



Kaakkois-Suomen
ammattikorkeakoulu



South-Eastern Finland
University of Applied Sciences

PLEASE NOTE! THIS IS PARALLEL PUBLISHED VERSION / SELF-ARCHIVED VERSION OF THE OF THE ORIGINAL ARTICLE

This is an electronic reprint of the original article.
This version may differ from the original in pagination and typographic detail.

Author(s): Kosonen, Miia ; Kirmo, Outi

Title: Kuinka arkistoida sosiaalista media?

Version: publisher´s PDF

Please cite the original version:

Kosonen, M., Kirmo, O., 2017. Kuinka arkistoida sosiaalista media? Faili 4 / 2017

HUOM! TÄMÄ ON RINNAKKAISTALLENNE

Rinnakkaistallennettu versio voi erota alkuperäisestä julkaistusta sivunumeroiltaan ja ilmeeltään.

Tekijät: Kosonen, Miia ; Kirmo, Outi

Otsikko: Kuinka arkistoida sosiaalista media?

Versio: publisher´s PDF

Käytä viittauksessa alkuperäistä lähdettä:

Kosonen, M., Kirmo, O., 2017. Kuinka arkistoida sosiaalista media? Faili 4 / 2017

Kuinka arkistoida sosiaalista mediaa?



Outi Kirmo



Miia Kosonen

Yksi digiajan suurista muutosvoimista on mahdollisuus luoda nopea ja suora yhteys muihin ihmisiin ohi virallisten hierarkioiden ja roolien. Sosiaalinen media on avannut ovet erilaisten ruohonjuuritason kansalaisliikkeiden, kampanjoiden, tapahtumien, avointen yhteisöjen ja oppimisen alustojen rakentamiseen jopa yhdessä yössä. Tieto virtaa eteenpäin erilaisissa verkostoissa ja niitä tukevissa sovelluksissa. Murre vaikuttaa perinteisten instituutioiden, asiantuntijoiden ja johtajien työhön. Päätä ei voi enää laittaa pensaaseen tai tarrautua vanhaan valta-asemaan ja toimintatapoihin.

Van Dijck (2013) rinnastaa internetin infrastruktuurin syntymisen sähkökaapeleiden ja vesijohtojen yleistymiseen. Veden ja sähkön sijaan paikasta toiseen siirtyykin tieto. Erityisen nopeasti ihmiset ovat löytäneet palvelut, joiden tuella on helppoa keskustella ystävien kanssa ja jakaa vaikkapa perheen lomakuvat omalle lähipiirille. Mutta samalla jaettu sisältö leviää usein laajemmalle yleisölle kuin sen

jakaja oli ajatellut – eikä käyttäjällä välttämättä ole mitään käsitystä siitä, mihin hänen tietojansa hyödynnetään. *Clarkin* (2015) mukaan sosiaalinen media onkin toteuttanut sen, mistä Stasi eli Saksan demokraattisen tasavallan DDR:n turvallisuuselin olisi voinut vain unelmoida: ihmiset luovuttavat itsestään tietoa täysin vapaaehtoisesti.

Sosiaalinen media on kattava ja koko ajan kasvava kokoelma ihmisten mielipiteitä, analysoitua tietoa, arkipäiväisiä tilannekuvauksia ja tunneilmauksia, mutta myös vaikuttamisyrittäjiä – sekä tietoisia että tiedostamattomia (kts. Jantunen, 2015). Jos tätä aineistoa käsitellään massana, on sillä kiistatta sellaista arvoa, jota yksittäisistä postauksista ja päivityksistä ei hahmota. Sosiaalisen median hyödyntäminen eri tieteenalojen tarkoituksiin onkin yleistynyt. Erinomainen esimerkki tällaisen tietosisällön merkityksellisyydestä on tutkimus, jossa twiittien perusteella pystyttiin mallintamaan influenssan leviämistä ajallisesti nopeammin kuin viranomaiset pystyivät omista virallisista kanavistaan saamaan

selville (Paul et al., 2014). Aineistot voivat tarjota aarreaitan myös ihmisten käyttäytymismallien tunnistamiseen ja siten vaikkapa ongelmien ennaltaehkäisyyn. Yritykset käyttävät sosiaalisen median dataa riskienhallinnassa jo nyt – olisitko itse valmis myöntämään lainaa ihmiselle, joka jakaa somessaan pelikasinoihin liittyvää sisältöä?

Arvailematta jälkikäteen, kuinka Stasi olisi sosiaalisen median tietomassaa hyödyntänyt, olisi koncisto ollut todennäköisesti kiinnostunut siitä, miten aineisto arkistoidaan ja ketkä aineistoon pääsevät käsiksi. Muistammeko me olla kiinnostuneita sosiaalisen median muodostaman aineiston säilymisestä? Osaammeko nähdä oman aikamme tuotokset tulevaisuuden historiantutkijoiden silmin?

Arkistokokoelmiin suhtaudutaan tällä hetkellä hyvin asiakirjalähtöisesti. Jos keskitymme yksinomaan asiakirjojen – vaikkakin sähköisten – säilyttämiseen, rakentuu ajastamme kovin hallintokulttuurikeskeinen kuva. Löytävätkö myöhemmät sukupolvet tällä tavoin tietoa siitä, miten ihmiset elivät, mitä ajattelivat, miten käyttivät aikaansa ja miten pitivät yhteyttä toisiinsa. Asiakirjat edustavat luotettavaa ja todistusvoimaista asiakirjallista kulttuuriperintöä, mutta historiantutkimuksessa niiden anti on rajallinen.

Moninaisuutta arkistointiin

Väitämmekin, että kansallista kulttuuriperintöä tulisi lähestyä nykyistä huomattavasti laajemmin. Arkistonhallinta ei kohdistu vain asiakirjoihin, vaan pitkäaikaisesti säilytettävien aineistojen valintaan, järjestämiseen, kuvailuun ja säilyttämiseen liittyväksi toiminnaksi (Henttonen, 2015). Meidän on pohdittava moni-

muotoisen tiedon merkitystä perinteisen asiakirjallisen tiedon rinnalla – osana yhteistä kulttuuriperintöä. Samalla on laajennettava arkistoalan ammattilaisten asiantuntemusta. Uudenlaiset aineistot pakottavat havaitsemaan, millaisessa tietoympäristössä tiedonhallinnan ammattilaisiksi luokiteltavat arkistoalan osaajat toimivat.

Sosiaalisen median arkistoinnissa on laajempia haasteita kuin mitä arkistoalalla on tähän asti nähty. Nämä liittyvät tekniikkaan, osaamiseen, lainsäädäntöön ja aineistoon itseensä. Yhdessä nämä kytkeytyvät arkistoinnin tarpeisiin; mitä arkistoidaan, kun arkistoidaan sosiaalista mediaa, ja kuinka arkistoitava kohde rajataan.

Kuinka arkistoida vuorovaikutusta?

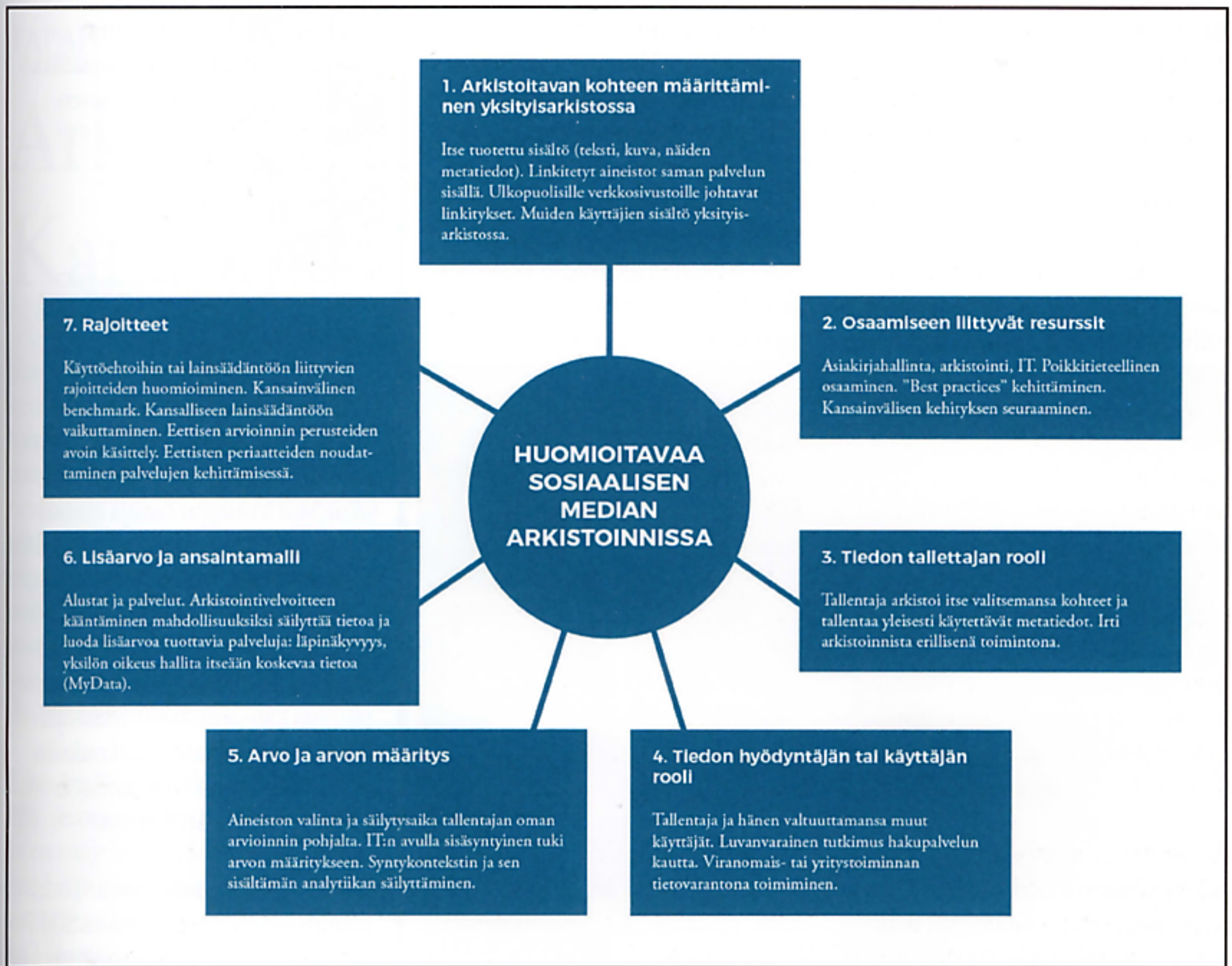
Otsikko tiivistää edellä mainitun problematiikan kolmeen sanaan. Lisäksi sosiaalisen median luonteeseen kuuluu muutos. Palvelut kehittyvät ja käyttöehdot muuttuvat. Myös ihmisten tapa käyttää sosiaalista mediaa muuttuu kaiken aikaa. Digitaalisessa muodossa olevan tiedon määrä kasvaa ja käsite ”tiedon hallinta” tuntuu ylipäättään istuvan huonosti sosiaalisen median kaltaiseen aineistoon. Arkistointi kuitenkin vaatii, että säilytettävään aineistoon kohdistetaan toimenpiteitä, joilla aineistoa voidaan hallita.

Asiakirjahallintaan kuuluvat oleellisenä osana sähköisen asiakirjan metatiedot, jotka kuvailevat aineistoa. Metatiedoilla on merkittävä rooli sosiaalisen mediankin arkistoinnissa. Esimerkiksi Twitterissä metatietoja ovat muun muassa päivämäärä, kellonaika, käyttäjätili, maininnat muista käyttäjistä sekä lyhytosoitteet. Näiden tietojen avulla voidaan osoittaa aineiston autenttisuus.

Pelkästään Facebookiin on integroitunut kymmeniä miljoonia ulkopuolisia sovelluksia ja verkkosivuja. Sosiaalisesta mediasta muodostuva arkistoitava kohde on pikemminkin **laaja tietokokonaisuus** kuin yksittäinen päivitys. Säilyttämällä tietyn postauksen tai kuvan säilyttämme vain ikään kuin teoksen nimen ja ehkä myös takakanteen kirjatun juonen, mutta itse sisältö jää puuttumaan.

Sosiaalisen median aineistoissa olemme yhtä lailla kiinnostuneita syntykontekstista – esimerkiksi siitä, ketkä sisällön luovat – mutta muutoin aineiston synty on varsin hallitsematonta. Twiitti, Facebook-päivitys tai blogikirjoitus on **sosiaalinen**: sitä voidaan kommentoida, siihen voidaan liittää linkkejä, päivitystä saatetaan muokata tai kirjoitus voidaan poistaa. Aktiiviaikana, jolloin aineistoa verkossa muodostuu, aineisto elää jokseenkin omaa elämäänsä palveluntarjoajan toimintakentällä. Palveluntarjoaja käyttää ja hyödyntää dataa haluamallaan tavalla ja voi vaikkapa myydä sitä eteenpäin. Tänä aikana aineisto kiinnostaa kyseisen palvelun muita käyttäjiä, mutta myös kaupallisia toimijoita tai viranomaisia. Entä milloin aineisto muuttuu passiiviseksi? On vaikea määritellä, milloin sosiaalisen median aineiston osalta päästään elinkaarissa todelliseen ”passiiviaikaan”.

Historiantutkijat toivovat tutkimuslähteidensä kuvastavan menneisyyttä tutkimuksensa kannalta mahdollisimman luotettavasti. Onkin paikallaan muistaa, ettei kukaan sosiaalista mediaa aktiivisesti käyttävä henkilö saa eteensä sisältöä ”sellaisenaan”. Elämme kaikki erilaisissa kuplissa. Uutisvirrat on profiloitu, näkymien taakse piiloutuu oma ekonomiansa linkitysten ja tykkäysten kautta jne.



Kaavio kokoaa muistilistan sosiaalisen median arkistointiin.

Verkko-osoitteisiin saadaan ujutettua analytiikkaa ja älyä.

Vaikka tämän logiikan taustalla on kaupallisia yrityksiä ja tavoitteita, sitä voidaan ajatella hyödynnettävän myös muihin tarkoituksiin. Tietojen analysoinnista voi olla hyötyä arkistoinnissa ja arkistoitavan kohteen määrittämisessä: sen sijaan, että pyrkisimme yksittäisinä asiantuntijoina määrittelemään mikä on säilyttämisen arvoista, annamme verkostojen ja analytiikan tehdä työtä puolestamme ja tarjota apuvälineitä päätöksenteon tueksi.

Sosiaalisen median arkistoinnin erityiskysymyksiä

Sosiaalisen median osalta merkittävä rajoite on käyttöehdoissa yleensä oleva rajaus, jonka mukaan avointen rajapintojen (API) avulla talteen otettua tietoa ei saa siirtää, jakaa tai tallettaa esimerkiksi pilvipalveluihin. Aivan näin yksinkertainen tuo linjaus ei kuitenkaan ole. Esimerkiksi julkiset tahot, kuten Iso-Britannian valtionvirastot tallettavat Twitter- ja YouTube -aineistonsa UKGWA-arkistoon, ja itsearkistotoiminnolla myös suomalaiset organisaatiot voisivat arkistoida omat julkaisunsa. Yksityishenkilöille tämä olisi selkeästi yksinkertaisin tapa.

Entä kysymys siitä, voidaanko sosiaalista mediaa arkistoida ilman käyttäjien nimenomaista suostumusta? Sosiaalisen median arkistopalvelujen kehittämisessä on otettava huomioon ihmisten tietoinen valinta toimia ja näkyä verkossa, ja toisaalta EU:n tietosuoja-asetuksen mukainen oikeus tulla unohdetuksi – oli digijälkemme kuinka vahva tahansa. Yksityisyyttä kunnioittava viesti on antaa mahdollisuus jättäytyä halutessaan kokoelmien ulkopuolelle. Henkilökohtaisen arkiston muodostaminen käyttäjän itsensä määrittelemistä kohteista voidaan nähdä eettisesti kestävimpana ratkaisuna.

Yhteenvetona voidaan todeta, että sosiaalisen median aineiston talteenottoon, arkistointiin ja aineiston tutkimuskäyttöön liittyvät dokumentoitavat eivät kansainvälisen tutkimuskirjallisuuden perusteella ole kovinkaan vakiintuneita. Sosiaalisen median arkistointi ei ole vain yksittäisen tekstin, kuvan tai postauksen säilyttämistä, vaan vuorovaikutuksen talteenottoa. Aineistoa tallennettaessa muodostuu myös metatietoa tallennusprosessista. Alan kirjallisuudesta on nähtävissä, että tutkijat toivovat arkisto- ja kirjastoalan ammattilaisilta apua nimenomaan dokumentoinnin standardisoinnin eteenpäin viemiseen. Kun aineisto ei ole alun perinkään luonteeltaan pysyvää, vaan sen tarkoitus on muuttua ja välittyä vaihtelevalle määrälle yleisöä, myös ajatus metadatan muuttumattomuudesta ja staattisuudesta on saanut haastajia. Joidenkin tutkijoiden mukaan metadattaa saatetaan tulevaisuudessa lähestyä tieteellisen kommunikoinnin välineenä eikä pysyvänä lopputulemana (Kinder-Kurlanda & Weller, 2015).

Lopuksi on vielä nostettava esiin osaamishaasteet ja koulutustarpeet. Nykyisen kaltaisen verkkoympäristön infrastruktuurin ymmärtäminen edesauttaisi arkistoalan toimijoita hahmottamaan, millainen ympäristö sosiaalinen media on ja toisaalta näkemään tämän ympäristön arvon kokonaisuutena. Arkistoammattilaisilla olisi paljon annettavaa esimerkiksi rakennettaessa palveluita MyData-aineistojen käsittelyyn; onhan monella arkistoalan asiantuntijalla kyky ymmärtää tiedonhallinnan prosesseja ja tietojen välisiä suhteita.

Uusien palvelujen kehittäminen vaatii kuitenkin aiempaa rohkeampaa otetta ja uskallusta visioida tulevaisuutta silloinkin, kun lainsäädäntö ei vielä edes tunnista visioinnin koh-

teena olevia tietojen käyttötilanteita tai kun niiden lainsäädännölliset tulkinnat ovat epäselviä. Sosiaalisen median arkistointi kytkettynä MyData-ajatteluun ja Kansalaisarkiston tyyppisten henkilökohtaisen tiedon säilytyspalvelujen kehittämiseen hyödyttäisi laajalti yksityisiä kansalaisia, yrityksiä, viranomaisia ja eri alojen tutkimustoimintaa.

Artikkeli pohjautuu Outi Kirmon yamk-opinnäytetyöhön Sosiaalinen media arkistolähteenä (2017), Kaakkois-Suomen ammattikorkeakoulu.

Kirjallisuutta

Clark, J. 2015. Social Media ultimate data mine. *The DataCenter Journal*, 36, February 2015.

Jantunen, S. 2015. *Infosota. "Iskut kohdistuvat kansalaisten tajuntaan."* Helsinki: Otava.

Henttonen, P. 2015. *Johdatus asiakirjahallinnan tutkimukseen.* Vantaa: Avain.

Kinder-Kurlanda, K. & Weller, K. 2015. Uncovering the Challenges in Collection, Sharing and Documentation: The Hidden Data of Social Media Research. *9th International AAAI Conference on Web and Social Media.* Oxford, UK, 26-29 May, 2015.

Paul, M.J., Dredze, M. & Broniatowski, D. 2014. Twitter Improves Influenza Forecasting. *PLOS Currents*, 28th October, 2014. <http://currents.plos.org/outbreaks/article/twitter-improves-influenza-forecasting/>

van Dijck, J. 2013. *The Culture of Connectivity. A critical history of Social Media.* Oxford University Press.