

## Ajankohtaista

*Elina Grundström, Lauri Haapanen & Sakari Ilkka*

# JSN määritteli algoritmit osaksi journalistista työtä

Muutama vuosikymmen sitten toimituksissa ei oikein osattu mieltää valokuvaajia ja graafikoita journalisteiksi. Sittemmin on hyvin ymmärretty, että myös he tekevät journalistisia valintoja siinä missä kirjoittajatkin. Ja oli aika, jolloin kaikki eivät pitäneet esimerkiksi juttuihin liitettyjä infografiikoita journalismina. Nykyään asiaa ei kiistä kukaan.

Nyt samantapaisen keskustelun kohteena ovat algoritmit, jotka viime vuosina – ehkä yllättävänkin myöhään – ovat löytäneet tiensä myös toimitustyön prosesseihin. Suomessa jonkinlaisena uutisautomaatiikan sisäänmarssina voi pitää kevään 2017 kunta-vaaleja, joissa tulosuutisia tuottivat ihmistoimittajien lisäksi *Helsingin Sanomien* Latoja, *Yleisradion* Voitto-robotti sekä Vaalikone Valteri, jonka rakensi Helsingin yliopiston ja VTT:n poikkitieteellinen tutkimuskonsortio. Vaalien lisäksi automaatiikan kohteena ovat olleet erityisesti liikenne-, sää-, pörssi- ja urheilu-uutiset.

Siinä missä niin kutsuttujen uutisrobottien käyttö on Suomessa vielä kokeiluluonteista, personointi eli sisältöjen kohdentaminen on jo monissa toimituksissa arkea – vaikka lukijat eivät sen kaikkia muotoja välttämättä tiedosta tai tunnista. Personoinnilla tarkoitetaan esimerkiksi henkilön käyttäytymishistorian perusteella muokkautuvaa verkkosivun tai mobiilisovelluksen etusivunäkymää tai muuta palvelun osaa, kuten juttusuosituslistaa.

Vastatakseen sekä algoritmien käytön että siihen liittyvän lainsäädännön nopeaan kehitykseen Julkisen sanan neuvosto (JSN) antoi lokakuun lopussa lausuman, jonka nimi on *Lausuma uutisautomaatiikan ja personoinnin merkitsemisestä*.<sup>1</sup> Kyseessä on JSN:n ensimmäinen aihetta käsittelevä lausuma – tietääksemme ensimmäinen laatuaan kansainvälisestikin. Lausuman keskeisenä tarkoituksena on määritellä algoritmisten apuvälineiden käyttö osaksi journalistista työtä. Lisäksi lausumassa annetaan – kuten sen nimestäkin käy ilmi – suosituksia uutisautomaatiikan ja personoinnin merkitsemisestä, joka on tähän asti ollut epäjohdonmukaista.

Konkreettisena liikkeellelähdenä lausumatyölle voi nähdä muutamien alalle suunnattujen keskustelunavauksien<sup>2</sup> jälkeen Euroopan medianeuvostojen kokouksen, jonka JSN isännöi Helsingissä lokakuussa 2018<sup>3</sup>. Algoritmien käyttö journalismissa oli yksi kokouksen aiheista.

Lausumaa valmisteleva työryhmä<sup>4</sup> kokoontui ensimmäisen kerran kesällä 2019 ja teki kyselyn, jossa kartoitimme algoritmien käytön nykytilaa ja tulevaisuutta suomalaisissa tiedotusvälineissä (tästä lisää alla). Tämän pohjalta työryhmä valmisteli lausumaluonnoksen, josta keskusteltiin useaan otteeseen myös neuvoston kokouksissa. Koska muutamat suuret tiedotusvälineet suhtautuivat kyselyssä varauksellisesti lausuman laatimiseen – joskaan tuossa vaiheessa vielä tietämättä sen sisällöstä – JSN:n puheenjohtaja esitteli luonnosta kyseisten toimitusten päälliköille lokakuun alussa. Vastaanotto oli myönteinen, mikä on lausuman legitimitietin kannalta tärkeää. Lausuma hyväksyttiin JSN:n kokouksessa 30. lokakuuta 2019.

Tässä kirjoituksessa esittelemme ensin kyselyn tuloksia siitä, millä tavoin algoritmeja suomalaisissa, JSN:n itsesääntelyn piiriin kuuluvissa tiedotusvälineissä tällä hetkellä käytetään. Tämän jälkeen kerromme lausuman keskeiset ohjeistukset ja sen, millä tavalla ne *Journalistin ohjeita* täydentävät. Lopuksi perustelemme, miksi lausuma on pääasiassa periaatteellinen ja siinä annettavat suositukset väljiä.

### Algoritmien rooli suomalaisessa journalismissa

Lähetimme verkossa toteutetun kyselyn 58 tiedotusvälineen päätoimittajalle, joilta saimme elokuun 2019 aikana 18 vastausta.<sup>5</sup> Joissain tilanteissa kyselyyn vastasi päätoimittajan sijaan toimituspäällikkö tai digikehitystä johtava henkilö. Vaikka vastausprosentti jäi alhaiseksi (31 %), mukana olivat lähes kaikki suuret mediat – ne, joilla luultavimmin on parhaimmat resurssit algoritmisia työkaluja kehittää ja joita lausuma siis eniten koskettaa. Tuloksia tulee tulkita tätä vinoumaa vasten.

Kyselyn aluksi kysyimme uutisautomaatiikasta ja personoinnista yleisluontoisesti. Vastaajista viisi sanoi käyttävänsä uutisautomaatiikkaa, joskin yksityiskohtaisemmista vastauksista selvisi, että kaksi oli siitä joko luopunut tai käyttö oli hyvin vähäistä:

*Olemme testanneet algoritmisia apuvälineitä sisältötuotannossa [...] totesimme, että laatu ei vielä vastaa odotuksiamme, joten apuvälineet eivät tällä hetkellä ole käytössä.*

Toisin kuin uutisautomaatiikka, personointi on suomalaismedioissa selvästi yleisempää; personointia kertoi käyttävänsä yli puolet vastanneista eli 10 tiedotusvälinettä 18:sta. Lisäksi eri tavoin tapahtuva personointi arvioitiin, kuten myöhemmin kerromme, lähitulevaisuuden tärkeimpänä algoritmien käytön kehityskohteena. Kysyttäessä yksityiskohtaisemmin algoritmien käyttösovellutuksista vastaukset jakautuivat yleisimmästä harvinaisimpaan seuraavasti:

1. Juttujen suosittelu käyttäytymisen perusteella (esim. klikkaukset) 13/18
2. Vaalikone 11/18
3. Otsikkovaihtoehtojen testaus 9/18
4. Juttujen suosittelu käyttäjän antamien tietojen perusteella 7/18
5. Lukijakommenttien algoritmiaivusteinen moderointi 3/18
6. ”Robotin” automaattisesti tekemä juttu tai jutun osa 3/18
7. Juttu, joka muuttuu lukijan syöttämien tietojen perusteella 3/18
8. Muu käyttötarkoitus (4/18).

Lausumatyön kannalta oleellinen tieto oli se, että tiedotusvälineet kertoivat lukijoilleen mahdollisesta uutisautomaatiikan tai personoinnin käytöstään vaihtelevasti. Erityisesti personoinnin osalta vastauksista heijastui käytäntöjen jonkinasteinen jäsentymättömyys ja tilannekohtainen tulkinta. Ensinnäkin personointia varten tehtävästä tiedonkeruusta kerrottiin vain lain vaatimassa tietosuojalausekkeessa.

*Meillä on GDPR:n mukaiset tietosuojaselosteet, joiden lukemiseen käyttäjää kehoitetaan. Cookieita koskevaa selostetta näytetään uudelle käyttäjälle popupina verkkopalvelun alalaidassa, kuten monissa muissakin verkkopalveluissa.*

*Ilmoitamme siitä normaalien GDPR-säännösten puitteissa, itse artikkelissa tai navigaatiassa emme mitenkään.*

*Pyrimme kertomaan aina push-ilmoituksilla / evästeillä + tietosuojalausekkeessa.*

Toisaalta siitä, että lukija on lukemassa hänelle personoitua sisältöä, ei kertonut vastanesta tiedotusvälineistä juuri kukaan.

*[kerromme personoinnista] Yleensä kun sillä on lukijalle merkitystä.*

*Suositteluelementtien otsikossa lukee "Suosittelemme sinulle". Eri asia on, osaako käyttäjä päätellä tästä, että kyseessä on hänen käyttäytymiseensä perustuva suosittelu. Muuten algoritmin käytöstä ei tässä yhteydessä kerrota, mutta tietosuojaselosteessa siitä toki mainitaan.*

Vaikka merkintäkäytännöt vaihtelevat ja ovat häilyviä, ei algoritmien käytöstä toisaalta myöskään suuremmin kysellä. Vain yksi vastanneista ilmoitti, että lukijoilta tulee viikoittain yhteydenottoja liittyen siihen, miten heistä kerättävää tietoa käytetään. Kaksi sanoi saavansa tällaisia kyselyjä kerran kuukaudessa, loput vielä harvemmin, jos ollenkaan. Kun kysyimme, onko algoritmisten apuvälineiden käyttöä varten laadittu sisäisiä ohjeistuksia, noin puolella algoritmeja aktiivisesti käyttävistä tällainen oli. Avovastausten perusteella ne ovat kuitenkin pääasiassa yleisiä eettisiä ohjeita.

*Meillä on sisällön merkitsemiseen ja muokkaamiseen liittyviä, yleisiä ohjeistuksia sekä luonnollisesti itse juttuprosessiin liittyviä ohjeita.*

*Asiasta ei ole erillisiä ohjeistuksia. Kehittäjät pitävät huolta siitä, että toiminnallisuudet täyttävät GDPR:n vaatimukset.*

Siinä missä personointia hyödynnettiin jo nyt varsin paljon, nousi se eri muodoissaan esiin myös, kun kysyimme, millaisia algoritmien käyttösovellutuksia kyseinen tiedotusväline saattaa ottaa lähitulevaisuudessa käyttöön. Personoinnin ohella useampaa vastaajaa kiinnosti automatisoitu juttutuotanto ja keskustelupalstojen moderointi eli käytösääntöjen vastaisten viestien tunnistaminen.

*Personoinnin käyttöönotto laajemmin koko uutistuotteessa lukijan kiinnostuksen/käyttäytymisen perusteella.*

*Käyttäjän käyttäytymiseen perustuva suosittelu ja moderointi ainakin.*

*Kommenttien automaattinen moderointi, suositusten tekeminen käyttäjän fyysisen sijainnin perusteella, sisällönkuvauksen muokkaaminen käyttöhistorian tai käyttäjän preferenssien mukaisesti.*

*Personoidut uutiskirjeet, parempia työkaluja, koneoppimiseen perustuvia artikkeleita.*

Lopuksi kartoitimme näkemyksiä siitä, miten tekoälyn ja algoritmien käyttö suomalaismedioissa lähivuosina kehittyy. Vastaajat olivat hyvin yksimielisiä kehityksen suunnasta, ja monet arvioivat sen edelleen kiihtyvän. Kehitykseen suhtauduttiin myönteisesti.

*Ne tulevat lisääntymään ja hyvä niin.*

*Käyttö laajenee nopeasti, samoin ymmärrys mahdollisuuksista.*

*Käyttö lisääntyy samaa tahtia kuin digisisältöjen osuus ja määrä kasvavat.*

*Uskon että tekoälyllä ja algoritmeilla tulee olemaan huomattavasti keskeisempi rooli tulevaisuudessa, etenkin suosittelun ja löydettävyyden saralla, mutta vähitellen myös tiedon hankinnassa ja sisällön tuotannossa.*

*Mediat tulevat käyttämään yhä enemmän valmiiksi tehtyjä, ulkoa ostettavia ratkaisuja (vrt. verkon mainontaratkaisujen kehitys).*

Muutamissa vastauksissa kehityksen nopeutta arvioitiin varovaisemmin ja tuotiin esiin esteitä algoritmien käytön yleistymiselle:

*[arvioin algoritmien käytön] Lisääntyvän alkuun hitaasti, mutta pitemmällä aikavälillä paljon.*

*Kehittyy, mutta ei niin nopeasti kuin alalla on arvioitu.*

*Niitä otetaan käyttöön kaikissa asioissa, joissa ihmistyövoimaa voidaan järkevästi korvata roboteilla. Tähän mennessä vain puhetta on ollut enemmän kuin käytännössä toimivia ratkaisuja. Esimerkiksi robotit eivät kirjoita vielä lähellekään järjellisen näköistä urheilujuttua suomeksi, vaikka kohta vuosikymmen on puhuttu siitä, että lähiaikoina tällaisia tulee. Journalistiset toimijat voisivat saada myös kilpailuedun Facebookia ym. suosittelualgoritmeihin perustuvia palveluita vastaan sillä, että ne rohkeasti tekisivätkin ammattilaisten kuratoimaa sisältöä sen sijaan, että yrittäisivät arvata, mikä käyttäjä milloinkin haluaa eniten klikata.*

Kyselyn perusteella kävi selväksi, että suurimmissa JSN:n itsesääntelyn piiriin kuuluvissa tiedotusvälineissä algoritmisia apuvälineitä käytetään jo laajasti ja niiden käyttö tulee lähiaikoina mitä ilmeisimmin merkittävästi lisääntymään. Akuuttia tarvetta lausumalle

ei näyttänyt kyselyn perusteella kuitenkaan olevan, ja JSN:n pöydälle on päätyneet toistaiseksi ainoastaan yksi algoritmien käyttöön liittyvä kantelu (Vapauttava 6887/SL/18, *Kauppalehti*). Tästä huolimatta – ja erityisesti, koska sekä teknologinen että lainsäädännöllinen kehitys algoritmien suhteen on juuri nyt kiivasta – JSN päätti edetä lausuman suhteen ennakoivasti: jos media ei sääntele itse itseään, joku muu tekee sen ennemmin tai myöhemmin, ja mikäli journalistit jättävät algoritmien käytön sääntelyn viranomaisille, EU:lle tai alustayhtiöille, se vaarantaisi lehdistönvapauden.

### Algoritmien käyttö on osa journalistista työtä

Algoritmisten apuvälineiden käyttö journalistisessa työssä vaikuttaa siihen, mitä julkaistaan, millaisin painoituksin ja kenelle. Nämä eivät ole mekaanisia valintoja vaan journalistisen päätöksenteon ydintä. Tämän vuoksi lausuman keskeinen tehtävä on määritellä algoritmisten apuvälineiden käyttö osaksi journalistista työtä. Tästä seuraa, että koska *Journalistin ohjeet* koskevat kaikkea journalistista työtä, on lukijoiden, katselijoiden ja kuuntelijoiden mahdollista kannella JSN:lle myös silloin, kun he epäilevät, että ohjeita on rikottu algoritmeja käyttämällä.

Lausumassa muistutetaan, ettei journalistista päätösvaltaa saa luovuttaa toimituksen ulkopuolisille, mikä koskee myös toimituksen ulkopuolisia algoritmien kehittäjiä. Tämä ei kuitenkaan edellytä, että kaikki apuvälineet tulisi koodata toimituksissa tai että päätoimittajan tulisi tuntea algoritmien toimintalogiikka perin pohjin. (Alla esitettävät kursivoidut katkelmat ovat otteita lausumasta.)

*Neuvosto toteaa, että tiedotusvälineillä tulee olla riittävä ymmärrys siitä, miten niiden käyttämät algoritmiset työkalut vaikuttavat sisältöön. Jos tiedotusväline esimerkiksi ostaa käyttöönsä talon ulkopuolisen tahon kehittämän työkalun, tiedotusvälineen on selvitettävä sekä hyväksyttävä sen keskeiset toimintaperiaatteet ja tarvittaessa pystyttävä reagoimaan ongelmatilanteisiin.*

*Neuvosto muistuttaa, että Journalistin ohjeet koskevat kaikkea journalistista työtä. Tiedotusvälineiden on siten varmistettava, että myös niiden digitaalisia palveluita kehittävät henkilöt noudattavat Journalistin ohjeita silloin, kun he tekevät itsenäisesti ratkaisuja, jotka vaikuttavat journalistiseen sisältöön.*

Se, että yleisö kokee tiedotusvälineiden toiminnan läpinäkyväksi, on tärkeää journalismin uskottavuudelle. *Journalistin ohjeet* edellyttävät, että tiedot on pyrittävä hankkimaan avoimesti ja että lähde mainitaan, kun käytetään toisen julkaisemia tietoja. Lausuma suosittaa, että jos tiedotusvälineen julkaisema journalistinen sisältö on ”olennaiselta osin tuotettu ja julkaistu automaattisesti”, tiedotusvälineet kertovat siitä merkinnällä:

*Tämä on automaattisesti tuotettu/päivittyvä juttu/artikkeli, jonka tiedot ovat peräisin [lähde].*

Personoinnin osalta JSN puolestaan suosittelee, että ”jos näkymässä on merkittävässä määrin käyttäjätietojen mukaan kohdennettua sisältöä, kohdentamisesta kerrotaan siten, että tieto on löydettävissä vähällä vaivalla ja ymmärrettävästi”.

Jo lausuman valmistelun aikana jotkut alan edustajat ihmettelivät, miksi lausuma käsittelee personointia, vaikka paljon puhuttu *yleinen tietosuoja-asetus* jo sääntelee henkilötietojen käyttöä varsin tiukasti (General Data Protection Regulation, GDPR). Vastauksemme on: juuri siksi. Mahdollisilta oikeudenkäynneiltä voidaan välttyä, kun yleisölle annetaan mahdollisuus kannella henkilötietojen keräämisestä ja käyttämisestä enemminkin neuvostolle kuin viedä tapaus oikeuteen. Tässä toimitaan samalla tavalla kuin esimerkiksi oikaisun ja vastineen kohdalla. Vaikka ne on mainittu sananvapauslaissa, ne löytyvät myös *Journalistin ohjeista*.

Käytännössä tulkintalinja siitä, mitä ”riittävä ymmärrys” ja ”olennainen osa” uutisautomaatiikassa ja ”merkittävä määrä”, ”vähä vaiva” ja ”ymmärrettävyys” personoinnissa ja sen merkitsemisessä tarkoittavat, hahmottuu vasta mahdollisten kanteluiden ja niistä annettavien ratkaisujen perusteella.

Lausumat täydentävät ja tulkitsevat *Journalistin ohjeita*. Toisin sanoen kantelussa tulee aina viitata siihen kohtaan tai niihin kohtiin *Journalistin ohjeista*, jota epäillään rikotun. Yksinomaan lausumiin ei voi kannellessa vedota. Tämä lausuma täydentää kolmea *Journalistin ohjeiden* kohtaa, joista 2. ja 7. kohdalla laajennus on ilmeinen.

*2. Tiedonvälityksen sisältöä koskevat ratkaisut on tehtävä journalistisin perustein. Tätä päätösvaltaa ei saa missään oloissa luovuttaa toimituksen ulkopuolisille.*

*7. Myös toisen työtä käytettäessä on noudatettava hyvää tapaa. Lähde on mainittava, kun käytetään toisen julkaisemia tietoja.*

Ohjekohdan 9 suhteen muutos on kategorisempi: siinä missä *tietojen hankkimisen avoimuus* on aiemmin viitannut jutun lähteenä käytettävien tietojen hankintaan, lausuman myötä tällä tarkoitetaan nyt myös sitä tiedonhankintaa, jossa yleisön edustajista kerätään henkilötietoja personointia varten.

*9. Työtä tehdessään journalistin on suositeltavaa ilmoittaa ammattinsa. Tiedot on pyrittävä hankkimaan avoimesti. Jos yhteiskunnallisesti merkittäviä seikkoja ei voida muutoin selvittää, journalisti voi tehdä haastatteluja ja hankkia tietoja myös tavallisuudesta poikkeavilla keinoilla.*

### **Lausuma on periaatteellinen ja kannustava**

Lausuma on luonteeltaan periaatteellinen. Sen antamat suositukset ovat varsin väljiä, ja voi olla, että monen tiedotusvälineen kohdalla se tapa, jolla käyttäjätietojen keräämisestä ja personoinnista on tietosuojalausekkeessa ja siitä informoivissa ponnahdusikkunoissa kerrottu, täyttää lausuman suositukset. Digitaalinen kehitys on juuri nyt nopeaa. Tällaisessa tilanteessa itsesääntelyä ei ole järkeä sitoa liian yksityiskohtaisiin tai velvoitaviin ohjeisiin, jotka saattaisivat haitata alan kehitystä ja joita ei ehkä voisi muutaman vuoden päästä enää soveltaa.

Digitaaliseen ympäristöön liittyy monenlaisia kiinnostavia kysymyksiä. Personointi ja uutisautomaatiikan käyttö ovat vain pieni osa kokonaisuutta, mutta niihin liittyviin käytäntöihin JSN voi toimivaltansa puitteissa konkreettisesti vaikuttaa. Lausumassa neuvosto suosittelee, että tiedotusvälineet käyttävät algoritmisia työvälineitä siten, että yleisö saa tulevaisuudessakin monipuolista tietoa maailmasta.

Jos tarvetta ilmenee, voidaan algoritmeja koskevaa ohjeistusta tarkentaa myöhemmin yksityiskohtaisemmillä lausumilla.

Keskustelua – ja tutkimusta – algoritmien käytöstä ja sen etiikasta ja vastuista lausuma toivottavasti entisestään lisää ja syventää. Millaisia ovat personoinnin ja uutisautomaatiikan tavoitteet ja vaikutukset ja minkälaisia ongelmia niihin yhtäältä tiedotusvälineiden ja toisaalta yleisön puolelta liitetään? Mitä uusia sisällön tuottamisen ja kohdistamisen tapoja ja kanavia on kehitteillä? Millaisia ovat henkilötietojen käyttöä ja algoritmeja koskevien valintojen heijastelemat arvot, ja kenen arvoja nämä ovat? Mikä on uutisautomaatiikassa käytettävän aineiston tuottajan vastuu? Kadottaako journalismi kykynsä korostaa, mikä on tärkeää ja mikä ei?

## Viitteet

---

<sup>1</sup> Lausuma on luettavissa osoitteessa <http://www.jsn.fi/lausumat/lausuma-uutisautomaatiikan-ja-personoinnin-merkitsemisesta-2019/>

<sup>2</sup> Esim. Grundström, Elina (2018). Tuntematon koodari. *Journalisti* nro 8 (17.8.2018). Saatavilla: <https://www.journalisti.fi/artikkelit/2018/9/tuntematon-koodari/>; Haapanen, Lauri (2017). Kirjoittaako robotti uutiset sadan vuoden päästä? *Kotimaisten kielten keskuksen blogi* (14.12.2017). Saatavilla: [https://www.kotus.fi/nyt/kolumnit\\_artikkelit\\_ja\\_esitelmat/artikkelit/sanoin\\_savutettavissa\\_2117\\_artikkelit/kirjoittaako\\_robotti\\_uutiset\\_sadan\\_vuoden\\_paasta](https://www.kotus.fi/nyt/kolumnit_artikkelit_ja_esitelmat/artikkelit/sanoin_savutettavissa_2117_artikkelit/kirjoittaako_robotti_uutiset_sadan_vuoden_paasta)

<sup>3</sup> Alliance of Independent Press Councils of Europe, AIPCE (<http://www.aipce.net/>).

<sup>4</sup> Lausumaa valmisteli työryhmä, johon kuuluivat JSN:n puheenjohtajan Elina Grundströmin lisäksi *Lapin Kansan* vastaava päätoimittaja Antti Kokkonen (JSN:n 1. vpj.), toimituspäällikkö Esa Mäkinen *Helsingin Sanomista*, vastaava päätoimittaja Jussi Tuulensuu *Aamulehdestä*, *Yle News Labin* päällikkö Jukka Niva sekä neuvoston jäsenet, tutkija Lauri Haapanen ja asiantuntija Heli Parikka. Työryhmän sihteerinä toimi JSN:n valmisteleva sihteeri Sakari Ilkka.

<sup>5</sup> Tässä luvussa esitettävät kyselyvastaukset ovat nimettömiä, koska ilmoitimme kyselylomakkeessa, että vastauksia referoidessamme ja lainatessamme emme nimeä, minkä tiedotusvälineen vastauksesta on kyse. Näin sen vuoksi, että vastaukset saattaisivat sisältää liikesalaisuuksia tai niihin verrattavissa olevia seikkoja.