

KULTTUURIHISTORIAN
TUTKIMUS
Lähteistä menetelmiin
ja tulkintaan



Toimittaneet Rami Mähkä, Marika Ahonen, Niko Heikkilä,
Sakari Ollitervo & Marika Räsänen

CULTURAL HISTORY – KULTTUURIHISTORIA 17

Kustantaja Kulttuurihistorian seura, Turku
ISBN 978-952-68776-8-6 (pehmeäkantinen)
ISBN 978-952-68776-9-3 (PDF)
ISSN 1458-1949

Kulttuurihistorian seura
c/o Kulttuurihistoria
20014 Turku
<http://kulttuurihistoria.net>

© kirjoittajat
Ulkoasu: Henri Terho ja Heli Rantala
Taitto: Heli Rantala

Kannen kuva: Valtionarkiston (nyk. Kansallisarkisto) tutkijasali tarjosi historian-
tutkijoille ajanmukaiset puitteet vuonna 1901. Kuvaaja: Atelier Nyblin A.B. Lähde:
Museovirasto, Historian kuvakokoelmat, HK19650527:2. CC BY 4.0.

Painopaikka: BoD – Books on Demand, Norderstedt, Saksa 2022



VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Digitaaliset tutkimusmenetelmät kulttuurihistorian tutkimuksessa

SEMANTTINEN WEB JA VERKOSTOANALYYSI

Kimmo Elo ja Harri Kiiskinen

Viimeisen vuosikymmenen aikana digitalisaatio on laajentunut myös tieteenaloille, joissa ei ole ainakaan laajassa mittakaavassa hyödynnetty suuria aineistomassoja, tilastollisia aineistoja tai menetelmiä eikä tietokoneavusteisia menetelmiä tai työkaluja. Tämän, tässä artikkelissa digitaalisiksi ihmistieteiksi (*digital humanities*) määritellyn paradigman juuret voidaan johtaa kielitieteisiin, jossa luotiin perusta muun muassa nykyisin laajasti sovelletulle tietokone-lingvistiikalle.¹

Digitaaliset ihmistieteet ovat myös herättäneet vilkasta keskustelua historiantutkijoiden piirissä. Vastakkain on asetettu käsitykset historiantutkimuksesta idiografisena – tapahtumafokusoituneena – ja nomoteettisena – yleisiä säännönmukaisuuksia löytämään pyrkivänä – tieteenä.² Idiografisen käsityksen mukaan formaalit mallit ja teoriat eivät kuuluisi historiatieteiden työkaluihin. Taustalla vaikuttaa myös syvempi tieteenalaidentiteetin jakolinja luonnontieteille ominaisen tut-

¹ Esim. Kirschenbaum 2012; Elo 2016a.

² Niiniluoto 2001.

kimusryhmäkeskeisyyden ja ihmistieteiden yksittäisen tutkijan roolia aineistojen tuntijana ja tulkitsijana korostavan näkemyksen välillä.

Yksi digitaalisten ihmistieteiden ”ikuisuuskymsyksistä” on, onko kyse menetelmä- vai ilmiötieteestä. Toisin kysyen: onko kyse aineistojen tietokoneavusteisesta analysoinnista vai tutkitaanko digitalisaatiota (historiallisena tai yhteiskunnallisena) ilmiönä ilman vaatimusta digitaalisten työkalujen käytöstä. Vaikka todellisuudessa jakolinja ei ole näin mustavalkoinen³, määrittelimme oman näkökulmamme menetelmätieteeseen painottuvaksi nojautumalla Fitzpatrickin määritelmään digitaalisista ihmistieteistä ihmistieteille ominaisten kysymysten tutkimisesta tietojenkäsittelytieteiden menetelmiä hyödyntämällä.⁴

Artikkelissa esittelemme kaksi digitaalisten ihmistieteiden menetelmää, tavoitteenamme avata lukijalle digitaalista tutkimusprosessia ja antaa eväitä soveltaa digitaalisia välineitä omissa tutkimuksissa. Valitut menetelmät ovat (1) *semanttisen webin* soveltaminen historiantutkijan työkaluna ja (2) *historiallinen verkostoanalyysi*. Artikkelin havainnollistaa konkreettisten empiiristen tapaustutkimusten avulla, millaisia tutkimuskysymyksiä voidaan tutkia, miten aineisto tulisi valita ja valmistella sekä millaisia tuloksia menetelmien avulla voidaan saada. Artikkelin päättävä loppukeskustelu kokoaa yhteen esiteltyjen menetelmien vahvuudet ja heikkoudet kulttuurihistoriallisessa tutkimuksessa.

Semanttinen web ja historiantutkimus

Tässä osiossa esitellään niin sanotun semanttisen verkon tai linkitetyn datan mahdollisuuksia historiantutkijan työkaluna. Mitä seuraavassa esitetään, poikkeaa jossain määrin esimerkiksi Eero Hyvösen esittämistä semanttisen webin käytön periaatteista.⁵ Hyvönen työryhmineen tuottaa semanttista mallinnusta aineiston pysyvään tallennukseen esi-

³ Elo 2016b.

⁴ Fitzpatrick 2012.

⁵ Laaja suomenkielinen yleisesitys semanttisesta verkosta on Hyvönen 2018, ja kirjaan tutustuminen on käytännössä paras reitti tässä esiteltyjen menetelmien käyttöönottoon.

merkiksi kulttuuriperintöaineistoa varten.⁶ Heidän tarkoituksenaan on tuottaa pysyviä malleja laajaan käyttöön. Sen sijaan tässä artikkelissa käsitellään semanttista webiä tutkimusprojektien näkökulmasta, ajallisesti, käyttäjällisesti ja institutionaalisesti rajattuun tarkoitukseen kehitettävänä työkaluna, jonka tarkoitukskaan ei ole olla yleispätevä malli vaan tehokas työväline.

Aineiston muuntaminen semanttisen webin tarkoittamaan muotoon tarkoittaa seuraavaa: aineisto esitetään yksinkertaisina väitteinä käyttäen käsitteitä, joiden merkitys on ennalta sovittu ja joiden väliset suhteet määritellään mahdollisimman tarkasti. Tarkoituksena on synnyttää väitelauseiden joukko, jota on mahdollista käsitellä ohjelmallisesti. Semanttisessa kuvailussa käytettävä kieli on hyvin yksinkertaista. Se rakentuu ”subjekti predikaatti objekti” -lauseille, joita toistamalla asia kuvataan. Toisaalta mitä tahansa näistä lauseenjäsenistä voidaan käyttää myös muissa rooleissa, esimerkiksi voimme myös sanoa ”predikaatti predikaatti2 objekti”. Voimme kuvata eri entiteettien välisiä monimutkaisiakin suhdeverkostoja itse valitsemillamme käsitteillä, joiden keskinäisiä suhteita voidaan kuvata standardoiduilla RDF/RDFS ja OWL-skeemoilla.⁷

Semanttisessa aineistonkäsittelyssä on tavallaan kyse universaalien kielen synnyttämisestä. Tämä ei ajatuksena ole varsinaisesti uusi, ja itse asiassa hyvinkin suuri osa ratkaisuista, joita tämän päivän semanttisen webin tarkoituksiin tehtyjen kielten ja merkintäjärjestelmien luomisessa on tehty, on jo löydettävissä 1600-luvun puolivälin ”universaalien kielten” kehittäjiltä.⁸ Mutta se, mikä erottaa tämän päivän semanttisen webin tarkoituksiin tehdyn työn aikaisempien vuosisatojen työskentelyistä, on työskentelyn tarkoitus. Tämän päivän semanttisen webin parissa työskentelevät eivät kuvittele luovansa tapaa kuvata asioita siten, että ne olisivat paremmin toisten ihmisten ymmärrettävissä; pyrkimys

⁶ Hyvösen työryhmä on Semantic Computing Research Group (<https://seco.cs.aalto.fi/>), ja heidän toteuttamistaan sovelluksista esimerkkinä mm. Sotasampo (<https://www.sotasampo.fi/>). Haettu 27.10.2022.

⁷ Skeemojen käytöstä ks. mm. Hyvönen 2018, luvut 8–12 ja Allemang & Hendler 2011.

⁸ Tästä teemasta hyvä esitys on mm. Knowlson 1975, joka esittelee niin luotuja skeemoja kuin niiden taustalla vaikuttanutta ajattelua.

on kuvata tietoa tavalla, joka mahdollistaisi sen merkityksien ja tietoyksiköiden välisten yhteyksien käsittelyn tietoteknisin välinein.

Fyysisten dokumenttien järjestämisessä perinteinen tapa on luokittelu. Keräämme yhteen paikkaan matkakertomukset, jotka sitten jaamme eri osiin esimerkiksi ajallisten tai alueellisten kriteerien perusteella. Perinteinen tapa yhdistää luokitteluun eri järjestelmää noudattavia aineiston kuvailuja on ulkopuolinen indeksi, joka voidaan järjestää omien tarpeiden mukaisesti esimerkiksi kuvaamalla tutkimuksen kannalta keskeisiä käsitteitä. Tässä vaiheessa olemme alkaneet jo rakentaa aineistoa kuvaavasta metadatatista semanttista verkkoa: meillä on joukko kuvattavia kohteita ja joukko asiasanoja, ja tallennamme kohteiden ja asiasanojen välisiä suhteita. Todellista semanttisen verkon rakentamista tästä tulee vasta, kun annamme myös näille suhteille muitakin merkityksiä kuin perinteisen luokittelun asiasanat.⁹

Seuraavassa taulukossa on kuvattu eräs lähde a ja sitä koskevia muistiinpanoja. Tällainen taulukko voisi hyvin olla taulukkolaskentaohjelmassa, jossa jokainen palsta kuvaa yhden lähteen tietoja.

Kenttä	Sisältö
Lähteen nimi	Tenore Viaggi
Tekijä	Michele Tenore
Nimi	Viaggio per diverse parti d'Italia, Svizzera, Francia, Inghilterra e Germania del Prof. Tenore
Julkaisuvuosi	1832
Teema	Matkakertomus
Aikajakso	1800-luku

⁹ Gartner 2016 on hyvä yleisesitys metadatatista ja sen kuvailun tekniikoista laajemmin; Gartner 2020 kuvaa miten temaattisesti fokusoitu käsitteistö on jäsennettävissä ontologisen ajattelun avulla.

Semanttisen verkon käyttöön edellä kuvatun taulukon sisältä voidaan esittää seuraavasti:

```
<TenoreViaggi> a <Lähde> ;
    <nimi> "Viaggio . del Prof. Tenore" ;
    <tekijä> "Michele Tenore" ;
    <julkaisuvuosi> "1832" ;
    <kohdema> <Italia>, <Sveitsi>, <Ranska> ;
    <teema> <Matkakertomus> ;
    <aikajakso> <1800-luku> .
```

Tämä niin sanotussa Turtle-muodossa kirjoitettu esitys sisältää saman informaation kuin itse taulukko, eikä sellaisenaan välttämättä ole sen helpompi käyttää. Käyttämällä tiettyjä semanttisen verkon käyttöön standardoituneita sanastoja voimme sanoa termeistä kuitenkin vielä lisää:

```
<Matkakertomus> a <Teema> .
<1800-luku> a <Aika> .
<Italia> a <Maa> .
<Sveitsi> a <Maa> .
<Ranska> a <Maa> .
<Teema> rdfs:subClassOf <Asiasana> .
<Aika> rdfs:subClassOf <Asiasana> .
<Maa> rdfs:subClassOf <Asiasana> .
<kohdema> rdfs:subPropertyOf <asiasana> ;
    rdfs:range <Maa> .
<teema> rdfs:subPropertyOf <asiasana> ;
    rdfs:range <Teema> .
<aikajakso> rdfs:subPropertyOf <asiasana> ;
    rdfs:range <Aika> .
<asiasana> rdfs:domain <Lähde> .
```

Ensin kerrotaan, minkä tyyppisiä käsitteitä eri kentissä käytetyt asiasanat ovat. Seuraavaksi, että kaikki nämä käsitetyypit ovat asiasanoja. Tämän jälkeen kerrotaan, että lähteen kuvailussa käytetyt ominaisuudet *kohdema*, *teema* ja *aikajakso*, ovat kaikki ominaisuuden *asiasana* alatyyppisiä, sekä se, että kohdemaan arvo on tyyppiä *Maa*, teeman arvo on tyyppiä *Teema* ja aikajakso kohde on tyyppiä *Aika*. Lopuksi

todetaan, että kaikkia asiasana-tyyppin ominaisuuksia – joihin siis yllä määrittelimme kolmen ominaisuuden kuuluvan – voidaan käyttää Lähde-tyyppisen tietoyksikön kuvailussa. Näin kuvattu tieto muodostaa verkoston, jossa ominaisuuksilla liitetään yhteen erilaisia tietoyksiköitä ja näin syntyy verkosto, joka kuvaa tutkimuksemme aineistoa.

Eri tietotyyppien välisiä suhteita ja näiden suhteiden ominaisuuksia kuvailevaa rakennetta kutsutaan *ontologiaksi*. Historioitsijalle ontologinen ajattelu on ehkä tutumpaa toisentyypisistä konteksteista, mutta tässäkin tapauksessa peruseriaate on sama: ontologialla pyritään kuvaamaan semanttisesta aineistosta sitä, mitä voi olla olemassa ja mitä tämä olemassaolo käytännössä tarkoittaa.¹⁰

Annotoinnin muodoista

Oma kysymyksensä on se, missä muodossa esimerkiksi tekstuaalisen lähdeaineiston *semanttinen annotointi* olisi kannattavaa tehdä. Mikäli käytössä on valmiiksi digitoitu tekstiaineisto, on eräs mahdollisuus käsitellä tekstejä kokonaisuutena. Esimerkki Tenoren matkakertomuksen sivulta 34:

```
Da Portella a <placeName ref="http://www.wikidata.org/entity/
Q128226">Terracina</placeName>tutto è incolto e deserto, ...
```

Tämä niin sanottua TEI-notaatiota¹¹ hyödyntävä teksti annotoi Terracinan paikannimeksi, ja antaa sille ulkoisen viitteen Wikidatan semanttiseen tietokantaan, josta on mahdollista hakea sinne tallennettuja lisätietoja kaupungista. Kaupungin sijaintitietojen merkintä tähän samaan yhteyteen on turhaa, koska tiedot on helposti tässä mainitun nimen avulla haettavissa muualtakin.

Toinen vaihtoehto on käyttää suoraan RDF-tyyppistä semanttista notaatiota esimerkiksi näin:

¹⁰ Erit. Allemang & Hendler 2011.

¹¹ TEI eli *Text Encoding Initiative* on tuottanut mm. tekstien kriittisiin editioihin tarkoitettua XML-pohjaista TEI-annotointimallia. Lisää aiheesta mm. <https://tei-c.org/>. TEI Consortium 2020.


```

<TenoreTerracinaViite> <lähde> <TenoreViaggi> ;
  <sivu> «32» ;
  <rivi> «8» ;
  <sananro> «4» ;
  <sana> "Terracina" ;
  <wikidata> <http://www.wikidata.org/entity/Q128226>.

```

Tässä annotoimme saman sanan kuin edellä ja annamme sille kohteen. Tällä tavalla kuvattuna voisimme antaa sille useampiakin kohteita, ja hieman merkintätapaamme laajentamalla taas mahdollistaa yksittäisten sanojen lisäksi pidempien tekstiosuuksien kuvaamisen.

Näiden kahden vaihtoehdon välinen valinta riippuu siitä, mitä yritämme aineistolla tehdä: ensimmäisessä tapauksessa liitämme tietoa, jota on helppo löytää aineiston täydennykseksi itse tekstiä luettaessa, kuten perinteisessä hypertextissä. Toisella menetelmällä teemme ulkoiseen tietokantaan muistiinpanoja, jotka liittyvät tekstiin. Molemmilla käyttötaivoilla on tarkoituksensa ja perinteensä.

Formaalin esittämisen tasot

Olemme edellä puhuneet formaalista historiatieteellisten argumenttien esittämisestä. Tätä ei kuitenkaan pidä ymmärtää siten, että pitäisimme toivottavana tai edes mahdollisena pyrkiä esittämään jonkinlaisia täydellisiä argumentteja tai luomaan jonkinlaista objektiivisen tai parhaimman mahdollisen esittämisen ideaalia. Kuten historioitsijoina tiedämme, kuvauksen yksityiskohtaisuutta on mahdollista syventää loppumattomiin, ja oleellinen tekijä tutkimuksen tekemisessä onkin hahmottaa, missä vaiheessa tutkimuksen tavoitteiden kannalta oleellinen yksityiskohtien taso on saavutettu. Sama rajoitus koskee myös tässä käsiteltyä formaalia kuvausta.

Esimerkkinä voimme ajatella edellisessä esimerkissä esiintynyttä Michele Tenoren¹² kertomusta matkastaan Napolista Englantiin. Kerromuksen avulla on mahdollista rekonstruoida reitin kulku paikkakuntineen ja maantieteellisine koordinaatteineen, päivämäärineen ja

¹²Tenore 1832.

kellonaikoinen sekä kuvata kaikki Tenoren näkemät yksittäiset kohteet. Huolellinen aineiston kirjaaminen mahdollistaisi myös ulkoisen data-aineiston linkittämisen omaan aineistoomme, esimerkiksi hyödyntäen Wikidataa ja muita jo valmiita semanttisen webin tarjoamia mahdollisuuksia. Mutta olisiko tämä tutkimukselle merkityksellistä tietoa? Riittäisikö tutkimuksellisen intressimme tyydyttämiseen tieto merkittävimmistä kohteista vai onko kysymyksillemme tärkeää dokumentoida myös kellonajat?

Samaten on myös tehtävä valintoja sen suhteen, miten kuvaamme ajallis-tilallisen järjestyksen aineistossa. Mikäli kuvaamme aineistoa semanttisen webin käytänteiden mukaisesti, vaihtoehtoja on (ainakin) kolme:

- 1) Kirjaamme Tenoren matkaan liittyviä yksilöityjä pysähdyspaikkoja siten, että kullekin pysähdykselle määritetään maantieteellinen sijainti, paikkakunnan nimi, ja lopuksi vielä pysähdykseen liittyvä kalenterimerkintä.
- 2) Kirjaamme Tenoren matkan aikana kohtaamat paikat järjestykseen listan muodossa.
- 3) Voimme yhdistää nämä kaksi muotoa, jossa esimerkiksi annamme yöpymispaikoille tarkat aikamääreet ja kuvaamme näiden välisen matkan järjestykseen asetettuna listana.

Ennen valintaa tulee pohtia, onko tällaisen dokumentaation tekeminen tutkimukselle tarpeellista. Jos haluamme matkasta täydellisen kuvauksen, vastaus on varmasti kyllä, mutta tätä ei ole pakko toteuttaa tarpeettoman yksityiskohtaisella tasolla, vaikka olemme kuvaamassa Tenoren matkaa formaalisti. Voimme tyytyä määrittelemään entiteetin ”Prof. [Michel] Tenoren matka”, jolle annamme lähteen lähtöpäivän, paluupäivän ja keskeiset matkakohteet. Jos sen sijaan käytämme matkakertomusta yhtenä lähteenä, jonka avulla tutkimme Pariisissa kävijöitä, kirjaamme ylös ne päivät, jolloin hän on ollut Pariisissa. Tällöin pystymme joka tapauksessa hyödyntämään matkakertomusta pohtiesamme, ketkä ovat Pariisissa olleet mihinkin aikaan.

Tarkoituksenmukaisuusanalyysi on tärkeää, kun lähteiden sisältöä kuvataan semanttisen webin työkaluilla, sillä mitä monimutkaisempaa lähteen tietojen kuvailu on, sitä enemmän sen tekemiseen kuluu aikaa. Ajankäyttö on perusteltua, jos lopputuloksena syntyy helpompi tai hyödyllisempi aineisto, mutta liiallinen yksityiskohtaisuus tuottaa tarpeettoman monimutkaisen aineiston. Sellaisen tehokas hyödyntäminen on vaikeampaa kuin yksinkertaisemman tietorakenteen ympärille rakennetun aineiston käyttö. Tutkimukselle käytetty aika voi olla siis täysin hukkaan heitettyä.

Historioitsijalle semanttisen webin suurin mahdollisuus lienee olemassa olevien aineistojen yhteensovittamisessa. Eri aineistoista on mahdollista kerätä omiin tarkoituksiin sopivat tietoyksiköt, jotka tallennetaan omien käsitteiden mukaan kuvattuna omaa tarkoitusta varten. Hyvinkin heterogeenisistä lähteistä peräisin oleva tieto on mahdollista integroida, koska tutkijalla on täysi vapaus päättää, mitä tietoa hän lähteistä haluaa ottaa käyttöönsä ja millä käsitteistöllä hän sitä haluaa kuvata sekä tallentaa näin tuotettu ”tulkinta” omaa käyttöä varten. Käytetty käsitteistö on mahdollista kuvata ja dokumentoida siten, että aineiston perusteella voidaan tehdä johtopäätöksiä ja etsiä ristiriitaisuuksia algoritmisesti. Lisäksi dokumentaatio mahdollistaa tutkimusryhmän kesken saman käsitteistön kehittämisen ja jakamisen. Näin jäsennetty tietoaaineisto on tehokkaasti jatkokäytettävissä, analysoitavissa ja muunnettavissa.

Verkostoanalyysi

Vaikka yhteiskuntatieteille ominaisen verkostoajattelun juuret ulottuvat aina 1800-luvun loppupuolelle,¹³ graafiteoriaan¹⁴ perustuvaa formaalia verkostoanalyysiä sovellettiin pitkään lähinnä sosiologisessa tutkimuksessa. Vasta verkostojen visualisointiin tarkoitettujen helppokäyttöisten tietokoneohjelmistojen myötä kiinnostus verkostoanalyysiä kohtaan on lisääntynyt.¹⁵ Nykyisin verkostovisualisoinnit ovat tehokas väline

¹³ Eriksson 2015.

¹⁴ Esim. Ruohonen 2013.

¹⁵ Schultz-Jones 2009.

havainnollistaa sosiaalisen ja kulttuurisen vuorovaikutuksen usein monimutkaisia rakenteita ja muutoksia.¹⁶

Verkosto rakentuu pisteiden eli solmujen (*nodes, vertices*) ja niitä yhdistävien linkkien (*edges, links*) muodostamana kokonaisuutena.¹⁷ Menetelmän joustavuus sekä tutkittavan ilmiön että käytettävän lähde- materiaalin suhteen tekee siitä kiinnostavan myös ihmistieteille. Ainoa selkeä rajoite on, että tutkimuskysymyksen tulee fokusoitua toimijoiden (ihmiset, organisaatiot, mutta myös esimerkiksi dokumentit, sanat ja vastaavat) välisten suhteiden analysointiin.¹⁸ Oletuksena siis on, että tutkittavaan ilmiöön liittyvien suhteiden perusteella mallinnetun verkoston rakenne selittää verkoston puitteissa havaittua toimintaa ja verkoston dynamiikkaa¹⁹. Verkostoanalyysi tarjoaakin tutkijalle välineitä tarkastella toimijoita osana huomattavasti laajempaa kokonaisuutta ja siten erottaa omalle tutkimusongelmalle merkitykselliset ja kiinnostavat elementit muista, vähemmän kiinnostavista elementeistä. Verkostoanalyysin systemaattisen soveltamisen avulla voidaan myös ymmärtää syvällisemmin toimijoiden välisten suhteiden todellista merkitystä.²⁰

Perusverkostoa voidaan rikastaa pisteisiin ja linkkeihin liitettävien laadullistavien attribuuttien avulla. Näin mallinnuksessa voidaan tutkimusprosessin edetessä askel askeleelta lähestyä kiinnostuksen kohteena olevaa tosimaailman ilmiöitä. Esimerkiksi tietyn yhteisön sosiaalisia verkostoja voidaan aluksi tutkia pelkkänä henkilö–henkilö-verkostonä, jossa suhde on tai ei ole olemassa. Tämän jälkeen henkilöitä (pisteitä) rikastetaan esimerkiksi tiedoilla sukupuolesta ja syntymä- ja kuolinajasta. Verkoston suhteita (linkkejä) puolestaan voidaan rikastaa tiedoilla yhteyden laadusta (esimerkiksi sukulaissuhteet, hierarkiat, aikatiedot). Verkoston pisteiden välille voi muodostua myös useita linkkejä, joista jokainen kuvaa tiettyä sosiaalisen todellisuuden ulottuvuutta. Attribuut-

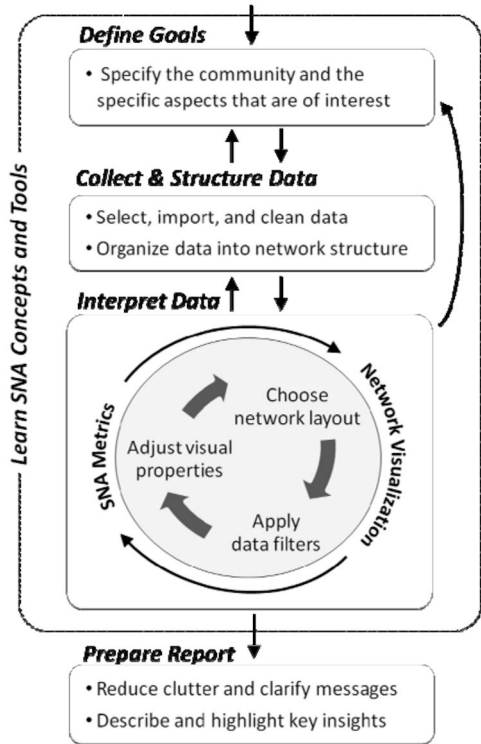
¹⁶ Esimerkkeinä von Keyserlingk-Rehbein 2018, Kaiser et al. 2018 sekä Lintunen & Elo 2019.

¹⁷ Verkostoanalyysin perusteista, ks. Prell 2012; Scott 2013.

¹⁸ <https://cvcedhlab.hypotheses.org/125>. Haettu 2.12.2020.

¹⁹ Esim. Bearman, Moody & Faris 2012; Brughmans 2013; Epple 2017.

²⁰ Morrissey 2015, 69–70.



Kuva 1. Verkostanalyysi tutkimusprosessin näkökulmasta. Lähde: Hansen et al. 2012, 5.

tien sisältämä data on myös mahdollista sisällyttää verkostovisualisointeihin visuaalisten tehosteiden (värit, muodot, koko) avulla.²¹

Kuva 1 esittää verkostanalyysille tyypillisen tutkimusprosessin Hansenin ja tutkijakollegoiden mukaisesti kuvattuna. Tutkimusprosessin formalisointi on osaltaan vastaus verkostanalyysin herkkyydelle aineiston puutteiden tai vinoumien suhteen. Tutkimusaineiston keruu tapahtuu tieteenalan käytänteitä noudattaen. Aineistoa valitessa tulee

²¹ Esimerkiksi Suomen sisällissotaan vuonna 1918 liittyvässä artikkelissa Lintunen & Elo (2019) eri kriteerein muodostuneet sosiaaliset suhteet – jäsenyydet yhdistyksissä tai ammattiosastoissa, sama asuinpaikka – erotellaan verkostovisualisoinneissa omilla väreillään, mikä sekä parantaa visualisoinnin luettavuutta että sisällyttää olennaista, tulkinnan kannalta olennaista tutkimustietoa jo itse visualisointiin.

varmistaa, että se mahdollistaa tutkittavan ilmiön kannalta merkityksellisten suhteiden mallintamisen varsinaiseksi verkostodataksi, johon mallinnukset ja visualisoinnit perustuvat. Tätä voidaan havainnollistaa pienen esimerkin avulla: ajatellaamme olevamme kiinnostuneita ”digitalisaatio”-termin yhteydessä *Turun Sanomissa* vuosina 1990–2015 esiintyvien käsitteiden muodostamasta kollokaatioverkostosta. Tutkimusaineistomme koostuu tällöin *Turun Sanomissa* vuosina 1990–2015 julkaistuista artikkeleista, joissa termi ”digitalisaatio” esiintyy vähintään kerran. Verkostodata puolestaan muodostetaan luomalla yhteydet kunkin ”digitalisaatio”-termin ja sitä esimerkiksi edeltävän ja seuraavan viiden sanan välille. Verkostodatan muodostaminen redusoi tutkimusaineistoa poimimalla sieltä tutkimuskysymykselle olennaiset entiteetit – esimerkissämme yksittäiset sanat – ja näiden väliset yhteydet.

Verkostodata muodostetaan siis strukturoimalla tutkimusaineisto ennalta määrättyjen kriteerien mukaiseksi, taulukkomuotoiseksi dataksi. Ihmistieteellisessä tutkimuksessa on vakiintunut tapa strukturoida aineisto linkkitaulukoksi (*edge table*)²². Yksinkertaisimmillaan linkkitaulu koostuu kahdesta sarakkeesta, jotka määrittävät yhteyden kohteena olevat pisteet. Linkkitaulukkoon voi sisältyä attribuuttsarakkeita rikastamaan kunkin lähtö- ja kohdepisteiden välillä vallitsevaa yhteyttä. Vaikka verkoston rakenne voidaan mallintaa ja visualisoida pelkän linkkitaulukon perusteella, usein luodaan myös erillinen taulu (*node table*) pisteiden attribuuteille. Tämän taulun jokainen rivi sisältää attribuutit kullekin verkoston pisteelle.

Verkostoanalyttisen tutkimusprosessin ”pihvi” on verkoston visualisointi eli verkoston rakenteen ja sen ominaispiirteiden esittäminen graafisesti. Tässä artikkelissa esitetyt visualisoinnit on toteutettu *open source* -pohjaisella *visone*-ohjelmistolla (<http://visone.ethz.ch/>). Muita vaihtoehtoja löytää helposti syöttämällä internetin hakukoneeseen hakulauseen ”network visualisation software”.

Verkostograafi rakennetaan verkostodatan perusteella pisteiden ja niitä yhdistävien viivojen muodostamana kokonaisuutena. Visualisointi perustuu ladontamalliin (*layout*), joka järjestää graafin elementit tietyn algoritmin mukaisesti siten, että verkoston rakenne on helpommin hahmotettavissa. Ihmistieteissä suositetaan pisteohjautuvia ladon-

²² Muista datarakenteista, ks. Scott 2013, 41ff.

tamalleja, jotka pyrkivät sijoittamaan toisiinsa vahvemmin yhteydessä olevat pisteet graafissa omiksi ryhmikseen eli klustereiksi. Ladontamallin valintaan ei ole mitään yhtä ja selkeää valintakriteeriä, vaan yleensä analyysivaiheessa selvitetään kokeilemalla, mikä ladontamalli tuo verkostosta esille tutkimuskysymykselle olennaisimmat piirteet. Usein graafeja elaboroidaan käyttämällä visualisointitehosteita solmujen ja linkkien värin, muodon tai koon esitystavan muuttamiseksi. Tehosteet sidotaan verkoston attribuutteihin, jolloin esimerkiksi eri sukupuolta olevat henkilöt voidaan esittää toisistaan erottuvilla muodoilla tai väreillä. Visualisointitehosteiden avulla graafiin lisätään informaatiota, jolloin jo graafin tarkastelu antaa katselijalle enemmän tietoa graafin taustalla olevasta ilmiöstä.

DDR-opposition luomat vastakulttuurin verkostot

Verkostanalyysin käytännön toteutusta havainnollistetaan seuraavassa kompaktilla analyysillä entisen Saksan Demokraattisen Tasavallan (DDR) oppositioryhmien verkostoihin liittyen. DDR:n sosialistisessa diktatuurissa oppositio muodosti merkittävän vastakulttuurin ja kanavoi yhteiskunnallista protestia. Oppositioon osallistuminen oli kuitenkin henkilökohtainen riski, kun DDR:n turvallisuusministeriö Stasi (*Ministerium für Staatssicherheit*) pyrki kaikin keinoin murskaamaan oppositionaaliset tahot.²³ Repressiosta huolimatta DDR-oppositio toimi käytännössä koko DDR:n olemassaolon ajan (1949–1990).²⁴ Vasta Berliinin muurin avautuminen marraskuussa 1989 ja sitä nopeasti seuranneet päätökset Saksojen yhdistämiseksi romahduttivat opposition aseman.²⁵

Esimerkkiverkostomme kuvaa Jenan alueen oppositiota 1970-luvun jälkimmäiseltä puoliskolta DDR:n murrosvuoteen 1989/1990 saakka.²⁶ Jena oli DDR-aikana keskisuuri yliopistokaupunki, josta vuoden 1968 ”Euroopan hullun vuoden” ja neuvostoarmeijan murskaaman ”Prahan

²³ Weber 1999, 130

²⁴ Neubert 1998.

²⁵ een et al. 2000, 8–9.

²⁶ Käytetyn tutkimushankkeen tuloksia on esitelty aiemmin, ks. esim. Elo 2018.

kevään” jälkimainingeissa muodostui DDR-opposition epävirallinen pääkaupunki. Kirkon suojissa toimineessa ”Nuoressa seurakunnassa” (*Junge Gemeinde*) sekä Jenan Rauhanyhdistyksessä (*Jenaer Friedensgemeinschaft*) nuoret dissidentit, eli toisinajattelijat, keskustelivat, verkostoituivat ja organisoivat poliittista aktivismia. Huolimatta Stasin toistuvista yrityksistä murskata dissidenttiverkostot Jena säilyi DDR:n poliittisen opposition yhtenä keskuspaikkana.²⁷

Jenan oppositioverkostojen ytimeen kuului kaksi nuorta dissidenttiä, Matthias Domaschk ja Roland Jahn. Domaschk menehtyi epämääräisissä olosuhteissa huhtikuussa 1981 Stasin tutkintavankeuden aikana. Jenan Rauhanyhdistys perustettiin osin jatkamaan Domaschkin aloittamaa työtä. Rauhanyhdistys järjestäytyi kirkon ulkopuolella, kun ”Nuoren seurakunnan” paljastuminen Stasin soluttamaksi lisäsi painetta luoda poliittiselle vastakulttuurille foorumi aiemmin varsin turvallisenä pidetyn kirkon ulkopuolelle.²⁸

Käytetty tutkimusaineisto koostettiin Robert Havemann -seuran hallinnoimaan DDR-opposition arkistoon kuuluvasta tietokannasta, johon kuuluu noin 60 000 digitoitua valokuvaa metatietoineen.²⁹ Tutkimusaineistoon valittiin Jenan oppositioaktiivien aktiviteetteja, kuten retkiä, illanviettoja ja mielenosoituksia dokumentoineet kuvat. Kuvien metatiedoista koottiin dataa muun muassa kuvausajankohtaan ja -paikkaan sekä kuvassa olleisiin henkilöihin liittyen. Koska kuvatietokanta on rakennettu jälkikäteen, metatiedoissa on jonkin verran puutteita ja selkeitä virheitä. Virheellisten tai puuttuvien tietojen osuus on kuitenkin melko pieni, eikä sillä havaittu olennaista merkitystä rekonstruoitujen verkostojen luotettavuudelle. Verkostoja tulkittaessa aineistoon mahdollisesti sisältyvät vinoumat on pyritty minimoimaan aiemman tutkimuksen kriittisen reflektoinnin kautta.

Verkostodata muodostettiin melko suoraviivaisella prosessilla linkitaulukoksi, jossa jokainen kuvatietueessa mainittu henkilö linkitettiin kaikkien muiden kuvatietueessa mainittujen henkilöiden kanssa. Henkilöiden nimet tarkistettiin ja tarpeen vaatiessa korjattiin. Linkkidataa

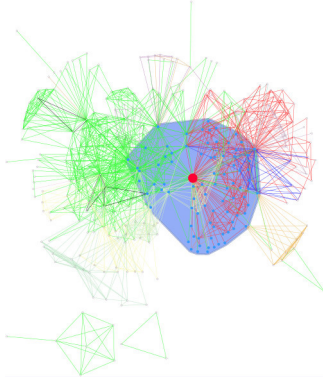
²⁷ Neubert 1998, 488; BStU 2013.

²⁸ Veen et al. 2000, 106; Scheer 1999, 201, 231.

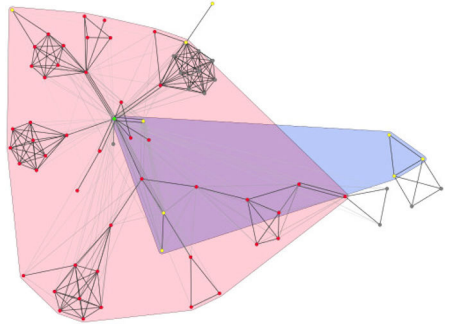
²⁹ <https://www.havemann-gesellschaft.de/archiv-der-ddr-opposition/bildarchiv/>. Haettu 2.12.2020.

rikastettiin vielä lisäämällä tarkistetut ja korjatut tiedot kuvausajasta sekä kuvauspaikasta (jos saatavilla). Näin luodun linkkitaulukon perusteella mallinnettu verkosto kuvaa oppositioaktiivien keskinäisiä yhteyksiä ja mahdollistaa verkoston tarkastelun ajan ja paikan funktiona.

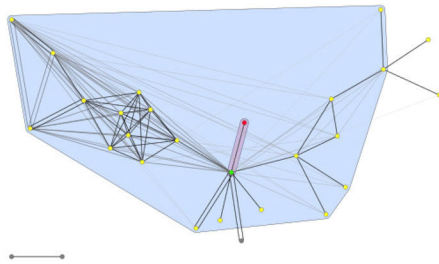
Kuvassa 2 on visualisoitu Roland Jahnin verkostot neljässä eri aikaikkunassa. Jahn päätyi Stasin ”tutkalle” jo 1970-luvun puolivälissä Jenan



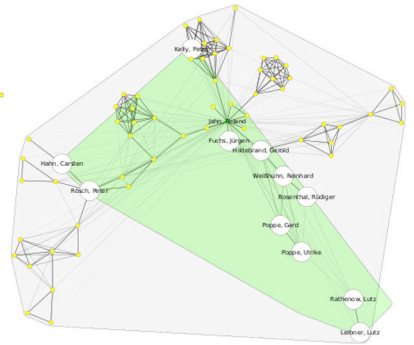
(a) Koko verkosto (1975–1990).



(b) Jahnin kontaktit Berliinissä ja Jenassa (1975–1983).



(c) Jahnin kontaktit Berliinissä ja Jenassa (1984–1990).



(d) Jahnin verkostot ennen ja jälkeen DDR:stä karkottamisen (1983).

Kuva 2. Neljä näkökulmaa Roland Jahnin verkostoihin.

oppositioverkoston aktiivina.³⁰ Poliittisen aktivisminsa vuoksi Jahn muun muassa erotettiin Jenan yliopistosta, pidätettiin lukuisia kertoja, sai vuonna 1982 puolentoista vuoden vankeustuomion ja lopulta karkotettiin vuonna 1983 Saksan liittotasavaltaan. Karkotuksensa jälkeen Jahn jatkoi aktiivista yhteydenpitoaan DDR:n oppositiopiirin kanssa ja loi yhteyksiä politiikan ja kulttuurin toimijoiden välille yli Saksojen välisen rajan.

Ensimmäinen verkostograafi (Kuva 2a) esittää Jahnin (punainen iso piste graafissa) osana rekonstruoitua sosiaalista verkostoa vuosina 1975–1990. Verkostossa on kaikkiaan 315 henkilöä. Ihmisiä yhdistää 1510 linkkiä eli näin monta kertaa verkostossa mukana olevat ihmiset ovat esiintyneet toistensa kanssa. Linkkien ihmisiä suurempi määrä antaa tiedon, että ainakin osa verkostoon kuuluneista ihmisistä on esiintynyt yhdessä useammin kuin kerran. Tosimaailman sosiaalisille verkostoille tyypillisesti verkosto jakaantuu tiheämpiin osaverkostoihin. Jahnin lähipiirin muodostava Ego-verkosto – kooltaan yhteensä 84 henkilöä – on rajattu sinisellä taustavärillä näkyväksi alueeksi. Ego-verkosto koostuu Jahnin suorista kontakteista sekä näiden kontaktien välisistä yhteyksistä.

Tiheä, monista linkeistä koostuva verkosto osoittaa myös ladontamallin valinnan tärkeyden verkoston tulkittavuudelle. Verkosto on ladottu ”quadrilateral Simmelian backbone”-algoritilla, joka on kehitetty erottamaan tiheissä verkostoissa piilevät klusterit ilman, että verkoston kokonaisrakenteen hahmottaminen merkittävästi kärsisi.³¹ Linkkien värit puolestaan kuvaavat kontaktien maantieteellistä jakautumista: vihreällä esitetyt linkit ovat yhteyksiä Berliinissä, punaiset Jenan alueella ja siniset linkit muilla alueilla toimineisiin oppositioaktiiveihin. Kiinnostava havainto on Berliinin ja Jenan verkostojen erillisyyden, joka aineiston perusteella selittyy lähes yksinomaan Jahnin karkottamisella länteen vuonna 1983. Näyttääkin siltä, että Jahn – joka dokumentoidusti vieraili 1984–1990 useasti laittomasti DDR:ssä – pyrki suo-

³⁰ Scheer 1999.

³¹ Ks. tarkemmin Nocaj, Ortman & Brandes 2015.

jelemaan Jenan aikaista verkostoaan välttelemällä esiintymistä heidän kanssaan samoissa aktiviteeteissa.³²

Verkostoissa 2b ja 2c Jahn on esitetty vihreänä pisteenä, minkä lisäksi paikkatietoa on hyödynnetty pisteiden väreissä siten, että Jahnin Berliinin-kontaktit on esitetty keltaisella, Jenan verkoston jäsenet puolestaan punaisella. Graafeissa näkyvät alueet ovat ryhmäsolmuja eli niiden alueelle sijoittuvat graafissa Jenan (punainen alue) ja Berliinin (sininen alue) verkostojen kontaktit. Näiden ulkopuolelle jäävät muissa paikoissa solmitut kontaktit. Graafien vertailu osoittaa verkoston muutoksen Jahnin karkottamisen jälkeen: Verkostossa 2c Jenan verkosto on kutistunut lähes olemattomiin, Berliinin verkoston ollessa tärkein maantieteellinen aliverkosto. Tämä muutos näkyy vielä selvemmin verkostossa 2d, jossa isoina valkoisina pisteinä on esitetty ydinryhmän henkilöt, jotka säilyivät Jahnin Ego-verkostossa läpi tarkastelujakson. Tämän ydinryhmän ympärille rakentui ennen karkottamista Jahnin muu sosiaalinen verkosto Jenan alueella. Karkottamisen jälkeen Jahn rakensi Berliinin alueella toimineiden aktiivien kanssa uuden sosiaalisen verkoston.

Jälkimmäisellä verkostolla näyttää kuitenkin olleen myös toisenlainen operatiivinen funktio. Kun Jahn ennen karkottamistaan rakensi Jenassa aktiivisesti toiminutta dissidenttiverkostoa, karkottamisen jälkeinen verkosto tuki Jahnin toimintaa linkkinä läntisen Saksan poliitisten ja kulttuuristen piirien ja DDR:n kulttuuriopposition välillä. Tätä tulkintaa tukee, että vuosien 1984–1990 verkostossa on mukana lukuisia DDR-opposition eri ryhmittymien ydinhenkilöitä, joiden keskinäistä vuorovaikutusta ja tiedonvaihtoa Jahn osaltaan näyttää tukeeneen.

Keskustelu

Tässä artikkelissa olemme esitelleet kaksi digitaalisen historiantutkimuksen menetelmää, jotka soveltuvat hyvin myös kulttuurihistorian piirissä sovellettavaksi. Molemmat menetelmät perustuvat ajatuk-

³² Vastaavaa toimintamallia on havaittu myös muissa, diktatuurien vastakulttuuria analysoineissa tutkimuksissa, ks. esim. Düring & Beer 2011.

selle asioiden liittämistä toisiinsa erilaisten määreiden avulla, tavoitteena mallintaa ja kuvata monimutkaisia asioita verkostoina. Erityisen lupaava käyttökohde molemmille on myös mahdollisuus yhdistää ne suurten aineistomassojen automaattiseen käsittelyyn esimerkiksi erilaisten *Named Entity Recognition* -työkalujen kanssa. Näitä hyödyntämällä teksti- tai kuva-aineistoista tunnistettavat entiteetit voidaan liittää niin aineistoon kuin toisiinsa.

Molemmilla menetelmillä on myös ongelmansa. Ne pyrkivät yksinkertaistamaan menneisyyttä pakottamalla ihmiset, tapahtumat tai tilanteet luokittelevan ja määrittelevän käsittearsenaalin kuvaamiksi. Tämä pakosta yksinkertaistettu käsitteistö, jonka varaan asiat ja niiden väliset liitokset rakennetaan, ei tavoita lähdeaineiston sisältämiä nyansseja tehden sekä ihmisistä että heidän välisistä suhteistaan vain osia koneistossa.

Tutkijan vastuulla on tiedostaa nämä ongelmat ja varmistaa, että hän käyttää menetelmiä siten, ettei tutkimuksen kohteelle tehdä vääryyttä. Miten? Esimerkiksi pitämällä huolen siitä, ettei näillä skemaattisesti määritetyillä käsitejärjestelmillä yritetä kuvata asioita, jotka eivät välttämättä siihen taivu. Esitettyjen tapaustutkimusten osalta olisi esimerkiksi virheellistä liittää samassa valokuvassa esiintyvät ihmiset toisiinsa ”ystävinä”, koska tämä on jo pitkälle menevä tulkinta. Sen sijaan on aivan hyväksyttävää tallentaa tietokantaan tieto, että ihmiset ovat todennettavasti esiintyneet samassa valokuvassa. Tällöin ei oteta kantaa siihen, mitä samassa valokuvassa esiintyminen tarkoittaa.

Missä sitten on semanttisen verkon ja graafianalyysien välinen ero? Ehkä keskeisin ero on siinä, että semanttinen verkko, ontologiat ja tietomallit eli skeemat ovat erityisen hyödyllisiä, kun dokumentoidaan tietoa hyödynnettäväksi toisissa käyttöyhteyksissä. Verkostoanalyysi puolestaan tarjoaa välineet mallintaa ja analysoida entiteettien merkitystä osana huomattavasti laajempaa kokonaisuutta sekä tuottaa ymmärrystä verkoston rakenteen merkityksestä havaitulle toiminnalle.

Vaikka digimenetelmät mahdollistavat hyvinkin suurten tietomasojen käsittelyn, ne eivät ole oikotie onneen. Mitä enemmän tehdään tietokoneavusteisesti, sitä suuremmaksi kasvaa tutkijan aihepiirin ja tutkimusprosessin hallinnan merkitys. Kyse ei ole vain analyysitulosten tulkinnasta, vaan myös aineiston ja analyysin valmistelusta ja ana-

lyysimenetelmien ”ohjeistamisesta”. Laskennalliset menetelmät antavat tuloksia myös huonosti valmistellulla, aukkoisella ja vinolla datalla. Menetelmän taakse ei siis voi mennä piiloon ja selittää, että ”tietokone nämä tuotti”. Kun aineistojen koon kasvaessa mahdollisuudet kontrolloida yksittäisiä seikkoja pienenevät, analyysit tulee aina testata pienillä ”dataseteillä”. Näin voidaan havaita ja korjata virhelähteet, joiden vaikutus moninkertaistuisi siirryttäessä suuriin aineistoihin.

Laskennallisen tutkimuksen oppimiskäyrä on paikoin jyrkkä, mutta me kirjoittajat uskomme artikkelin osoittavan, että haasteeseen tarttuminen kannattaa. Laskennalliset menetelmät avaavat tutkijalle mahdollisuuksia, joita ei ole mahdollista toteuttaa perinteisin lähiluentaan perustuvien menetelmin. Kyse ei ole pelkästään mahdollisuudesta käsitellä huomattavan suuria aineistokokonaisuuksia, vaan niihin kuuluvat myös esimerkiksi mahdollisuus tutkia aiempia aineistoja uusista näkökulmista tai soveltaa menetelmiä uusien tutkimuskysymysten löytämiseksi. Tutkimuksellisesti voidaankin puhua ”win-win”-tilanteesta, josta koko kulttuurihistoriallisen tutkimuksen kenttä voi parhaimmillaan hyötyä.

Lähteet ja tutkimuskirjallisuus

- Allemang, Dean & Jim Hendler: *Semantic web for the working ontologist: Effective modeling in RDFS and OWL*. Morgan Kaufmann Publishers 2011.
- Bearman, Peter S., Moody, James & Faris, Robert: *Networks and History. Complexity* 8, (2012), 61–71.
- Brughmans, Tom: Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory* 20:4 (2013), 623–662.
- BstU: *Aktion ”Gegenschlag”: Die Zerschlagung der Jenaer Opposition 1983*. Gera 2013. <http://www.bstu.bund.de/SharedDocs/Downloads/DE/20130304-gera-handout-operation-gegenschlag.pdf>
- Düring, Marten & Beer, Susanne: Hilfe für jüdische Verfolgte im Nationalsozialismus. *Medaon – Magazin für jüdisches Leben in Forschung und Bildung* 9 (2011).
- Elo, Kimmo (toim.): *Digitaalinen humanismi ja historiatieteet*. Turun historiallinen yhdistys 2016a.

- Elo, Kimmo: Digitaalisen historian tutkimuksen kenttää louhimassa. Teoksessa Kimmo Elo (toim.) *Digitaalinen humanismi ja ihmistieteet*. Turun historiallinen yhdistys 2016b, 11–35.
- Elo, Kimmo: Geospatial Social Networks of East German Opposition (1975–1989/90). *Journal of Historical Network Research* 2:1 (2018), 143–165.
- Epple, Ruedi: Flucht und Unterstützung: Die Auskunftsstelle für Flüchtlinge in Zürich und ihre Netzwerke. *Journal of Historical Network Research* 1:1 (2017), 106–145.
- Eriksson, Kai: *Verkostot yhteiskuntatutkimuksessa*. Gaudeamus 2015.
- Fitzpatrick, Kathleen: The Humanities, Done Digitally. Teoksessa Matthew K. Gold (toim.) *Debates in Digital Humanities*. University of Minnesota Press 2012.
- Gartner, Richard: *Metadata: Shaping Knowledge from Antiquity to the Semantic Web*. Springer 2016.
- Gartner, Richard: Towards an ontology-based iconography. *Digital Scholarship in the Humanities* 35:1 (2020), 43–53.
- Hansen, Derek L., Rotman, Dana, Bonsignore, Elizabeth, Milić-Frayling, Mendes Rodrigues, Nataša Eduarda, Smith, Marc, Shneiderman, Ben & Capone, Tony: Do You Know the Way to SNA?: Process Model for Analyzing and Visualizing Social Media Network Data. *Proceedings of the Social Informatics Conference 2012*, 304–313. DOI: 10.1109/SocialInformatics.2012.26
- Hyvönen, Eero: *Semanttinen web: Linkitetyn avoimen datan käsikirja*. Gaudeamus 2018.
- Kaiser, Maximilian, Schlögl, Matthias & Rumpolt, Peter Alexander: Artist Migration through the Biographer’s Lens: A case study based on biographical data retrieved from the Austrian Biographical Dictionary. *Journal of Historical Network Research* 2:1 (2018), 76–108.
- von Keyserlingk-Rehbein, Linda: *Nur eine “ganz kleine Clique”? Die NS-Ermittlungen über das Netzwerk vom 20. Juli 1944*. Lukas Verlag 2018.
- Kirschenbaum, Matthew: What Is Digital Humanities and What’s It Doing in English Departments? Teoksessa Matthew K. Gold (toim.) *Debates in Digital Humanities*. University of Minnesota Press 2012.
- Knowlson, James: *Universal language schemes in England and France, 1600–1800*. University of Toronto Press 1975.
- Lintunen, Tiina & Elo Kimmo: The Networks of Revolutionary Workers: Socialist Red Women in Finland in 1918. *International Review of Social History* 64:2 (2019), 279–307.
- Morrissey, Robert Michael: Archives of Connection. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 48:2 (2015), 67–79.
- Neubert, Ehrhart: *Geschichte der Opposition in der DDR 1949–1989*. Ch. Links 1998.
- Niiniluoto, Ilkka: Tieteiden ykseys. *Tieteessä tapahtuu* 19:4, 2001. <https://journal.fi/tt/article/view/57993/19717>.

- Nocaj, Arlind, Ortmann, Mark & Brandes, Ulrik: Untangling the Hairballs of Multi-Centered, Small-World Online Social Media Networks. *Journal of Graph Algorithms and Applications* 19:2 (2015), 595–618.
- Prell, Christina: *Social Network Analysis: History, theory and methodology*. SAGE 2012.
- Ruohonen, Keijo: *Graafiteoria*. Tampereen teknillinen korkeakoulu 2013.
- Scheer, Udo: *Vision und Wirklichkeit: Die Opposition in Jena in den siebziger und achtziger Jahren*. Ch. Links 1999.
- Schultz-Jones, Barbara: Examining information behavior through social networks: An interdisciplinary review. *Journal of Documentation* 66:4 (2009), 592–631.
- Scott, John P.: *Social Network Analysis*. Sage Publishing 2013.
- TEI Consortium: *Guidelines for Electronic Text Encoding and Interchange*. 2020. <http://www.tei-c.org/P5/>. Haettu 2.12.2020.
- Tenore, [Michele]: *Viaggio per diverse parti d'Italia, Svizzera, Francia, Inghilterra e Germania del Prof. Tenore*. 3 vol. Lorenzo Sonzongo 1832.
- Weber, Hermann: *Geschichte der DDR*. Dtv 1999.
- Veen, Hans-Joachim, Eisenfeld, Peter Kloth, Hans Michael, Knabe, Hubertus, Maser, Peter, Neubert, Ehrhart & Wilke, Manfred (toim.): *Lexikon Opposition und Widerstand in der SED-Diktatur*. Propyläen 2000.