukaan sisällönanalyysia voidaan lähestyä kolmella eri tavalla: aineistolähtöisesti, teoriaohjaavasti ja teorialähtöisesti. Aineistolähtöisessä sisällönanalyysissa keskitytään ainoastaan aineistosta nouseviin havaintoihin ja analyysia lähdetään rakentamaan puhtaasti aineiston pohjalta. Tällöin aineistosta pyritään luomaan teoreettinen kokonaisuus ja tutkimuksen metodologiset sitoumukset ohjaavat analyysia vain sen toteuttamisen näkökulmasta. Teorialähtöisessä analyysissa taas luokittelu perustuu jo olemassa olevaan käsitejärjestelmään, kuten teoriaan tai malliin. Siinä siis kategoriat pohjautuvat aikaisempaan tietoon ja aineistosta etsitään sisältöjä kategorioita kuvaavista asioista. Teoriaohjaavassa lähestymistavassa aikaisemmat teoriat toimivat apuna, mutta analyysi ei suoraan pohjaudu teoriaan. Aikaisemman tiedon vaikutus tunnistetaan analyysia tehdessä, mutta se ei suoranaisesti toimi perustana analyysille. (Tuomi & Sarajärvi, 2018, 95-97.) Tässä tutkimuksessa sisällönanalyysi on teoriaohjaavaa ja analyysi perustuu aineiston läpikäynnin myötä nousseisiin toistuviin havaintoihin, jotka on koostettu seuraavassa luvussa esitettäviin teemoihin.

5 Tutkimustulokset ja niiden tulkintaa

Tässä osiossa käyn läpi haastatteluissa esiin nousseita havaintoja tutkimuskysymyksittäin. Alussa avaan haastateltavien suhtautumista sosiaalisen median palveluihin ja algoritmeihin, seuraavana käyttäjien toiminnan vaikutuksia algoritmeihin ja personointiin ja viimeisenä asioita, jotka asiantuntijoiden mukaan jokaisen sosiaalisen median palveluiden käyttäjien olisi hyvä ymmärtää palveluista ja niiden algoritmeista.

5.1 ”Some on tosi-tv:tä pahimmillaan” – Haastateltavien suhtautuminen sosiaalisen median palveluihin

Kaikki haastateltavat kuvailivat suhtautumistaan sosiaalista mediaa kohtaan jossain määrin kriittisenä. Haastateltava 1 nosti esiin monia sosiaalisen median hyviä puolia, mutta kertoi silti suhtautuvansa moniin kehityskulkuihin kriittisesti. Hän näkee oman suhtautumisensa noudattelevan pitkälti yleistä suhtautumista sosiaalista mediaa kohtaan, joka on muodostunut viimeisen 10 vuoden aikana:

”Ensinhän sosiaalinen media nähtiin Arabikevään ja muiden kansannousujen yhteydessä todella voimakkaasti demokratisoivana ja positiivisena muutosvoimana, siihen linkittyy tällainen voimakas demokratialupaus jostakin uudesta ja tällaiset ylisuuret odotukset, mitä teknologian kehitykseen linkittyy. Sittenhän tämä kuva on monipuolistunut hyvin paljon. Edelleenkin mä ajattelen, että sosiaalinen media ja algoritmit, niin siellä on se demokratiapotentiaali, niin kun me nytkin nähdään, esimerkiksi Valko-Venäjän tilanne todennäköisesti ei olisi kehkeytnyt noin pitkälle, se demokratia liikehdintä, ilman sosiaalisen median kaltaisia alustoja. Mut sitten taas samaan aikaan, jos katsotaan vaikka jonnekin autoritaariseen valtioihin tai ihan vaikka Yhdysvaltoihin niin on hirveän paljon sellaista pimeää vaikuttamista, jota tehdään muun muassa erilaisten datan hyödyntämiseen ja kalasteluun erikoistuneiden välikäsien kautta, joka on sitten taas sellaista vaikuttamista, jossa mä nään hirmu paljon negatiivisia piirteitä. [...] Sinänsä mä en ole totaalinen teknopessimisti, mutta kyllä suhtaudun moniin kehityskulkuihin myös aika kriittisesti.” (H1)

Haastateltava 2 kuvailee myös omaa suhtautumistaan sosiaalista mediaa kohtaan kriittisenä ja sanoo, että vaikka hän työskentelee sosiaalisen median kouluttajana, hän kouluttaa mieluummin tavallisia käyttäjiä kuin esimerkiksi yrityksiä:

”Ei niin hirveesti kiinnosta kertoa yrityksille kaikkia niitä keinoja, mitä sosiaalisessa voisi käyttää [vaikuttamiseen], vaan enemmänkin niin päin, että avataan normikäyttäjille sitä, että miten teihin vaikutetaan ja osaatte lukea näitä alustoja ja niitä vaikutuskeinoja.” (H2)

Haastateltava 3 kuvailee suhtautumistaan ”kriittisenä mielenkiintona”. Hän tutkii sosiaalista mediaa työkseen ja siksi on niistä kiinnostunut, mutta suhtautuu kriittisesti moniin palveluihin liittyviin tekijöihin ja kehityskulkuihin:

”Jos minun pitäisi jollain ehkä asteikolla vastata, että olenko puolesta vai vastaan, en mä oikein tiedä mitä siihen laittaisin. Et kyl ne tavallaan tekee myös paljon hyvää.” (H3)

Haastateltava 4 näkee, että sosiaalisesta mediasta on tullut jo niin iso osa ihmisten elämää, *”ettei maailmassa ole muunlaista vaikutusvaltaa kuin digitaalinen vaikutusvalta” (H4)*. Hän näkee, että vaikka elämä ilman sosiaalista mediaa ja internetiä voisikin olla onnellinen ratkaisu, tarkoittaa niiden ulkopuolella eläminen kaikenlaisten vaikuttamismahdollisuuksien ulkopuolelle jäämistä.

Haastatteluissa asiantuntijat nostivat esiin useita huolia liittyen sosiaalisen median algoritmeihin ja personointiin. Ensimmäinen usein toistunut huoli, joka mukailee myös julkisessa keskustelussa usein esiin nousevaa sosiaaliseen mediaan liittyvää ongelmaa, on disinformaation leviäminen sosiaalisen median palveluiden kautta (vrt. esim. Difranzo & Gloria-Garcia, 2017). Kun sosiaalisen median kautta yksittäisen henkilön julkaisemilla sisällöillä on potentiaali tavoittaa suuriakin yleisöjä, sekä tahattomasti, että tahallisesti esitettävä ja levitettävä virheellinen tieto voi saavuttaa isoja yleisöjä nopeasti. Haastateltava 1 näkee perinteisen median portinvartijuuden puuttumisen johtaneen esimerkiksi salaliittoteorioiden leviämiseen ja nousuun (vrt. esim. Seppänen & Väliverronen, 2014). Haastateltava 2 taas kuvaa algoritmien mahdollistaneen valeuutisten laajan leviämisen (vrt. esim. Difranzo & Gloria-Garcia, 2017). Palveluiden kautta kerättävän datan avulla disinformaatiota, kuten salaliittoteorioita, levittävien tahojen on mahdollista kohdistaa viestinsä sellaisille palveluiden käyttäjille, jotka ovat aikaisemmin olleet kiinnostuneet vastaavista teemoista tai jotka muutoin käyttäytymisensä perusteella voisivat olla helpommin alttiita uskomaan tarkoituksellisesti harhaanjohtavaa disinformaatiota (Benkler ym., 2018). Vastaavaa logiikkaa on hyödynnetty disinformaation levittämisen lisäksi esimerkiksi poliittisessa kampanjoinnissa, kuten Haastateltava 1 huomauttaa:

”Brexitistäkin tiedetään, että viimeisten viikkojen aikana ”Vote leave” -kampanja A/B testasi noin 450 erilaista Facebook-mainosta, josta sitten valikoitui hyvinkin matemaattisen tarkasti poliittiset valheet mitä lähdettiin ajamaan eteenpäin. Ja sitten se kohdennus tehtiin ydinkannattajaryhmiin ”suostuteltavat” erottaen. Elikkä se ei välttämättä näkynyt koko valtakunnan tasolla, vaan tehtiin hienovaraisesti suoraan näille kohderyhmille”

Haastateltava 2 kiteyttää ongelman olevan siinä, että aiheet ovat sosiaalisen median palveluiden kannalta arvottomia. Sillä onko jokin totta ei ole yritysten liiketoiminnan kannalta merkitystä vaan sillä, kuinka paljon se kiinnostaa palvelun käyttäjiä. Pullinen (2019) kuvaa, kuinka Facebook arvotti aiemmin sisältöjä sen mukaan, kuinka usein niitä klikattiin. Paljon klikkauksia saaneet sisällöt saivat algoritmien kautta entistä enemmän näkyvyyttä riippumatta siitä, mihin linkki johti. Tämän seurauksena palvelussa levisi trollaustarkoituksessa tehtyjä klikkiotsikoilla varustettuja linkkejä, joiden tarkoituksena oli ainoastaan klikkien kerääminen. Ilmiön seurauksena ja käyttäjiltä saadun palautteen perusteella Facebook muutti algoritmiaan. Myös 2016 Yhdysvaltojen presidentinvaalien aikaan disinformaatio levisi sosiaalisen median kautta vauhdilla, mistä seurasi erilaisia ”some gateja”. Rärkein esimerkki tästä on kenties ”pizzagate”, jossa sosiaalisessa mediassa levisi salaliittoteoria, jonka mukaan vaaleissa ehdolla ollut Hillary Clinton pyöritti ravintolan kellarissa lasten hyväksikäyttö ja ihmiskaupparinkiä yhdessä muiden demokraattipuolueen edustajien kanssa. ”Pizzagate” johti lopulta siihen, että joulukuussa 2016 yhdysvaltalainen mies hyökkäsi aseensa kanssa washingtonilaiseen pizzeriaan pelastamaan vankina pidettäviä lapsia, vain huomatakseen ravintolan tavanomaiseksi pizzapaikaksi. (Zadrozny, 2019.) ”Pizzagate” on viimeisen neljän vuoden aikana jalostunut Qanon -salaliittoteorialiikkeeksi, joka on erityisesti Yhdysvalloissa, mutta myös muualla maailmassa ja esimerkiksi Suomessa kerännyt viime aikoina paljon näkyvyyttä ja huomiota (ks. esim. Vehkoo, 2020; Zadrozny, 2019). Osa sosiaalisen median palveluista, kuten YouTube ja Reddit ovat tehneet tietoisien päätöksen salaliittoteorioihin liittyvien ryhmien kitkemisestä palveluissa, kun taas Facebookin perustaja Mark Zuckerberg on puolustanut erilaisten ryhmien oikeutta palveluissa, joskin kertonut Facebookin pyrkivän aktiivisesti misinformaation vähentämiseen alustalla. (Zadrozny, 2019.)

Disinformaation leviämiseen liittyen Haastateltava 1 huomauttaa, että ilmiön voidaan odottaa kasvavan lähitulevaisuudessa, kun sisältöä tuottavat tekoälyt kehittyvät vauhdilla, mikä tulee mahdollistamaan alustoille spesifin disinformaation tehotuotannon ja levittämisen. Myös Haastateltava 4 nostaa esiin tekstiä itsenäisesti generoivien tekoälyjen kehittymisen ja ennustaa näiden tulevan merkittävästi muuttamaan sosiaalisen median käyttöämme jo seuraavan vuoden aikana:

”Me ollaan nyt siinä pisteessä, että sisällöntuottaminen boteilla täysin rajattomassa mittakaavassa täysin uskottavalla sisällöllä on mahdollistumassa seuraavien muutamien kuukausien aikana. Siinä ei mene kuin hetki sen jälkeen, että noissa isoissa työkaluissa on generointityökalut, joita kaikki käyttää, koska se on vaan niin älyttömän helppo oikotie siihen mitä yrität tehdä [vaikuttaa]. Mutta sillä on myöskin väkisin vaikutuksia. Koska nyt jos kaikki kanavat menee ihan täyteen boteilla tuotettua sisältöä, niin mitä sitten? Kenen ääni kuuluu ja missä? Ketä seurata? Kaikki kanavat vaan menee sillai ”PUFF” täyteen.” (H4)

Hän ennustaakin sosiaalisen median käyttöämme muuttuvan seinäpohjaisesta (esim. Facebook, Twitter, LinkedIn) suljettuihin ryhmiin, jotka perustuvat tiettyihin kiinnostuksen kohteisiin ja joihin liittymisen edellytys on jonkinlainen validointi ihmiseksi.

Toinen haastatteluissa toistunut sosiaaliseen mediaan liittyvä ongelma oli tunnevetoisuus, jota myös Hatakka (2019) ja Matikainen ja kumppanit (2020) ovat nostaneet esiin tutkimuksissaan. Haastateltava 4 huomauttaa, että sosiaalisessa mediassa reaktio syntyy sekunneissa ja perustuu intuitioon. Erilaisia sisältöjä ei jäädä pohtimaan ja analysoimaan, vaan johtopäätökset tehdään nopeasti ja niihin reagoidaan impulsiivisesti tunteisiin pohjautuen. Hän toteaa, että *”on paljon kiinnostavampaa kuka sanoo ja miten sanotaan kuin että mitä sanotaan”* (H4). Haastateltava 2 näkee ongelmana tässä sen, että tunnevetoisten sisältöjen herättämät reaktiot ohjaavat sisällöntuottajia palveluiden parissa sellaiseen viestintään, joka tarkoituksenhakuisesti herättää voimakkaita tunteita, kuten esimerkiksi ärtymystä tai pettymystä. Hän näkee tämän erityisen ongelmallisena siksi, että eniten suuttumusta ja sen myötä myös reaktioita herättävät sisällöt tuottavat palvelun tarjoajille eniten voittoa, sillä ne

saavat käyttäjän viettämään palveluiden parissa enemmän aikaa, minkä seurauksena näille voidaan näyttää enemmän mainoksia. Myös Haastateltava 1 nostaa esiin tunnevetoisuuteen perustuvan viestinnän faktapohjaisten argumenttien sijaan ja näkee huolena erityisesti sen, kuinka sosiaalisen median kautta yleistynyt konfliktihakuisuus tulee läikkymään myös reaali maailmaan ja laajempaan keskustelukulttuuriin myös sosiaalisen median ja internetin ulkopuolella.

Kolmas haastatteluissa esiin noussut huoli liittyi intressiristiriitaan palvelun tarjoajien ja käyttäjien välillä, mikä nivoutuu myös osaksi 2. luvussa esiteltyä Hildenin (2017) nostamaa keskustelua käyttäjien oikeuksista. Haastateltava 2 muistuttaa, että algoritmit pyrkivät aina maksimoimaan algoritmin tekijän edun ja toteaa, että:

”Facebookin slogan on että ”voit pitää yhteyttä sinulle läheisiin ja tärkeisiin ihmisiin”, milloin mitenkään muotoiltuna. Että ikään kuin nämä alustat haluaisi tarjota meille sen yhteydenpito kanavan ja hyviä kokemuksia ystävien ja tuttavapiirin kanssa. Tai että voidaan seurata maailman menoa ja oppia siitä. Kaikkea tällaista positiivista. Sitten jos katsotaan miten algoritmit toimii, niin nehan tavallaan toimii juuri sitä vastaan. Algoritmi ihan käytännössä tekee sen, että et näekään kaikkien seurattavien tai ystäväpiirin julkaisuja, vaan näet niitä suosituimpia sisältöjä, jotka algoritmi sinulle valitsee. Ja tietenkin sitten mainoksia. Et tavallaan se toimii juuri vastoin kuin miten nämä alustat markkinoi itseensä.” (H2)

Hän näkee algoritmien tarkoituksena olevan käyttäjien toiminnan muuttamisen niin, että se tuottaa enemmän voittoa sosiaalisen median alustalle. Toiminta, jota pyritään muuttamaan, on kouruttaa käyttäjät pysymään palveluiden parissa mahdollisimman pitkään, jonka seurauksena näille voidaan näyttää mahdollisimman sopivia mainoksia. Haastateltava 3 huomauttaa, että *”alun perin esimerkiksi Facebook on luotu ”college-kirja sähköiseksi korvikkeeksi” (H3).* Kuitenkin kaupallistuminen ja sen mukana tulleet mahdollisuudet on ajaneet palveluita suuntaan, joka haastateltujen asiantuntijoiden näkemyksen mukaan on vienyt palvelun tarjoajien intressit kauas käyttäjien intresseistä. Kuten 2. luvussa esitettiin, tämä on ongelmallista erityisesti tietojen keruun ja käsittelyn laillisuuden

kannalta, mikäli käyttäjät eivät ole täysin tietoisia siitä, kuinka heidän tietojiaan käytetään.

Haastateltava 4 näkee, että lähtökohtainen ongelma sosiaalisen median palveluissa on se, että yrityksillä on nykyisen liiketoimintamallin kautta intressi myydä käyttäjien dataa markkinoille ja markkinoijille. Sen lisäksi, että tämä itsessään jo voidaan nähdä ongelmallisena (vrt. 2. luvussa esitelty EU:n yleinen tietosuoja-asetus), problemaattisuutta lisää se, että kattavan datan keräämiseksi palveluiden intressi on saada käyttäjä viettämään palveluiden parissa mahdollisimman paljon aikaa. Haastateltava 4 huomauttaakin, että siinä missä palvelut pyrkivät maksimoimaan palveluiden parissa vietetyn ajan, käyttäjien intressi olisi saada mahdollisimman laadukasta ymmärrystä ympäröivästä maailmasta mahdollisimman vähällä vaivalla, liittyä ympäröivä maailma sitten yhteiskunnalliseen tilanteeseen tai käyttäjän läheisten elämään. Myös Haastateltava 2 toteaa, että *”sosiaalisen median palvelujen näkökulmasta, on aika sama mitä sä teet siellä palvelun sisällä, kunhan sä oot siellä”* (H4). Kuitenkaan vaihtoehtoista liiketoimintamallia, joka mahdollistaisi palveluiden maksuttomuuden käyttäjille ei ole toistaiseksi löytynyt.

Milton Friedman määritteli 1970-luvulla, että yritysten tärkein tehtävä on tuottaa voittoa omistajilleen (Kanninen, 2014). Jos ajatellaan sosiaalisen median yritysten liiketoimintamallia, joka perustuu mainosrahoitukseen, voitaisiin nähdä, että näiden yritysten tehtävä on maksimoida mainostulot, mikä tarkoittaisi alustalla mainostavien mainostajien edun ajamista. Tästä näkökulmasta ristiriita palveluiden käyttäjien ja palveluiden tarjoajien välillä on erityisen ongelmallinen, kun mietitään tutkielman alussa (s. 13) esitetyn yleisen tietosuoja-asetuksen asettamaa vaatimusta siitä, että henkilötietoja käsiteltäessä rekisteröitävälle tulisi olla selvää mihin tästä kerättäviä tietoja käytetään. Jos palvelun käyttäjät ajattelevat tietojen keruun ja käsittelyn tarkoituksena olevan Haastateltavan 2 kuvaamat parempi yhteydenpito läheisiin ja tuttaviiin, mutta tosiasiasa tietoja kerätään erityisesti markkinoinnin paremman kohdentamisen mahdollistamiseksi, on tietojen keräämisen laillisuus perusteltua kyseenalaistaa. Tämä keskustelu kytkeytyy nähdäkseni hyvin keskeisesti siihen, miksi esimerkiksi Facebookin perustaja Mark Zuckerberg on ollut useita kertoja erinäisissä kuulemisissa kertomassa Facebookin tavasta kerätä käyttäjien tietoja. Onko palvelun olemassa olon tarkoitus tehdä bisnestä vai tarjota

käyttäjilleen mahdollisuuksia parempaan yhteydenpitoon? Kysymys ei ole yksiselitteinen ja kytkeytyy vahvasti laajempaan keskusteluun yritysten yhteiskuntavastuusta (ks. esim. Kanninen, 2014).

Neljäs esiin noussut algoritmeihin liittyvä ongelma koskee algoritmien epäselvää toimintalogiikkaa, mikä tekee niiden arvioimisesta ja vaikutusten seuraamisesta haastavaa. Haastateltava 3 huomauttaa, että sosiaalisen median algoritmeihin liittyy sekä palveluiden toimesta salattua tietoa että tietoa, jota palveluiden käyttäjät eivät ymmärrä, vaikka se olisi saatavilla. Hän näkee ymmärryksen lisääntyneen vuoden 2016 Yhdysvaltain Presidentinvaalien jälkeen nousseen Cambridge Analytica -kohun (Benkler ym., 2018) seurauksena:

"Ennen Cambridge Analyticaa se on ollut todella unknown-unknown, mutta nyt se on ehkä niinku known-unknown" (H3).

Tällä hetkellä siis tilanne on kenties se, että ymmärrämme, mitä emme ymmärrä. Haastateltava 3 näkee sosiaalisen median algoritmeihin liittyvän myös paljon "paisuttelupuhetta" tai niin sanottuja "kansantaruja":

"Just tyyliin nämä kuinka puhelimet kuuntelee meitä ja algoritmi tekee sitä ja tätä, joista tulee sellaisia kansan uskomuksia, joita edelleen levitellään, vaikka kuinka joku kaivaisi todeksi, että se ei mene näin" (H3).

Lisäksi hän huomauttaa, että algoritmeista puhuttaessa on hyvä ymmärtää, että palveluissa on päällekkäin usein tuhansia koodinpätkiä ja näkee ongelmana sen, että näiden toiminnasta on vaikea muodostaa kokonaisuutta. Monet haastateltavat korostivat, että sovelluksen toimimisen kannalta algoritmit ovat välttämättömiä, sillä ilman sisältöjen jäsentämistä olisi palveluita mahdotonta käyttää. Mutta se, mikä on palveluiden liiketoiminnan kannalta paras tapa jäsentää sisältöjä, ei haastateltavien mukaan ole parasta käyttäjien kannalta.

Viidentenä ja omasta mielestäni kenties huolestuttavimpana ongelmana algoritmeihin liittyen haastateltavat nostivat esiin sen, että algoritmit muuttavat meitä ilman, että itse tajuaamme sitä. Haastateltava 4 kuvaa algoritmien ja käyttäjien vuorovaikutusta kehänä, jossa käyttäjän tekemät valinnat johtavat tietynlaiseen sisältöön, joka taas vaikuttaa käyttäjään:

”Nuori tulee sosiaaliseen mediaan, sillä ei vielä ole vahvoja mielipiteitä. Sitten se kiinnostuu jonain päivänä vaikka QAnonista, lukaisee siihen liittyviä juttuja ja sen jälkeen sosiaalisen median filtteri adoptoituu niin, että ne syöttää sille melkein pelkästään sitä QAnon schaissee. Sen jälkeen sen maailmankuva rupee adaptoituu siihen Qanoniin sen seurauksena, et se joskus oli vähän aikaa kiinnostunut siitä. Filtreri adaptoitu sen käyttäytymisen kautta ja rupes suosittelee lisää ja siitä tuli kaninkolo, johon menttiin. Ja sit sä oot lopulta konspiraatioekstremisti. Se on vähän niinku otat jonkun gateway-huumeen ja ajaudut heroiniin.” (H4)

Haastateltava 4 korostaa, että kaikki algoritmiset optimoinnit vaikuttavat maailmankuvaan, sillä ne rajoittavat vaihtoehtoja ja tekevät valintoja käyttäjien puolesta. Käyttäjien tekemät valinnat eivät enää ole hänen omiaan, jos vaihtoehtoista on karsittu jo iso osa pois algoritmien toimesta. Algoritmien suodattaman sisällön kautta muodostunut kuva maailmasta aiheuttaa Haastateltavan 4 mukaan käyttäjille vinouman, jonka läpi maailmaa tarkastellaan. Suurimpana ongelmana algoritmeihin liittyen ei hänen mukaansa tässä ole niinkään se, että vinoumia syntyy (sillä vinoumia on ollut aina), vaan se, että ihmiset eivät ymmärrä vinoutumien syntyneen algoritmien vaikutuksen seurauksena, vaan he luulevat tehneensä itse tietoisia päätöksiä, jotka ovat johtaneet tiettyihin maailmankatsomuksiin ja mielipiteisiin:

”Ihmiset osaa identifioida missä sosiaalisessa kuplassa ne on. Se on selkee: mulla on tällöinen maailmankuva. Mutta se ei oo selkee miten siihen ollaan tultu. Heille omasta mielestä se on selkeä, he ajattelee että se on heidän oma valintansa. Mutta ei olla ymmärretty sitä, että siinä on saattanut tapahtua jotain muutakin kuin omia valintoja. Tai omat valinnat on vaikuttaneet algoritmeihin, jotka on ohjanneet käyttäytymistä ja sun identiteetti on nyt tämä. [...] Kukaan ei sano, että algoritmit teki musta natsin, vaan että mä ajattelin omia valintoja ja nyt mä nään asiat kirkkaammin ja oon nyt natsi. Eikä nää sitä, että ne on natseja sen takia, että ne aikaisemmin rupesi katsomaan jotain ylilautaforumilla tai jostain ja ajautui sieltä eteenpäin. Se polku on vienyt syvemmälle sinne kaninkoloon ja ollaan päädytty tohon. Ja sit ne kuvittelee, että mä aina halusinkin tätä ja mä tein itse päätöksen olla natsi tässä tapauksessa.” (H4)

Haastateltava 4 siis uskoo, että ihmiset ajattelevat itse päättäneensä millaisia he haluavat olla, vaikka hänen näkemyksensä mukaan tosiasiallisesti algoritmi on muokannut heistä sellaisia kuin he ovat. Ympäristö on aina muokannut valintojamme ja maailmankatsomustamme: aiemmin uskonto määräytyi perheen

mukaan ja puoluekanta ehkä ympäristön ja koulutuksen perusteella, mutta tällöin position syyt ja seuraukset oli helposti havaittavissa. Nyt tilanne on Haastateltavan 4 mukaan eri:

”Algoritmien maailmassa voit olla QAnon kannattaja ja et tiedä miksi. Luulet, että teit omia päätöksiä sinne menemisestä, mutta et tajua sitä polkua. Nykyisessä maailmassa olet jotain mieltä ja luulet että teet vapaan tahdon päätöksiä, mutta et välttämättä ymmärrä, että oot tätä mitä oot siksi, että teit 5 vuotta sitten valintoja sosiaalisen median sisältöjen suhteen, jotka rupesi pitkällä aikavälillä vaikuttamaan sun ajatteluun ja filtteröimään sun maailmankuvaa. Kunnes se sitten ajautui tähän pisteeseen ja nyt sä olet tämmöinen QAnon radikaali. Eli sä oot manipuloituna siinä paikassa, etkä vapailla valinnoilla.”

Haastateltava 4:n mukaan ihmiset ei hyväksy ajatusta, että heitä on ohjailtu ulkoapäin, miksi tätä on monien vaikea tunnustaa. Myös Haastateltavat 2 ja 3 nostavat esiin huolen algoritmien vaikutuksesta käyttäjiin. Haastateltava 3 toteaaakin:

”Mitä noi tavallaan tekee, ne jotenkin ajaa ihmiskuntaa sellaiseen, mitä me pahimmillaan ollaan. Jotenkin tosi teatraalisesti ilmaistu, mutta se tukee sellaisia perus inhimillisiä asioita, joista valitettavasti monet on myös huonoja.” (H3)

Sosiaalisen median tapa vahvistaa olemassa olevia vinoumia liittyy olennaisesti keskusteluun sosiaalisista kuplista, jota esittelin tämän tutkielman alussa (s. 4). Keskeinen tutkijoiden esittämä kritiikki sosiaalisen median vaikutuksista sosiaalisten kuplien syntymiseen oli, että palveluiden kautta ihmiset altistuvat tahtomattaankin erilaisille sisällöille, eikä siksi sosiaalisia kuplia pääse syntymään. Kuitenkin sosiaalisen median kautta radikaalisti leviävät salaliittoteoriat kuvastavat hyvin sitä, kuinka ne ajavat meitä ekstriimimpään ajatteluun, kuten myös haastateltavat yllä kuvasivat. Vaikka sosiaalisen median palveluiden kautta altistuisikin erilaisille sisällöille, esimerkiksi salaliittoteorioihin liittyvissä ryhmissä teorian kanssa ristiriidassa olevat uutiset käännetään aatteeseen sopivaksi. Osaltaan tätä tapaa kieltää oman mielipiteen kanssa ristiriidassa olevat näkemykset ja leimata ne valheellisiksi on edistänyt Yhdysvaltojen presidentti Donald Trump, joka leimaa ”vale uutisiksi” kaiken, mikä ei vastaa hänen näkemyksiään (Lee, 2018). Se, että uutiset, jotka ovat oman

mielipiteen kanssa ristiriidassa leimataan valeutisiksi ei tietenkään ole algoritmien syytä. Mutta kuten haastateltavat kuvaavat, algoritmit mahdollistavat sen, että käyttäjien on helpompi löytää omille, toisinaan valheellisiin käsityksiin perustuville näkemyksille laajempaa kaikupohjaa ja sitä kautta vahvistusta (vrt. esim. Vehkoo, 2020). Keskustelun sosiaalisista kuplista ei siis nähdäkseni pitäisikään liittyä siihen onko kuplien syntyminen sosiaalisen median palvelujen syytä vaan enemmänkin siihen, kuinka sosiaalinen media mahdollistaa olemassa olevien kuplien voimistumisen ja toisistaan erkaantumisen.

Lisäksi Haastateltava 1 nosti esiin huolen kerätyn datan päätyemisestä väärin käsiin. Hän näkee, että esimerkiksi autoritaarisemmissä valtioissa sekä valtioissa, joissa datan hallintaan ja yksityisyyteen liittyvät lait ovat Eurooppaa löyhemmät, mahdollisuus datan väärinkäytöksille aiheuttaa uhkia niin yksilö- kuin yhteiskunnallisella tasollakin. Tätä näkökulmaa nostaa myös Hilden (2017). 2. luvussa esitelty Yleinen tietosuoja-asetus koskettaa Euroopan unionia, eikä siksi päde esimerkiksi Yhdysvalloissa. Vaikka tiettyä sosiaalisen median palvelua käyttäisikin esimerkiksi vain EU alueella, teknisesti nämä yritykset saattavat sijaita paikassa, jossa datan keruuta ja käsittelyä ei ole yhtä tarkasti säännelty. Viimeinen Haastateltavan 1 esiin nostama ongelma sosiaalisen median palveluihin liittyen on huoli datayhtiöiden kasaantuvasta vallasta yhteiskunnassa, joka haastateltavan 1 mukaan on jo ”*todella ongelmallisissa mittasuhteissa*” (H1).

5.2 Käyttäjän toiminnan vaikutus personointiin

Monet edellisessä luvussa kuvatuista haastatteluissa esiin nousseista sosiaalisen median palveluiden ongelmista on luonteeltaan sellaisia, joihin palveluiden käyttäjä ei juurikaan itse voi vaikuttaa. Kuitenkin se, mistä haastateltavat olivat yleisesti yhtä mieltä on, että käyttäjien ymmärryksen kasvattaminen on tärkeää. Vaikkakin monet heistä näkivät sen yksistään riittämättömänä ratkaisuna koko tilanteeseen ja peräänkuuluttivat esimerkiksi poliittisen päätöksen teon tärkeyttä datan hallinnan ja datatalouden pelisääntöjen määrittelyssä. Kuten aiemmin mainitsin, yleisesti haastatteluissa nousi esiin huoli siitä, että käyttäjillä ei ole riittävästi ymmärrystä siitä, mihin sisältöjen personointi sosiaalisessa mediassa perustuu. Tässä luvussa käyn läpi

haastatteluissa esiin nousseita käyttäjistä kerättäviä tietoja, sekä sitä, kuinka nämä tiedot haastateltavien mukaan vaikuttavat sisältöjen personointiin sosiaalisen median palveluiden parissa. Tarkoitukseni on avata käyttäjille ymmärrettävällä tavalla sitä, kuinka heidän tietojaan palveluissa käsitellään ja mihin tietoja hyödynnetään.

5.2.1 Käyttäjistä kerättävät tiedot - ”Kaikki mitä irti lähtee”

Luvussa 2 (s. 16) esittelin Hildenin (2017) tekemän jaon personointimenetelmistä ja niihin vaikuttavista käyttäjistä kerättävistä tiedoista. Haastatteluissa esiin nousseet käyttäjistä kerättävät tiedot voidaan jakaa karkeasti seuraaviin ryhmiin: demografiset tiedot, tekniset tiedot, käyttöön perustuva data ja verkosto, jotka jossain määrin mukailevat myös Hildenin jaottelua. Kuten myös 2 luvussa mainittiin, haastatteluissa nousi vahvasti esiin, että se mitä kaikkea tietoa käyttäjistä kerätään ja miten sitä arvotetaan personoinnissa, riippuu vahvasti siitä, mistä palvelusta on kyse. Esimerkiksi Google kerää Haastateltavan 4 mukaan huomattavasti tarkempaa tietoa käyttäjistään kuin muut alustat. Samoin personointiin vaikuttavat painotukset vaihtelevat algoritmista ja palvelusta riippuen ja esimerkiksi TikTokin algoritmi personoi sisältöjä huomattavasti vahvemmin käyttöön perustuvan datan pohjalta, kun taas Facebookissa ja Instagramissa käyttöön perustuvan datan lisäksi personointiin vaikuttaa se, ketä tämän verkostoon kuuluu.

Kenties selkeimmät käyttäjistä kerättävät tiedot ovat demografiset tiedot, joista suurin osa ovat sellaisia, jotka syötetään palveluun kirjautumisen yhteydessä. Tällaisia tietoja ovat esimerkiksi ikä, sukupuoli, asuinpaikka, ammatti, koulutustaso, siviilisääty ja perhesuhteet (esimerkiksi Facebookissa voi merkitä käyttäjiä puolisoiksi, sisaruksiksi tai vanhemmiksi). Demografisten tietojen mukaan tehty personointi voitaisiin nähdä 2. luvussa esiteltyinä vihjepohjaisena personointina (Winter, Maslowska & Vos, 2020). Teknisten tietojen keruulla haastateltavat viittasivat esimerkiksi sähköpostiosoitteeseen, IP-osoitteeseen, laitetunnisteeseen, evästeisiin ja selain sormenjälkeen, joita kaikkia voidaan hyödyntää käyttäjän tunnistamiseen eri sivustoilla, vaikka tämä ei olisi kirjautunut palveluun. Selain sormenjälki on Haastateltavan 2 mukaan näistä kehittynein ja sitä voidaan hyödyntää seuraamiseen myös esimerkiksi

silloin, kun käyttäjä on kytkenyt evästeet pois päältä, minkä voisi kuvitella estävän käyttäjän verkkoselaamisen seurannan. Selain sormenjäljellä tarkoitetaan erilaisia tietoja, joita yhdistelemällä selain voidaan yksilöidä. Tällaisia tietoja ovat esimerkiksi laitteen käyttöjärjestelmä, resoluutio, kieliasetukset, laitteella olevat ohjelmistot, fontit, selaimen asetukset ja monet muut, joista mikään yksinään ei mahdollistaisi tunnistamista, mutta joiden yhdistelmä on uniikki. Teknisiä tietoja itsessään voidaan käyttää personointiin suosittelemalla esimerkiksi laitetunnisteen perusteella tunnistetun iPhone-puhelimen käyttäjälle muita Applen tuotteita. Yleisemmin teknisiä tietoja käytetään kuitenkin käyttäjän tunnistamiseen eri sivustoilla, jolloin verkkoselauksesta kerääntyvä data voidaan yhdistää tiettyyn käyttäjään. (Briz, 2018.)

Verkostolla viitataan siihen, keitä käyttäjän kavereihin kuuluu tai keitä hän seuraa sekä mitä tietoja näistä verkoston jäsenistä palvelulla on, kuten Haastateltava 2 huomauttaa:

”Keneen sinä olet yhteydessä tai keitä sun verkoston jäsenet on, kertoo siitä kuka sä olet... Ajatus siitä, että se ryhmä tai joukko määrittelee myöskin sinua. Ja tämähän näkyy myöskin tosi selvästi, että jos sinulla on kännykässä näitä somealustojen appseja, niin nehan jatkuvasti ehdottaa että tuoppa sun kontaktit tänne. Ja sekin toki myydään meille näin, että ”tee facebookista parempi ja löydä kaverisi ja löydä sinua kiinnostavat ihmiset Instagramista” ja siksi toisit niitä sinne. Vaikka sillä nimenomaan yritetään lisätä sitä personointidataa.” (H2)

Eniten tietoa käyttäjistä palveluille kerääntyy kuitenkin käyttöön perustuvana datana, eli käyttäjän toiminnan kautta. Käyttöön perustuva data voidaan jakaa palveluiden parissa tehtävästä toiminnasta kertyvään dataan sekä palveluiden ulkopuolella tehtävästä toiminnasta kertyvään dataan, joita molempia hyödynnetään palveluiden personoinnissa. Haastateltavilla oli hyvin yhtenäinen näkemys siitä, mistä toiminnasta palveluiden parissa tallennetaan tietoa, kuten Haastateltava 3 tiivistää:

”Kaikki mitä irti lähtee” (H3)

Haastateltavien mukaan käyttöön perustuvaan dataan liittyy esimerkiksi tiedot reaktioista, tykkäyksistä ja kommentteista, mutta sen lisäksi algoritmit tallentavat tietoa muun muassa sisältöjen parissa vietetystä ajasta,

”skrollauksen” eli sivun selaamisen nopeudesta sekä *jokaisesta* klikkauksesta, joka palveluiden parissa tehdään. Haastateltava 3 kuvaa käyttöön perustuvan datan tallentamista seuraavasti:

”Kaikki se mitä sä teet siellä somealustalla vaikuttaa, se data kerätään. Minkä tyyppisiin julkaisuihin olet reagoinut, kommentoinut, jakanut. Ja nekin sisältää aika monta asiaa: kuka sen julkaisun julkaisija oli, oliko se sun joku kaveri vai oliko se jossain ryhmässä, se vaikuttaa. Että mikä se julkaisija taho on suhteessa sinuun. Sit tietenkkin se itse julkaisu, että mitä aihetta se käsittelee ja minkä muotoista sisältöä siinä on, tekstiä, kuvaa, videota. Ja sitten jos on linkki niin mikä on se sivusto mihin se linkki johtaa.” (H3)

Käyttöön perustuvan datan keräämisen tarkoitus on auttaa algoritmia ymmärtämään, millaisesta sisällöstä käyttäjä on kiinnostunut ilman, että tätä kysytään käyttäjältä suoraan (ks. esim. Skrubbeltrang Mahnke, 2019; Thurman & Schifferes, 2012). Haastateltava 4 kertoo TikTokin algoritmin olevan kaikista kerättävistä tiedoista eniten käyttöön perustuvaan dataan pohjautuva ja kuvaa sen toimintaa seuraavasti:

”Esimerkiksi TikTokissa ne mittaa jokaisen teon. Jos kelaat sen videon eteenpäin ennen kuin se loppuu, niin se on negatiivinen signaali sen videon kiinnostavuudesta sinulle. Ja jos sä katsot sen loppuun tai katot useita kertoja loopilla, niin se on todennäköisesti sinulle kiinnostavaa sisältöä. Ja jos sä klikkaat sen ”discover music” napin siellä alalaidassa, niin se päättelee, että sä katsot videota sen musiikin takia. Tai jos sä klikkaat profiilia, niin ne päättelee et sä katot sen henkilön takia. Jos sä klikkaat jako tai kommentti tai tykkää, niin siitä vedetään kans omat johtopäätökset sen sisällön vaikuttavuudesta sinulle. Ja koska TikTokissa ne näyttää sulle vaan yhden videon kerrallaan siinä for you-pagessa, se kaikki interaktio kohdistuu vain siihen yhteen videoon. Jokainen sun teko sen videon kanssa merkitsee sille algoritmille.” (H4)

Hän sanoo TikTokin algoritmin olevan tällä hetkellä yksi maailman parhaista, sillä siinä algoritminen optimointi on viety *”täydelliseen maksimiin”*, jonka on mahdollistanut se, että TikTokissa ruudulla näkyy kerralla vain yksi sisältö (pystyvideo puhelimen ruudulla), jolloin kaikki interaktio voidaan yhdistää kyseiseen sisältöön. Esimerkiksi Facebookissa ja Twitterissä taas käyttäjä näkee kerrallaan vähintään 3 eri sisältöä, jolloin algoritmi ei voi varmasti tietää mikä sisältö sai käyttäjän pysähtymään.

Palvelusta riippumatta Haastateltava 4 näkee, että jokainen datapiste tallennetaan ja se vaikuttaa käyttäjälle suositeltavaan sisältöön. Myös se, ettei tee mitään ja ohittaa sisällön on algoritmin kannalta kiinnostava datapiste. Haastatteluissa nousi esiin myös se, että sivusta tai sisällöstä tykkäämisen on helppo ymmärtää vaikuttavan personointiin, sillä tällöin käyttäjä eksplisiittisesti ilmaisee pitävänsä kyseisestä asiasta. Se, millainen vaikutus sivujen selausnopeudella tai sisältöjen katseluajalla on personointiin ei kuitenkaan haastateltavien mukaan ole välttämättä käyttäjille yhtä selkeää, sillä tällöin käyttäjä ei varsinaisesti eksplisiittisesti ilmaise mielipidettään, vaan tulkinta tehdään käyttäjän implisiittistä toimintaa analysoimalla.

Sosiaalisen median palveluiden parissa kerättävän datan lisäksi palveluiden ulkopuolella tapahtuvasta toiminnasta kerätään dataa, joka vaikuttaa yhtä lailla sisältöjen personointiin palveluiden parissa. Esimerkiksi Facebook tarjoaa mainostileille ”Facebook pikseliä”, eli koodinpätkää, joka lisätään mainostajan omille verkkosivuille ja jonka avulla mainostaja pystyy seuraamaan ketkä Facebook-käyttäjät vierailevat yrityksen verkkosivuilla. Haastateltava 2 kuvaa Facebookin verkkoselailusta kertyvän datan tallennusta seuraavasti:

”Jos sä menet jollekin uutissivustolle tai verkkokauppaan tai jonkun ihan yrityksenkin sivustolle, niin siellä hyvin usein on joku Facebookin seuranta scripti, siis Facebook pikseli tai Facebook kirjautuminen tai upotettu se Facebook sivu sinne vaikka yrityksen kotisivuille. Mikä tahansa Facebookista ladattava scripti tai tiedosto mahdollistaa sen, että Facebook pääsee seuraamaan sua siellä sivustolla. Sillä saadaan toisin sanoen sun verkkoselailusta kertyvä data myöskin talteen. Ja siis pelkästään Facebookin tykkäysnappula, jos se on tehty Facebookin omalla scriptillä, niin niitähän on miljoonilla sivustoilla. Ja niistä kertyvän datan määrä on valtava. Ja vaikka sä et olisi Facebookissa kirjautunut sillä hetkellä, jos sulla on koneella Facebookin eväste tallentunut aikaisemmin, niin se seuranta saadaan sinusta, että missä olet käynyt. Ja vaikka siellä ei olisi evästettäkään, nykyään ne on jo niin kehittyneitä, puhutaan selain sormenjäljistä, niin jokainen tietokone voidaan yksilöidä sen perusteella mitä ohjelmia, fontteja, mikä resoluutio, kaikki tällaiset asiat, mikä nettiselain siellä on. Ne muodostaa ikään kuin tämmöisen sormenjäljen sitä käyttäjän laitteesta ja selaimesta. Niin silläkin voidaan tunnistaa että sä oot ollu sillä samalla laitteella aikaisemmin Facebookissa, niin sitten tietää että sinä olet se joka nyt kävi tällä sivustolla.” (H2)

Toisin sanoen siis riippumatta siitä, onko käyttäjä kyseisellä hetkellä kirjautunut Facebookiin vai ei, Facebook tallentaa tietoa siitä, millä verkkosivuilla käyttäjä vierailee ja mitä näillä sivustoilla tekee. Haastateltava 3 näkee

palveluiden ulkopuolella tapahtuvan datankeruun ongelmallisempänä kuin palveluiden parissa tapahtuvan datankeruun:

”Toi on ehkä mun mielestä kaikkein näätämäisin täs somen seuranta puolesta. Että käyttäjänä olisin valmis allekirjoittamaan sen, että no kytätkää mitä mä teen täällä, mutta älkää kytätkö mua kaikkialla muualla missä olen. [...] Pari vuotta sitten kohkattiin siitä, että Facebook saattaa olla tallentanut tiedon että sä olet soittanut tällaiseen numeroon siihen ja siihen aikaan, koska sillä applikaatiolla on pääsy tällaiseen tietoon. Että se saatto tietää myös meidän puhelimen käyttötavoista ja kaikkee muuta ihan friikkiä.” (H3)

5.2.2 Miten sisältöjen personointi tapahtuu?

Käytännössä personointi sosiaalisessa mediassa tarkoittaa haastateltavien mukaan sitä, että kerätyn datan perusteella käyttäjälle tarjotaan sen kaltaisia sisältöjä, joista tämä on aiemmin kiinnostunut. Haastateltava 3 kuitenkin muistuttaa, että personointiin vaikuttaa käyttäjän lisäksi koko palvelun toimintalogiikka, eli liiketoimintamalli ja siihen liittyvät asiat. Personointi asettuukin kiinnostavaan välimaastoon aiemmin mainitun käyttäjien ja palvelun tarjoajien intressien välisessä ristiriidassa: algoritmien tarkoitus on edistää sen omistajan, eli palvelun tarjoajien intressiä ja liiketoimintaa, jonka edellytys on, että käyttäjä viettää palveluiden parissa mahdollisimman paljon aikaa. Sillä mitä enemmän aikaa palveluiden parissa vietetään, sitä enemmän dataa kertyy ja sitä enemmän käyttäjälle voidaan esittää mainoksia, joka tietää palveluntarjoajille mainostuloja. Palvelut siis pyritään rakentamaan käyttäjille mahdollisimman kiinnostaviksi ja koukuttaviksi.

Haastateltava 4 kuvaa personoinnin tapahtuvan isossa kuvassa seuraavasti:

”Sulla on x määrä sisältöjä, joita sä voit nähdä tällä hetkellä. Facebook esimerkiksi nappaa niistä top 300 ja pistää ranking järjestykseen erilaisilla kriteereillä: milloin on postannut, kuka on postannut suhteessa sinuun, kuinka paljon ne on saanut engagemanttia muualta ja mikä niiden sisältö on suhteessa siihen, mihin sinä olet aikaisemmin osallistunut. Nämä kaikki asiat kun summautuu yhteen niin siitä tulee joku ranking arvo ja sitten se näyttää sulle sen top 300 skrollattavana feedinä.” (H4)

Haastateltava 1 kuvaa käyttäjistä kerätyn tiedon muodostavan pilven asioita, jotka jollain tapaa linkittyvät käyttäjän kiinnostuksen kohteisiin. Ja mitä kattavampi pilvi on, sitä monipuolisempaa sisältöä käyttäjille voidaan tarjota.

Haastateltava 2 näkee, että algoritmit pyrkivät laajentamaan näkymää siihen, mistä kaikista aiheista käyttäjä on kiinnostunut, mutta tällä hetkellä algoritmit muodostavat hyvin yksipuolisia kuvia käyttäjistä:

”Sä vietät enemmän aikaa ja sitä kautta sinulle voidaan nimenomaan näyttää sen tyyppisiä julkaisuja, kun mitä sä olet aikaisemmin kuluttanut, katsonut, lukenut, klikannut. [...] Se lopputulos on se, että se kaventaa myöskin sitä mitä sä saat jatkossa siitä alustasta. Ja sitä kautta se ohjaa sinua keskittymään vielä voimakkaammin niihin asioihin, mihin sinä olet reagoinut määrällisesti eniten. Ja lopulta se on sekä sosiaalinen että tiedollinen kupla. Se luo harhan ja oppii sinusta enemmän joitakin puolia kuin toisia puolia.” (H2)

Haastateltavan 2 mukaan personoinnin seurauksen käyttäjän näkökulmasta voi tiivistää sillä, että käyttäjille annetaan yhä vähemmän vaihtoehtoja tehdä valintoja itse ja yhä enemmän valmiina se, mikä näitä pitäisi kiinnostaa ja mitä näiden pitäisi lukea tai katsoa. Samalla personointi johtaa siihen, että näemme yhä vähemmän sisältöjä henkilöiltä, jotka ovat kanssamme eri mieltä, mikä lopulta johtaa siihen, että emme enää ymmärrä toisiamme ja eriävien mielipiteiden taustoja:

”Se vaikuttaa niin monella tasolla jo sekä niihin sosiaalisiin että tiedollisiin asioihin, mitä me saadaan verkon kautta, että se pahimmillaan voi lopettaa luovan ajattelun ja oppimisen. Että sulle vaan syötetään sitä samaa, mihin olet jo kallistunut. Sinulle annetaan vähemmän uusia virikkeitä ja uusia tietoja.” (H2)

Tämä näkökulma puoltaa vahvasti ajatusta sosiaalisista kuplista, joka esiteltiin luvussa 2 (ks. esim. Seppänen & Väliaverron, 2014). Yleisesti sosiaaliset kuplat nähtiin haastateltavien keskuudessa lähes itsestäänselvyytenä. Kuitenkin esimerkiksi Haastateltava 3 huomautti useampaan otteeseen, että palveluiden kautta saatavaan tietoon vaikuttaa muutkin asiat kuin algoritmit, jatkaen esimerkiksi Brunsin (2019) ajatuksia.

Haastateltava 2 antaa lisäksi konkreettisen esimerkin palvelun ulkopuolisen datan keruun hyödyntämisestä sosiaalisen median palveluiden personoinnissa: esimerkiksi Facebook kerää tietoa myös sellaisista käyttäjistä, jotka eivät ole vielä rekisteröityneet Facebookiin.

”Puhutaan tällaisista varjoprofiileista, eli Facebook kerää tietoa myöskin käyttäjistä, jotka eivät ole vielä Facebookin käyttäjiä. Eli sinusta on kertynyt seuranta sinne, vaikka et vielä olisi rekisteröitynyt sinne palveluun ja kun joskus rekisteröidyt niin sinusta on jo valmiiksi hirveä määrä tietoja, jolla sulle voidaan personoida sitä sisältöä.” (H2)

Sen lisäksi, että palvelut keräävät käyttäjistä tietoa omaan käyttöönsä, eri datan kerääjät yhdistelevät lisäksi keräämiään tietoja palveluiden välillä, kuten Haastateltava 2 kuvaa:

”Se ei ole vaan se mitä se yksittäinen sivusto tai palvelu susta kerää, vaan näiden tunnisteiden eli esim. nimi, sähköpostiosoite, laitetunnisteet, mainostunnisteet ja se selain sormenjälki ja sitten jopa ip-osoite jossakin tilanteessa, niin kaikki tällaiset mahdollistaa sen, että nämä eri datan kerääjät sitten vielä synkronoi tiedot keskenänsä. Toisin sanoen, jos sä oot vaikka johonkin sivustolle antanut sinun sähköpostiosoitteen uutiskirjeen tilaamista varten, niin sen sähköpostiosoitteen perusteella suhun voidaan yhdistää valtava määrä tietoa muista datan kerääjistä.” (H2)

Haastateltavan 3 mukaan tämä on ongelma erityisesti Yhdysvalloissa, jossa dataa on ylipäänsä enemmän kaupan.

Sen lisäksi, että palvelut itse hyödyntävät käyttäjistä kerättyä dataa muiden käyttäjien tuottamien sisältöjen suodattamiseen, dataa myydään myös mainostajille. Esimerkiksi aiemmin mainitun Facebook pikselin avulla kerätyn tiedon perusteella Facebook tarjoaa mainostajalle mahdollisuuden kohdentaa mainontaa suoraan niille käyttäjille, jotka ovat jo vierailleet näiden sivuilla ja siten ilmaisseet kiinnostuksensa mainostajan palvelua kohtaan. Käyttäjistä kerätyn datan perusteella voidaan tehdä hyvinkin tarkkoja päätelmiä esimerkiksi tämän poliittisista mielipiteistä, joita sitten voidaan hyödyntää poliittisessa kampanjoinnissa, kuten myös aiemmin esitetty Brexitin ”Vote Leave” -kampanjan A/B-testaukseen liittyvä esimerkki osoitti. 2. luvussa esitetty Cambridge Analytica hyödynsi nimenomaan tätä mallia sekä Donald Trumpin presidentinvaalikampanjassa, että Brexitin ”Vote Leave” -kampanjassa profiloimalla käyttäjiä näistä kerätyn tiedon mukaan ja sitten syöttämällä näille kohdennettuja viestejä sen mukaan, minkä kulman uskottiin purevan kyseiseen vastaanottajaan parhaiten (Benkler ym., 2018).

Kuten aiemmin mainittiin, se, miten käyttäjistä kerättyä tietoa personoinnissa arvotetaan, riippuu palvelusta ja algoritmista. Verkostoa painottavat algoritmit

suosittelevat sisältöjä sen mukaan, joihin käyttäjän verkoston keskeiset jäsenet (eli ne käyttäjät, joiden jakamiin sisältöihin käyttäjä on reagoinut) ovat reagoineet. Tällaisilla alustoilla verkoston jäsenten toiminta on siis yksi personointiin merkittävästi vaikuttava tekijä. Haastateltavan 4 mukaan esimerkiksi Facebook ja Instagram painottavat personoinnissaan verkostoa huomattavasti enemmän kuin esimerkiksi TikTok tai Reddit. Haastateltavan 3 mukaan Facebook on kertonut personointinsa perustuvan pääasiassa ”merkityksellisiin interaktioihin” (*meaningful interactions*, vapaa suomennos), ja tässä sisältöjen kommentointi on erityisen merkittävää. Mutta kuten aiemmin mainittua, aiheet ovat sosiaalisessa mediassa arvottomia ja siksi ei ole sinänsä väliä mitä kommentoi, vaan se, että kommentoi. Tämän seurauksena käyttäjät saattavat saada lisää sellaisia sisältöjä, joista he eivät pidä, jos he ovat provosoituneena jättäneet esimerkiksi vasta-argumentin tiettyyn sisältöön, kuten Haastateltava 3 kuvaa:

”Tää kommentoi näitä, tää haluaa nähdä tällaista lisää, lisää tätä tuubaa.”
(H3)

Haastateltavan 2 mukaan ongelma tiedostetaan myös Facebookin sisällä:

”Saatetaan reagoida sillä vihanaamalla, koska me ei pidetä siitä ja sit me juuri sen takia saadaan sitä lisää. [...] Facebook tiedostaa tämän myös. Heidän yks operatiivinen johtaja olikin sanonut haastattelussa, että ihmiset reagoi kaikista eniten näihin provosoiviin sisältöihin, vaikka he sitten jälkikäteen sanovat, että he eivät halua nähdä näitä.” (H2)

Haastateltavan 2 mukaan ärsyyntymistä ja suuttumusta herättävien sisältöjen on havaittu sitouttavan käyttäjiä tehokkaasti ja koska algoritmit suosivat sitouttavia sisältöjä, ne suosittelevat näitä eteenpäin ja kauemmin näin laajentaen sisällön tavoittamaa yleisöä entisestään. Tämä puolestaan on johtanut siihen, että sisältöjä tuotetaan tarkoituksellisesti provosoivasti laajemman yleisön tavoittamiseksi, mikä lisää riitoja ja suuttumusta sosiaalisen median palvelujen parissa. Esimerkiksi populistiset puolueet ovat Hatakan (2019) mukaan hyödyntänyt tätä logiikkaa viestinnässään sosiaalisessa mediassa.

Haastateltava 1 nostaa kuitenkin datankeruusta esiin myös kiinnostavan pointin: vaikka suurin osa käyttäjistä on hänen mukaansa tietämättömiä heihin liittyvän datan keräämisestä, suuri osa on myös välinpitämättömiä. Myös

Haastateltava 3 huomauttaa, että sisältöjen personointi ja siitä seuraava kuplautuminen ei välttämättä kaikille näyttäytyä ongelmallisena:

”Et siinä on sitten helposti se ongelma, että me voidaan täällä huudella korkeakoulutetusta norsunluutornista, että me haluttaisiin erilaista sisältöä, mutta sit voi olla että 90% käyttäjistä ei halua ja se näkyy koko ajan siellä metriikoissa, että ne on tosi tyytyväisiä siihen nykytilaan. Ja et ne ei todellakaan haluakaan nähdä kaikkea mahdollista vaan sitä heille kiinnostavaa sisältöä.” (H3)

5.3 Mitä sosiaalisen median palveluiden käyttäjien tulisi palveluista ymmärtää

Kuten asiantuntijoiden suhtautumisesta sosiaaliseen mediaan voi päätellä, palveluihin liittyy paljon tekijöitä, joista käyttäjien olisi hyvä olla tietoisia. Haastateltava 1 näkee, että sosiaaliseen mediaan liittyviin ongelmiin ratkaisun löytäminen on fundamentaalisesti poliittinen kysymys, mutta koska poliittiset ratkaisut saadaan huomattavasti hitaammin kuin teknologia kehittyä, on käyttäjien tietoisuuden ja medialukutaidon lisääminen tällä hetkellä tärkeää. Yleisesti käyttäjien tietoisuuden kasvattaminen nähtiin haastateltavien keskuudessa tärkeänä.

Vaikka sosiaalisen median algoritmeihin liittyy paljon asioita, joihin käyttäjien tulisi suhtautua kriittisesti, on hyvä muistaa, että pohjimmiltaan ne kuitenkin vastaavat konkreettiseen tarpeeseen. Kuten luvussa 2. mainittiin, sosiaalisen median algoritmit syntyivät tarpeesta järjestää sisältöjä (ks. esim. Montaner, Lopez & de la Rosa, 2003). Ilman algoritmeja palveluja olisi mahdotonta käyttää, sillä sisältöä on niin paljon, kuten Haastateltava 3 toteaa:

”En mä haluaisi nähdä sitä mun raakaa Twitter fiidiä myöskään. Koska mähän kuolisin siihen. Että tavallaan sieltä tulee niin paljon informaatiota, että mä myös ymmärrän sen, että tuo on ainoa ratkaisu tehdä niistä palveluista käytettäviä.” (H3)

Lisäksi kaikki haastateltavat muistuttivat, että sosiaalinen media ja siihen liittyvä teknologia ovat mahdollistaneet paljon hyvää. Kansannousut ja muut poliittiset liikehdinnät, viimeisimpänä esimerkkinä Valko-Venäjän viimeaikaiset tapahtumat (Kerttula, 2020) eivät olisi nykyisessä mittakaavassa olleet

mahdollisia ilman sosiaalisen median mahdollistamaa kansalaisliikehdintää ja organisoitumista. Haastateltava 3 toivookin, että myös näistä puhuttaisiin julkisessa keskustelussa enemmän:

”Kyllä siitäkkin tavallaan voisi puhua myös julkisuudessa vähän enemmän eikä aina vain kauhistella. Kyllä se on myös mahdollistanut paljon erilaista sosiaalista ja poliittista toimintaa.” (H3)

Algoritmeille ei voi myöskään yksistään sysätä vastuuta sisällöistä alustoilla, sillä kokonaisuuteen liittyy paljon muutakin: palvelun tarjoajien liiketoimintamallit, sääntely sekä sisällöntuottajien toiminta alustoilla. Kuten luvussa 2 esitettiin, tätä myös useat tutkijat kyseenalaistivat sosiaaliin kupliin liittyen (ks. esim. Ananny & Crawford, 2018). Kuten aiemmin todettua, käyttäjät kuitenkin itse vaikuttavat algoritmeihin myös ja siksi omaa toimintaa palveluissa on hyvä tarkastella kriittisesti, kuten Haastateltava 3 toteaa:

”Sekä teknologia että ne ihmiset itse vaikuttaa siihen mitä he näkee fiidissä. Et tavallaan sen monimutkaisuuden ymmärtäminen ja juuri se, että ei aina kannata syyttää teknologiaa, voi myös miettiä niitä omia tekemisiään. [...] Olis hirveen helppoa meille ettei tarvis puhua itsestään, vaan näistä algoritmeista ja teknologiasta. Mut et ei voi syyttää Facebookkia siitä, kun maailma on paska tyyppisesti. Et ehkä voi olla, että meillä on jotain omia asenteita.” (H3)

Sosiaaliin kupliin liittyen haastateltavat näkevät, että käyttäjät voivat itse vaikuttaa siihen millaisiin kupliin nämä sosiaalisessa mediassa ajautuvat ymmärtämällä tiettyjä asioita palveluista ja niiden toimintalogiikasta. Haastateltava 1 nostaa tärkeimpänä asiana sosiaalisen median ymmärrykseen liittyen sen, että käyttäjän omassa virrassa näkyvät asiat eivät välttämättä näy kumppanin tai parhaan ystävän virrassa. Hän uskoo, ettei tämän seuraukset välttämättä ole merkittäviä, mutta ne voivat olla, jos sen seurauksena mielletään esimerkiksi yhteiskunnalliset teemat tai ilmiöt täysin eri tavoin:

”Yhteisen tilannekuvan luominen voi olla sen seurauksena hankalaa. Mutta joka tapauksessa ajatus siitä, että ei ole mitään neutraalista uutisvirtaa, joka jotenkin muodostuisi kronologisesti siihen, vaan se on sinulle räätälöityä. Se on sinulle tehtyä sisältöä.” (H1)

Myös Haastateltava 4 nostaa tämän tärkeimpänä asiana, mikä sosiaalisen median palveluista pitää ymmärtää:

”Tieto mitä sä saat sun feedistä ei ole neutraalia. Se vaan pitää ymmärtää, että se on vinoutunut johonkin suuntaan ja se muokkaa sun maailmankuvaa.”
(H4)

Oma toiminta vaikuttaa merkittävästi siihen, millaista sisältöä algoritmit käyttäjälle suosittelevat ja siksi olisikin tärkeää kiinnittää huomiota siihen, millaisiin sisältöihin reagoi. Haastateltava 4 kehottaa ennen mihinkään sisältöön reagointia pohtimaan omaa suhtautumista sisältöön. Sisältöjen tulkintaan vaikuttaa aina merkittävästi oma tausta ja sosiaaliset suhteet, ja erilaisista lähtökohdista tulevat ihmiset tulkitsevat viestejä eri tavoin (Turow, 2011, 16). Siksi Haastateltava 4 kehottaa pohtimaan miltä sama viesti näyttää toisesta kuplasta käsin:

”Tee semmoinen toisinpäin ajattelu: jos mä olisin jossain toisessa kuplassa, yritä miettiä joku päinvastainen kupla itsellesi, mitä mä ajattelin tästä samasta asiasta silloin? Ainakin se pakottaisi sinut miettimään asiaa toiseltakin kannalta kuin siitä sun omasta kuplasta katsoen. Ja jos se näyttää hyvin erilaiselta sen kuplan silmin, sä pystyt näkemään samalla sun omia vinoumia ja mihin suuntaan ne menee.” (H4)

Ongelmana tässä on kuitenkin Haastateltavan 4 mukaan se, että sosiaalisessa mediassa reaktiot perustuvat impulsseihin. Maailmankuvan avautumisen kannalta tulisi olla aikaa tutkia erilaisia asioita ja punnita niiden eri puolia. Lisäksi erilaisten sisältöjen kohdalla tulisi pohtia, ”miksi näen tämän?”, sillä esimerkiksi mainosten tunnistaminen sosiaalisen median alustoilla saattaa olla haastavaa kuten Haastateltava 2 huomauttaa:

”Esimerkiksi ei se tubettaja oikeasti ole kiinnostunut siitä asiasta, vaan se sinulle mainostaa sitä, koska siinä on se maksaja taustalla ja sinä näet sen videon koska sä oot ollut ennestään kiinnostunut niistä aiheista. Et suhtautuu periaatteessa aina silleen kriittisesti ja epäluuloisesti, että onko tämä nyt mainos.” (H2)

Kuplien rajojen rikkomiseksi Haastateltava 4 suosittelee lisäksi pohtimaan, keitä omaan verkostoon kuuluu:

”Frendasin LinkedInissä lähtökohtaisesti kaikista mahdollisista kuplista olevia ihmisiä tarkoituksella. Halusin varmistaa, että minulla ei ole minkään näköistä filttterivinoumaa. Mulla on siis kaikista geografioista, kaikista ammatillisista senioriteettitasoista ja kaikista mahdollisista systeemeistä porukkaa. – Koitan pitää Twitter-feediä semmoisena neutraalina, että siellä on monenlaisista suunnista tulevia mielipiteitä ja viestejä. Että mulla löytyy

sieltä liberaaleja, republikaaneja, demokraatteja Jenkkien osalta. Ja sitten on bisnestä ja muuta erilaisista näkökulmista tulevia asioita.” (H4)

Lisäksi hän kehottaa tietoisesti tutustumaan eri kuplissa käytäviin keskusteluihin:

”Meen Redditiin ja katon sieltä, että mitä republikaanikuplassa jutellaan. Ja sit mä käyn kattomassa, että mitä täällä toisella puolella jutellaan. Mä katson mitä konspiraatioteoristit juttelee tällä hetkellä. Mä teen tuollaisia tsekkejä vaan, että mitä täällä tapahtuu. Ja se on aika mielenkiintoista. Se tuo aika selkeästi esille, että miten eri kuplat keskustelee.” (H4)

Lisäksi haastatteluissa nousi esiin muutamia teknisiä asioita, jotka haastateltavien mukaan voivat auttaa käyttäjiä omien sosiaalisten kuplien rikkomisessa. Haastateltavat 2 ja 4 kehottivat tutustumaan Facebookin asetusten alta löytyviin tietoihin siitä, kuinka Facebook on käyttäjän profiloinut. Halutessaan tiettyjä, esimerkiksi mainosten personointiin vaikuttavia, Facebookin tallentamia kiinnostuksen kohteita voi poistaa. Samoin Googlen tallentamia tietoja käyttäjästä voi käydä tarkastelemassa osoitteessa *myactivity.google.com* (Pullinen, 2019). Haastateltava 4 nostaa lisäksi esiin, että eniten maailmankuvaamme vaikuttaa hakukoneiden tarjoamat hakutulokset, jotka personoidaan käyttäjälle samalla tavalla kuin sisällöt sosiaalisessa mediassa ja antaa hakukoneen käyttöön liittyen konkreettisen vinkin, jolla välttää personointia:

”Googلهan personoi tuloksia sillä perusteella mitä sä oot tehnyt aikasemmin, mitä videoita sä katot YouTubesta, mitä hakuja on sun systeemissä, se tallentaa ne ja antaa sulle personoituja tuloksia. Mä aluksi siirryin sellaiseen hakukoneeseen, kun DuckDuckGo, nykyään käytän sellaista hakukonetta kun StartPage, joka itseasiassa renderöi Googlen tulokset, joka poistaa niistä kaiken personoinnin. Eli mulle näkyy kaikki hakutulokset netistä neutraalina. Se on niin valtavan iso osa meidän tiedon keräämistä ja käyttöä se hakukone, ja tolla saa sen ensimmäisen filtterin pois siitä systeemistä. Toi oli mulle aika tärkeä moovi. Jos mulle kaikki tieto tulee jonkun filtterilinssin läpi, niin se on aika nihkeä piste lähtee rakentaa jotain maailmankuvaa.” (H4)

Hakukoneiden personoinnin vaikutusta tiedon välityksen kannalta ovat tutkineet tarkemmin esimerkiksi Epstein ja Robertson (2015).

Lisäksi Haastateltava 4 nostaa esiin kiinnostavan asian omasta sosiaalisen median käytöstään, jota suosittelee myös muille; hän ei käytä aikaa sosiaalisen

median seinien tai syötteiden selailuun vaan osallistuu keskusteluihin sosiaalisessa mediassa päätyemällä niihin yksityisissä ryhmissä jaettujen linkkien kautta.

5.4 Johtopäätökset

Sosiaalisen median tutkijat ja asiantuntijat suhtautuvat kriittisesti palveluihin ja näkevät näihin liittyvän monia ongelmallisia kehityskulkuja. He kuitenkin myös korostavat palvelujen tuoneen paljon hyvää, eivätkä siksi halua tuomita palveluja kokonaan. Keskeisimpänä ongelmana palveluihin liittyen he näkivät algoritmisen personoinnin vaikutuksen disinformaation leviämiseen sekä identiteetin tiedostamattoman muokkautumisen algoritmien personoiman sisällön seurauksena.

Käyttäjistä kerätään sosiaalisessa mediassa hyvin kattavasti tietoa, jotka voidaan haastattelujen perusteella luokitella demografisiin tietoihin, teknisiin tietoihin, käyttöön perustuvaan dataan ja verkostoon. Demografiset tiedot ovat usein sellaisia, jotka käyttäjät itse syöttävät palveluun sinne rekisteröityessään, kuten nimi, ikä ja asuinpaikka. Teknisillä tiedoilla viitataan käyttäjän yksilöiviin tietoihin, joiden avulla käyttäjää voidaan seurata eri sivustoilla ja palveluissa, vaikka tämä ei olisi kirjautunut palveluihin sisään. Tällaisia tietoja ovat esimerkiksi laitetunnisteet ja selain sormenjälki. Käyttöön perustuvalla datalla viitataan kaikkeen käyttäjän toimintaan verkossa ja se voidaan jakaa palveluissa kerättyyn käyttöön perustuvaan dataan ja palveluiden ulkopuolella kerättyyn käyttöön perustuvaan dataan. Palveluiden parissa kerätyllä käyttöön perustuvalla datalla tarkoitetaan esimerkiksi klikkauksia, tykkäyksiä ja kommentteja, jotka tehdään sosiaalisen median palveluissa, kuten esimerkiksi Facebookissa. Palveluiden ulkopuolella kerätyllä käyttöön perustuvalla datalla taas viitataan käyttäjän toimintaan muilla verkkosivuilla, joita esimerkiksi Facebook tallentaa mainostajille tarjottavan Facebook pikselin avulla. Sekä palveluiden parissa, että palveluiden ulkopuolella kerätty käyttöön perustuva data vaikuttaa sisältöjen personointiin sosiaalisessa mediassa. Verkostolla viitataan taas henkilöihin, joita käyttäjän verkostoon palvelussa kuuluu eli ketä tämä esimerkiksi seuraa tai kenet on lisännyt tai hyväksynyt kaverikseen. Se, millä perusteella sisältöjä suositellaan, riippuu palvelusta ja sen algoritmeista. Toiset arvottavat esimerkiksi verkostoa

suosittelussaan enemmän, kun taas toiset perustavat suosittelun lähes täysin käyttöön perustuvaan dataan.

Haastateltavat näkevät, että käyttäjät voivat itse vaikuttaa siihen millaista sisältöä he sosiaalisen median kautta saavat. Tärkeintä on pohtia omaa toimintaa alustoilla: algoritmeille tykkäys, kommentti tai sisällön jakaminen, riippumatta siitä missä mielentilassa se on tehty tai mitä käyttäjä on sisällöstä mieltä, on aina indikaattori sisällön kiinnostavuudesta käyttäjälle ja siksi saman tyyppisiä sisältöjä tullaan suosittamaan käyttäjälle myös jatkossa. Jos siis oma Facebook tai Twitter seinä täyttyy esimerkiksi maahanmuuttokriittisistä sisällöistä, joista käyttäjä itse ärsyyntyy, suuttuu tai ahdistuu, syynä on todennäköisesti se, että käyttäjä on aiemmin reagoinut, tai ainakin pysähtynyt lukemaan saman tyyppisiä sisältöjä. Vaikka reaktio olisi ollut sisällön jakaminen pöyristyneenä tai vastaargumentin jättäminen. Ja jos tällaisia sisältöjä ei jatkossa halua saada, paras tapa on lopettaa kyseisen henkilön seuraaminen tai esimerkiksi Facebookissa merkata sisällön jakanut taho käyttäjäksi, jolta ei halua saada sisältöjä jatkossa. Palveluiden asetuksissa pääsee lisäksi tutustumaan siihen, kuinka palvelut ovat käyttäjän ja tämän kiinnostuksen kohteet luokitelleet. Halutessaan näitä tietoja voi myös poistaa. (Hilden, 2017, 15).

Lisäksi haastatteluissa korostui, että käyttäjien tulisi käyttää enemmän aikaa sisältöjen pohtimiseen ja kyseenalaistamiseen. Ensinnäkin tulisi miettiä ”miksi näen tämän” ja toiseksi ”mitä joku eri tavalla ajatteleva henkilö ajattelisi tästä sisällöstä”. Asioiden pohtiminen eri kulmasta käsin auttaa käyttäjää tunnistamaan omat sosiaaliset kuplansa ja niiden vaikutukset omaan ajatteluun. Ja tiedostamalla oman kuplansa, voi tietoisesti hakea myös tietoa muista kuplista, mikä auttaa ymmärtämään laajemmin ympäröivää maailmaa ja ajankohtaisia aiheita. Ongelmana tässä kuitenkin on, että se vaatisi aikaa. Ja yksi sosiaalisen median personoinnin suurimmista hyödyistä on nimenomaan se, että sisältöjä personoimalla käyttäjien ei tarvitse käyttää aikaa tiedon kaivamiseen, kun kutakin käyttäjää kiinnostavat asiat annetaan valmiina.

Tutkimuksen tarkoituksena oli kartoittaa asiantuntijoiden näkemyksiä sosiaalisesta mediasta, algoritmeista ja personoinnista, erityisesti siitä näkökulmasta, kuinka nämä vaikuttavat tiedonkulkuun. Tutkimustulosten perusteella voidaan sanoa, että haastateltavat ovat monilta osin huolissaan sosiaalisen median vaikutuksista, vaikkakin näkevät palveluissa edelleen myös

potentiaalia esimerkiksi demokratian edistämisen kannalta. Lisäksi tutkimustulosten perusteella voidaan todeta, että käyttäjien omalla toiminnalla sosiaalisen median palveluissa on merkittävä vaikutus siihen, millaista sisältöä nämä palveluiden kautta saavat ja sitä kautta siihen, millainen kuva näille maailmasta, yhteiskunnasta ja ajankohtaisista asioista muodostuu.

6 Luotettavuus

Tutkimukseni tarkoitus oli kartoittaa sosiaalisen median ja algoritmien parissa työskentelevien asiantuntijoiden ja tutkijoiden näkemyksiä sosiaalisen median algoritmeista sekä palveluiden käyttäjien mahdollisuuksista vaikuttaa algoritmeihin ja sitä kautta sisältöjen personointiin. Tutkimus toteutettiin haastattelemalla neljää asiantuntijaa ja perustui näiden asiantuntijoiden mielipiteisiin ja näkemyksiin.

Laadullisena haastattelututkimuksena tutkimustuloksia ei voida pitää yleistettävänä, sillä ne edustavat vain pienen joukon näkemyksiä aiheesta. Yleistettävien tutkimustulosten keräämiseksi tulisi aineiston olla huomattavasti laajempi ja edustavampi. Laadulliselle tutkimukselle tyypillisesti tarkoitukseni oli avata tutkimuksen kohteena olevaa aihetta kattavasti, mutta on tärkeää huomioida, ettei tutkimuksessa esiin nousseet välttämättä edusta kaikkien sosiaalisen median ja algoritmien asiantuntijoiden näkemyksiä. Vaikka haastateltavia ei valittu sen perusteella mikä näiden suhtautuminen sosiaalista mediaa kohtaan on, kaikki kuvasivat suhtautumistaan jossain määrin kriittisenä. Näin tuskin kuitenkaan on kaikkien aihepiirin asiantuntijoiden parissa. Koska kyseessä on kuitenkin opinnäytetyö, ei sen laajuuden kannalta aineiston kasvattaminen ollut mielestäni kuitenkaan tarpeellista.

Lisäksi on hyvä muistaa, että vaikka tutkimus toteutettiin haastattelemalla asiantuntijoita, joilla oletettavasti on hyvin kattava ymmärrys sosiaalisesta mediasta ja sen toiminnasta, ei myöskään näillä asiantuntijoilla ole pääsyä algoritmeihin. Kaikki vastaukset haastatteluissa perustuvat siis tutkijoiden näkemykseen siitä, kuinka palvelut *todennäköisesti* toimivat, mutta täysin varmaa tietoa ei myöskään heillä ole.

Olen pyrkinyt kuvaamaan tutkimuksen toteutusta, eli esimerkiksi aineistoa, menetelmää ja haastateltavia mahdollisimman kattavasti, jotta tutkimuksen lukijalla olisi perusteelliset tiedot tulosten luotettavuuden arviointiin ja niiden soveltamiseen. Jos peilataan tutkimuksen tuloksia suhteessa tutkimuskysymyksiin, koen aineiston vastanneen niihin monipuolisesti. Tutkimustulokset mukailivat suurilta osin aikaisempaa tutkimusta aiheesta.

Tutkimuksen luotettavuutta arvioidessa on hyvä huomioida tutkijan ennakoasenteet ja odotukset. Kuvasin luvussa 4 omaa ennakoasennettani tutkimuskohdetta kohtaan kriittisenä ja tämä on hyvä pitää mielessä tutkimuksen tuloksia arvioitaessa. Vaikka pyrin tarkastelemaan tutkimuskohdettani objektiivisesti, ei oman asenteeni vaikutusta voi täysin sulkea pois.

Haastatteluiden toteuttaminen etänä vaikutti haastattelutilanteisiin jonkun verran, mutta koen päässeeni haastateltavien kanssa hyvään keskusteluyhteyteen etähaastatteluista huolimatta. Kaksi haastateltavaa oli minulle ennakkoon tuttuja, minkä näen vaikuttaneen haastatteluihin positiivisesti erityisesti keskusteluyhteyden luomisen kannalta. Toisaalta tämä saattoi vaikuttaa haastatteluissa niin, että jotkin asiat, jotka olisivat voineet olla tutkimuksen kannalta olennaisia jäivät sanomatta, sillä olemme käyneet heidän kanssaan keskustelua aiheesta myös aiemmin. Pyrin haastattelujen aikana esittämään tarkentavia kysymyksiä, mikäli tällaisia tilanteita havaitsin, mutta on kuitenkin mahdollista, että joitain ”itsestäänselvyyksiä” jäi sanomatta.

7 Pohdintaa

Tämän tutkimuksen tarkoituksena oli kartoittaa sosiaalisen median ja algoritmien parissa työskentelevien asiantuntijoiden ja tutkijoiden näkemyksiä sosiaalisen median algoritmeista sekä palveluiden käyttäjien mahdollisuuksista vaikuttaa algoritmeihin ja personointiin. Tutkimuksen lähtökohtana oli sosiaalisen median muuttama tapamme vastaanottaa tietoa maailmasta ja ajankohtaisista aiheista. Haastateltujen asiantuntijoiden näkemykset sosiaalisesta mediasta ja algoritmeista sekä suhtautuminen niitä kohtaan mukaillee vahvasti yleistä asiantuntijoiden suhtautumista: he näkivät palveluissa paljon hyvää ja paljon potentiaalia, mutta olivat huolissaan ja suhtautuivat kriittisesti moniin vallitseviin suuntauksiin ja kehityskulkuihin.

Haastateltavien mukaan palveluiden käyttäjillä on merkittävä vaikutus siihen, millaista sisältöä algoritmit näille personoivat. Haastateltava 4 kertoo hämmästelevänsä toisinaan, kun kuulee ihmisten sanovan, että jokin sosiaalisen median palvelu on täynnä pelkkää vihaa ja negatiivisuutta. Hän itse ei näe sosiaalisessa mediassa tätä puolta, sillä on tietoisesti päättänyt olla reagoimatta sen tyllysiin sisältöihin. Reagoimalla sisältöihin kerromme palveluille haluavamme tätä lisää ja tämä olisi jokaisen käyttäjän hyvä pitää mielessä.

Haastatteluissa nousi esiin erilaisia visioita siitä, mihin suuntaan haastateltavien mielestä palvelut tulevat kehittymään tai mihin niiden pitäisi kehittyä. Haastateltava 4 näkee, että tekstiä itsenäisesti generoivien tekoälyjen seurauksena tulemme jo lähiaikoina siirtymään sosiaalisen median palveluihin, jotka pohjautuvat ”seinien” tai ”feedien” sijaan ryhmiin, joihin mukaan pääsy vaatii ihmiseksi validoitumista. Hän suosittelee sosiaalisen median käyttäjiä pohtimaan jo nyt, mitkä ovat näiden ammatillisten tai muiden tavoitteiden kannalta sellaisia ryhmiä, joihin olisi hyvä kuulua. Sillä muutoksen jälkeen ryhmiin pääsy tulee Haastateltavan 4 näkemyksen mukaan olevan rajatumpaa. Sosiaalisten kuplien muodostumisen kannalta on nähdäkseni pelottava ajatus, että siirtyisimme sosiaalisessa mediassa vielä enemmän suljettuihin ryhmiin kuin missä tällä hetkellä olemme.

Haastateltava 3 näkee yhtenä tärkeimpänä asiana sosiaalisten median tulevaisuuden kannalta sen, että käyttäjillä tulisi olla oikeus vaikuttaa siihen, kuinka personoitua sisältöä he palveluissa näkevät:

”Se oleellisin asia mitä pitäisi tehdä olisi tuoda enemmän vaikutusmahdollisuuksia käyttäjälle nähdä ja vaikuttaa, ikään kuin ymmärtää sitä, että minkä takia mä nään täällä tällaisia asioita kun mä nään ja mahdollistaa se, että voisin myös kytkeä sen personoinnin pois päältä. [...] En mä tiedä, olisiko joku liukukytkin vaikka, jossa voisit säätää sitä kuinka paljon algoritmi lajittelee. Että saatko sä nyt täysin algoritmin tekemää feediä vai täysin sen vaikuttamatonta vai ehkä jotain aivan vastakkaista. Et sä voisit ottaa vaikka semmoisen, en mä tiedä, vastapuoli viikonpäivän niinku joku semmoinen erikoinen perjantai, jolloin sä saatkin jotain ihan muuta sisältöä kuin mitä sulle todellakaan pitäisi näyttää.” (H3)

Mahdollisuutta personoinnin kytkemiseen pois päältä on esitetty yleensä eri mediatalojen verkkosivujen personointiin liittyen (ks. esim. Hilden, 2017, 29), mutta sen tuominen sosiaalisen median palveluihin antaisi käyttäjälle huomattavasti lisää mahdollisuuksia vaikuttaa siihen, kuinka sosiaalisen median kautta saadut sisällöt vaikuttavat tähän.

Kaiken kaikkiaan tärkein havainto tutkimuksesta liittyy mielestäni siihen, että vaikka sosiaalisen median vaikutusta sosiaalisten kuplien muodostumiselle on monissa tutkimuksissa kyseenalaistettu, ei ajatusta niistä kannata kuitenkaan haudata ja hylätä. Erityisesti, koska sosiaalisen median palvelut ja niihin liittyvä teknologia kehittyvät huomattavasti nopeammin kuin tutkimuksia pystytään tekemään, on kaikenlainen sosiaaliseen mediaan liittyvä tutkimus ja niiden vaikutuksiin liittyvä keskustelu ajankohtaista ja tärkeää. Oman tutkimukseni näkökulmasta kiinnostavaa olisi erityisesti tutkia lisää sitä, kuinka hyvin käyttäjät tuntevat tiedonkeruuseen liittyviä asioita, niiden mahdollisia uhkia ja ongelmia ja kuinka paljon käyttäjät hyödyntävät mahdollisuuksiaan vaikuttaa algoritmeihin ja personointiin.

Kuten luvussa 2 kuvasin, tietojen keruuseen liittyvän lainsäädännön mukaan käyttäjillä tulee olla oikeus tietää miten ja miksi heistä kerätään tietoa. Tämä on pakottanut sosiaalisen median palvelut avaamaan toimintaansa käyttäjille ja antamaan käyttäjille mahdollisuuksia vaikuttaa heistä kerättävään ja jo kerättyyn tietoon. Kuten luvussa 5.3 mainitsin, esimerkiksi Facebookissa käyttäjän on mahdollista tutustua siihen, kuinka Facebook on hänet profiloinut ja halutessaan poistaa tiettyjä kiinnostuksen kohteita, joiden perusteella esimerkiksi mainontaa kohdennetaan. Vaikka tämä mahdollisuus tutustumiseen on olemassa, harva sitä kuitenkaan tekee. Ehkä keskeinen ongelma sosiaalisen median algoritmeihin

liittyen ei siis olekaan se, että käyttäjät eivät ymmärrä kuinka palvelut personoivat heille sisältöä, vaan se, että he eivät välitä.

Kaikissa haastatteluissa haastateltavat korostivat sosiaalisen median kasvanutta merkitystä ihmisten arjessa. Palveluista on tullut niin iso osa elämää, että niistä pois lähteminen tuntuu todella vaikealta. Henkilökohtaisesti havaitsin tämän ensimmäisenä opiskeluvuotena yliopistolla, kun ilmoitukset erilaisista opiskelijatapahtumista tuli pääasiassa Facebookin kautta. Samalla vuosikurssilla ollut opiskelija, joka ei käyttänyt Facebookia jäi pois monesta tapahtumasta vain siksi, ettei tiennyt niistä. Sosiaalisen median hyödyt yksittäisille käyttäjille ovat niin suuret, että mahdollisia haittoja on helppo ylenkatsoa. Sen sijaan, että käyttäjien tietoisuutta potentiaalisista ongelmista pyrittäisiin lisäämään, ehkä tärkeämpää olisikin saada käyttäjät välittämään.

Lähteet

Ananny, M. & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>

Atkins, L., Wallace, S. & British Educational Research Association. (2012). *Qualitative Research in Education*. SAGE Publications Ltd.

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, 20:1, 1-13, DOI: 10.1080/1369118X.2016.1216147

Benkler, Y., Faris, R. & Roberts, H. (2018). *Network Propaganda: Manipulation, disinformation, and radicalization in American politics*. New York, NY: Oxford University Press.

Briz, N. (2018). This is Your Digital Fingerprint. *Internet Citizen*. (Viitattu 6.11.2020) <https://blog.mozilla.org/internetcitizen/2018/07/26/this-is-your-digital-fingerprint/>

Bruns, A. (2019). *Are filter bubbles real?* Medford, Ma: Polity Press.

Bucher, T. (2018). *If...Then: Algorithmic Power and Politics*. Oxford University Press. (Viitattu 16.3.2020) <https://www-oxfordscholarship-com.libproxy.helsinki.fi/view/10.1093/oso/9780190493028.001.0001/oso-9780190493028-chapter-1>.

Clement, J. (2020). Number of Facebook users worldwide 2008-2019. *Statista*. (Viitattu 16.3.2020) <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>

Difranzo, D. J. & Gloria-Garcia, K. (2017). Filter bubbles and fake news. *Crossroads* 23(3):32-35 DOI: [10.1145/3055153](https://doi.org/10.1145/3055153)

Dunkle, J. (2016). *Digital Trends. Tell us something we didn't know: Social media is America's favorite news source*. (Viitattu 2.2.2020) <https://www.digitaltrends.com/social-media/social-media-main-news-source-for-american-adults-pew-study/>

Epstein, R. & Robertson, R. E. (2015). The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *PNAS*, 112(33):E4512–E4521.

Eskola, J. & Suoranta, J. (1998). *Johdatus laadulliseen tutkimukseen*. Tampere: Vastapaino.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Euroopan parlamentin ja neuvoston asetus (EU) 2016/679. Annettu 27.4.2016. Luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (yleinen tietosuoja-asetus). (Viitattu 2.11.2020) <http://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX%3A32016R0679>.

Frey, D. (1986). Recent research on selective exposure to information. *Advances in Experimental Social Psychology*, 19, 41-80.

Hatakka, N. (2019). Populism in the hybrid media system: populist radical right online counterpublics interacting with journalism, party politics, and citizen activism. Turku: Turun Yliopisto

Hildebrandt, M. (2016). *The New Imbroglio. Living with Machine Algorithms*. Teoksessa Janssens, L. (2016). *The Art of Ethics in the Information Society*. Mind You. Amsterdam: Amsterdam University Press

Hilden, J. P. (2017). Julkisen palvelun yleisradiotoiminnan personointi. Viestinnän tutkimuskeskus CRC, Helsingin yliopisto
<https://drive.google.com/file/d/0B-v6fV6ewetZQWoycmVUb1dzQkU/view>

Hern, A. (2017) How social media filter bubbles and algorithms influence the election. *The Guardian*. (Viitattu 31.10.2019)
<https://www.theguardian.com/technology/2017/may/22/social-media-election-facebook-filter-bubbles>

Hosanagar, K. (2019). *A human's guide to machine intelligence: How algorithms are shaping our lives and how we can stay in control*. Viking: New York

Kanninen, V. (2014). Yritysten yhteiskuntavastuu normina: normin epävakaus ja normia tukevat mekanismit. TTS-luento. Poliitiikan ja talouden tutkimuksen laitos. Helsingin Yliopisto. (Viitattu 4.11.2020)
<https://blogs.helsinki.fi/hponka/files/2013/12/YritystenYhteiskuntavastuu2014Opiskelijaluento.pdf>

Kerttula, A. (2020). Valko-Venäjällä poliisi ampui ilmaan varoituslaukauksia ja pidätti yli 200 mielenosoittajaa – massaprotestit jatkuvat jo 12:tta viikkoa. Yle.
<https://yle.fi/uutiset/3-11625452>

Koliska, M. & Diakopoulos, N. (2018). *Disclose, Decode and Demystify: An Empirical Guide to Algorithmic Transparency*. Teoksessa Eldridge S. & Franklin, B. (toim.) *The Routledge Handbook of Developments in Digital Journalism Studies*. London: Routledge

Lee, D. (2018). How President Trump took 'fake news' into the mainstream. How President Trump took 'fake news' into the mainstream. BBC. (Viitattu 6.11.2020) <https://www.bbc.com/news/av/world-us-canada-46175024>

Matikainen, J., Ojala, M., Horowitz, M., Jääsaari, J., Aapio, F., Grigor, I., Husu, E., Laine, O. & Ojanperä, H. (2020). Media ja yleisön luottamuksen ulottuvuudet: instituutiot, journalismi ja mediasuhde. Helsinki: Helsingin Sanomain Säätiö.

Media-alan tutkimussäätiö. (2019). 2019 – Verkko pääasiallinen uutislähde jo yli puolella suomalaisista. (Viitattu 3.2.2020)
<https://www.mediaalantutkimussaatio.fi/tutkimukset/reuters/2019-verkko-paaasiallinen- uutislahde-jo-yli-puolella-suomalaisista/>

Merrill, J. B. (2016). Liberal, Moderate or Conservative? See How Facebook Labels You. New York Times. (Viitattu 3.11.2020)
<https://www.nytimes.com/2016/08/24/us/politics/facebook-ads-politics.html>

Montaner, M., Lopez, B. & de la Rosa, J. L. (2003). A Taxonomy of Recommender Agents on the Internet. The Artificial Intelligence Review, 19(4), 285–330. <https://doi.org/10.1023/A:1022850703159>

Munson, S.A. & Resnick, P. (2010). Presenting diverse political opinions: how and how much. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1457–1466. <https://doi.org/10.1145/1753326.1753543>

Obama, B. (2009). Freedom of information memorandum. (Viitattu 16.3.2020)
<https://obamawhitehouse.archives.gov/the-press-office/freedom-information-act>

Pullinen, J. (2019) Mitä meille tapahtui? Helsinki: HS Kirjat.

Roth, C. (2019). Algorithmic Distortion of Informational Landscapes. arXiv preprint arXiv:1907.10401.

Seppänen, J. & Väliverronen, E. (2014). Mediatyhteiskunta. Tampere: Vastapaino.

Skrubbeltrang Mahnke, M. (2019). "Please leave my news feed alone" Exploring user protest against algorithmic personalization. Teoksessa Mortensen, M., Neumayer, C. & Poell, T. (toim.) (2019). Social Media Materialities and Protest. London: Routledge.

Smyth, B. (2008). Personalization-Privacy Tradeoffs in Adaptive Information Access. Teoksessa Uchyigit, G. & Ma, M. Y. (2008). Personalization techniques and recommender systems. Singapore: World Scientific.

Sundman, R. (2019). Sosiaalisen median palvelut tietävät meistä paljon, ja siksi algoritmien ymmärtämisen tulisi olla "kansalaistaito". Yle. (Viitattu 6.11.2020)
<https://yle.fi/uutiset/3-10867375>

Thurman, N. & Schifferes, S. (2012). The Future of Personalisation at News Websites: Lessons from a Longitudinal Study. Journalism Studies, 13(5-6), 775-790. doi: 10.1080/1461670X.1462012.1664341.

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

Turow, J. (2011). The daily you: how the new advertising industry is defining your identity and your worth. New Haven: Yale University Press

Vehkoo, J. (2020). Helsinkiläinen joogaharrastaja kertoo, miten joutui QAnonin pauloihin – salaliittoteoria kerää kannattajia valaistumisen retoriikan avulla. Yle. Viitattu 4.11.2020. <https://yle.fi/aihe/artikkeli/2020/10/25/helsinkilainen-joogaharrastaja-kertoo-miten-joutui-qanonin-pauloihin>

Vehkoo, J. (2019) Valheenpaljastaja: Valeuutiset uppoavat iäkkäisiin verkkokäyttäjiin, jotka ovat oppineet luottamaan painettuun sanaan. Yle. (Viitattu 31.10.2019) <https://yle.fi/aihe/artikkeli/2019/01/27/valheenpaljastaja-valeuutiset-uppoavat-iakkaisiin-verkkokayttajiin-jotka-ovat>

Willson, M. (2016). Algorithms (and the) everyday. Information, Communication & Society, DOI: 10.1080/1369118X.2016.1200645

Winter, M., Maslowska, E. & Vos, A. L. (2021). The effects of trait-based personalization in social media advertising. Computers in Human Behavior, 114, 106525–. <https://doi.org/10.1016/j.chb.2020.106525>

Zadrozny, B. (2019). Fire at 'pizzagate' shop reignites conspiracy theorists who find a home on Facebook. NBC News. (Viitattu 4.11.2020) <https://www.nbcnews.com/tech/social-media/fire-pizzagate-shop-reignites-conspiracy-theorists-who-find-home-facebook-n965956>

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H. & Helberger, N. (2016). Should we worry about filter bubbles?. Internet Policy Review, 5(1). DOI: 10.14763/2016.1.401

Liitteet

Haastattelurunko

1. Taustaa
 - a. Miten sosiaalinen media ja algoritmit liittyvät omaan työhösi?
 - b. Miten suhtaudut algoritmeihin ja sosiaaliseen mediaan?
2. Sosiaalinen media
 - a. Mihin sosiaalisten median palveluiden suosio mielestäsi perustuu?
 - b. Mitä käyttäjät toivovat saavansa sosiaalisesta mediasta/miksi niitä käytetään?
 - c. Mitä termi "personointi" sosiaaliseen mediaan liittyen mielestäsi tarkoittaa?
3. Algoritmit
 - a. Miten kuvailisit sosiaalisen median algoritmeja lyhyesti?
 - b. Millaista tietoa käyttäjistä kerätään?
 - c. Miten algoritmit prosessoivat tietoa käyttäjistä?
4. Eksplisiittisen toiminnan vaikutus algoritmeihin
 - a. Kuinka käyttäjä voi eksplisiittisesti vaikuttaa algoritmien tarjoamaan sisältöön omalla toiminnallaan (esim. sivun seuraaminen, tykkäykset)?
 - b. Kuinka hyvin mielestäsi käyttäjät tuntevat nämä eksplisiittisen toiminnan vaikutukset algoritmeihin?
5. Implisiittisen toiminnan vaikutus algoritmeihin
 - a. Kuinka käyttäjä voi implisiittisesti vaikuttaa algoritmien tarjoamaan sisältöön omalla toiminnallaan (esim. tiettyyn kuvaan pysähtyminen, sivun tarkastelu ilman seuraamista)?
 - b. Kuinka hyvin mielestäsi käyttäjät tuntevat nämä implisiittisen toiminnan vaikutukset algoritmeihin?
6. Millaisia mahdollisuuksia käyttäjien ja algoritmien dialogiin mielestäsi liittyy?
7. Millaisia ongelmia käyttäjien ja algoritmien dialogiin mielestäsi liittyy?
8. Mikä yksi asia sosiaalisen median palveluiden käyttäjien tulisi ymmärtää algoritmeista?