

Avoimesti tieteestä!: Kirsi Hänninen – Vastuulliseksi aineistohallitsijaksi voi opetella

Äänitteen kesto: 17 min

Litterointimerkinnät

Haastattelija: Henriikka Mustajoki
Vastaaja: Kirsi Hänninen

sa- sana jää kesken
(sana) epävarmasti kuultu jakso puheessa tai epävarmasti tunnistettu puhuja
(-) sana, josta ei ole saatu selvää
(--) useampia sanoja, joista ei ole saatu selvää
, . ? : kieliopin mukainen välimerkki tai alle 10 sekunnin tauko puheessa

Intro: Avoimesti tieteestä. Keskustelua tieteen, tutkimuksen ja oppimisen avoimuudesta ja vastuullisuudesta.

Henriikka Mustajoki: Tervetuloa Avoimesti tieteestä-podcastin pariin. Minä olen Henriikka Mustajoki. Tänäpäin kanssani keskusteleee Kirsi Hänninen, joka on Turun yliopiston historian, kulttuurin ja taiteiden tutkimuksen arkiston arkistovastaava, folkloristi ja arkistoalan ja asiakirjahallinnon yliopisto-opettaja. Tervetuloa, Kirsi!

Kirsi Hänninen: Kiitos, Henriikka. Ilo olla täällä.

Henriikka Mustajoki: Ensimmäisenä mä mietin sellaista, että historian, kulttuurin ja taiteiden tutkimuksen parissa, niin miten sellaiset sanat kuin "data" ja "aineisto" siellä konkretisoituu? Mitä ne tarkoittaa sen alan tutkijoille?

Kirsi Hänninen: Aineisto on meidän alan tutkijoille se helpommin lähestyttävä sana. Puhutaan tutkimusaineistosta, lähdeaineistosta. "Data" monelle viittaa heti sähköiseen, määrälliseen aineistoon. Mutta yhä enemmän tutkimuksen ja rahoituksen suunnittelussa puhutaan datasta, niin data on meille myös tuttu. Oma taustani kun on folkloristiikassa ja kun työskentelen HKT-arkistossa, niin se tutkimusdata on laadullista, ihmistieteellistä, etnografisin menetelmin tuotettua. Eli haastatteluilla, havainnoinnilla, pyytämällä ihmisiä kirjoittamaan kokemuksestaan, valokuvaamalla, videoimalla, pyytämällä ihmiseltä aineistoja. Näistä menetelmistä johtuen se aineisto voi olla hyvin yllättävää, kaikkeen ei osata varautua ja se on hyvin rikasta. Siinä on paljon henkilötietoja ja se antaa sille sen hetkiselte tutkimukselle ja jatkotutkimukselle paljon erilaisia käyttömahdollisuuksia. Ja ne muodot jossa tämä on vuosikymmeniä ollu hyvin fyysisiä, analogisia eli on paperia, erilaisia nauhoitteita, äänitteitä, valokuvia, dioja, negatiiveja, kirjeitä, postikortteja,

karttoja, piirustuksia, päiväkirjoja. Nykyäänhan toki aineisto on syntysähköistä yhä enemmän ja enemmän ja meillä arkistossa on käynnissä digitointihanke. Eli myös tämä fyysinen, analoginen aineisto tulee olemaan sähköisessä muodossa.

Henriikka Mustajoki: Tuollaisen hyvin rikkaan ja usein ilmeisesti aika henkilökohtaisen datan ja aineiston parissa, niin miten siellä mietitään datan avaamisen prosessia ja tarvetta?

Kirsi Hänninen: Se nähdään tärkeenä. Ihmisten kanssa tuotettu aineisto, ja mä nyt korostan sitä että se tehdään ja tuotetaan yhdessä niitten ihmisten kanssa joita haastatellaan tai joita ,pyydetään kirjoittamaan, piirtämään tai jotka vastaavat kyselyihin, niin tämmösen aineiston tuottaminen ja se aineisto ymmärretään, että se palvelee paitsi sitä tutkimusta, niin se palvelee myös niitä tulevia tutkijoita, opiskelijoita ja opettajia. Ja tämmösen ymmärryksen päälle, myös se arkisto missä mä työskentelen on 50-luvulla perustettu, eli on varta vasten silloisessa kansatieteen laitoksessa ajateltu ja myöhemmin uskontotieteessä, folkloristiikassa, museologiassa, historian oppiaineissa ja taiteen tutkimuksen oppiaineissa ajateltu, että se on sen tiedeyhteisön käytössä myöhemminkin. Eli se perus suhtautuminen tutkimusdatan avaamiseen on positiivinen. Ja siinä on selkeästi hyötyjä sekä sen tutkimuksen sen hetkisen tekijän ja jatkokäyttäjän näkökulmasta.

Henriikka Mustajoki: Kertoisitko, Kirsi, enemmän siitä mikä on sen datan avaamisen hyöty?

Kirsi Hänninen: Toki sen sen hetkisen tutkijan näkökulmasta se, että ihmiset pystyy tarkistamaan sen, että tutkimuksen analyysi ja tulokset pohjautuu tosiaan aineistoon joka on olemassa. Aineistoviittaukset pystytään tarkistamaan sieltä. Mutta myös sitten jatkossa uuden tutkimuksen mahdollistaminen. Eli uusien menetelmien kehittäminen, uusien menetelmien kokeilu aineistoon johon niitä ei oo aikaisemmin käytetty, katsominen mitä uutta uusilla menetelmillä syntyy, uusien aiheiden, teemojen tutkimus, ongelmien keksiminen olemassa olevan aineiston perusteella, ajallinen, paikallinen tai muitten tekijöille vertailu vanhempaan, olemassa olevaan aineistoon. Ja onhan aineiston jatkokäyttö taloudellista ja säästää aikaa.

Tässä kohtaa mä aina muistan sanoa sen, että avoin data ei oo pikatie onneen, vaan on tiedettävä sen aineiston elinkaari, on tiedettävä ne vaiheet ja taustat. Eli miksi, miten, kuka, milloin, kenen rahoituksella, kenen aloitteesta on sen tuottanut, miksi tutkittavat on osallistuneet, miten se data on muodostunu yhdessä tutkijan ja tutkittavan välillä. Tämän ymmärtäminen ja tietäminen on tärkeitä siinä jatkokäytössäkin. Ja sitten on hyvin tärkeä asia, mihin viitattiin tuossa taannoisessa etiikan päivässäkin, niin tutkimusdatan avaaminen vähentää tutkimuspainetta joka voi kohdistua aktiivisesti tutkittuja kohteita kohtaan. Esimerkiksi normaalikouluja tai sitten semmosten kohteitten kohtaan, joitten etnografinen paikalle menevä, arkeen osallistuva tutkiminen voi olla eettisesti ja

laillisesti haasteellista, kuten muistisairaitten hoitokodeissa tehtävä tutkimus. Tällöinen tutkimusdata mahdollisuuksien mukaan avataan, niin se vähentää painetta joita näe paikat muuten kokis. Ja vähemmistöryhmät jotka kokemustensa tai taustojensa takia kiinnostaa tutkimusta, mut joille ei välttämättä ole resursseja olla mukana siinä tutkimuksessa. Tällöisiä aihepiirejä koskevan datan avaaminen myös vähentää sitä tutkimuspainetta heitä kohtaan.

Henriikka Mustajoki: Kun puhutaan tästä hyödyistä, ja sä, Kirsi, hirveen hyvin osaat selittää, että tällä on todella moninaisia hyötyjä, mut sit samalla se aineisto jossa viitataan on hyvin rikasta ja täynnä henkilötietoja. Niin mitä vastuullinen avaaminen sitten vaatii tutkijalta?

Kirsi Hänninen: Se vaatii sitä, että opiskelijat koulutetaan vastuulliseen tieteen tekemiseen ja vastuulliseen aineistonhallintaan ja -avaamiseen. Ja että opiskeluvaiheessa jo opiskelijoille tulee ymmärrys siitä, että datan avaaminen on yksi osa vastuullista aineistonhallintaa. Ja vastuullinen aineistonhallinta on yksi osa heidän ammattitaitoaan ja vastuullista tutkimuksen tekemistä. Eli mä näen sen hyvin tärkeänä, että se lähtee ihan fuksivaiheesta jo se koulutus siihen. Ja myös ymmärrys siihen, että lainsäädännön kehittymistä ja muutosta on seurattava aktiivisesti ja samoin minkälaisia eettisiä ohjeistuksia omalla alalla on.

Henkilötietojen käsittely on sisäänrakennettua siihen ihmistieteelliseen laadulliseen aineistoon jota itse olen tuottanut ja jota nyt työssäni käsittelen, koska kun ihmiseltä haastattelussa tai kyselyssä kysytään avoimia kysymyksiä, että kerro kokemuksestasi X, niin ihminen voi aloittaa lapsuudestaan tai jostain hyvin henkilökohtaisesta piirteestään. Voi kertoa perheenjäsenistään, naapureistaan, eli tulee muittenkin ihmisten henkilötietoja. Tällöisessä aineistossa henkilötiedot kuuluu siihen aineistoon, se tekee siitä rikkaan. Historian, kulttuurin ja taiteen tutkimuksen laitoksella on myös oppiaineita, kuten uskontotiede, jotka käsittelee ihmisten vakaumusta, joka jo itsessään on niin sanottu erityinen henkilötietoryhmä. Eli aineistoissa on syntyjään paljon sensitiivistä dataa, on erityisiä henkilötietoryhmiä. Tai puhutaan ihmisten terveydentilasta, sairauksien parantamisesta.

Se asettaa haasteen, että jo ennen kuin lähdet tuottamaan aineistoa pitää tietää mitä tarvitsee kysyä ja mitä ei tarvitse kysyä henkilötiedoista. On päätettävä etukäteen, että onko tarpeen edes harkita anonymisointia, jos on, niin miten se käytännössä tehdään. Tarvitaanko eettistä ennakoarviointia yliopiston taholta, jos tarvitaan niin miten se suoritetaan. Haasteet alkaa jo silloin kun mietitään tutkimusaihetta. Ja sitten se, että miten henkilötietoja käsitellään sen tutkimuksen aikana, mutta varsinkin tutkimuksen jälkeen. Tuhotaanko, anonymisoidaanko vai mitä minä suositan opiskelijoille, tutkijoille ja jo ammattini ja oman tutkija- ja opettajataustani takia on se, että pitkäaikaissäilytetään. Ja arkistointi on siihen hyvä ratkaisu.

No sitten mietitään mihin arkistoon, millä arkistolla on sopiva profiili, resursseja, mikä arkisto vastaanottaa, minkälaisia vaatimuksia se arkisto asettaa sille tuotettavalle tutkimusdataalle esimerkiksi tiedostomuodoissa. Ja sitten tämä kaikki on kerrottava sille tutkittavalle, eli informoitava siihen. Ja tutkimuksesta kertomisen lisäksi kysyttävä lupa siihen aineiston pitkäaikaissäilyttämiseen ja jatkokäyttöön ja kuvailtava millaista jatkokäyttö on. Ja sitten pitää myös ymmärtää, että on luonnollista ja olla loukkaantumatta jos tutkittava sanoo, että ei, en halua, että tätä pitkäaikaissäilytetään tai avataan ja se on ihan ok. Ja sen jälkeen päästään sitten kysymään tutkittavalta niitä asioita joita alun perin on ajateltu, että lähetään tutkimuksessa kysymään. Tämän prosessin valmistelu, eli se aineistonhallinnan, datanhallintasuunnitelman tekeminen, tarkistuttaminen, pitkäaikaissäilytyspaikan valinta, tietosuojailmoituksen tekeminen, informointikirjeen tekeminen, niin se vie aikaa. Ja se ei oo semmonen äidinmaidossa, alma materista, tuleva, vaan siihen pitää kouluttaa.

Henriikka Mustajoki: Vaikka tuntuu ehkä pitkältä se suunnitteluvaihe, niin tällaisten eettisten kysymysten pohdinta on kuitenkin paljon helpompaa kuin eettisten ongelmien ratkaiseminen jos sitä suunnitteluvaihetta ei oo hyvin tehty.

Kirsi Hänninen: Ehdottomasti. Mielestäni on hirveen tärkeätä tehdä se kunnolla, paitsi tutkijan näkökulmasta, että se oma työ tai tutkimusryhmän työ etenee eikä tule ongelmia jälkikäteen ratkottavaksi, niin myös se, että minkälainen kuva me annetaan tieteellisestä tutkimuksesta ihmiselle. Yks mätä omena voi pilata koko korin. Eli jos tulee huonoja kokemuksia tutkittaville, niin ei näitä ihmisiä hirvittävästi innosta vastata seuraaville tutkijoille. Eli me ollaan vastuussa itsellemme, rahoittajille, tutkittaville, mutta mielestäni myös seuraaville tutkijoille eli meidän ammattiyhteisölle.

Henriikka Mustajoki: Mä mietin semmosta ihan käytännön kysymystä, että kun meillä on ajatuksena, että on pitkäaikaissäilytyksessä olevia aineistoja, niin kuka niitä pääsee näkemään? Miten on määritelty se avoimuuden aste arkistoiduissa aineistoissa? Onko siellä valtava kirjo vai onko siellä joku standardimalli vai miten se toimii?

Kirsi Hänninen: Ohjenuora, joka on hyvä ja johon monessa kohtaa viitataan, eli tutkimusdata on niin avointa kuin mahdollista ja niin suljettua kuin välttämätöntä toimii meillä niin, että meidän aineistoja voi käyttää tutkimus-, opetus- ja museotoimintaan. Ja tapauskohtaisesti, harkinnan varaisesti muihin tarkoituksiin. Sama pätee myös soveltavissa määrin metadataan eli kuvailutietoihin. Osa aineistojen luetteloista ja kuvailuista on avoimia kokoelmakohtaisesti, osa taas on suljettuja niin, että niitä kuten tutkimusaineistojakin pääsee katsomaan ja tutkimaan käyttö lupaa hakemalla. Ja tämä käyttö lupa on se millä ihmiset pääsee niitten aineistojen pariin. Eli tällä hetkellä ottamalla arkistoon yhteyttä ja kertomalla mihin tarkoitukseen ja minkälaista aineistoa tarvitsee ja me sitten katsomme, että vastaako se niitä meidän käyttösääntöjämme se tarkoitus ja sitten

myönnetään tai kielletään se käyttöluva. Ja nämä on arkisto- ja tietovarantokohtaisia. Näitten arkistojen ja varantojen sisälläkin voi olla vielä eroja siinä, että onko joku tutkija asettanut vielä erityisiä rajoituksia aineistolle vai ei.

Henriikka Mustajoki: Sä mainitsit, että osa metatiedoista on kaikille avoimia, mutta osa jopa metatiedoista vaatii lupajärjestelyä. Onko siihen joku erityinen syy?

Kirsi Hänninen: Metadatatalla tässä arkistoyhteydessä voi tarkoittaa eri asioita. Eli metadatatahan on yhtä lailla se, että missä muodossa ja milloin luotu joku aineisto tai sen digitoitu tiedosto on. Mutta se metadatat mihin mä viittasin tässä suljetussa on kuvailutieto siitä esimerkiksi ketä ja minkä ikäistä ja sukupuolista ihmistä on haastateltu, missä osoitteessa. Eli henkilötiedot ei oo avoimia.

Henriikka Mustajoki: Miten sä näät nyt sun aikalaisesta näköalapositiosta sen, että mihinkä tämä suunta on menossa arkistoalalla? Eli just avoimuuden suhteen. Onko näkyvissä muutoksia tai jotain vastavoimia? Millainen kehityskaari, jos sä mietit seuraavaa 5 tai 10 vuotta, arkistonäkökulmasta?

Kirsi Hänninen: Kun mä katson eteenpäin, niin mä ensin katson taaksepäin. Eli mä olen jo opiskelijana haastatellut ihmisiä, pyytänyt ihmisiä kirjoittamaan kokemuksistaan, kuvannut ihmisiä. Meidät opetettiin arkistomaan tähän samaiseen HKT-arkistoon meidän aineistoja. Mutta silloin se opiskelijana tuntui todella kaukaiselta ajatukselta ajatella, että joku näitä vielä joskus käyttäisi. Silloin 90-luvulla, 2000-luvun alussa kun opiskelin, niin samalla tavalla ei ollut se aineiston avaaminen ja metadatan avaaminen, se totta kai koettiin tärkeäksi ja sitä edistettiin, mutta se ei ollut samalla tavalla pinnalla keskustelussa ja päätöksissä kuin nykyään. Mä nään nyt sen tilanteen niin, että se tulee opiskelijoille jo kenttätöitä harjoitteleville opiskelijoille ja nuorille tutkijoille tutuksi. Ja mä nään, että siitä tulee yksi oleellinen osa heidän ammattitaitoaan, päätyi he sitten tekemään tutkimusta työkseen tai ei.

Ja arkistoalalla, kyllä samoin mä nään, että se aineistojen avaamisen, keinoja varsinkin siihen aineistojen käyttöön saataville saamiseen, aineistojen pitkäaikaissäilytykseen, niitä etsitään ja niitä kehitetään. Ja hyvin tärkeänä myös se, että siihen saavutettavuuteen kiinnitetään huomiota ja etsitään resursseja, että millä saavutettavuus ratkaistaan. Miten ja millä resursseilla videot tekstitetään. Ja mun mielestäni se ymmärrys tulee menemään yhä enemmän siihen, että avoimuus ei pelkästään tarkoita sitä, että ennen analoginen aineisto on nyt digitoitu ja jollain tapaa sähköisesti jossain ns. avoimena, vaan avoimuus tarkoittaa sitä, että se oikeasti on löydettävissä, käytettävissä ja saavutettavissa. Aika näyttää sitten sen, että miten se tulee vaikuttamaan siihen, että miten arkistoalalla tietopalvelu ja asiakaspalvelu tulee muuttumaan ja kehittymään.

Henriikka Mustajoki: Musta ajatus siitä, että se ei oo vaan digitoitua, vaan se on aidosti löydettävissä ja saavutettavissa on aika keskeinen ydinsuunta ja todella

hieno ajatus siitä mihin ollaan menossa. Se arvokas tieto joka siellä on ja, kuten sä aikaisemmin sanoit, sen laajemman hyödyntämisen edut tulevat silloin vielä enemmän esille. Se on aika positiivinen tulevaisuus johon arkistoala on menossa ja katsomassa. Ja mä haluaisin kiittää sua, Kirsi Hänninen Turun yliopistosta, erinomaisesta näkökulmasta siihen miten tällaisen hyvin erityisen ja henkilötiedoista rikkaankin aineiston ja datan avaamiselle on työkalut, mahdollisuudet ja halua. Kiitoksia, Kirsi!

Kirsi Hänninen: Kiitoksia, Henriikka ja kiitoksia kuulijoille mielenkiinnosta.

[Lopputunnusmusiikki]