

Digitalisaation haasteet historiantutkimuksessa

Digitalisaation vaikutuksia ei voi nyky-yhteiskunnassa välttää. Digitalisaatio muuttaa ja sen nimenomaan tähdätäänkin muuttavan myös akateemisen opetuksen ja tutkimuksen lähtökohtia. Painetta tai myönteisesti ajateltuna tukea tähän tulee muun muassa Suomen Akatemian ja yliopistojen myöntämän rahoituksen ja strategisen ohjauksen kautta, mikä vaikuttaa luonnollisesti myös historian alan opetukseen ja tutkimukseen. Suomen Akatemialla on Digitaaliset ihmistieteet -akatemiaohjelma (DIGIHUM) vuosina 2016–2019¹ ja 2020–2022.² Eri yliopistojen strategioissa digitalisaatio on nostettu esiin erilaisin painotuksin. Esimerkiksi Helsingin yliopiston vuosien 2017–2020 strategiaan kuuluu oppimisympäristöjen digitalisaatio,³ Itä-Suomen yliopiston strategian yksi neljästä maailmanlaajuisesta haasteesta on oppiminen digitalisoituneessa yhteiskunnassa,⁴ ja Oulun yliopiston strategiaan kuuluu fokusalueena digitalisaatio ja yhteiskunta.⁵ Lisäksi ainakin Helsingin yliopistossa on ollut ja on yhä eri oppiaineissa, mukaan lukien historian maisteriohjelmassa, niin sanottuja digiloikkoprojekteja, joissa digitalisaatiota edistetään oppiaineen omista lähtökohdista käsin.⁶

Strategiataso ei välttämättä avaa käytännön tasolla sitä, miten digitalisaatio vaikuttaa ja näkyy

historia-alan opetuksessa ja tutkimuksessa yliopistoissa. Onkin osuvasti todettu, että digitalisaatio ei ole mikään palikka, jonka voi vain tuoda alan ulkopuolelta.⁷ Vaikka digitalisaatio – niin eri tavoin kuin se voidaan ymmärtää eri konteksteissa – ei ole mikään täysin uusi ilmiö, näyttää se saaneen uudenlaiset koko yhteiskunnan läpikäyvät mittasuhteet. Aikaisemmin historian alalla painetta muutokseen, tietokoneavusteiseen historiantutkimukseen (nykytermein digitaaliseen historiantutkimukseen) syntyi lähinnä tieteensaisesti, mutta nyt käynnissä olevan murros on erilainen. Muutospaine tulee laajalta rintamalta koko yhteiskunnassa, mikä näkyy yliopistojen strategioissa, rahoituksessa ja uusina työelämätaitojen vaatimuksina.⁸

Digitalisaation historian alaan ulottuvan vaikutuksen pohdintaa vaikeuttaa se, että digitalisaatiolla ei ole tarkkaa määritelmää ja se voidaan ymmärtää eri konteksteissa eri tavoilla. Digitalisaatioksi voidaan esimerkiksi määritellä tiedon ja tietotekniikan hyödyntäminen toiminnan muuttamiseen tai uuden mahdollistamiseen.⁹ Tällainen yleinen määritelmä sopisi myös historiantutkimukseen ja -opetukseen, mutta se ei avaudu käytännön tasolla. Keskeinen yleinen piirre digitalisaatiossa kuitenkin on, että kvantitatiivisten menetelmien käytön painoarvo kasvaa

1. Digitaaliset ihmistieteet (DIGIHUM), Akatemiaohjelma 2016–2019 ohjelmamuistio, Suomen Akatemia, http://www.aka.fi/globalassets/32akatemiaohjelmat/digihum/ohjelmamuistio_digi_suomeksi_final.pdf (22.1.2019).
2. Digitaaliset ihmistieteet (DIGIHUM), Akatemiaohjelma 2016–2019 ja 2020–2022, Suomen Akatemia, <http://www.aka.fi/DIGIHUM/> (5.5.2019)
3. Helsingin yliopiston strategia 2017–2020, Helsingin yliopisto, <http://strategia.helsinki.fi/#kehittamiskohde3/oppimisymparistojen-digitalisaatio> (22.1.2019).
4. Itä-Suomen yliopiston strategia, Itä-Suomen yliopisto, <https://www.uef.fi/documents/10184/833133/UEF+Strategia+2015-2020.pdf/976838b5-5ff9-4665-b436-4fe6045c75bb> (22.12.2019).
5. Oulun yliopiston strategia, Oulun yliopisto, <https://www.oulu.fi/yliopisto/node/194817> (21.12.2019).
6. Digitalisaation avulla kohti parempaa yliopistoa, Turun yliopiston ylioppilaskunta, <https://www.tyy.fi/fi/medialle/ajan-kohtaista/digitalisaation-avulla-kohti-parempaa-yliopistoa> (28.1.2019); Jyväskylän yliopisto on yksi digitalisaation kärkiyliopistoista, Jyväskylän yliopisto, <https://www.jyu.fi/ajankohtaista/arkisto/2016/02/tiedote-2018-02-28-10-55-33-020547> (28.1.2019).
7. Anu Lahtinen, *Historian digiloikka. Havaintoja matkan varrelta*, Helsingin yliopiston Historia-blogi, <https://blogs.helsinki.fi/historia/2019/02/10/historia-ja-digi-valivaiheita-ja-kokemuksia/> (17.3.2019).
8. Tietokoneavusteisten menetelmien käyttö ei siis ole täysin uusi asia historiantutkimuksessa, mutta nyt muutospainetta ja myös kannustimia digitalisaation hyödyntämiseen tulee monelta taholta. Kuten artikkelin alussa jo todettiin, mukaan voi lukea eri muistiorganisaatioiden digitalisaatioon liittyvät hankkeet.
9. Jyrki. J. Kasvi, Digi digi digi. Digitalisaatiossa on kyse organisaatiokulttuurin muutoksesta, Tietoyhteiskunnan kehittämiskeskus ry, <https://tieke.fi/digi-digi-digi/> (19.12.2019).

hyödynnettäessä tietoa ja tietotekniikkaa uudella tavalla. Tutkimuksen ja opetuksen alalla tästä voi mainita esimerkkinä erilaiset digityökalut ja -menetelmät, jotka pohjautuvat laskennallisiin menetelmiin. Juuri tätä ulottuvuutta voi pitää digitalisaation aiheuttamana keskeisimpänä haasteena historiantutkijoille, joiden koulutuksessa ja tutkimuksessa kvalitatiiviset menetelmät ovat olleet hallitsevia.¹⁰

Tutkijat Inés Matres, Mila Oiva ja Mikko Tolonen ovat tarkastelleet digitaalisia ihmistieteitä yhteiskuntatieteiden ja humanististen tieteiden viimeaikaisen kehityksen valossa, mutta tässä katsauksessa pohditaan digitalisaatiota nimenomaan historiantutkimuksen haasteiden näkökulmasta.¹¹ Myös Tolosen ja Leo Lahden tutkimuksessa tarkastelunäkökulma suuntautuu enemmän digitaalisten ihmistieteiden suunnasta historiantutkimukseen kuin historiantutkimuksen näkökulmasta digitaalisiin ihmistieteisiin.¹² Uudenlaista lähestymistapaa edustaa Petri Pajun ja Mila Oivan kuuteen suomalaiseen yliopistoon järjestämä ”Roadshow”-opetuskiertue digitaalisesta historiantutkimuksesta. Kiertueen taustana oli heidän aikaisemmin tekemänsä kysely digitaalisen historiantutkimuksen keskeisistä haasteista, joita olivat digitaalisten menetelmien opetuksen puute ja historiantutkijoiden vajavaiset taidot tietokoneiden sekä ohjelmistojen hyödyntämisestä, ja toisaalta historiantutkijoiden heikot yhteydet it-alan osaajiin.¹³

Tekstissä ei ole tarkoitus kritisoida digitaalisia ihmistieteitä tai digihistoriaa, vaan tarkastella niitä reunaehtoja ja edellytyksiä, joita digitaalisuuteen ja digitaalisiin ihmistieteisiin liittyvät tietokoneavusteiset ja datatieteen menetelmät nostavat esiin historiantutkijoiden näkökulmasta. Näitä näkökohtia on pohdittava, jos historiantutkijat itse haluavat olla määrittelemässä tieteenalansa ehdoista käsin, miten digitalisaatio tulee vaikuttamaan historiantutkimukseen ja

-opetukseen. Pikemminkin artikkelin tarkastelu saattaa edistää datatieteen ja digitaalisten menetelmien hyödyntämistä historiantutkimuksessa siten, että historioitsijat tiedostavat niihin liittyviä mahdollisia esteitä, jäykkyyksiä ja ongelmia oman alansa näkökulmasta, jolloin ne ovat myös paremmin ylitettävissä. Lisäksi on todettava, että historiantutkijoiden olisi oleellista käydä tällaista keskustelua juuri nyt, kun digitalisaatioon liittyviä rakenteellisia muutoksia tehdään yliopistoissa ja akateemisessa yhteisössä koskien niin infrastruktuurihankkeita kuin opetuksen ja tutkimuksen käytäntöjä.¹⁴

Yleisenä historian alaa koskevana kysymyksenä on, ajautuuko historiantutkimuksen rahoitus ahtaalle muiden sellaisten tieteenalojen paineessa, joihin datatieteen ja ylipäätään digitalisaation synnyttämät mahdollisuudet soveltuvat suoraviivaisemmin. Miten se vaikuttaa muun muassa projekti- ja infrastruktuurirahoituksen saamiseen? Esimerkiksi Suomen Akatemian tutkimusinfrastruktuurien kartalla historiantutkimus, nimenomaan historiantutkimuksen mahdollistavat digitaaliset menetelmät, eivät ole vahvasti edustettuina. FIN-CLARININ Kielipankin kautta tarjoama digitaalisten aineistojen ja työkalujen palvelukokonaisuus on ennen kaikkea kielitieteilijöille suunnattu. ESS Suomi, European Social Survey, taas on selvästikin yhteiskuntatieteilijöille kohdistettu tutkimusinfrastruktuuri. Sen sijaan Finna, mikä on suomalaisten museoiden, kirjastojen ja arkistojen yhteinen alusta eri aineistoille, on historiantutkijoiden kannalta hyödyllinen, mutta aineistojen ja niiden metatietojen hajanaisuuden takia varsin perustason digitaalinen alusta, vaikka sitä voikin yrittää hyödyntää samaan tapaan kuin Kansalliskirjaston Digin hakua. Suomen Akatemian tiekartan tutkimusinfrastruktuurit yhteiskunta- ja humanistiset tieteet -kategoriassa ovat FIN-CLARIN, CESSDA, ESS Suomi, FinElib, FMAS ja Finna.¹⁵

10. Muista digitalisaation aiheuttamista haasteista historiantutkijoille esimerkiksi Lara Putnam, *The Transnational and the Text-Searchable: Digitized Sources and the Shadows they cast*. *American Historical Review*, huhtikuu 2016.
11. Inés Matres, Mila Oiva & Mikko Tolonen, *In Between Research Cultures. The State of Digital Humanities in Finland*. *Informaatiotutkimus* 37 (2018), 37–38. Informaatiotutkimus arkisto, <https://doi.org/10.23978/inf.71160> (26.3.2019).
12. Mikko Tolonen & Leo Lahti, *Digitaaliset ihmistieteet ja historiantutkimus*. Teoksessa Matti O. Hannikainen, Mirkka Danielsbacka & Tuomas Tepora (toim.) *Menneisyyden rakentajat. Teoriat historiantutkimuksessa*, Gaudeamus 2018, 235–258.
13. Petri Paju & Mila Oiva, *Digitaalisen historiantutkimuksen opetuskiertue*. *Historiallinen Aikakauskirja* 117 (2019), 89–94.
14. Vrt. Matres, Oiva & Tolonen 2018, 43. Informaatiotutkimus arkisto, <https://doi.org/10.23978/inf.71160> (26.3.2019).
15. Suomen tutkimusinfrastruktuurien strategia ja tiekartta 2014–2020, Väliarviointi 2018, Suomen Akatemia, https://www.aka.fi/globalassets/tiedostot/aka_infra_tiekartta_raportti_fi_030518.pdf (5.3.2019).

Kirjoituksen punaisena lankana on ajatus siitä, että historian tutkijoiden on syytä itse osallistua ja ottaa paikkansa digitalisaation aikaansaamassa murroksessa historian alalla. Katsauksessa nostetaan pohdinnan kohteeksi sellaisia asioita, joita historian tutkijat tulevat kohtaamaan yhä voimistuvan digitalisaation myötä, ja joihin heidän on otettava kantaa tavalla tai toisella ammattikuntansa edustajina – joko aktiivisesti keskustelemalla tai tutkimus- ja opetuskäytäntöjä muokkaamalla ja rakentamalla. Historiantutkijat eivät toki ole mikään yhtenäinen ryhmä, joka kohtaisi digitalisaation haasteet samankaltaisina. Pelkästään historian eri tutkimusalueiden kirjo on niin laaja ja lähteet ja tutkimusmenetelmät niin erilaisia, että ratkaisuja digitalisaation hyödyntämiseen ja haasteisiin on käytännön tasolla haettava monin eri tavoin.

Digitaaliset ihmistieteet ja historia

Digitaaliset ihmistieteet on tunnettu jo vuosikymmeniä, vaikkakin yleensä eri termein ilmaistuna. Englanninkielisessä akateemisessa maailmassa on käytetty muun muassa ilmaisia *humanities computing* ja *humanist informatics*.¹⁶ Suomessa on esimerkiksi historian professori Viljo Rasilan jälkeensä jättämä tietokoneavusteisten menetelmien perintö.¹⁷ Viime vuosina vallitsevaksi käsitteeksi on noussut *digital humanities*, suomeksi digitaaliset ihmistieteet.¹⁸ Paikoittain suomeksi käytetään myös termiä digihumanismi tai digitaalinen humanismi.¹⁹ Digitaaliset ihmistieteet -käsitteellä ei ole vakiintunutta määritelmää, ja sen määrittelystä käydään alan sisällä jatkuvaa

keskustelua, mikä kuvastaa sitä, että digitaaliset ihmistieteet ovat hakemassa entistä selvempää identiteettiä akateemisella kentällä. Määrittely ei ole yhdentekevä kysymys, koska sillä on käytännön seurauksia suhteessa tutkimuksen rahoittajiin, yliopistohallintoon ja niin sanottuun suureen yleisöön. Samalla on keskeistä, miten määrittely vaikuttaa digitaalisia ihmistieteitä koskevaan tutkijayhteisöön. Miten kyseinen yhteisö määritellään, ketkä siihen kuuluvat ja mitkä ovat kyseisen yhteisön sisäiset jännitteet? Entä mikä on suhde muihin tieteenaloihin ja missä rajat näihin mahdollisesti kulkevat?²⁰

Digitaaliset ihmistieteet sisältävät Suomen Akatemian ohjelman mukaan esimerkiksi korpuslingvistiikan ja sen laskennalliset menetelmät.²¹ Helsingin yliopiston humanistisen tiedekunnan verkkosivuilla Digitaaliset ihmistieteet -otsikon alta löytyy kolme tutkimusala: foneetiikka, kieliteknologia ja kognitiotiede.²² Digitaalisten ihmistieteiden alalta nostetuista tutkimusaloista syntyy vaikutelma, etteivät ne tule lähelle historian tutkimusta. Käytännössä digitaalisten ihmistieteiden ala on huomattavasti kattavampi ja Suomessa tehdään toki digitaalista historian tutkimusta digitaalisten ihmistieteiden merkityksessä.²³ Historiantutkimuksen kannalta saatakin olla mielekkäämpää puhua nimenomaan digitaalisesta historiasta tai digihistoriasta, mikä sekin kattaa laajan kirjon erityyisiä tutkimuksia niin ajallisesti, aineistoiltaan kuin kysymyksenasetteluiltaan.²⁴ Digihistoria ei ole kuitenkaan kovin yleisesti käytetty käsite tai ainakaan sillä ei ole yhtä vakiintunutta asemaa kuin digitaalis-

16. Julianne Nyhan, Melissa Terras & Edward Vanhoutte, Introduction. Teoksessa Julianne Nyhan, Melissa Terras & Edward Vanhoutte (toim.) *Defining Digital Humanities. A Reader*. Taylor and Francis Group 2013, 2.

17. Petri Paju, The evolution of IT use by a field of research. Case: (digital) history. DH pizza, Otaniemi 2.5.2018; Viljo Rasila, Tietokone historian tutkimuksessa. *Historiallinen Aikakauskirja* 65 (1967), 140–146.

18. Kts. esimerkiksi Digitaaliset ihmistieteet (DIGIHUM) Akatemiaohjelma 2016–2019 ohjelmamuistio, Suomen Akatemia, http://www.aka.fi/globalassets/32akatemiaohjelmat/digihum/ohjelmamuistio_digi_suomeksi_final.pdf (22.1.2019); Digitaaliset ihmistieteet, Helsingin yliopisto, <https://www.helsinki.fi/en/helsinki-centre-for-digital-humanities> (22.1.2019); Nyhan, Terras & Vanhoutte 2013.

19. Kimmo Elo (toim.) *Digitaalinen humanismi ja historiatieteet*. Historia mirabilis. Turun historiallinen yhdistys 2016.

20. Nyhan, Terras & Vanhoutte 2013, 1–2, 6; Matres, Oiva & Tolonen 2018, 39. Informaatiotutkimus arkisto, <https://doi.org/10.23978/inf.71160> (26.3.2019).

21. Digitaaliset ihmistieteet (DIGIHUM) Akatemiaohjelma 2016–2019 ohjelmamuistio, Suomen Akatemia, http://www.aka.fi/globalassets/32akatemiaohjelmat/digihum/ohjelmamuistio_digi_suomeksi_final.pdf (23.1.2019).

22. Digitaaliset ihmistieteet, Helsingin yliopisto, <https://www.helsinki.fi/fi/humanistinen-tiedekunta/tutkimus/tieteenalat/digitaaliset-ihmistieteet> (23.1.2019).

23. Esim. Computational history and the Transformation of Public Discourse in Finland, 1640–1910, Turun yliopisto, <https://www.utu.fi/fi/yksikot/hum/yksikot/kulttuurihistoria/tutkimus/Sivut/comhis.aspx> (27.3.2019); Translocalis -tietokanta, Tampereen yliopisto, <https://research.uta.fi/hex-fi/esittely/translocalis/> (27.3.2019).

24. Kts. Digitaalinen historian tutkimus, *Ennen ja nyt* 2/2019 <https://agricolaverkko.fi/tietosanomat/ennen-ja-nyt-2-2019-digitaalinen-historiantutkimus/> (lainattu 10.2.2020).

ten ihmistieteiden käsitteellä suomenkielisessä tutkimuskentässä, vaikka myös digihistoriaa on nostettu esiin omana tutkimusalueenaan.²⁵ Alan suomenkielisen teoksen otsikko *Digitaalinen humanismi ja historiatieteet* kuvaa hyvin tilannetta.²⁶ Digihistorian legitimointi ja oikeutus ovat paremmin perusteltavissa, kun se nähdään osana digitaalisia ihmistieteitä. Kuuluuko digihistoria historian alaan, digitaalisiin ihmistieteisiin vai molempiin? Ja minkälaisia käytännön seurauksia kategorisoinnista voi mahdollisesti nousta tutkimuksen ja opetuksen kannalta?

Tarkastelun tekee haastavaksi se, ettei alan käsitteistö ole varsinkaan suomeksi vakiintunutta.²⁷ Historiantutkimuksen suhde digitaalisiin ihmistieteisiin riippuu siitä, miten digitaaliset ihmistieteet ja digihistoria kulloinkin eri yhteyksissä määritellään. Monet historiantutkijat ovat astuneet digihistorian tielle käytännössä huomaamattaan, jos digihistoria määritellään laajasti siten, että ”digitaalinen historiantutkimus on mitä tahansa menneisyyden ilmiöitä tai muutosprosesseja käsittelevää tutkimusta, jossa kaikki tai osa lähteistä on digitaalisessa muodossa.”²⁸ Jane Winters on todennut digihistorian olevan vaikea määritellä, sillä se kehittyi yhä ja merkitsee oikeutetusti eri ihmisille eri asioita erilaisissa asiayhteyksissä.²⁹ Jos digihistoria taas määritellään digitaalisissa ihmistieteissä yleisemmin käytetyn määritelmän mukaan siten, että käytetään tietokoneavusteisia menetelmiä, kuten Suomen Akatemian digitaalisten ihmistieteiden määri-

telmässä näytetään edellytettävän, astutaan jo selvästi historiantutkijoille vieraammalle maaperälle. Tosin akatemiaohjelmassa todetaan lisäksi, että ”ohjelmaan kuuluu myös digitalisaatiokehityksen tarkastelu kulttuurisena ja yhteiskunnallisena ilmiönä”, mikä viittaisi siihen, että myös digitalisaation tutkiminen ilmiönä ilman tietokoneavusteisia menetelmiä olisi mahdollista.³⁰ Digitalisaation tutkiminen yhteiskunnallisena ja kulttuurisena ilmiönä, niin sanottuna kolmantena teollisena vallankumouksena, luontuisikin varsin saumattomasti historiantutkimuksen aikaisempaan tutkimuskenttään ja -traditioon.³¹ Ilmiselvistikään tällainen näkökulma ei kuitenkaan ole Akatemiaohjelmassa pääpainona, vaan tutkimus, jossa keskiössä ovat tietokoneavusteiset menetelmät eli ”uudet menetelmät ja tekniikat, joissa digitaalista teknologiaa ja modernin datatieteen menetelmiä hyödynnetään” aineiston keruussa, hallinnassa ja analyysissa. Nimenomaan tällaisten menetelmien ja tekniikoiden käyttö asettaa haasteita historiantutkijoille.³²

Kun digitaalisten ihmistieteiden keskiöön nostetaan tietokoneavusteiset menetelmät, avainasemaan nousee tietojenkäsittelytiede. Tietojenkäsittelytiede taas voidaan määritellä yleisesti siten, että kyse on laskennallisista menetelmistä – vaikka tietojenkäsittelytieteen määrittelemisenäkään ei ole aivan yksinkertaista.³³ Yleisesti ymmärrettynä ”tietojenkäsittelytiede tutkii sitä, mitä ylimalkaan voidaan laskea ja miten se käy parhaiten päinsä käytettävissä

25. Esim. Turku Group for Digital history, Turun yliopisto, <https://digitalhistory.fi/> (12.5.2019); Paju & Oiva 2019, 89–94.

26. Elo (toim.) 2016. Vrt. myös Heli Rantala, Hannu Salmi, Asko Nivala, Petri Paju, Reetta Sippola, Aleksi Vesanto & Filip Ginter, Tekstien uudelleenkäyttö suomalaisessa sanoma- ja aikakauslehdistöissä 1771–1920, Digitaalisten ihmistieteiden näkökulma. *Historiallinen Aikakauskirja* 117 (2019), 53.

27. Kts. Elo (toim.) 2016.

28. Jaakko Suominen & Anna Sivula, Digisyyntysten ilmiöiden historiantutkimus. Teoksessa Kimmo Elo (toim.) *Digitaalinen humanismi ja historiatieteet*. Historia mirabilis. Turun historiallinen yhdistys 2016, 105–106.

29. Jane Winters, Digital history. Teoksessa Marek Tamm & Peter Burke (toim.). *Debating New Approaches to History*. Bloomsbury Academic 2019, 288.

30. Digitaaliset ihmistieteet (DIGIHUM) Akatemiaohjelma 2016–2019 ohjelmamuistio, Suomen Akatemia, <http://www.aka.fi/fi/akatemia/media/Tiedotteet/2019/suomen-akatemiaalta-12-miljoonaa-euroa-tutkimusinfrastruktuureiden-kehittämiseen/> (24.1.2019).

31. Kati Katajisto, Digimurroksen unelmia ja uhkakuvia – ja mitä tapahtuu historian alalla, <https://blogs.helsinki.fi/historia/tag/digihistoria/> (5.2.2019).

32. Digitaaliset ihmistieteet (DIGIHUM) Akatemiaohjelma 2016–2019 ohjelmamuistio, Suomen Akatemia, <http://www.aka.fi/fi/akatemia/media/Tiedotteet/2019/suomen-akatemiaalta-12-miljoonaa-euroa-tutkimusinfrastruktuureiden-kehittämiseen/> (24.1.2019); Myös uudessa ohjelmassa todetaan, että ohjelman yksi ulottuvuus on digitalisaation tutkiminen kulttuurisena ja sosiaalisena ilmiönä. Digital Humanities (DIGIHUM) 2016, 2019, Suomen Akatemia, <https://www.aka.fi/en/research-and-science-policy/academy-programmes/current-programmes/digihum/> (5.5.2019).

33. Kimmo Raatikainen, Mietteitä tietojenkäsittelytieteestä. Helsingin yliopisto, Tietojenkäsittelytieteen laitos, <https://www.cs.helsinki.fi/group/nodes/Papers/tktiede.pdf> (5.2.2019).

olevien resurssien puitteissa”³⁴ Laskennallinen tiede taas ”voidaan nähdä kolmantena tutkimusmenetelmänä kokeellisen ja teoreettisen rinnalla”, kuten opetusministeriön Laskennallisen tieteen kehittäminen Suomessa -selvityksessä todetaan. Laskennallinen tiede jakaantuu tällöin kolmeen osa-alueeseen: mallinnus- ja simulointiosaamiseen, algoritmeihin, ohjelmistoihin ja tietotekniikkaan liittyviin tieteihin sekä tietotekniseen infrastruktuuriin. Laskennallisessa tieteessä pystytään tutkimaan ilmiöitä, ”joita ei voi tutkia tai ymmärtää perinteisellä teoreettisella tai kokeellisella metodilla suuren datamäärän sekä ilmiöiden laajuuden ja kompleksisuuden vuoksi”. Laskennallisen tieteen kehittäminen on nähty strategisesti tärkeäksi asiaksi Suomen kilpailukyvyllä, ja siksi siihen perustuvaa osaamista on tähdennetty kautta linjan, ”peruskoulusta tutkijoihin ja professoreista toimitusjohtajiin.”³⁵

Historiantutkijoiden näkökulmasta tietokoneavusteiset menetelmät näyttävät ulottuvan moniin eri suuntiin, joissa heillä ei perinteisesti ole kompetenssia tai koulutusta, vaikka tavoitteessa tutkia monimutkaisia ilmiöitä, joita ei voi tutkia perinteisellä teoreettisella tai kokeellisella metodilla, onkin samoja ulottuvuuksia kuin historiantutkimuksessa. Erilaiset laskennalliset menetelmät, matematiikka ja tilastotiede, eivät ole historiantutkimuksen valtavirtaa, vaikka toki myös kvantitatiivista historiantutkimusta tehdään, ja jo 1970-luvulla historiallinen ekonometria (cliometrics) eli kansantaloustieteelle ominainen mallintaminen nousi Suomessa taloushistoriassa hallitsevaksi suuntaukseksi.³⁶ Lisäksi monissa kvalitatiivisissa historiantutkimuksissa hyödynnetään erilaisia laskelmia ja tilastoja. Pääasiallisena metodina laskennalliset menetelmät ovat kuitenkin melko harvinaisia niin historian tutkimuksessa kuin opetuksessa.

Digihistoria suuntautuu tällä hetkellä tietokoneavusteisten laskennallisten menetelmien käyttöön digitaalisten ihmistieteiden tavoin, mikä on

perinteiselle historiantutkimukselle erityinen haaste, ja mihin myös yleinen yhteiskunnallinen digitalisaation paine näyttäisi ohjaavan. Toisaalta taas digihistoriaksi voidaan määritellä jo digitoitussa muodossa olevien lähteiden käyttäminen. Tällainen määritelmä tulisikin käytännössä tarkoittamaan, että monet historiaa tutkivat olisivat tulevaisuudessa automaattisesti digihistorioitsijoita, sillä yhä suurempi määrä lähteitä syntyy (*born digital*) ja säilytetään vain ja ainoastaan digitaalisessa muodossa.³⁷

Historiantutkimus ja laskennalliset menetelmät – haasteiden tarkastelu käytännössä

Historiantutkijat eivät voi vältellä tai välttää haastetta, joita laskennallisten menetelmien lisääntyvä käyttö yhteiskunnan eri osa-alueilla digitalisaation myötä aiheuttaa. Olisi myös kyettävä kriittisesti arvioimaan, miten esimerkiksi algoritmit tai päätöksenteossa käytettävä numerodata ohjaavat ihmisten ajattelua tiettyyn suuntaan tai mitkä näkökulmat puuttuvat. Samalla kun historiantutkijoiden olisi hyvä oppia hyödyntämään tietokoneavusteisten laskennallisten menetelmien käyttöä tutkimuksissaan, vielä tärkeämpää olisi kyky pystyä nostamaan esiin niihin liittyvät rajoitukset ihmisen toiminnan analyysin kannalta. Molemmat päämäärät edellyttävät, että historiantutkijoiden olisi kyettävä ymmärtämään laskennallisten menetelmien perusteita, mikä taas on haastava ja vieras alue useille historiantutkijoille. Tarkastelen tätä seuraavaksi esimerkinomaisesti tilastotieteen kautta, joka on yksi keskeinen sovelusala digitalisaation taustalla. Tarkastelu tapahtuu kieltämättä yhden, kvalitatiivisesti suuntautuneen historiantutkijan positiosta, mutta tällaisia astumisia vieraammille alueille tarvitaan, jotta pystytään paremmin hahmottamaan digitalisaatioon liittyviä haasteita.

Tilastotiede on menetelmätiede, jonka teoria pohjaa todennäköisyyslaskentaan.³⁸ Tilastotieteen suhdetta historiantutkimukseen voi pitää

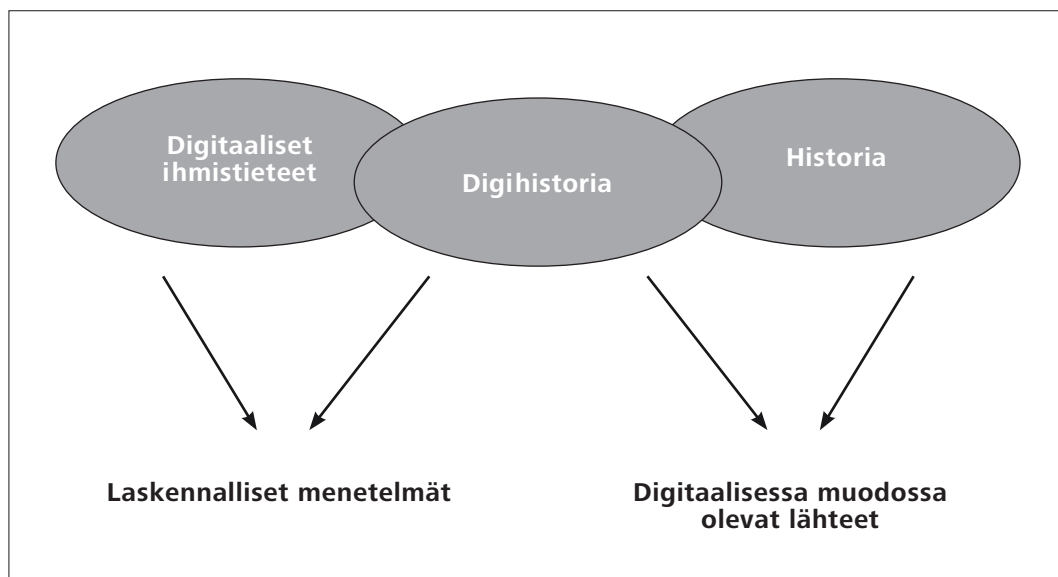
34. Tietojenkäsittelytiede, Wikipedia, <https://fi.wikipedia.org/wiki/Tietojenk%C3%A4sittelytiede> (7.1.2019). Vrt. myös Erkki Mäkinen, Johdatus tietojenkäsittelytieteisiin, 3-11. Informaatiotieteiden yksikön raportteja 13/2012, Tampereen yliopisto, http://www.uta.fi/sis/reports/index/Ru3_2012.pdf (5.2.2019).

35. Laskennallisen tieteen kehittäminen Suomessa, Opetusministeriön työryhmämuistioita ja selvityksiä 23 (2007), <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/79715/tr23.pdf?sequence=1> (lainattu 26.1.2019).

36. Pertti Haapala, *Sosiaalishistoria. Johdatus tutkimukseen*. Suomen Historiallinen Seura 1989, 17.

37. Jessica Parland-von Essen & Kenneth Nyberg, *Historia i en digital värld*, version 2, <https://digihist.se/hdv2/1-inledning/> (lainattu 27.12.2019).

38. 1 Tilasto ja sitä tutkiva tiede, Johdatus tilastotieteeseen, Tilastokeskus, https://tilastokoulu.stat.fi/verkkokoulu_v2.xql?course_id=tkoulu_tilaj&lesson_id=1&page_type=sisalto (18.3.2019).



varsin etäisenä. Aalto-yliopiston Ilkka Mellinin laatimassa Tilastollisten menetelmien johdannossa ei mainita historian tutkimusta lainkaan tilastotieteen sovelluskohteena. Sen sijaan ihmistieteistä mainitaan arkeologia, kielitiede ja psykologia ja yhteiskuntatieteistä sosiaalitieteet ja taloustiede.³⁹ Tilastotieteen ja historian tutkimuksen välinen heikko yhteys on yllättävä siitä näkökulmasta, että sana tilasto yhdistetään usein ensimmäiseksi nimenomaan yhteiskuntaa ja sitä kuvaaviin numeerisiin tilastoihin, ja myös historian tutkimuksessa yhteiskunta ja ihmisten toiminta siinä ovat olleet tutkimuksen keskiössä.

Historiantutkimuksen ja tilastotieteen lähtökohdat ovat mitä ilmeisemmin hyvin erilaiset. Melliniä lainaten: "Tilastotieteen kannalta mikä tahansa reaali maailman ilmiöitä kuvaava numeeristen tai kvantitatiivisten tietojen järjestelmällinen kokoelma muodostaa tilastollisen aineiston ja siten tilastollisen tutkimuksen mahdollisen kohteen." Historiantutkija taas pysähtyy tällaisessa määritelmässä heti alkuun kysymään, miten reaali maailma pitäisi määritellä. Syventymättä tästä laajemmin aukeavaan problematiikkaan keskeistä on, että tilastotiede näyttää suun-

tautuvan ajallisesti toisenlaiseen perspektiiviin kuin historian tutkimus, nykyaikaan ja mahdollisimman ennustettavaan tulevaisuuteen. Tilastotieteen lähtökohtaotuksiin kuuluu, ettei ihminen voi toimia "maailmassa järkevästi, ellei hän pysty muodostamaan oikeata kuvaa maailmasta ja sen tilasta":

Nykyaikana oikeata kuvaa varten tarvitaan maailmaa ja sen tilaa *merkityksellisesti ja oikein kuvaavia, ajantasaisia (tilasto-) tietoja*. [...] Yhteiskunnan kaikilla sektoreilla toiminnan seuranta, päätöksenteko ja ennakointi perustuvat sekä yhteiskunnan eri sektoreita kuvaaviin (tilasto-) tietoihin että tilastollisiin menetelmiin.⁴⁰

Historiantutkimuksessa katse suuntautuu menneisyyteen ja suurin osa relevanteista kysymyksistä on katsottu sellaisiksi, ettei vastauksia ole mahdollista saavuttaa laskennallisin tai tilastollisin menetelmin. Raja on tosin alkanut hämärtyä siinä mielessä, että digitaalisista lähdeaineistoista voi tehdä sanahakuja hakukoneilla, jotka perustuvat laskennallisiin ja mahdollisesti tilastollisiin

39. Ilkka Mellin, Tilastolliset menetelmät: johdanto, Aalto-yliopisto, <https://math.aalto.fi/opetus/sovtoda/oppikirja/Johdanto.pdf> (18.3.2019), 9; Tilastokeskuksen Johdatus tilastotieteeseen -sivulla mainitaan lääketiede, yhteiskuntatieteet ja biologia. Johdatus tilastotieteeseen, Tilastokeskus, https://tilastokoulu.stat.fi/verkkokoulu_v2.xql?page_type=sisalto&course_id=tkoulu_tilaj&lesson_id=1&subject_id=2 (5.2.2019).

40. Ilkka Mellin, Tilastolliset menetelmät: johdanto, Aalto-yliopisto, <https://math.aalto.fi/opetus/sovtoda/oppikirja/Johdanto.pdf> (18.1.2019), 10–11. Lainausten lihavoimittu ja kursivoimittu Mellinin.

malleihin. Esimerkiksi Kansallisarkisto on ollut mukana hankkeessa, jossa on kehitetty käsinkirjoitetun tekstin tunnistusalustaa Transkribusta, jossa sanahaut (*Key Word Spotting*) perustuvat luotettavuusmatriisien tekemään merkkikohtaiseen indeksointiin.⁴¹ Laskennallisilla menetelmillä datasta voidaan saada esiin aineistoja ja tuloksia – vaikka tähän liittyy yhä yleensä isoja haasteita –, joita voidaan analysoida tarkemmin kvalitatiivisesti.

Historiantutkijoiden laskennallisia ja tilastollisia menetelmiä koskevat taidot ovat usein puutteellisia. Vastaukseksi ongelmaan on esitetty monitieteisiä tutkijaryhmiä,⁴² mikä on sinänsä oiva ratkaisu, mutta ei liene realistista olettaa, että useimmilla historiantutkijoilla olisi mahdollista hyödyntää tällaisia resursseja. Lisäksi laskennallisiin menetelmiin liittyy muitakin kuin taidollisia esteitä. Näitä voisi kutsua tieteenfilosofiksi esteiksi. Historiantutkimuksessa ja muilla humanistisilla aloilla on suosittu hermeneuttista menetelmää, jossa pyritään eläytyvään ymmärtämiseen.⁴³ Kun tutkimuksen osaksi otetaan laskennalliset datatieteen menetelmät, tutkija saattaa osin menettää otteensa tutkimuskohteesta.⁴⁴ Hän ei ymmärrä niitä matemaattisia ja tilastollisia malleja ja algoritmeja, jotka tuottavat aineiston tai tulokset hänen tulkittavakseen, jolloin niin sanottu hermeneuttinen kehä katkeaa.⁴⁵ Lisäksi tilastollisten menetelmien toteuttamistapaan liittyvät vaatimukset esimerkiksi kyselytutkimuksessa voivat tuntua joustavaan hermeneuttiseen metodiin tottuneesta jäykiltä verrattuna moniulotteiseen ja kompleksiseen historialliseen men-

neisyyteen. Esimerkiksi Kimmo Vehkalahti toteaa, että on usein ”yksinkertaistamisenkin uhalla syytä vaatia, että luokittelut ovat yksikäsitteisiä eli että eri vaihtoehdot ovat toisensa poissulkevia”⁴⁶ Korostan, ettei kyse ole hyökkäyksestä tilastollisia menetelmiä vastaan, vaan ennemminkin niistä rajoitteista ja näkökulmaeroista, joita historiantutkija voi kokea. Näitä eroja on tarkasteltava avoimesti, jotta historiantutkijat voivat paremmin sopeutua digitalisaation tuomiin muutoksiin. Sitä paitsi nimenomaan kyselytutkimuksen juuret ovat yhteiskuntatieteissä, joten sen voisi olettaa olevan historiantutkijoiden kannalta helpommin lähestyttävissä ja ymmärrettävissä oleva tilastollisten menetelmien soveltamisala, ainakin verrattuna tilastotieteen muihin osa-alueisiin.⁴⁷ Silti kyseisestä teoksesta löytyy useita kohtia, joita historiantutkija voi kyseenalaistaa tai ainakin ihmetellä, vaikka varmasti tässä on historioitsijakohtaisia eroja.⁴⁸ Toisaalta Vehkalahti tuo hyvin esiin myös sen, että tilastolliseen tutkimukseen voi liittyä keskeisellä tavalla kvalitatiivisia elementtejä, mikä on hyvä muistaa.⁴⁹

Historiantutkimuksen ontologiset lähtökohdat ja tiedonintressi ovat varsin erilaiset kuin tilastotieteessä. Tämä on hyvin olennaista tiedostaa, kun pohditaan digitaalisten ihmistieteiden menetelmien käyttöä historiantutkimuksessa, sillä jos tutkimuksellisiin lähtökohtiin liittyy tällaisia eroavaisuuksia, on selvää, että laskennallisiin menetelmiin perustuvia menetelmiä voi olla vaikea hyväksyä, omaksua ja käyttää. Kyse ei ole vain taidollisista esteistä vaan myös perustavanlaatuisista tieteenfilosofisista kysymyksistä.

41. Sebastian Colutto, Transkribus, University of Innsbruck, [https://arkisto.fi/uploads/Ajankohtaista/Presentation_on_Transkribus%2520\(Colutto\).pdf](https://arkisto.fi/uploads/Ajankohtaista/Presentation_on_Transkribus%2520(Colutto).pdf) (lainattu 28.12.2019).

42. Digitaaliset ihmistieteet (DIGIHUM) 2016–2019 ja 2020–2022, Suomen Akatemia, <http://www.aka.fi/DIGIHUM/> (26.3.2019); Matres, Oiva & Tolonen 2018, 43. Informaatiotutkimus arkisto, <https://doi.org/10.23978/inf.71160> (26.3.2019); myös Vehkalahti 2008, 40.

43. Hermeneutiikka, Tieteen termipankki, <https://tieteentermipankki.fi/wiki/Filosofia:hermeneutiikka> (lainattu 8.1.2019).

44. Lara Putnam, The Transnational and the Text-Searchable: Digitized Sources and the Shadows they cast. *American Historical Review*, huhtikuu 2016, 392.

45. Lisäksi laskennalliset menetelmät edellyttävät tietoa, joka on koodattu rakenteisesti tietokantoihin ja niin sanotusti putsattu rönsyistä, jotka ovat tyypillisiä historiallisille lähteille.

46. Vehkalahti 2008, 27–28.

47. Kimmo Vehkalahti, *Kyselytutkimuksen mittarit ja menetelmät*. Tammi 2008, 7; vrt. Ilkka Mellin, Tilastolliset menetelmät: johdanto, Aalto-yliopisto, <https://math.aalto.fi/opetus/sovtoda/oppikirja/Johdanto.pdf> (18.3.2019), 8.

48. Vehkalahti 2008. Esimerkiksi järjestämiseen perustuva mittautus, 30; miten perusjoukko ja otos -menetelmää, varsinkin satunnaisuuden vaatimusta, tulisi mahdollisesti yrittää soveltaa historiallisessa aineistossa, 43; ohjelmistojen oletusarvoinen toimintaperiaate, että ohjelmistot jättävät havainnon kokonaan pois, jos kohdasta uupuu yksikin analyysiin valittu muuttuja eli vain täydelliset havainnot käyvät, 69; miten puutteellisia tietoja voisi paikata tilastotieteellisiin säännöin historiantutkimusta koskevissa aineistoissa, 84.

49. Vehkalahti 2008, 17–19.

Miten tietokoneet vastaisivat paremmin esitettyihin kysymyksiin

Digitaalisen historian professori Tim Hitchcock toteaa digitaalisen historian symposiumissa Turussa vuoden 2018 lopulla, että nykyhetkellä digitaaliset menetelmät ohjaavat osaltaan tutkimustyötä. Hän kannustikin historiantutkijoita miettimään, miten tietokoneet saataisiin vastaamaan sellaisiin kysymyksiin, jotka kiinnostavat heitä itseään.⁵⁰ Hitchcockin näkemykseen on helppo yhtyä – ainakin periaatteessa. Ei ole vaikea keksiä toiveita ja ideoita, miten tietokoneiden toivoisi helpottavan ja nopeuttavan lähteiden käsittelyä ja läpikäymistä. Sen sijaan historian tutkijan on varsin vaikea pohtia käytännössä sitä, miten tietokoneet saataisiin todella vastaamaan häntä kiinnostaviin kysymyksiin, varsinkaan kun pohjalla ei ole tietojenkäsittely- ja tilastotieteen osaamista. Sitä paitsi historiantutkijoita kiinnostavat kysymykset ovat usein niin monimutkaisia ja -ulotteisia, ettei ole kovin realistista olettaa koneiden antavan niihin vastauksia, edes osaan niistä – ilman mittavaa datatieteen ammattilaisten vaivannäköä.

Jos lähdetäänkin pohtimaan Hitchcockin kehoitusta historiantutkimuksen (perus)lähtökohdista käsin, alkuun on nostettava esiin kysymys tutkimusaineistoista. Perusedellytyksenä voi pitää sitä, että tutkimuksessa käytettävien lähdeaineistojen tulisi olla digitoidussa muodossa. Eri muistiorganisaatiot, kuten arkistot, kirjastot ja museot, ovat digitoineet yhä laajemmissa määrin aineistojaan viime vuosina, mutta yhä suurin osa aineistoista ei ole saavutettavissa verkon kautta. Lisäksi digitoitujen lähteiden käytettävyys vaihtelee koskien laatua, metatietoja sekä sitä, onko lähde myös koneluettavassa muodossa. Histori-

antutkijoiden kannalta olisi oleellista, että muistiorganisaatioiden hakukoneet toimisivat mahdollisimman hyvin ja luotettavasti ja että aineistoja koskevat luettelot ja puunäkymät sekä metatiedot olisivat mahdollisimman kattavat, jotta tutkijat pystyisivät selvästi näkemään, ovatko kaikki tutkimuksessa vaadittavat lähdeaineistot digitaalisesti saavutettavissa – ja mieluiten koneluettavassa muodossa, mikä ei tietenkään ole realistinen oletus kaikkia digitoituja aineistoja koskien.⁵¹

Toinen edelliseen liittyvä perusedellytys koskee aineistojen digitointia niissä tapauksissa, joissa lähdeaineistot eivät ole saatavissa valmiiksi digitoidussa formaatissa. Nykyään aineistojen digitointi onnistuu varsin helposti myös tutkijan omilla laitteilla, kuten älypuhelimella tai digikameralla, mutta niistä muodostuu yleensä yksityiskokoelmia tutkijoiden omaan käyttöön. Työekonomisesti ja tutkimusaineistojen laajemman jatkokäytön kannalta tätä ei voi pitää ideaalina ratkaisuna, ja historiantutkijat voisivat yrittää kehittää digitointia koskien yhteisiä suuntaviivoja yhdessä muistiorganisaatioiden kanssa. Muistiorganisaatioista näyttäisi löytyvän valmiutta tähän, mutta vakiintuneita käytänteitä ei vielä ole muodostunut.⁵² Voitaisiinko esimerkiksi tutkimuksessa tarvittavia aineistoja digitoida yhteistyössä niin, että aineisto tallennetaan muistiorganisaation ja myöhemmin myös muiden tutkijoiden käyttöön? Entä miten metatiedot tulisi laatia? Ja miten erityiset tiettyyn tutkimuskysymykseen liittyvät datakokoelmat tulisi tallentaa ja säilyttää, jos ne on kerätty useammasta muistiorganisaatiosta?⁵³

Aineistonhallinnan ratkaisuihin liittyy monia kysymyksiä ja monia erilaisia ratkaisuvaihtoehtoja. Käytännön ongelmana on, että vain harvalla historiantutkijalla on käytettävissään riittävän

50. Heidi Kurvinen, Digihistoriantutkijat tapasivat Turussa. Tiivis kavalkadi uusimmista tuloksista ja käynnistyneistä projekteista. Turun yliopiston kulttuurihistorian blogisivusto. <https://kulttuurihistoria.wordpress.com/2018/12/11/digihistoriantutkijat-tapasivat-turussa-tiivis-kavalkadi-uusimmista-tuloksista-ja-kaynnistyneista-projekteista/> (12.3.2019); vrt. myös kysymyksenasettelun tärkeydestä Matres, Oiva & Tolonen 2018, 43. Informaatiotutkimus arkisto, <https://doi.org/10.23978/inf.71160> (26.3.2019).

51. Datan koneluettavuudella tarkoitetaan, että se on tietokoneella helposti käsiteltävissä. Tietokoneilla on yleensä ongelmia käsinkirjoitetun datan kanssa, mutta myös digitaalisissa formaateissa on muotoja, joita tietokone ei pysty käsittelemään, kuten esimerkiksi pdf-dokumenteissa. Keskeinen kriteeri koneluettavuudessa on tiedon rakenteellisuus, ja koneluettavina formaatteina pidetään csv-, json- ja Excel-tiedostoja. Avoin data, Tampereen yliopisto, <https://blogs.uta.fi/datajournalismi/datan-hankkiminen/avoin-data/> (28.2.2019).

52. Esim. Kumppanuudet, Kansalliskirjasto, <https://www.kansalliskirjasto.fi/fi/kumppanuudet> (27.3.2019).

53. Ratkaisu viimeksi mainittuun kysymykseen saattaisi olla Tietoarkisto, joka on yhteiskuntatieteellinen tutkimuksen ja opetuksen valtakunnallinen palveluinfrastruktuuri. Tietoarkisto, Tampereen yliopisto, <https://www.fsd.uta.fi/fi/> (12.5.2019).

ammattimaisia digitointivälineitä ja vain hyvin harvojen historiantutkijoiden taidot riittävät datan muokkaamiseen koneluettavaan muotoon.⁵⁴ Muistiorganisaatiot taas tuskin pystyvät irrottamaan riittävästi resurssejaan ja tietojaan tutkijoiden käytettäväksi digitoinnin ja datanhallinnan eri näkökulmien ratkaisemiseksi, joten ratkaisuja on haettava laajemmin tiedeyhteisön piirissä, kuten jo esimerkiksi DMPTuuli-aineistonhallintatyökalulla pyritään tekemään, mutta myös spesifimmin historiantutkimuksen lähtökohdista käsin.⁵⁵

Lähdeaineistoihin liittyvät digitoitinkysymykset ovat perusedellytyksiä, mutta myös varsinaisen kysymys siitä, miten tietokoneet vastaisivat paremmin historiantutkijoita kiinnostaviin kysymyksiin, on kiinteästi sidoksissa lähdeaineistojen digitoimisratkaisuihin; miten tietokoneavusteisesti ja datatieteen menetelmin pystytään hakemaan vastauksia tutkimuksen kysymyksenasetteluihin. Lähdeaineistojen tiedostomuoto vaikuttaa osaltaan siihen, mitä digitaalisia työkaluja voidaan käyttää ja millaisten kysymysten esittäminen mahdollistuu.⁵⁶

Jäykkyydestä notkeuteen – pragmaattisesti pienten askelten kautta

”Perinteinen taloustiede sijoittaa uskomattoman älykkäitä ihmisiä uskomattoman yksinkertaisiin tilanteisiin, mutta todellista maailmaa kuvaa paremminkin tilanne, jossa uskomattoman yksinkertaiset ihmiset koettavat selvittää uskomattoman monimutkaisista tilanteista.”⁵⁷ UCLA:n ekonomisti Axel Leijonhufvudin ironinen lausahdus tuntuu saavuttavan jotain olennaista myös tilanteesta, jossa menneisyyden monimutkaista todellisuutta yritetään raamittaa datatieteen ja ylipäättään digitaalisuuden mahdollistamiin menetelmiin. Lisäksi edelliseen liittyen voi vielä

lisätä sen, että historiantutkimukselle on ollut tyypillistä, että samassa tutkimuksessa käytetään useita erilaisia lähteitä, jolloin mahdolliseen digitointiin, datan putsaukseen ja eri menetelmiin liittyvät haasteet ja työmäärä luonnollisesti keräytyvät useampaan otteeseen.

Tim Hitchcockin haastetta siitä, miten tietokoneet vastaisivat paremmin historiantutkijoita kiinnostaviin kysymyksiin, saattaisikin olla käytännöllisintä pohtia historiantutkimuksen toteuttamisen lähtökohdista käsin eikä yrittämällä abstrahoida tietokoneavusteisten laskennallisten menetelmien soveltamista. Näin voi todeta ainakin niin sanotun tavallisen eli datatieteeseen perehtymättömän historiantutkijan näkökulmasta, jolla ei ole mahdollisuutta hyödyntää monitieteisen tutkimusryhmän osaamista. Tällainen eteneminen mahdollistaa sen, että kysymyksenasettelu määrittää tutkimuksessa käytettäviä aineistoja ja mahdollisia tietokoneavusteisia menetelmiä. Alkuun voisi esittää niinkin triviaalin kysymyksen kuin mitä hyötyä kysymyksenasettelun kannalta on, että lähdeaineisto on digitaalisessa muodossa. Ja jos aineisto on jo valmiiksi digitaalisessa muodossa, niin millaisilla työkaluilla sitä voi käsitellä sen alkuperäisessä verkkoympäristössä. Muistiorganisaatioiden hakutyökalut saattavat olla niin pitkälle kehitettyjä, että on vaikea vetää rajaa siihen, onko kyseessä vain niin sanottu hakukone vai digitaalinen työkalu. Esimerkiksi Kansalliskirjaston Digin haku on todennäköisesti tällä hetkellä monien historiantutkijoiden näkökulmasta toimivampi ja käytännöllisempi kuin monet varsinaisiksi digityökaluiksi katsotut sovellukset.⁵⁸ Voidaan siis suorittaa empiiristä kokeilua, miten esimerkiksi tietokantojen hakukoneet toimivat, sillä hakukoneen toiminnan oppii parhaiten käyttämällä, kokeilemalla eri hakutapoja.⁵⁹

54. Kts. Arkistolaitoksen suositus digitoinnin laatukriteereiksi 15.8.2008, Arkistolaitos, <https://www.arkisto.fi/uploads/normit/valtionhallinto/suosituksset/digikriteerit.pdf> (25.2.2019).

55. DMPTuuli on aineistonhallinnan suunnittelua tukeva verkkopalvelu, joten se ohjaa vain suunnittelemaan aineistonhallintaa, eikä sinänsä anna historiantutkijoille valmiita välineitä tai ratkaisuja aineistonhallintaa koskeviin päätöksiin, kuten esimerkiksi aineiston pitkäaikaissäilytykseen. Data Management Planning tool DMPTuuli, Helsingin yliopisto, <https://datasupport.helsinki.fi/services/data-management-planning-tool-dmptuuli> (15.3.2019).

56. Vrt. Tolonen & Lahti 2018, 239.

57. Sitaatti on lainaus Juha Siltalan teorioita historiantutkimuksessa käsittelevästä artikkelista. Juha Siltala, Teoriat historiantutkimuksen työkaluina ja tutkimustuloksina. Teoksessa Mirka Danielsbacka, Tuomas Tepora & Matti Hannikainen (toim.) *Menneisyyden rakentajat. Teoriat historiantutkimuksessa*. Gaudeamus 2018, 72.

58. Kts. Haku digitaaliset aineistot, Kansalliskirjasto, <https://digi.kansalliskirjasto.fi/search> (20.3.2019).

59. Verkon rakenne, Jyväskylän yliopisto, https://kielikompassi.jyu.fi/resurssikartta/netro/pankki/parametrit_digi_rak.shtml (26.2.2019).

Historiantutkimuksessa pystyttäisiin ottamaan merkittävä digiloikka pelkästään, mikäli eri muistiorganisaatioiden tietokannoista pystyisi tekemään luotettavia ja kattavia asiasanahakuja – niin metatiedoista kuin sisällöistä. Entistä paremmat hakutyökalut edesauttaisivat ja nopeuttaisivat tutkimuksenteon edellytyksiä merkittävästi.⁶⁰ Näin pitkällä ei kuitenkaan olla, vaikka lupaavia työkaluja on kehitteillä, kuten edellä mainittu Transkribus-alusta, jossa Kansallisarkisto on mukana.⁶¹ Käytännössä tulosten saavuttaminen vaatii vielä varsin pitkällistä datan käsittelyä ja muokkaamista, mikä puolestaan edellyttää valtavasti aikaa ja resursseja.⁶² Tästä taas seuraa, että digitaaliseen historiantutkimukseen sisältyvä käytännön tasolla monenlaisia niin sanottuja jäykkyselementtejä, jotka rajoittavat yksittäisten historiantutkijoiden tutkimusmahdollisuuksia. Pragmaattisesti kokeilemalla, kuten eri hakusanoja ja yhdistelmiä tietokantoihin soveltamalla, osa näistä jäykkyyksistä lienee mahdollista ylittää. Etuna on myös se, että tällaisessa menetelmässä historiantutkija pystyy itse kontrolloimaan haku- ja tutkimusmenetelmien käyttöä eikä hänen tarvitse turvautua tietojenkäsittely- tai tilastotieteen ammattilaisiin. Ideaalitapauksissa historiantutkijalla olisi tietysti mahdollisuus halutessaan turvautua datatieteen menetelmiin, käytännössä datatieteen ammattilaisten tukeen, mutta resursseja tällaiseen ei usein ole. Pragmaattinen, käytännönläheinen digitaalisten aineistojen lähestyminen voikin olla yksi keino kuroa umpeen kuilua his-

toriantutkimuksen ja digitaalisten ihmistieteiden välillä, mutta sekään ei ole oikotie onneen vaan vaatii aikaa, kokeiluja, hakuyhdistelmien dokumentointia ja jälkitarkastuksia.

Digitalisaatio ja siihen liittyvät eri ilmiöt, kuten tietokoneavusteisten tutkimusmenetelmien nousu, ovat monella tapaa haasteellisia asioita historiantutkimuksen ja -opetuksen näkökulmasta, mutta siihen liittyy myös mielenkiintoisia mahdollisuuksia. Vielä on liian aikaista sanoa, miten voimakkaasti erilaisten digitaalisten, laskennallisten metodien käyttö vaikuttaa tai yleistyy historiantutkimuksessa, tai miten pitkälle datatieteen menetelmät sulautuvat niin sanotuksi normaaliksi osaksi historiantutkimuksen kenttää. Selvää kuitenkin on, että digitalisaatio tulee vaikuttamaan voimakkaasti akateemisen historiantutkimuksen ja -opetuksen käytäntöihin. Siksi on äärimmäisen tärkeää, että historioitsijat itse osallistuvat sekä digitalisaatiota koskeviin akateemisiin keskusteluihin että oman alansa käytäntöjen muokkaamiseen. Tällöin se voi tapahtua historiantutkimuksen ja -opetuksen ehdoista käsin eikä niiden tieteenalojen ja toimijoiden ehdoilla, joille tietojenkäsittelytieteen ja ylipäätään datatieteen menetelmät soveltuvat käytännössä ja myös tieteenfilosofisessa mielessä paremmin.

FT **Kati Katajisto** on historiantutkija Helsingin yliopistossa. **Sähköposti:** kati.katajisto@helsinki.fi

60. Kimmo Kettunen, Tuula Pääkkönen & Mika Koistinen, Kansalliskirjaston digitoitu historiallinen lehtiaineisto 1771–1910, <https://journal.fi/inf/article/view/59433/22433> (27.2.2019).

61. Recognition and Enrichment of Archival Documents, Kansallisarkisto, <https://www.arkisto.fi/fi/kansallisarkisto/hankkeet/read> (10.5.2019).

62. Tolonen & Lahti 2018, 239.