

# Tietokannoista tulkintoihin: digitaalisen historiantutkimuksen käytäntöjä

 ennenjanyt.net/2019/08/tietokannoista-tulkintoihin-digitaalisen-historiantutkimuksen-kaytantaaja/

Petri Paju, Heli Rantala ja Hannu Salmi

Digitaalinen historiantutkimus voidaan nähdä uutena tai vanhana, näkökulmasta riippuen. Yhtäältä historiantutkijat ryhtyivät käyttämään tietokoneita apuvälineinään jo 1960-luvun lopulla myös Suomessa, toisaalta laskennallisiin menetelmiin perustuva uusin digitaalinen historia on kokenut voimakkaan nousun 2010-luvulla. Viime vuosina Suomessa on tutkittu menneisyyttä laskennallisin menetelmin useassa hankkeessa sekä analysoitu myös digitaalisen historian taustoja ja muotoutumista.<sup>[1]</sup> Tämä *Ennen ja nyt* -lehden erikoisnumero keskittyy alan tuoreisiin tutkimustuloksiin. Millaisia näkökulmia hankkeet ovat tuottaneet ja mitä on kenties luvassa jatkossa? Erikoisnumeron tavoitteena on avata historiantutkimuksen uusien menetelmien – ja samalla digitaalisten aineistojen – tuottamia tuloksia ja niissä edelleen piileviä mahdollisuuksia niin historiantutkijoille kuin kaikille kiinnostuneille. Tässä tarkoituksessa pidämme erityisen tärkeänä julkaista uusien menetelmien tuloksia suomen kielellä.

Tietotekniikalla tuotettujen ratkaisujen kohdalla oleellista on katsoa tarkkaan ns. konepellin alle. Digitaalisen aikakauden historiantutkimus perustuu yhä enemmän ”konehuoneille”. Taustalla näkyvä valokuva edustaa Yleisradion varhaista televisiotoimintaa monimutkaisine radioputkilaitteineen 1950-luvun lopulla. Nykyajan verkkomaailman konehuoneet jäävät useimmille näkymättömiin, mutta on tärkeä muistaa, että niissä tehdyt päätökset ohjaavat aineiston käyttöä ja käyttäjiä. Siksi tarvitaan julkaisuja, jotka käytännöllistävät ja havainnollistavat digitaalisen historian menetelmiä ja mahdollisuuksia.

Historiantutkijat ovat toki jo aiemmin kirjoittaneet digitaalisista ihmistieteistä, ja näiden keskeisten termien kattama laeva tutkimuskenttä sisältää paljon muutakin kuin yhteen erikoisnumeroon mahtuu. Hyviä lähtökohtia tarjoavat esimerkiksi kotimaiset teokset *Digitaalinen humanismi ja historiatieteet* ja verkkokirja *Historia i en digital värld*.<sup>[2]</sup> Lehden artikkelit ehdottavat polkuja moniin muihin tutkimuksiin.

Erikoisnumeron artikkelien lähtökohdissa painottuvat digitoidut sanomalehdet ja suomalainen lähdemateriaali. Useissa maissa juuri vanhat sanoma- ja aikakauslehdet edustavat kaikkein ahkerimmin verkon kautta käytettyä historiallista aineistoa, jonka hyödyntämiseen digitoiminen sekä digitaaliset menetelmät ovat viime vuosikymmeninä tuoneet aivan uusia ulottuvuuksia. Samalla lehdet ja Suomen tapaus ovat malliesimerkki siitä, että etenkin digitaalinen kokoelma tarvitsee jatkuvaa kehitystyötä vastatakseen käyttäjien muuttuviin tarpeisiin. Tällaisesta pitkäjänteisestä työstä on kiittäminen monia asiaan vihkiytyneitä mutta erityisesti Kansalliskirjastoa.<sup>[3]</sup>

Tämän numeron tutkimusartikkeleissa on hyödynnetty Kansalliskirjaston sekä Yleisradion digitaalisia aineistoja ja niissä yhdistetään laskennallisia menetelmiä laadulliseen analyysiin. Numeron kantavana teemana on pyrkimys ylittää raja kaukolukemisen (*distant reading*) ja lähilukemisen välillä. Suomen Akatemian rahoittaman konsortion Computational History and the Transformation of Public Discourse in Finland, 1640–1910 (COMHIS) Turun yliopiston työryhmä on tutkinut erityisesti sanomalehdistössä ilmenevää tekstien kopioimista digitaalisin menetelmin. Työn tuloksiin lukeutuu avoin COMHIS-tietokanta ([comhis.fi/clusters](http://comhis.fi/clusters)), jonka pohja-aineistona toimii Kansalliskirjaston sanomalehtikorpus vuosilta 1771–1920. Tietokannassa on nähtävillä tämän aineiston louhinnan tulos, lehdistöaineistosta löydetty tekstintoistotapaukset. Artikkeleissa ”Tekstien pitkä elämä” ja ”Jäkälän paluu” esitellään tämän tietokannan avulla tehtyä tekstien uudelleenkäytön tutkimusta.



Insinöörit tarkastelevat Yleisradioon kesällä 1956 saapunutta RCA-yhtiön televisiolähetintä. Tulevan televisiotoiminnan valmistelu alkaa. Kuvaaja Ruth Träskman, Yle. Lähde: Yle Archives – Yle Elävä Arkisto, Flickr.

”Tekstien pitkä elämä: Ajassa liikkuvat tekstit suomalaisessa sanomalehdistössä 1771–1920” -artikkelissa (Heli Rantala, Hannu Salmi, Aleksis Vesanto, Filip Ginter) tarkastellaan tekstien uudelleenkäyttöä varhaisessa suomalaislehdistössä. Artikkelissa esitellään tämän uudelleenkäytön tai kopioimisen yksi muoto, pitkän aikavälin toisteisuus. Tämä tarkoittaa saman tekstin tai tekstikatkelman julkaisemista useissa eri lehdissä vuosikymmenten tai jopa yli sadan vuoden aikavälillä. Tekstien pitkä käyttöikä kertoo paitsi lehtien toimittajien historiatietoisuudesta myös siitä miten sanomalehdistö toimii eräänlaisena muistiorganisaationa.

Artikkelissaan ”Jäkälän paluu” Petri Paju tutkii 1800-luvun alusta alkaen toistuneita yrityksiä saada suomalaiset käyttämään jäkälää ruoka-aineena – ihmisille erityisesti leivän ja velliin teossa. Posthumanistisen ajattelun kannustamana tarkastelussa huomioidaan myös se, kuinka jäkälää tarjottiin eläinten appeeksi ja miten nämä erilaiset käyttötarkoitukset riitelivät keskenään erityisesti vuosien 1917–1918 debatissa kriisiajan hätäravinnosta. Oppineiden kuten tunnetusti Elias Lönnrotin toteuttama jäkälävalistus toimii samalla esimerkkitapauksena, jonka avulla avataan ja tarkastellaan COMHIS-projektin tuottamaa tekstintoistojen tietokantaa, sen mahdollisuuksia ja rajoituksia. Johtopäätös on, että tietokannan mittavan aineiston avulla voidaan entistä paremmin hahmottaa monenlaisia pitkän tai lyhyemmän ajan kestäneitä, sanomalehtiin päätyneitä ilmiöitä ja kampanjoita.

Satu Sorvalin artikkelissa ”Onko syytä suuttua? Ärtymyksen ilmaisu vuoden 1895 yleisönosastoissa” tarkastelussa ovat suomalaislehtien yleisönosastokirjoitukset. Tutkimus hyödyntää Kansalliskirjaston digitaalista sanomalehtiarkistoa, mutta varsinainen kirjoitusten analyysi perustuu tutkijan laatimaan omaan tietokantaan. Sorvali käsittelee yleisönosastokirjoituksia tunnetutkimuksen näkökulmasta ja osoittaa, miten mielipidekirjoitusten aiheet kytkeytyivät aikakauden ajankohtaisiin yhteiskunnallisiin kysymyksiin.

Artikkelissa ”Kekkonen, Euroviisut ja Helsinki – kansallinen audiovisuaalinen perintö NER-analyysin tunnistamana” Maiju Kannisto ja Pekka Kauppinen tarkastelevat Ylen Elävää Arkistoa. Sen verkkosivuille on koottu yleisradioyhtiön historiallista audiovisuaalista lähetyksmateriaalia taustatietoineen kenen tahansa luettavaksi ja palvelua täydennetään jatkuvasti. Nimentunnistukseksi (*Named-entity recognition*, NER) kutsutun tietokoneavusteisen menetelmän avulla Kannisto ja Kauppinen tutkivat Elävän Arkiston aineistoon liittyvää metatietoa. Tällä tavoin he saavat luotua yleiskuvan esimerkiksi siitä, keitä ihmisiä ja mitä tapahtumia sekä historiallisia aikakausia sivustolla on edustettuna ja missä suhteessa, ja edelleen analysoitua esimerkiksi Urho Kaleva Kekkonen ja Helsingin keskeisyyttä aineistossa.

Heikki Kokko esittelee katsausartikkelissaan 1800-luvun suomenkielisiä paikalliskirjeitä yhteen kokoavaa Translocalis-tietokantaa. Usein maaseutukirjeiksi kutsutut paikalliskirjeet ovat sanomalehtien toimitusten ulkopuolisten tahojen laatimia ja lehtiin lähettämiä lukijakirjeitä. Tällä hetkellä tietokannasta löytyvät 1850–1870-lukujen lehdissä julkaistut paikalliskirjeet, mutta tietokantaa on tarkoitus täydentää sekä varhaisemmalla että myöhemmällä aineistolla. Tietokanta parantaa huomattavasti tämän aineiston käytettävyyttä, sillä aiemmin lukijakirjeitä on pitänyt etsiä lukuisten eri lehtien yksittäisistä numeroista. Tietokannan kokoamisesta vastaa Tampereen yliopistossa toimiva Suomen Akatemian Kokemuksen historian huippuyksikkö.

Pohjoismainen digitaalisen humanistisen tutkimuksen tapaaminen (DHN 2019) tarjosi jälleen maaliskuussa 2019 läpileikkauksen uusinta tutkimusta ja hanke-esittelyjä. Matti La Mela kertoo raportissaan tästä järjestyksessään sarjan neljännessä konferenssista,

joka kokoontui Kööpenhaminassa. Siellä esiteltiin monia suomalaisia historia-alan hankkeita ja useat osanottajat myös julkaisivat tuloksiaan konferenssin vertaisarvioidussa julkaisusarjassa. Kannustamme tutustumaan.

*Teemanumeron toimittajat Petri Paju (FT, dosentti), Heli Rantala (FT) ja Hannu Salmi (akatemiaprofessori) toimivat Turun yliopiston kulttuurihistorian oppiaineessa ja ovat työskennelleet COMHIS-hankkeessa.*

## Lähteet

---

Elo, Kimmo. (toim.) *Digitaalinen humanismi ja historiatieteet*. Historia mirabilis 12. Turun Historiallinen Yhdistys, Turku 2016.

Haverinen, Anna & Suominen, Jaakko. Koodaamisen ja kirjoittamisen vuoropuhelu? – Mitä on digitaalinen humanistinen tutkimus. Katsaukset. *Ennen ja nyt* (2015).

Kettunen, Kimmo, Pääkkönen, Tuula & Koistinen, Mika. Kansalliskirjaston digitoitu historiallinen lehtiaineisto 1771–1910: sanatason laatu, kokoelmien käyttö ja laadun parantaminen. *Informaatiotutkimus* 35:3 (2016), 3–14.

Onnela, Tapio. Historiantutkimus Internetin ja digitaalisen kumouksen aikakaudella. *Agricolan Tietosanommat* 5/1998. <http://agricola.utu.fi/julkaisut/tietosanommat-1998-2005/numero5-98/historiantutkimus.html> (haettu 13.8.2019).

Paju, Petri. International collaboration and Finland in the early years of computer-assisted history research: Combining influences from Nordic and Soviet Baltic historians. Teoksessa Costanza Navarretta, Manex Agirrezabal & Bente Maegaard (toim.) *Proceedings of the 4th Conference of the Association Digital Humanities in the Nordic Countries. March 6-8, 2019*. CEUR Workshop Proceedings 2364, <http://ceur-ws.org/Vol-2364/>. Copenhagen 2019, 349–357.

Parland-von Essen, Jessica & Nyberg, Kenneth. *Historia i en digital värld*. Verkkokirja, 1. (2014) ja 2. versio. <http://dighist.se/> (haettu 13.8.2019).

Tolonen, Mikko & Lahti, Leo. Digitaaliset ihmistieteet ja historiantutkimus. Teoksessa Matti O. Hannikainen, Mirkka Danielsbacka, & Tuomas Tepora (toim.) *Menneisyyden rakentajat: teorian historiantutkimuksessa*. Gaudeamus, Helsinki 2018, 235–258.

---

<sup>[1]</sup> Onnela 1998; Tolonen & Lahti 2018; Paju 2019.

<sup>[2]</sup> Parland-von Essen & Nyberg 2014-; Elo 2016. Ks. myös Haverinen & Suominen 2015.

<sup>[3]</sup> Ks. Kettunen, Pääkkönen & Koistinen 2016.