

HANNU SALMI

# Viraalisuus – kulttuurihistoriallinen näkökulma

**Viraalisuus on ollut niin julkisessa keskustelussa kuin kulttuurintutkimuksessakin voimakkaan kiinnostuksen kohteena 1990-luvulta lähtien. Julkisessa keskustelussa sana ”viraali” on alkanut tarkoittaa viruksen tapaista tai viruksen tavoin leviävää. Se voi viitata myös konkreettisesti erittäin nopeasti suosituksi tulevaan viestiin tai ilmiöön, joka kiertää henkilöltä toiselle, erityisesti internetissä. Videoiden lisäksi viraalisia voivat olla nopeasti jaetut Twitter-viestit tai Facebookissa leviävät kuvalliset meemit. Tämän artikkelin tavoitteena on syventää viraalisuuden käsitettä pohtimalla sen vaikutuksia ja mahdollisuuksia historiallisessa tutkimuksessa.**

”Viraali” on itsessään virus, yhä uusille kulttuurin ja elämän aloille levinnyt – ja leviävä – käsitteellinen muodostelma. Nopeudesta kertoo sanan vakiintuminen niin puhekieleen kuin tieteeseenkin. Ilmaisun tausta on mikrobiologiassa ja lääketieteessä eli viruksissa, jotka replikoituvat isäntäsoluissa ja edetessään aiheuttavat sairauksia. Virukset voivat kätkeytyä soluun, siirtyä jakautumalla isäntäsolusta tytärsoluihin ja muuttua myöhemmin virulenteiksi, taudinaiheutuskykyisiksi. Sanakirjamäärittelyn mukaan ”viraali” voi viitata tähän mikrobiologiseen taustaan eli johonkin ”viruksen aiheuttamaan tai virukseen liittyvään”, mutta sanan käyttö on levinnyt paljon laajemmalle<sup>1</sup>.

*Oxford Dictionary of English* otti ”viraalisuuden” tietokantaansa vuonna 2015. Sanakirjan mukaan viraalisuus on ”viraaliseksi tuleminen ominaisuus”. Samalla *Oxford Dictionary of English* kytki sanan internetin läpimurtoon, sillä viraalisuus merkitsi ”kuvan, videon tai tiedon nopeaa ja laajaa leviämistä internetissä käyttäjältä toiselle”.<sup>2</sup>

Viraalisuus on noussut parrasvaloihin yhteiskunnan ja kulttuurin kysymyksissä 1990-luvulla. Sairausten metaforia on kuitenkin käytetty nopean leviämisen kuvauksissa paljon kauemmin, vuosisatojen ajan. Ilmiön nopeaa, vääjäämätöntä leviämistä saatettiin kutsua rutoksi tai syöväksi, ja useimmiten ilmiö haluttiin tällöin nähdä vaarallisena ja uhkaavana. Myös kulttuurin leviäminen saatettiin esittää tuhoisana kuin sairaus. Tätä voi kuvittaa suomalaisella esimerkillä kansallisvaltion rakennustyön käyttövoimasta, fennomaniasta. Kirjailija K. A. Järvi käsitteli suomalaisuuden edistämisen maniaa romaanissaan *Harry* (1910), joka kuvaa suomenruotsalaisen päähenkilön kääntymystä suomenmieliseksi. Romaanissa on kohta, jossa nuorukainen tilittää ruotsin kielen opiskelun

aiheuttamia kärsimyksiä ja ikääntynyt opettaja vastaa rinnastamalla fennomanian vesikauhuun:

– Koko kouluaikani on minua rääkätty sillä kielellä ... veristä pilkkaa ... pilkkaa ... pilk-kaa sen tähden olen saanut silmät korvat täyteen ... öitä olen sen vuoksi valvoa saanut. Rehtori Heideman ... hän itse suomalaisen rahvaan lapsi ... oli pahin pilkkaajani.

– Ja sen vuoksi syntyi maisterissa mania ruotsinkieltä kohtaan eli sanokaamme paremmin rabies.”<sup>3</sup>

Tässä sairausmetaforaa käytetään välillisesti, sillä kirjailija esittää sen fiktiivisen hahmonsa argumenttina fennomaniaa vastaan. Katkelmassa fennomania on paitsi herkästi tarttuvaa myös yksilölle kohtalokasta ja voimakaita reaktioita aiheuttavaa.

Menneisyydessä sairaudet ja niiden metaforinen käyttö toisinaan myös risteytyivät. Esimerkiksi 1800-luvulla kolera herätti paljon huomiota julkisuudessa – jo siksikin, että vuosisadan kuluessa koettiin kuusi globaalia pandemiaa. Kun tauti 1830-luvulla lähestyi Länsi-Eurooppaa idästä, Venäjän ja Puolan kautta, Ranskassa sairautta kuvattiin sokeana jättiläisenä, joka aiheuttaa barbariaa levitessään läntisen sivistyksen pariin<sup>4</sup>. Tässä sairaus näyttäytyi kulttuurisena kamppailuna.

Tämän artikkelin tarkoitus ei ole niinkään jatkaa viraalisuuden metaforista tai representatiivista tulkintaa. Sen sijaan tavoitteena on pohtia, miten ja missä mielessä kulttuurista tartuntaa, tai kulttuurisen välittymisen mekanismeja laajemmin, voitaisiin tarkastella viraalisena prosessina. Onko viraalisuudella 2000-luvun kontekstista irrotettuna laajempaa historiallista relevanssia? Missä mielessä ja millä ehdoilla 1800-luvun mediamaailmaa voitaisiin tarkastella viraalisuuden kannalta? Artikkelissa

## ”Viime kädessä taudin kantajia eivät ole vain tietokoneet ja mobiililaitteet vaan niitä ohjaavat ihmiset, jotka jakavat viestejä eteenpäin.”

pohditaan myös, millaisia metodologisia vaikutuksia tällä olisi. Käsittelen aluksi mediaviruksen käsitettä ja sen historiaa, minkä jälkeen tarkastelen viraalisuuden tutkimuksen keinoja 1800-luvun aineiston, sanomalehdistön, valossa. Lopuksi pohdin, millaisilla muuttujilla viraalisuutta voisi arvioida.

### Mediavirus ja kommunikaation historia

Ajatus viestinnän viraalisuudesta kytkeytyy internetin läpimurron aiheuttamaan kommunikatiiviseen muutokseen. Aihetta käsitteli pioneeriluontoisesti Douglas Rushkoff vuonna 1994 teoksessaan *Media Virus!*, joka lähtee liikkeelle tietoyhteiskunnan muutoksista. Rushkoff aloittaa kirjansa esittelemällä keskivertoamerikkalaiskotia, jossa on yhtä paljon tai oikeastaan enemmän kommunikaatioteknologiaa kuin menneiden vuosikymmenten uutisstudioissa: satelliittiantennit, kaapeliboksit, videokamerat, modeemit ja mikrotietokoneet olivat tehneet kodeista interaktiivisen multimedian keskuksia.<sup>5</sup> Hänen mukaansa kommunikaatioteknologian yleistyminen lisäsi räjähdysmäisesti myös median välittämää liikettä. Siksi mediavirus ei ollut virus vain metaforisesti. Verkon välittämät ”mediatapahtumat eivät ole kuin viruksia; ne ovat viruksia”.<sup>6</sup> Rushkoffin ennustuksessa mediakulttuuria hallitsisivat tulevaisuudessa *online*-infektiot.

Suomessa viraalisuuden historiallisen tutkimuksen pioneereja on vuonna 2007 väitellyt Jussi Parikka, jonka tutkimus julkaistiin samana vuonna nimellä *Digital Contagions. A Media Archaeology of Computer Viruses*. Jo Rushkoff kuvasi virusta Troijan hevosenä: biologisella viruksella on proteiiniukuori, jonka avulla se kiinnittyy

terveeseen soluun, tunkee sinne koodinsa ja ryhtyy leviämään ympäröiviin soluihin. Parikka analysoi 1990-luvulta lähtien rakennettua ”tietoyhteiskuntaa” enemmän digitaalisten tartuntojen kuin tiedon leviämisen foorumina. Tutkimuksen keskiössä olivat Troijan hevoset, madot, virukset ja haittaohjelmat, jotka leviävät tartuntoina ja joiden läsnäoloon tietoverkoissa toimivien on täytynyt sopeutua.<sup>7</sup>

Ajatus digitaalisen kulttuurin viraalisuudesta on Rushkoffista lähtien ollut laajempi kuin vain käsitys tietokoneita saastuttavista tai kaappaavista haittaohjelmista. Tietoverkoissa kiertävät sähköpostit, kuvat, videot tai oikeastaan mikä tahansa ilmiö itsessään voi olla luonteeltaan viraalista, nopeasti tarttuvaa ja leviävää. Tällä hetkellä esimerkiksi nettimeemit voivat levitä minuuteissa globaaleiksi ilmiöiksi sosiaalisessa mediassa. Viime kädessä taudin kantajia eivät ole vain tietokoneet ja mobiililaitteet vaan niitä ohjaavat ihmiset, jotka jakavat viestejä eteenpäin. Infektiota kiihdyttävät algoritmit, jotka näyttävät viestejä otollisille vastaanottajille ja potentiaalisille jakajille. Samoin ohjelmistot osallistuvat tiettyjen ilmiöiden ”tykkäämiseen” automatisoidusti ja siten kasvattavat niiden painoarvoa ja verkkonäkyvyyttä. Kulttuuristen virusten tartunta ei ole lineaarinen prosessi vaan pikemminkin odottamaton, oikukas ja rihmastomainen ketjureaktioiden sarja, jossa inhimillisten ja ei-inhimillisten toimijoiden vaikutukset sekoittuvat erottamattomaksi kokonaisuudeksi.<sup>8</sup>

Viraalisuuden kiihtyminen on ollut osa tietoverkojen rakentumisen historiaa 1990-luvulta lähtien. Vaikka viraalisuus näyttäytyy ei-inhimillisenä tai ainakin ei-yksilöllisenä prosessina, jossa tekstit, kuva ja muut si-

## ”Onko viraalisuus sittenkään vain viimeaikainen ilmiö?”

sällöt leviävät nopeasti, tietoverkkojen rakentuminen on herättänyt myös keskustelun yksilön merkityksestä ja mahdollisuuksista osallistujana, kulttuurin luojana. Henry Jenkins tuli vuonna 2006 tunnetuksi kirjoillaan *Convergence Culture. Where Old and New Media Collide* ja *Fans, Bloggers and Gamers. Exploring Participatory Culture*. Teoksissa korostettiin, että globaalit tietoverkot ja informaatioteknologia mahdollistavat yleisöille mediasisältöjen kommentoinnin, luokittelun ja arkistoinnin, niin henkilökohtaisesti kuin ryhmissäkin, sekä sisältöjen luovan hyväksikäytön ja kierrätyksen.<sup>9</sup> Tämän osallistuvan kulttuurin (*participatory culture*) ilmentymiä ovat sosiaalisen median monet muodot, kollektiivinen luovuus meemien synnyttämisessä ja yksilön osallisuus kulttuuristen merkkien ja symbolien muovaamisessa. Voi väittää, että nämä käytännöt ovat samalla viraalisen kulttuurin kasvualusta, sen ontologinen perusta. Osallistuva kulttuuri ilmenee haluna paitsi tuottaa itse mediasisältöjä myös jakaa muiden käyttäjien jakamia sisältöjä ja ilmaista niiden herättämiä tunteita. Markkinoinnin tutkimuksessa ja kehitystyössä on jo pitkään keskusteltu keinoista, joilla voitaisiin edistää tuotteen viraalistumista, käyttäjien lähes huomaamatonta hyödyntämistä uusien käyttäjien löytämisessä.<sup>10</sup> Viraalisuus perustuu käyttäjien ja verkoston väliseen dynamiikkaan, johon samanaikaisesti syötetään sisältöjä ja josta nousee esiin odottamattomia aiheita ja ilmiöitä.

### Menneet viraalisuudet

Henry Jenkinsin jälkeen osallistuvan kulttuurin ajatusta on kehitetty useissa teoksissa, kuten Aaron Delwichen ja

Jennifer Jacobs Hendersonin toimittamassa kokoelmassa *The Participatory Cultures Handbook* vuodelta 2012. Teoksen johdannossa toimittajat jakavat osallistuvan kulttuurin historiallisesti neljään vaiheeseen, joista ensimmäinen sijoittuu vuosiin 1985–1993, globaalin kommunikaatiomaisen muutoksen mikrotietokoneiden yleistymisen ja internetin alkuaskeleiden seurauksena.<sup>11</sup> Ajankasjo vastaa Douglas Rushkoffin näkemystä viraalisuuden alkulähteistä, jotka ulottuvat tietoverkkojen rakentamisen aikaan 1980- ja 1990-luvuille. Parikka on korostanut lisäksi kulttuurin medikalisoitumista 1980-luvulla: kulttuurisessa keskustelussa esiintyivät yhä useammin virukset ja parasiitit, ja HIV-epidemian vanavedessä tietokonevirusiakin alettiin kutsua tietokone-AIDSiksi.<sup>12</sup>

Epäilemättä 1900-luvun viimeiset vuosikymmenet ja sitä seurannut vuosituhanen vaihe olivat voimakkaan mediamuutoksen aikaa. Siitä huolimatta on syytä kysyä, onko viraalisuus sittenkään vain viimeaikainen ilmiö, sillä sen mahdollisuusehto, viestien nopea leviäminen, toteutui jo aiemmin sanomalehdistön laajenemisen yhteydessä. Rotaatiopainojen yleistyminen teki 1800-luvulla painotyöstä huokeaa ja merkittävästi aiempaa nopeampaa. Lennätinlinjat, niin maanpäälliset kuin valtameriä alittavat, tekivät viestinnästä paitsi vilkasta myös maanosien rajat ylittävää.<sup>13</sup> Jos 2000-luvulla digitaalisen median sisällöt kiertyvät yhteen printtimedian kanssa, samaan tapaan 1800-luvulla kommunikaatio oli transmediaalista: sanomalehtien ohella tekstien kierrätykseen osallistuivat arkkiveisut ja kirjeet, sähkösanomat ja pienpainatteet, aikakauslehdet ja kirjat.

Jo ennen lennätinkaapeleiden aikaa uutiset levisivät laajalle. Esimerkiksi uutinen Turun palosta kopioitui

satoihin ellei tuhansiin sanomalehtiin Ruotsissa, Tanskassa, Venäjällä, Saksassa, Iso-Britanniassa, Ranskassa... Kun onnettomuus tapahtui 4. syyskuuta 1827 Turussa, uutinen kerrottiin vielä 15. maaliskuuta 1828 Hobartin kaupungissa Tasmaniassa, minne se oli päätyneet purje-laivojen mukana.<sup>14</sup> Viralisuudesta ei tässä tapauksessa voi vielä puhua, sillä uutinen, joka 2000-luvulla olisi levinnyt muutamassa minuutissa tai tunnissa, tarvitsi yli kuusi kuukautta matkustusaikaa. Onnettomuus uutinen levisi huomattavasti rivakammin jo 1800-luvun lopulla mannertenvälisten lennätinyhteyksien ansiosta. Kun Krakataun tulivuori alkoi purkautua Rakatan saarella Indonesiassa 26. elokuuta 1883, *Hufvudstadsbladet*, *Nya Pressen* ja *Suomen Wirallinen Lehti* kertoivat tiedon jo 28. elokuuta. Kaikki kolme lehteä julkaisivat samanaikaisesti Lontoossa edellisenä päivänä ilmestyneen sähkösanoman.<sup>15</sup> Vuosien 1827 ja 1883 välillä sanomalehtien ja lennätinjärjestelmän muodostama kommunikaatioverkosto oli nopeuttanut tiedonkulkua ennennäkemättömällä tavalla. Jo 1800-luvun puolivälissä sellaiset taiteilijat kuin Niccolò Paganini ja Franz Liszt tunnettiin, ainakin nimeltä, kaikissa maanosissa, ja heidän konserteistaan ja elämänvaiheistaan kerrottiin niin meksikolaisissa kuin australialaisissa sanomalehdissä, alueilla, joissa he eivät koskaan varsinaisesti esiintyneet.<sup>16</sup> Heidän maineensa rajat eivät olleet maantieteelliset vaan sosiaaliset.

Sanomalehtien 1800-luvulla muodostamaa verkostoa on jo tutkittu viralisuuden näkökulmasta. Ryan Cordell ja David A. Smith käynnistivät vuonna 2013 Northeastern Universityssä Bostonissa hankkeen *Viral Texts*.

Projektissa on tutkittu yhdysvaltalaisessa sanomalehtiaineistossa tapahtunutta tekstien kopiointia ja – mikäli oletetaan, että tekstien siirtymät eivät aina ole intentionaalisia – kopioitumista. Hanketta varten Smith kehitti algoritmin, joka etsii digitoitujen lehtien korpuksesta toistuvia tekstejä tai tekstifragmentteja. Smithin rakentama ohjelma nimeltä *passim* pohjautuu n-grammien etsimiseen tekstijoukosta<sup>17</sup>. Menetelmä mahdollistaa sellaisten yhteyksien löytymisen, joista ei ole etukäteistietoa tai -oletusta<sup>18</sup>.

Suomessa olemme tutkineet tekstien uudelleenkäyttöä vuonna 2016 käynnistyneessä COMHIS-hankkeessa. *Computational History and the Transformation of Public Discourse in Finland, 1640–1910* on osa Suomen Akatemian rahoittamaa digitaalisten ihmistieteiden tutkimusohjelmaa. Olemme analysoineet vuosien 1771–1910 suomalaisten sanomalehtien ja aikakauslehtien sisällä tapahtunutta kopiointia.<sup>19</sup> Samaan tapaan kuin *Viral Texts*-hankkeessa lähtökohtana ovat skannatut sanomalehdet, joiden teksti on optisesti tunnistettu. Teksteissä on runsaasti tunnistusvirheitä, mikä osin johtuu Suomessa laajasti käytetystä fraktuurafontista mutta myös paperinlaadusta, käytetystä fonttikoosta, skannauksen laadusta ja monista muista tekijöistä. Tämän vuoksi hankkeessa ei sovellettu *passim*-ohjelmaa vaan uutta, Filip Ginterin johdolla kehitettyä menetelmää, joka nojautuu bioinformatiikassa kehitettyyn ja käytettyyn BLAST-työkaluun<sup>20</sup>. Basic Local Alignment Search Tool on algoritmi, jota käytetään erityisesti kohinaa sisältävän aineiston kuten aminohappojen tai DNA-sekvenssien vertailuun ja sa-

## Digitaalinen historiantutkimus ja julkisuuden muutos Suomessa 1640–1910

Suomen Akatemia käynnisti vuonna 2016 digitaalisten ihmistieteiden tutkimusohjelman. Sen hankkeisiin lukeutuu neljän partnerin muodostama konsortio *Computational History and the Transformation of Public Discourse in Finland, 1640–1910* (COMHIS). Projektissa ovat mukana Kansalliskirjaston Digitointikeskus (PI Kimmo Kettunen), Helsingin yliopiston humanistinen tiedekunta (PI Mikko Tolonen), Turun yliopiston tulevaisuuden teknologian laitos (PI Tapio Salakoski) ja Turun yliopiston kulttuurihistoria (konsortion PI Hannu Salmi).

Hanke tutkii ja arvioi uudelleen suomalaisen julkisen keskustelun laajuutta, luonnetta ja yllirajaisia kytkentöjä vuosina 1640–1910. Hanke yhdistää kaksi

toisiaan täydentävää lähestymistapaa ja pohjautuu yhtäältä kirjastojen metadatan analyysiin, toisaalta digitoitujen suomalaisten sanoma- ja aikakauslehtien tekstinlouhintaan. Konsortio analysoi kirja- ja julkaisuotannon muutoksia, kielirajan merkitystä sekä tekstien kierrätystä suomalaisessa lehdistössä.

COMHIS-hankkeen osana Turun yliopiston tutkijat ovat toteuttaneet tietokannan *Text Reuse in Finnish Newspapers and Journals, 1771–1910*. Se sisältää suomalaisten sanoma- ja aikakauslehdissä kopioituja ja toistettuja tekstejä alkaen lehdistön synnystä, vuodesta 1771. Tietokanta on avoimesti kaikkien kiinnostuneiden käytettävissä, ja siitä voi tehdä hakuja osoitteessa: [comhis.fi/clusters](http://comhis.fi/clusters)

manlaisuuksien tunnistamiseen. Menetelmä osoittautui erittäin tehokkaaksi: kun minimilajajuudeksi määriteltiin 300 merkkiä, suomalaisten sanoma- ja aikakauslehtien korpuksista löytyi lopulta 49 miljoonaa samuudeksi luokiteltavaa osumaa<sup>21</sup>. Luku sisältää kaikki ne osumat, joissa sama teksti tai tekstifragmentti esiintyy vähintään kaksi kertaa. Aineisto koostuu siis vaihtelevan pituisista toistoketjuista: joitakin tekstejä on toistettu vain kerran tai kaksi, joitakin tekstejä kymmeniä tai satoja kertoja. Seuraavassa vaiheessa löydetty aineisto ryhmiteltiin toistoketjuiksi eli klusteroitiin. Näin saatiin lopulta kahdeksan miljoonaa klusteria, vähintään kaksi kertaa esiintynyttä yli 300-merkkistä tekstiä.

Vaikka absoluuttiset määrät ovat epätarkkoja ja epäluotettavia, ne ovat kuitenkin suuntaa-antavia. Suomalainen lehdistö sisältää niin paljon kopioitua materiaalia, että sen volyymi on lähes musertava. On syytä kysyä, missä määrin historiakäsityksemme ylipäätään perustuu ajatukselle alkuperäisestä, ainutkertaisesta tai ”ensimmäisestä”. Väitän, että tietyssä historian hetkessä kulttuuri sisältää ”pimeää ainetta”, toisteisuutta, joka tutkijoiden on ollut helppo sivuttaa ei-originaalina. Jos toiston massa on suuri, aineistossa on mitä todennäköisimmin myös viraalisuutta. Samalla toiston massa viittaa siihen, että lehdistö on ”painovoimallaan” amplifioinut ajankoh- taisia ilmiötä ja lisännyt niiden kulttuurista painoarvoa. Voisiko viraalisuus olla keino ymmärtää menneisyyttä paitsi toiston ja kierrätyksen myös kopioinnin kautta tapahtuneen ilmiöiden voimistumisen näkökulmasta? Suomalaisessa sanomalehdistössä osa toistoketjuista oli niin lyhyitä, esimerkiksi vain kaksi tai kolme kopiota, että niiden leviämistä voi tuskin pitää viraalisuutena, mutta jo 1800-luvun Suomessa joitakin tekstejä kopioitiin nopealla rytmillä kymmeniä kertoja.

## Viraalinen dynamiikka

Sanomalehdistö kasvoi erittäin voimakkaasti 1800-luvulla: media eli räjähdysten tai jatkuvan synnytyksen tilassa vuosikymmenien ajan. Lehdistön muutosta voi verrata internetin läpimurtoon 1990-luvulla. Internetiä ei teknologiana ikinä suunniteltu sellaiseksi kuin se lopulta muodostui. Ei-keskitetty, rihmastomainen teknologia eteni odottamattomasti verkon rakentajien, sisältöjen tuottajien ja niiden käyttäjien yhteisen toiminnan tuloksena. Sama pätee 1800-luvun lehdistöön, joka leviittyi rihmastomaisesti ensin Euroopassa ja Yhdysvalloissa, lopulta kaikissa maanosissa.

Suomessa sanomalehdistön synty sijoittuu vuoteen 1771, jolloin *Tidningar Urigfne Af et Sällskap i Åbo* aloitti ilmestymisensä. Hyvin pitkään lehtiä julkaistiin vain muutamalla paikkakunnalla, Turun lisäksi muun muassa Helsingissä, Viipurissa ja Oulussa. Kiihkeä kasvun vaihe alkoi 1860-luvulla, kun se Keski-Euroopassa sijoittui vuosikymmeniä aikaisemmaksi: 1820-, 1830- ja 1840-luvuille. Tuona aikana ei ollut sellaista tekijänoikeuslakia, joka olisi estänyt tai hillinnyt uutisten kopiointia, vaan lehdet saattoivat vapaasti ottaa aineistoa toisiltaan. Uu-

delleenkäyttöön vaikuttivat joko hillitsevästi tai edistävästi monet tekijät kuten kielirajat, lehtien fyysinen kuljettaminen alueilta toisille, paikalliset sensuuriolosuhteet sekä lehtien jakaantuminen laajalevikkisempiin ja paikallisempiin lehtiin.

Lehtien välillä oli erittäin paljon sisällöllistä liikettä, mutta välttämättä tätä kaikkea ei voi pitää viraalisuutena. *Viral Texts* -hankkeessa rajattiin pois uudelleenjulkaisut samassa lehdessä, sillä mainokset ja ilmoitukset olivat raaka-ainetta, jota lehdet kierrättivät paljon, numerosta toiseen. COMHIS-hankkeessa olemme tutkineet kaikkea toisteisuutta sillä ajatuksella, että emme voi lähtökohtaisesti tietää, onko kopioitu aineisto journalistista sisältöä vai muuta aineistoa kuten mainoksia, viranomaistiedotteita, hinnastoja tai säätietoja, koska sanomalehti kulttuurin tuotteena etsi olomuotoaan 1800-luvulla. Toistot samassa lehdessä tuovat näkyviin myös sellaisen tekstuaalisen infektioisuuden, joka ei ilmene äkillisenä kopiointina vaan vasta pidemmän ajan kuluessa. Tekstejä painettiin uudelleen vuosikymmenien jälkeen: ne saattoivat olla kopioita aiemmista lehdistä, mutta ne tulivat myös kirjallisuudesta tai muusta aineistosta.<sup>22</sup> Jos nykypäivän näkökulmasta viraalisuuden edellytys on kopioinnin ja uudelleenkäytön nopeus, tämä aineisto ei täytä viraalisuuden kriteeriä mutta on silti merkittävää kulttuurista infektiivisyyttä.

Digitoitujen lehtien arkistot ovat usein kansallisia kokoelmia, mutta 1800-luvulla kommunikaatio oli ylijärästä. Algoritmisen menetelmän voidaan tunnistaa samuutta vain korpuksen sisällä. Sekä *passim*-ohjelma että kehittämämme BLAST-menetelmä tunnistavat samuutta kielestä riippumatta, mutta ne eivät löydä kielirajan yli siirtyneitä, käännettyjä tekstejä. Toki aineisto kertoo tarkemmassa analyysissä, että kopiointia tehtiin runsaasti ulkomaisista lähteistä. Mutta algoritmisesti kootun aineiston ulkopuolelle jäävät sellaiset yksittäiset tekstit, jotka ovat kansainvälisen kopiointiketjun osia mutta jotka eivät Suomessa ole päätyneet jatkokopioinnin kohteiksi.

Viraalisuutta on 1990- ja 2000-luvulla lähestytty, jo määritelmällisesti, ennen kaikkea nopeuden ja volyymin mittapuilla. Jos ilmiöstä tai tapahtumasta tulee viraalinen, se synnyttää hyvin nopean reaktioiden ketjun. Replikoinnin volyymi synnyttää myös informaatiomassaa, joka ei lopulta niinkään *leviä* verkossa vaan generoituu – tai kemiallista vertausta käyttäen saostuu – eri palvelimille. Viraalisuuden määrittelyissä viitataan usein siihen, miten ilmiöt leviävät maanosasta toiseen, globaalisti. Tämän perusteella viraalisuutta voitaisiin lähestyä kolmella ekstensiivisellä parametrilla: nopeus, volyymi ja levinneisyys.

Jos 1800-luvun sanomalehdistöä tulkitaan viraalisuuden näkökulmasta, nämä kolme muuttujaa ovat keskeisiä. Uutisten nopeasta leviämisestä kertovat esimerkiksi Suomalaisen Kirjallisuuden Seuran liikkeelle laittaman kirjeen reitit elokuussa 1864. ”Suomenmaan yleisölle” osoitetun kirjeen tavoite oli kertoa Henrik Gabriel Porthanin kunniaksi pystytettävän patsaan julkistamisesta. Kirje ilmestyi ensimmäisenä Helsingissä 5.

## ”Valtioiden väliset rajat, konfliktit, suhdanteet, sääilmiöt, henkilökohtaiset suhteet ja monet muut tekijät muovasivat kytkettyvyyden mahdollisuutta.”

elokuuta 1864 samanaikaisesti neljässä sanomalehdessä, jotka olivat *Suometar*, *Finlands Allmänna Tidning*, *Helsingfors Dagblad* ja *Helsingfors Tidningar*. Ruotsinkielisten lehtien julkaisut olivat poikkeuksellisesti kaksikielisiä. Tämän jälkeen uutisen julkaisi eri paikkakunnilla kuusi ruotsinkielistä ja kahdeksan suomenkielistä lehteä. Uutisointi kesti yhteensä 15 päivää, mutta valtaosa teksteistä ilmestyi viiden päivän kuluessa.<sup>23</sup> Nopeuden lisäksi Porthan-uutisella oli volyyymiä, sillä tekstin julkaisseita lehtiä kertyi lopulta 18. Suomessa julkaistiin vuonna 1864 yhteensä 22 sanomalehteä, mutta näistä elokuun aikana ilmestyivät vain nuo 18. Uutinen tavoitti siis kaikki sillä hetkellä aktiivisesti julkaistut lehdet.<sup>24</sup> Volyyymi on toki suhteellista sanomalehdistön myöhempään kehitykseen verrattuna, mutta ajankohdan kapasiteetilla uutinen sai maksimaalisen näkyvyyden. Jos tarkastelee uutisketjun levinneisyyttä, maantieteellinen kattavuus oli laaja: uutinen eteni Helsingistä Turkuun, Viipuriin, Mikeliin, Porvooseen, Ouluun, Hämeenlinnaan, Kuopioon ja Vaasaan. Levinneisyys oli vuoden 1864 kontekstissa poikkeuksellinen, sillä se tavoitti kaikki ne paikkakunnat, joissa lehtiä julkaistiin.

Nopeuden, volyymin ja levinneisyyden näkökulmasta suomalainen lehdistö sai tulevina vuosikymmeninä mitasuhteet, joissa viraalisten tekstien mahdollisuus kasvoi. Kun Suomi itsenäistyi vuonna 1917, maassa julkaistiin jo 142 sanomalehteä. Kapasiteetti oli yli kuusinkertais- tunut vuodesta 1864.

Vuonna 2011 singaporelainen tutkimusryhmä koetti mallintaa Twitterin viraalisuutta matemaattisesti. Nopeus ja volyyymi olivat analyysin ytimessä, sillä viraalisuuden kriteerinä oli kyky uudelleentwiitata mahdollisimman

paljon, mahdollisimman nopeasti.<sup>25</sup> Twitterissä kyky uudelleentwiitata liittyy tietoverkkojen lähes rajattomaan tarjontaan twiittejä, joista voi potentiaalisesti tulla jakamisen kohteita ja ennen pitkää myös viraalisia ilmiöitä. Historiallisessa kontekstissa tämä tekijä on vaikeammin artikuloitavissa. Ryan Cordell on käsitellyt sanomalehdistöä viraalisuuden näkökulmasta ja korostanut 1800-luvun mediajulkisuuden spesifisyyttä. Sanomalehdillä oli Yhdysvalloissa oma vaihtojärjestelmänsä: lehdet lähetivät uusimmat numerot toisilleen tai tilasivat toistensa julkaisuja. Tämä verkosto muodosti pohjan uudelleenjulkaisemiselle.<sup>26</sup> Koska tekijänoikeuden suoja ei ollut, sisällön jakaminen oli mahdollista, mutta käytössä oleva aineisto asetti sille rajoituksensa, mikä poikkeaa merkittävästi nykypäivän Twitter-julkisuudesta.

Viraalisuutta käsittelevässä teoreettisessa kirjallisuudessa nostetaan usein esiin ajatus kytkettyvyydestä tai kytkettyvyydestä (*connectivity*)<sup>27</sup>. Kun 1800-luvun mediaa tarkastellaan viraalisuuden kannalta, pitäisi pystyä rakentamaan kokonaiskuva niistä kytkentäkaavioista, joilla lehdet materiaalisesti linkittyivät toisiinsa. Tässä oli epäilemättä alueellisia eroja. Samoin valtioiden väliset rajat, konfliktit, suhdanteet, sääilmiöt, henkilökohtaiset suhteet ja monet muut tekijät muovasivat kytkettyvyyden mahdollisuutta.

Singaporelaisten tutkijoiden Twitter-analyysi nostaa viraalisuuden dynamiikan tutkimukseen myös kysymyksen viraalisuuden topiikasta. Projekti pyrki tutkimaan, millaiset aiheet tulivat todennäköisimmin runsaasti jaetuiksi.<sup>28</sup> Saman kysymyksen voi esittää historiallisessa mielessä ja pohtia niitä aihepiirejä, jotka tulivat aikanaan runsaasti kopioituiksi. *Viral Texts*

-projektissa havaittiin, että vuosien 1830–1860 yhdysvaltalaisissa sanomalehdissä viraalisiin teksti oli presidentti James Buchanin virkaanastujaispuhe vuodelta 1857 (150 uudelleenjulkaisua), seuraavaksi viraalisiin oli Washingtonin jäähyväispuhe (53 uudelleenjulkaisua). Suosituimpien aiheiden luettelo sisälsi presidenttien puheiden lisäksi muun muassa reseptin, novellin ja rauhansopimuksen.<sup>29</sup> Ryan Cordell on Karine Nahonin ja Jeff Hemsleyn ajatuksiin viitaten pohtinut myös viraalisten tapahtumien (*viral events*) merkitystä. Nykypäivän mediakulttuurissa on mielekästä yksittäisen tekstin tai videon leviämisseuranta sijasta pohtia, miten viraalisten tapahtumien ympärille rakentuu kokonainen ekosysteemi: esimerkiksi samaa tekstiä voidaan jakaa eteenpäin, mutta samalla jaetaan siihen liittyviä kommentteja, vastauksia, muunnelmia ja parodioita.<sup>30</sup> Tällaisia tapahtumia voidaan pitää viraalisuuden topiikan osana: tapahtuma on teema, jonka ympärille muodostuu viraalisten tekstien tai vaikkapa kuvien moneus. Viraaliset ketjut muodostavat laajemman rihmaston, jota on mielekästä tutkia kokonaisuutena.

Myös 1800-luvulla voidaan nähdä viraalisia tapahtumia, jotka saivat lukemattomia inter- ja transmediaalisia ilmenemismuotoja. Franz Lisztin kaltainen musiikkielämän tähti oli oikeastaan viraalinen tapahtuma, johon kytkeytyi sanomalehti uutisia, novelleja, piirroksia ja pilakuvia. Toinen esimerkki voisivat olla kolerapandemiat, jotka 1800-luvulla vyöryivät yli Euroopan. Sairaus sai yksittäisiä tekstejä laajemman viraalisen luonteen mediakulttuurissa. COMHIS-hankkeen koossa tietokannassa on yli 10 000 koleraan liittyvää toistoklusteria.

Olen edellä käsitellyt viraalisuutta nopeuden, volyymin, levinneisyyden ja topiikan kannalta. Nämä näkökulmat korostavat kuitenkin lähinnä viraalisuuden ekstensiivisiä ulottuvuuksia. Viraalisuutta voidaan tarkastella myös intensiivisten voimien näkökulmasta.<sup>31</sup> Tony D. Sampson on kirjassaan *Virality. Contagion Theory in the Age of Networks* jakanut viraalisuuden kahteen, toisistaan poikkeavaan kategoriaan: molaariseen ja molekulaariseen viraalisuuteen.<sup>32</sup> Molaarisuus ilmenee pyrkimyksenä rakentaa kokonaisuuksia, tuoda singulaarisia ilmiöitä yhteen. Se voi luoda affektiivisia siteitä niin pelon ja kauhun kuin ilon, toivon ja empatiankin avulla. Molekulaarisuus taas on spontaania, äkillistä tai satunnaista. Se rikkoo sosiaalisia koosteita, joita esimerkiksi yhteiset käsitykset ja uskomukset ovat rakentaneet. Ajatus on erityisen relevantti 1800-luvun Euroopassa, jossa voidaan nähdä molaarisia voimia: pyrkimyksiä korostaa perheen, suvun ja kansallisvaltion kaltaisia kokonaisuuksia. Samaan aikaan yhteiskunnallisia ja kulttuurisia rajoja kuitenkin pirstoivat uusien liikkumisen teknologioiden, modernin rahatalouden ja tiedotusvälineiden ohella äkilliset katastrofit, pandemiat ja poliittiset konfliktit. Topiikka itsessään ei ratkaise viraalisuuden luonnetta: esimerkiksi kolerauutisten nopea leviäminen saattoi olla molaarista viraalisuutta, joka pelon ja kauhun tunteiden välityksellä rakensi affektiivisesti yhteisöllisyyttä. Se

saattoi silti olla myös molekulaarista viraalisuutta, joka hajotti yhteisöjä ja loi epävarmuutta.

## Kohti kulttuurista seismografiaa

Olen käsitellyt viraalisuutta historiallisena ilmiönä ammentamalla 1990-luvun ja 2000-luvun teoreettisista lähtökohdista ja nojautumalla 1800-luvun sanomalehtiaineiston empiiriseen tutkimukseen. Viraalisen dynamiikan parametrit voivat olla antoisia ja hedelmällinen historiallinen näkökulma, mutta samalla tutkimuksessa avautuu laajempi perspektiivi siihen, miten medioiden väliset suhteet ovat muodostuneet, miten kulttuuriset ilmiöt ovat välittyneet ja miten media on osallistunut menneisyyden aatteiden, tunteiden ja käsitteiden siirtämiseen ja vahvistamiseen. 1800-luku oli monella tapaa modernin maailman syntysija: tietotuotannon volyymi kasvoi voimakkaasti vuosisadan loppua kohti, tieto liikkui nopeammin ja samalla mediakapasiteetti moninkertaistui. Suomi on tästä erinomainen esimerkki, sillä lehdistön volyymi nousi lähes eksponentiaalisesti.

Viraalisuus rakensi ja rakentaa yhä edelleen kulttuurisia tihentymiä. Ilmiöt voimistuvat odottamattomasti ja purkautuvat yhtä nopeasti. Viraalisuus ei ole hyvä eikä paha. Se voi olla molaarista tai molekulaarista, kokoavaa tai hajottavaa. Dan Sperber halusi vuonna 1985 korostaa kulttuurin leviämisen mekanismeja ja ehdotti perustettavaksi tutkimusalaa, jota hän kutsui ”representaatioiden epidemiologiaksi”<sup>33</sup>. Ajatus epidemiologiasta voisi sinänsä sopia viraalisuuteen, mutta nämä virukset eivät ole metaforia tai representaatiota vaan menneisyyden, nykyisyyden ja tulevaisuuden maailmassa vaikuttavia voimia, inhimillisten ja ei-inhimillisten tekijöiden yhteenkietoutumia. Viraalisuuden kulttuurista tutkimusta voisi pikemminkin kutsua kulttuuriseksi seismografiaksi, joka tutkii kommunikaation liikevoimia. Paikallisesti nämä värähdykset voivat liikuttaa raskaita taakkoja, mutta ne voivat olla myös seismisiä liikkeitä, jotka hiipuvat mitä kauemmas alkulähteestään ne päätyvät.

## Viitteet

- 1 ”Viraaliti” merkityksestä, ks. *Suomisanakirja*: [www.suomisanakirja.fi/viraali](http://www.suomisanakirja.fi/viraali), haettu 10.11.2017.
- 2 Ks. lähemmin *Oxford Dictionary of English*. 3. verkkop. Oxford University Press, Oxford 2015. DOI: 10.1093/acref/9780199571123.001.0001
- 3 K. A. Järvi, *Harry. Romaani kansallisen herätyksen ajoilta*. Sanomalehti O.-Y. Keski-Suomi, Jyväskylä 1910. Verkossa: [www.gutenberg.org/ebooks/48544](http://www.gutenberg.org/ebooks/48544)
- 4 Ranskan kansalliskirjaston kolera-aiheiset kuvat, ks. esim. ”La barbarie et le choléra morbus entrant en Europe” (1831). Verkossa: [gallica.bnf.fr/ark:/12148/btv1b53006141r](http://gallica.bnf.fr/ark:/12148/btv1b53006141r), haettu 10.11.2017.
- 5 Rushkoff 1994, 3.
- 6 Sama, 9.
- 7 Parikka 2007.
- 8 Parikka 2007, 1–11; Lister, Dovey, Giddings, Grant & Kelly 2003, 200–202.
- 9 Jenkins 2006a, 3; Jenkins 2006b, 1.

- 10 Lescovec, Adamic & Huberman 2007, 1–39.
- 11 Delwiche & Jacobs Henderson 2012, 4.
- 12 Parikka 2007, 8.
- 13 Osterhammel 2014, 31–37.
- 14 *The Hobart Town Courier* 15.3.1828. Turun palon mediakäsittelystä, ks. Salmi 2017, 121–138.
- 15 *Hufvudstadsbladet* 28.8.1883; *Nya Pressen* 28.8.1883; *Suomen Wirallinen Lehti* 28.8.1883.
- 16 Salmi 2016b, 135–153.
- 17 Kielitieteessä n-grammi viittaa merkkien, morfeemien tai sanojen jakssoon, jossa alkioita on n kappaletta. Jos jakso koostuu merkeistä, kolmen merkin jakso on trigrammi (n=3). N-grammeja voidaan käyttää esimerkiksi tekstin kielen koneelliseen tunnistamiseen laskemalla, kuinka usein n-grammit esiintyvät. N-grammeja voidaan käyttää myös merkkijonojen samanlaisuuden tunnistamiseen.
- 18 Cordell 2015a, 417–445. *Passim*-ohjelma on ladattavissa Githubista osoitteesta: [github.com/dasmiq/passim](https://github.com/dasmiq/passim).
- 19 Kun hanke käynnistyi vuonna 2016, se sai tutkimusaineistokseen kaikki Kansalliskirjaston digitoimat sanoma- ja aikakauslehdet vuosilta 1771–1910. Aineisto ulottuu ensimmäisen suomalaisen sanomalehden *Tidningar Utgifne Af et Sällskap i Åbo* ilmestymisvuoteen 1771. Projektin alkaessa viimeinen tutkijoiden käytettävissä oleva vuosi oli 1910, mutta tätä kirjoittaessa projekti on saamassa käyttöön myös vuosien 1911–1920 aineiston.
- 20 BLASTin toiminnasta, ks. Vesanto, Nivala, Salakoski, Salmi & Ginter 2017, 330–333.
- 21 Arvioimme, että 300 merkkiä pienemmissä samanlaisuuksissa optisen tekstintunnistuksen ongelmat korostuivat. Jos merkkien tunnistuksessa on virheitä, lyhyiden tekstien samuuden tunnistaminen on epäluottavampaa. Samalla myös korostuivat lehdistön vakioilmoitukset, otsikoinnit ja muu säännöllinen materiaali.
- 22 Pitkällä aikavälillä julkaistuista teksteistä voi mainita esimerkkinä Henrik Gabriel Porthanin kuoleman muistoksi vuonna 1804 julkaistun runon, joka painettiin myöhemmin useita kertoja uudelleen. Ks. lähemmin Salmi 2016a, 93–94.
- 23 *Suometar* 5.8.1864, *Finlands Allmänna Tidning* 5.8.1864 (kaksikielinen), *Helsingfors Tidningar* 5.8.1864, 6.8.1864 (kaksikielinen), *Helsingfors Dagblad* 5.8.1864 (kaksikielinen), *Päivätär* 6.8.1864, *Suomen Julkisia Sanomia* 8.8.1864, *Åbo Underrättelser* 9.8.1864 (kaksikielinen), *Wiborgs Tidning* 10.8.1864, *Folkvännan* 10.8.1864, *Mikkelin Wiikko-Sanomia* 11.8.1864, *Sanomia Turusta* 12.8.1864, *Björneborgs Tidning* 12.8.1864 (kaksikielinen), *Tähti* 12.8.1864, *Tapio* 13.8.1864, *Vasabladet* 13.8.1864, *Borgåbladet* 13.8.1864, *Hämäläinen* 19.8.1864, *Oulun Wiikko-Sanomia* 20.8.1864. Analysoinnissa olen tukeutunut COMHIS-projektin tuottamaan klusteritietokantaan, joka julkaistaan kuluvana vuonna osoitteessa: [comhis.fi/clusters](https://comhis.fi/clusters)
- 24 Ks. Historiallinen sanomalehtiarkisto: [digi.kansalliskirjasto.fi/sanomalehti/titles?year=1864](https://digi.kansalliskirjasto.fi/sanomalehti/titles?year=1864)
- 25 Hoang ym. 2011, 212–221.
- 26 Cordell 2015b, 37.
- 27 Sampson 2012, 1–3.
- 28 Hoang ym. 2011, 212–221.
- 29 Smith, Cordell & Dillon 2013, 86–94.
- 30 Cordell 2015b, 34; Nahon & Hemsley 2013, xii–xiii.
- 31 Viitataan tässä Manuel DeLandaan ekstensiivisten ja intensiivisten tilojen eroteluun, jonka pohja on Gilles Deleuzen ajattelussa. Ekstensiivisiä ulottuvuuksia ovat esimerkiksi volyymi, pituus, leveys tai nopeus, intensiivisiä taas esimerkiksi lämpötila, paine ja gravitaatio. Ks. DeLanda 2005, 80–88.
- 32 Sampson 2012, 5–6.
- 33 Sperber 1985, 73–89.

## Kirjallisuus

- Cordell, Ryan, Reprinting, Circulation, and the Network Author in Antebellum Newspapers. *American Literary History*. Vol. 27, No. 3, 2015a, 417–445.
- Cordell, Ryan, Viral Textuality in Nineteenth-Century US Newspaper Exchanges. Teoksessa *Virtual Victorians. Networks, Connections, Technologies*. Toim. Veronica Alfano & Andrew Stauffer. Palgrave Macmillan, New York 2015b, 29–55.
- DeLanda, Manuel, Space. Extensive and Intensive, Actual and Virtual. Teoksessa *Deleuze and Space*. Toim. Ian Buchanan & Gregg Lambert. Edinburgh University Press, Edinburgh 2005, 80–88.
- Delwiche, Aaron & Jennifer Jacobs Henderson, Introduction. What is Participatory Culture? Teoksessa *The Participatory Cultures Handbook*. Toim. Aaron Delwiche & Jennifer Jacobs Henderson. Routledge, New York 2012, 3–9.
- Hoang, Tuan-Anh, Ee-Peng Lim, Palakorn Achananuparp, Jing Jiang & Feida Zhu, On Modeling Virality of Twitter Content. *Digital Libraries. For Cultural Heritage, Knowledge Dissemination, and Future Creation. 13th International Conference on Asia-Pacific Digital Libraries, ICADL 2011*. Proceedings. Vol. 7008. Research Collection School Of Information, Singapore 2011. Verkossa: [ink.library.smu.edu.sg/sis\\_research/1440](https://ink.library.smu.edu.sg/sis_research/1440)
- Jenkins, Henry, *Convergence Culture. Where Old and New Media Collide*. New York University Press, New York 2006a.
- Jenkins, Henry, *Fans, Bloggers, and Gamers. Exploring Participatory Culture*. New York University Press, New York 2006b.
- Leskovec, Jure, Lada A. Adamic & Bernardo A. Huberman, The Dynamics of Viral Marketing. *ACM Trans. Web*. Vol. 1, No. 1, 2007, 39 sivua. DOI 10.1145/1232722.1232727.
- Lister, Martin Jon Dovey, Seth Giddings, Iain Grant & Kieran Kelly, *New Media. A Critical Introduction*. Routledge, New York 2003, 200–202.
- Nahon, Karine & Jeff Hemsley, *Going Viral*. Polity, Cambridge 2013.
- Osterhammel, Jürgen, *The Transformation of the World. A Global History of the Nineteenth Century* (Verwandlung der Welt. Eine Geschichte des 19. Jahrhunderts, 2009). Käänt. Patrick Camiller. Princeton University Press, Princeton 2014.
- Parikka, Jussi, *Digital Contagions. A Media Archaeology of Computer Viruses*. Peter Lang, New York 2007.
- Rushkoff, Douglas, *Media Virus! Hidden Agendas in Popular Culture*. Ballantine Books, New York 1994.
- Salmi, Hannu, Big data ja Porthan. *Auraica. Scripta a Societate Porthan edita*. No. 7, 2016a, 91–95. Verkossa: [journal.fi/aur/article/view/60329](https://journal.fi/aur/article/view/60329).
- Salmi, Hannu, Catastrophe, Emotions and Guilt. The Great Fire of Turku 1827. Teoksessa *Catastrophe, Gender and Urban Experience, 1648–1920*. Toim. Deborah Simonton & Hannu Salmi. Routledge, New York 2017, 121–138.
- Salmi, Hannu, Viral Virtuosity and the Itineraries of Celebrity Culture. Teoksessa *Travelling Notions of Culture in Early Nineteenth-Century Europe*. Toim. Hannu Salmi, Asko Nivala & Jukka Sarjala. Routledge, New York 2016b, 135–153.
- Sampson, Tony D., *Virality. Contagion Theory in the Age of Networks*. Polity, Cambridge 2012.
- Smith, David A., Ryan Cordell & Elizabeth Maddock Dillon, Infectious Texts. Modeling Text Reuse in Nineteenth-Century Newspapers. *2013 IEEE International Conference on Big Data 2013*, 86–94. DOI: 10.1109/BigData.2013.6691675
- Smith, David A., Ryan Cordell, Elizabeth Maddock Dillon, Nick Stramp & John Wilkerson, Detecting and Modeling Local Text Reuse. *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*. Piscataway, New Jersey 2014, 183–192.
- Smith, David A., Ryan Cordell & Abby Mullen, Computational Methods for Uncovering Reprinted Texts in Antebellum Newspapers. *American Literary History*. Vol. 27, No. 3, 2015, E1–E15.
- Sperber, Dan, Anthropology and Psychology. Towards an Epidemiology of Representations. *Man*. Vol. 20, No. 1, 1985, 73–89.
- Vesanto, Aleks, Asko Nivala, Tapio Salakoski, Hannu Salmi & Filip Ginter, A System for Identifying and Exploring Text Repetition in Large Historical Document Corpora. *Proceedings of the 21st Nordic Conference of Computational Linguistics*. Gothenburg 23.–24.5.2017. Linköping University Electronic Press, Linköping 2017, 330–333. Verkossa: [www.ep.liu.se/ecp/131/049/ecp17131049.pdf](https://www.ep.liu.se/ecp/131/049/ecp17131049.pdf)