

REETTA EIRANEN

Kirjallisuudentutkija, joka haluaa rikkoa kirjan

Laura Mandell ja itsekriittinen ihmistiede

”Ihmisten pitäisi pystyä psykologisesti mukautumaan digitaalisten multigraafien kirjoittamiseen”, visioi englannin kielen professori Laura Mandell. Teknologia ei ole kynnyskysymys yhteisöllisten työskentelytapojen omaksumisessa, mutta digitaalisten projektien vertaisarvioinnin puutteet ovat iso ongelma: vaaditaan lisää avoimuutta, jotta tulokset voidaan toistaa ja kyseenalaistaa. Mandell johtaa Texasin A&M-yliopistossa digitaalisiin ihmistieteisiin, mediaan ja kulttuuriin keskittyvää hanketta ja ihastelee ”kansalaistutkijoiden” aktiivisuutta kulttuuriperinnön dokumentoinnissa. Uusimmassa teoksessaan *Breaking the Book* (2015) hän käsittelee kirjan historiaa ja digitaalista vallankumousta pohtien, miten väline vaikuttaa ajatteluun ja ihmistieteelliseen tutkimukseen¹.



Akateeminen taustanne on kirjallisuuden ja 1700-luvun tutkimuksessa. Kumppaako kiinnostus digitaalisiin ihmistieteisiin tutkimusteemoistanne, vai oletteko aina ollut innostunut digitaalisesta ja sen mahdollisuuksista?

Kaltaistani 1700-luvun ja brittiromantiikan naiskirjoittajien tutkijaa kiinnosti internetissä mahdollisuus levittää naisten kirjoittamia tekstejä, joihin oli muuten vaikea päästä käsiksi. Ongelmaa ratkovista hankkeista varhaisimpia ja parhaita oli Women Writers Project Online². Projekti julkaisi tieteellisiä editioita avoimesti verkossa ja oli edelläkävijä digitaalisten tekstien merkintäkoodaamisen menetelmien kehittämisessä. Niiden avulla kirjastot saattoivat arkistoida editioita, jolloin teknologiset muutokset eivät hävittä tekstejä eivätkä tee niistä käyttökelvottomia. Osallistuin kirjastojen järjestämiin työpajoihin, ja nyt minulla on riittävä tietämys yhdestä merkintäkoodista ja ohjelmointikielestä, jota käytetään koodin muuntamiseen useille alustoille.

Olette kirjoittanut historiallisesta näkökulmasta siitä, miten viestintäväline vaikuttaa ajattelutapaamme. Miten vertailisitte siirtymää käsinkirjoituista teksteistä painettuun joukkoviestintään ja nykyistä digitaalista murrosta tavassamme omaksua ja käsitellä tietoa?

Nähdäkseni painetut joukkoviestimet toivat kirjoittajan tyyliin tiettyä anonyymisyyttä, josta sittemmin tuli suorastaan välttämätöntä humanistisilla tieteenaloilla. Anonyymien asiantuntijan äänellä kirjoittaminen poisti esteitä ajattelulta. Myönteisten vaikutusten lisäksi muutoksella oli ainakin yksi kielteinen seuraus: tunnis-

Kuva: Laura Mandell

tamaton arvovaltainen ääni ei ollut vastuussa omasta näkökulmastaan.

Nykyinen siirtymä digitaaliseen kulttuuriin ei ole juurikaan muuttanut asetelmaa: suurin osa digitaalisesta tutkimuksesta levitetään pdf-muodossa, joka on pelkkä kuva kirjasta tai journalista. Vaihtoehtoisia tutkimustapoja voidaan kuitenkin kehittää ja tukea: esimerkiksi monografijulkaisuja voidaan korvata virtuaalisilla tutkimusympäristöillä. Digitaalisissa ympäristöissä anonyymien äänien määrä voi moninkertaistua, jolloin ne voitaisiin ainakin luonnontyyppitellä, jos ei yksilöidä (ääni opittaisiin tuntemaan vaikkapa ”skeptikkona” vaikkei sillä olisi kutsumanimeä).

En tiedä, millaista ajattelua tämä mahdollistaa, mutta toivoakseni sisäiset ristiriitaisuudet voitaisiin rutiininomaisesti välttää pitämällä yllä itsetietoisuutta. Yksinkertaisemmin sanottuna akateeminen kirjoittaminen ja ajattelu voisi muuttua maanläheisemmäksi. Voi kuulostaa hieman paradoksaaliselta, mutta kun kirjoittamisesta tulee virtuaalisempaa – kun digitaalisen median mahdollisuudet lunastetaan täysimääräisesti – ajattelusta tulee maanläheisempää.

Olette itse perustanut kaksi digitaalista arkistoa, The Poetess Archiven ja 18thConnectin³. Millainen tehtävä tällaisilla tietokannoilla on tutkimuksessa? Helpottavatko ne vain lähdeaineistojen löytämistä, vai vaikuttavatko ne perustavammin itse tutkimukseen, esimerkiksi tutkimuskysymyksiin ja -sessiin?

Ne vaikuttavat molemmiin tavoin, sillä on vaikeaa erottaa lähteiden saatavuutta tutkimuksesta: syvälinen tutkimus ei voi mennä eteenpäin ilman korkealaatuisia, helposti saavutettavia aineistoja. Ihannetilanteessa The Poetess Archive edistäisi laadukkaiden tekstien saatavuutta, mutta olemme pahasti jäljessä julkaisemisessa ja kehittämisessä. Parhaassa tapauksessa The Poetess Archive ja vastaavat sivustot tuovat tekstejä sekä digitaalisten toimittajien että tiedonlouhijoiden saataville. Tekstin merkintäkoodaus on vapaasti saatavilla niin, että kuka tahansa voi ottaa sen käyttöön ja luoda valmiin tieteellisen edition nimekkeistä, jotka on hienovaraisesti transkriptoitu, tai tehdä hakuja aineistosta ohjelmointirajapinnan avulla.

18thConnectin tarkoitus on samanlainen. Sen välityksellä pääsee käsiksi teksteihin, mutta lisäksi se pyrkii Early Modern OCR Projectin avulla parantamaan koelmien tekstien laatua. TypeWright Toolia käyttämällä toimittajat tai toimitusryhmät voivat korjata tekstiin jääneitä virheitä, joita optinen merkintunnistus (*optical character recognition*, OCR) voi tuottaa käsiteltäessä koelmien kuvia mekaanisesti.⁴

Miten itse uskotte kirjan muuttuvan digitaalisella aikakaudella? Missä mielessä se säilyy ennallaan ja korvaako jokin muu painotuotteet? Miten kirjaa voisi kehittää ja parantaa esimerkiksi digitaalisilla editioilla, arkistoinneilla ja datan visualisoinneilla?

Ajattelevalla maailmalle olisi eduksi, jos tieteelliseen kirjoittamiseen sitoutuneet tutkijat työskentelisivät yh-

dessä virtuaalisissa tutkimusympäristöissä, joissa argumentteja kehitellään parhaiden vertaisarviointiprosessin tapaan. Tällaisten virtuaalisten tutkimusympäristöjen pitäisi luoda edellytykset monilukuiselle tekijyydelle laajan, viitteistetyn kommentoinnin avulla. Näin syntyisi muutakin kuin pelkkiä huomautuksia, jotakin jo itsessään vertaisarvioitua. Ympäristön täytyisi myös ehdottomasti mahdollistaa monenlaiset lainaukset, jotka sisältävät viitatun materiaalin: tekstejä, videoita, jopa tiedonlouhinnan interaktiivisia tuloksia, joka sallisi niiden muuntelun ja kommentoinnin eri tallennusversioina. Geoffrey Rockwell ja Stéfan Sinclair ovat jo onnistuneet sisällyttämään edelleen käsiteltävää dataa artikkeleihin⁵. On myös olemassa varsin kehittyneitä muistiinpanojärjestelmiä⁶.

Ihmisten pitäisi pystyä psykologisesti mukautumaan digitaalisten multigraafien kirjoittamiseen, mikä on itse asiassa paljon työläämpää kuin teknologian kehittäminen sitä varten. Sopeutuminen vaatisi, että tekijyys kuvitellaan pikemminkin osallistumiseksi tiedon rakentamiseen kuin jo tiedetyn lausumiseksi. Tämä edellyttäisi vertaisarvioivia yhteisöjä. Tarvittaisiin myös ammatillinen palkkiojärjestelmä, jossa panostukset vertaisarvointiin arvostettaisiin yhtä korkealle kuin laadukkaat artikkelit tai kirjat. Toivon yhä tämän mahdollisuuden toteutumista.

Digitaalisia ihmistieteitä on toisinaan kritisoitu digitaalisten ja laskennallisten menetelmien epäkriittisen ja pelkästään välineellisen käytön edistämisestä. Toisaalta olette huomauttanut, että esimerkiksi kriittiset editiot (vaikkapa runoudessa) eivät toimi vanhassa painetussa muodossa ja uusien ohjelmistojen kehittäminen auttaisi lievittämään ongelmia. Voisivatko digitaaliset ihmistieteet mahdollistaa uudenlaista kriittistä tutkimusta? Voisivatko ne jopa tarjota menetelmiä tai työkaluja kriittiselle ajattelulle ylipäätään?

D. F. McKenzien kaltaiset bibliografisen tutkimuksen asiantuntijat ovat osoittaneet, että tekstieditioiden epäkriittinen hyväksyminen on johtanut kirjallisuuden tutkimuksessa virhetulkintoihin. Kriittiseen ajatteluun pitää kuulua lähteiden tarkastelu. Tieteelliset digitaaliset laitokset voivat parantaa perustavanlaatuisesti eri lähteiden saavutettavuutta ja helpottaa niiden vertailua. Käsikirjoitusten vertaaminen editioihin tai useisiin eri laitoksiin on kirjojen kohdalla edellyttänyt tieteellisen merkintätavan kehittämistä. Järjestelmää voitaisiin merkittävästi parantaa pyörittämällä sitä tekstin pintatason takana sen sijaan, että se muodostaa tekstin esillepanon. Toisin sanoen: kun luetaan alaviitteitä, joissa esitetään sanojen eroavuudet eri editioissa, lukija joutuu luomaan yhteydet päässään. Verkkosivulla voi sen sijaan esittää rinnakkain teksteistä useita versioita, jotka ovat kaikki näkyvissä luettavina kokonaisuuksina⁷.

Kuten McKenzie tunnetusti osoitti analyysillaan Wimsattin ja Beardsleyn epäkriittisestä tukeutumisesta yhteen editioon, lähilukeminen hyötyy editoiden vertailutyökalujen saatavuudesta⁸. Etälukemisella vaikuttaa

kuitenkin olevan eniten annettavaa kriittisen ajattelun välineiksi⁹. Väittäisin silti, että etälukemisen täytyy aina hyödyntää lähilukemista analyysin ratkaisevilla hetkillä. Kuten Toril Moi esittää teoksessaan *Revolution of the Ordinary*, ihmistieteiden kohde ovat yksittäistapaukset¹⁰. Oma tutkimukseni data-analyysin ja tapaushistorioiden suhteesta toivoakseni osoittaa, että kirjallisessa analyysissä on välttämätöntä yhdistää lähi- ja etämenetelmä¹¹. Työni myös näyttää, miten se tyypillisesti tehdään.

Voivatko digitaaliset ihmistieteet edistää valistuksen ihannetta jaetusta julkisuudesta tuomalla akateemiset keskustelut lähemmäs ihmisiä tai parantamalla kaikkien kansalaisten kykyjä ja mahdollisuuksia osallistua järjen julkiseen käyttöön?

Kyllä, ehdottomasti. Yleisölle suunnattujen painettujen julkaisujen ongelma on, että kirjoittaja-lukijasetelma ylläpitää hierarkiaa, jossa intellektuelli hallitsee yleisöä: lukijat myöntyvät passiivisesti kirjoittajan auktoriteettiin. Uudet merkittävät joukkoistamishankkeet edellyttävät aktiivisia lukijoita: ne värväävät kansalais-tutkijoita kulttuurisen perinnön säilyttämiseen ja analysoimiseen. Ensimmäinen iso ilmiö tällä alalla oli Rose Holleyn vetämä australialaisten sanomalehtien digitoitihanke, jossa yleisö kirjoitti puhtaaksi kolme miljoonaa lehtijuttua projektin kahden ensimmäisen vuoden aikana¹². Lontoon UCL-yliopiston hanke ”Transcribe Bentham” puolestaan osallisti kansalaistutkijoita kirjoittamaan puhtaaksi 4 364 täysimittaista käsikirjoitusta kahdessa ja puolessa vuodessa¹³. Vaadittava työ ei ole ulkokohtaista, mekaanista toistoa. Omaa projektiaan Irlannissa pyörittänyt Susan Schreibman on todennut, että puhtaaksikirjoittajat kirjoittavat uudelleen menneisyyden tapahtuman elvyttämällä julkisista muisteluritualeista häivyttetyt ”eletyn elämän sotkuisen ja monimutkaisen mosaikin”¹⁴. Yleistäen: kun kansalaistutkijat osallistuvat menneisyyden dokumenttien säilyttämiseen, he ovat tekemisissä alkuperäisten historiallisten aineistojen kanssa samalla tavalla kuin tutkijat. Näin he kohtaavat sotkuisen menneisyyden, joka täytyy pelkistää yleisiksi historiallisiksi väittämiksi. He näkevät itse, miten historiaa pelkistetään. Mikään ei ole valeutisten aikakaudella yhtä tärkeää kuin osallistaa yleisö analysoimaan, miten historiallisia tosiasioita tehdään ja miten moninaisista eri näkökulmista niitä voidaan tuottaa eri tavoin.

Miten digitaalisten ihmistieteiden kenttä on muotoutunut niinä liki 25 vuotena, joina olette ollut aktiivisesti mukana? Mitä sopii odottaa seuraavalta neljännesvuosisadalta? Onko humanististen tieteiden institutionaalille ja tutkimukselliselle rakenteiden uudistamiselle mahdollisuutta, lupausta tai jopa tarvetta?

Vaikka itse vihkiydyin ”digitaalisiin ihmistieteisiin” internetin alkuaikoina 1990-luvulla, sitä edeltäneellä ja muovanneella alalla, ”laskennallisilla ihmistieteillä”, on rikas historia¹⁵. Internet avasi kentän kaikenlaisille osanottajille. Viimeisen viiden tai kymmenen vuoden aikana tiedonlouhinta ja kvantitatiivinen analyysi ovat olleet trendi digitaalisissa ihmistieteissä – ala saattaa keskittyä jälleen ohjelmointiin. Minulle tekstinlouhintaan osal-

listuminen on erityisen vaikeaa: pääasiallisesti toimitan digitaalisia tekstejä, ja käytän vain tieteellisten digieditoitoiden luomiseen tarvittavia ohjelmointikieliä¹⁶. Tämän uuden kehityksen ydin on, että kunnollisen vertaisarvioinnin varmistamiseksi prosesseista täytyy tehdä avoimia. Prosessien täytyy olla saatavilla, jotta tuloksia voidaan toistaa ja kyseenalaistaa. Prosesseja ja dataa pitää voida uudelleenarvioida, jotta vasta-argumentteja voidaan esittää.

Vertaisarvioinnin mekanismien kehittäminen digitaaliseen tutkimukseen on ollut hidasta ja vaivalloista. Raportit tekstinlouhinnan tuloksista jäävät liian usein teknisesti tarkastamatta: villin lännen lait ovat jossain määrin edelleen vallalla. Kirjallisuuden ja kulttuurin tutkimuksen alalla merkittävässä *Representations*-lehdessä on julkaistu artikkeleita, jotka esimerkiksi esittävät *Eighteenth-Century Collections Online* tiedonlouhintaan perustuvia väitteitä, vaikka tietokannan kokotekstihaku ei ole vielä teknisesti riittävän kehittynyt luotettavien tulosten saamiseksi.

Institutionaalisia mekanismeja digitaalisten projektien vertaisarviointiin tarvitaan siksikin, että tohtoreita palkataan ”digitaalisina humanisteina” mutta heitä ei vakinaisteta sen työn perusteella, josta heille maksetaan. Toisin sanoen palkkaamisen ja vakinaistamisen välillä on ylittämätön kuilu. Nämä kaksi ongelmaa kertovat humanististen alojen epätasaisesta uudistumisesta digitaalisen tutkimuksen alkuvaiheissa.

Olen luonnehtinut 1900-luvulle tyypillisissä instituutioissa harjoitettuja humanistisia tieteitä ”printti-ihmistieteiksi”. Totta kai esitetään videoita elokuvakritiikkiä, puhutaan maalauksista maalauksin ja niin edelleen, mutta viimeisten vuosisatojen ajan tieteellisen kommunikaation tärkein kanava on ollut painettu sana. 2000-luvun ihmistieteet ovat digitaalisia, halusimme sitä tai emme. Olen huolissani siitä, että jos perinteiset humanistit eivät liittoudu digitaalisuuden tutkijoiden kanssa, tutkijat eivät itse pääse valitsemaan käyttämäänsä teknologioita. Digitaalisten teknologioiden omaksumisen pitää olla itsekriittistä. Täytyy kääntyä koodareiden ja ohjelmoijien puoleen ymmärtääksemme, mitä olemme tekemässä ja miksi.

Vitteet & Kirjallisuus

- 1 Laura Mandell, *Breaking the Book. Print Humanities in the Digital Age*. Wiley, Hoboken 2015.
- 2 wwp.northeastern.edu
- 3 Ks. idhmcmain.tamu.edu/poetess; 18thconnect.org
- 4 18thConnectissa ovat mukana esimerkiksi Early English Books Online (EEBO) ja Eighteenth-Century Collections Online (EEO). Early Modern OCR Projectista ks. emop.tamu.edu. TypeWright Toolista ks. 18thconnect.org/typewriter/documents.
- 5 Ks. Geoffrey Rockwell & Stéfan Sinclair, *Hermeneutica. Computer-Assisted Interpretation in the Humanities*. MIT Press, Cambridge, Mass. 2016. Projektin nettisivuilta voi ladata verkkoversioina kirjan käytännölliset luvut, joihin sisältyy tekstin analysointiin kehitetyllä Voyant-työkalulla tehtyjä dynaamisia paneeleja. Ks. hermeneuti.ca.
- 6 Ks. hypothes.is.

- 7 Ks. Juxta Commons, juxtacommons.org.
- 8 Kirjallisuustieteilijät William K. Wimsatt (1907–1975) ja Monroe C. Beardsley (1915–1985) kuuluivat uuskritiikin perustajiin, joille oli keskeistä tekstien lähilukeminen niiden tekijän biografian tai intentioiden rekonstruoinnin sijasta. McKenzie osoittaa kuitenkin teoksessaan, että Wimsattin ja Beardsley'n kuuluisin essee ”The Intentional Fallacy” (1946) sisältää nolon filologisen kömmähdyksen, joka johtuu hänen mukaansa heidän taipumuksestaan ylenkatsoa typografian merkitystä. Essee alkaa epigrammilla, jossa lainataan väärin William Congrevea. Muiden virheiden ohella Wimsatt ja Beardsley jopa sekoittavat keskenään verbit *wrought* ja *wrote*. Tämä kaikki vie pohjan heidän tulkinnahtaan. D. F. McKenzién, *Bibliography and the Sociology of Texts*. Cambridge University Press, Cambridge 1999, 18–20.
- 9 Ks. esim. *Journal of Cultural Analytics*. Verkossa: culturalanalytics.org.
- 10 Toril Moi, *Revolution of the Ordinary: Literary Studies after Wittgenstein, Austin, and Cavell*. University of Chicago Press, Chicago, 2017.
- 11 Gender and Cultural Analytics: Finding or Making Stereotypes? Teoksessa *Debates in Digital Humanities* (tulossa); Quantitative Literary Analysis and Case Histories. Esitelmä ”Limits of the Numerical” -tapaamisessa, University of Chicago Franke Institut, 26.–30.7.2017.
- 12 Yhteistyöprojektista kansalliskirjaston kanssa, ks. Australian Newspapers Digitisation Program: www.nla.gov.au/content/newspaper-digitisation-program; ks. myös Rose Holley, *Many Hands Make Light Work. Public Collaborative OCR Text Correction in Australian Historic Newspapers*. National Library of Australia, 2009.
- 13 Hanke kokoa brittifilosofi Jeremy Benthamin (1748 – 1832) ennen julkaisemattomia käsikirjoituksia koottujen teosten pohjaksi. Ks. blogs.ucl.ac.uk/transcribe-bentham/
- 14 Muistiorganisaatioiden ja yksityiskokoelmien hallussa olevaa kirjeenvaihtoa vuosisadan takaa tallentanut hanke ”Letters of 1916” (2013–) on edennyt vuoteen 1923 asti, ks. lisää: letters1916.maynoothuniversity.ie. Lainauksesta, ks. Julia Schreibman, Vinayak Das Gupta & Neale Rooney, Notes from the Transcription Desk. Visualising Public Engagement. *English Studies*. Vol. 98, No. 5, 2017, 506–525.
- 15 Ks. erit. Humanist-keskustelulista, jossa *humanities computing* oli vahvasti edustettuna: dhumanist.org
- 16 XSLT eli eXtensible Stylesheet Language Transformations.