

Yhteenvetoa keskusteluista

Toivon, että asiat ovat tulleet kirjatuksi tuohon alas oikein. Jos joku tunnistaa virheitä, korjatkoon asian suoraan tekstiin. Myös lisäyksiä, tarkennuksia ym. voi tehdä.

Alkukeskustelu

Laskujeni mukaan paikalla oli 22 historian tutkijaa tai asiasta kiinnostunutta, jotka jakautuivat aloittain tai organisaatioittain seuraavasti:

- Suomen ja Pohjoismaiden historia: 6
- yleinen historia: 4
- poliittinen historia: 3
- musiikkitiede: 2
- taidehistoria: 2
- kirkkohistoria: 1
- Pohjois-Amerikan tutkimus: 1
- taloushistoria: 1
- kasvatustieteet: 1
- Kansalliskirjasto: 1

Paikalle saapuneet olivat tulleet joko erityisestä käytännön syystä, kuten akatemiahakemuksen takia (10), tai yleisemmän kiinnostuksen takia (12). Osa akatemiahakijoista ilmaisi myös yleisempää kiinnostusta DMP-teemaa kohtaan.

Poimintoja alkukeskustelusta:

- Tutkijan havahtuminen siihen, että entisen, kaoottisen mallin mukaan tehty datankeruu ei mahdollista toisten tutkijoiden käyttöä.
 - Mikko Ojanen (MO) tuo esille, miten datan laatu vaikuttaa esim. toistettavuuteen.
- Historiantutkimusta ei voi verrata luonnontieteisiin. Historiantutkimuksessa lähteet johtavat eteenpäin toisiin lähteisiin – etukäteen tutkimusprosessin hallitseminen (mahdollisine tutkimusaineistoinen) ei ole realistinen ajatus. Historiantutkimuksen organisuutta korostettiin myös ryhmä- ja loppukeskusteluissa.
- Historiantutkimuksessa tekstilähteet ovat hallitsevia, ja tutkijat ovat tottuneet työskentelemään tekstikorpuksen parissa. Milloin tutkimuksessa hyödynnetty tekstikorpus muuttuu tutkijan tuottamaksi tutkimusdataksi? Missä menee raja lähteen ja tuotetun datan välillä? Tähän teemaan palattiin myös loppukeskustelussa.
- Historiassa joillakin aloilla syntyy dataa (näytteet ym.), mutta usein historioitsijat käyttävät aineistoja, jotka ovat julkisesti saatavilla. Tutkijat eivät koe mielekkäänä sitä, että tällaista aineistoa kuvaisi aineistonhallintasuunnitelmaa varten. Riittääkö, että toteaa, että aineistot ovat saatavilla? Kysymys on siitä, syntyykö oikeasti uutta dataa.
 - MO tuo esiin DMP-ohjeissa esitetyn datan kolmijaon: (1) tätä tutkimusta varten kerättävä aineisto, (2) tutkimusprosessin tuloksena syntyvä aineisto, (3) aiemmin kerätty aineisto
- Viime vuonna aineistonhallintasuunnitelman tehnyt tutkija koki DMP:uulin "kauheaksi ja kömpelöksi kuin mikä" – ja koko homman pointiksi sen, että hän luettelee datan olevan arkistossa. Hän koki harmilliseksi sen, ettei osannut kirjoittaa suunnitelmaan enempää.
- Tutkija pohti, pitäisikö aineistonhallintasuunnitelmaan kuvata omia työmetodeja siitä näkökulmasta, ettei vaaranna alkuperäistä (esim. arkistossa olevaa) dataa.

Loppukeskustelu

Seuraavaksi poimintoja loppukeskustelusta ryhmittäin. Nämä saattavat mennä osittain tai paljonkin päällekkäin ryhmän sihteerin merkintöjen kanssa.

Ryhmä 1: DATA

- Milloin tekstikorpus muuttuu dataksi eli milloin siitä tulee omaa dataa – missä raja kulkee? Tämä pitäisi pystyä jotenkin konkretisoimaan ohjeissa.
- Muistiinpanojen asema herätti kauhunsekaisia ajatuksia siitä, että ne pitäisi avata muille.
 - MO muistuttaa, että DMP ja avoin tiede eivät ole synonyymeja. DMP on vastuullisuuden työkalu, joka takaa sen, että se, mitä avataan, avataan tietoisesti ja asianmukaisesti.
 - MO esittää kysymyksen, voisivatko tutkijoiden muistiinpanot olla relevanttia tutkimusmateriaalia esim. 50 vuoden päästä.
- Omaan työhön tukevien muistiinpanojen hallinta voidaan nähdä metadatanä, ehkä tärkeänäkin sellaisena.
 - MO viittaa verifiointiin ja toistettavuuteen, siihen, että voi omien muistiinpanojensa kautta todistaa päätelmänsä.
- MO korostaa, että eräs DMP:n tarkoituksista on se, että tutkija osaisi kuvata oman työnsä arvon. Tällöin tutkijan tulee arvioida, mikä osa muistiinpanoista on arvokasta säilyttää ja millä voisi olla uudelleenkäyttöarvoa.

Ryhmä 2: ETIIKKA

- Historiantutkimuksessa voi olla perusteltua, että tavalliset ihmiset esiintyvät omalla nimellä. Tämä pitäisi pystyä selkeästi muotoilemaan sopimuksissa – ryhmäkeskustelussa viitattiin ns. pakoklausuuliin: milloin tutkittavan on mahdollista vetäytyä tutkimuksesta.
 - MO: Tarvitaan konkreettista työnkulun yksityiskohtiin (yliopiston juristit).
- Salaiseen aineistoon liittyen: pitäisikö DMP-ohjeissa sopimusten lisäksi kertoa tarkemmin avoimuudesta – kaikki eivät ehkä voi käyttää sitä aineistoa, jota tutkija käyttää. Tutkijalla on myös hyvin vähän mahdollisuuksia vaikuttaa käytäntöihin (aineistoon pääsyyn).
 - MO: Riittää, että kertoo DMP:ssä tekevänsä sopimuksen (ja noudattavansa sitä). Sopimukset on myös syytä arkistoida.
- MO tuo esiin ERC-rahoituksen kolmivaiheisen DPM:n (ensin idea datasta ennen kuin tutkimus on käynnistynyt, sitten kuvaus datasta prosessin ollessa käynnissä, ja lopuksi kuvaus syntyneestä datasta). Suomessa aineistonhallintasuunnitelmia ei tiettävästi tsekata tutkimuksen jälkeen. Dataan liittyvät kysymykset ovat usein vaikeita tutkimuksen alkuvaiheessa – ideana on kuitenkin laittaa tutkija ajattelemaan dataan liittyviä asioita ennen kuin ne (ml. ongelmat) ovat edessä.
- Yrityksaineistot: Pitäisikö ohjeissa korostaa, että on tämäntyyppisiä ongelmia aineiston jakamisen suhteen?

Ryhmä 3: METADATA

Varsinaista keskustelua ei syntynyt. Ryhmän pohdinnat on kirjattu.

Ryhmä 4: SÄILYTYS

- Miten kauan aineistoa pitää säilyttää, jos ottaa arkistoaineistosta kuvia?
- Pitääkö suunnitelmaan kirjoittaa itsestäänselviä asioita auki, esimerkiksi se, että näitä aineistoja säilytetään arkistossa, ne ovat arkiston vastuulla ja käytettävissä siellä.
 - MO: Pitää luokitella, millaista data on; vrt. ohjeen kolmijako: (1) tätä tutkimusta varten kerättävä aineisto, (2) tutkimusprosessin tuloksena syntyvä aineisto, (3) aiemmin kerätty aineisto.
- Lyhyt, ei-sanatarkka dialogi antiikintutkijan (AT) ja MO:n välillä:
 - AT: Mitä jos uudelleenkäytetty (aiemmin kerätty) data on ainoa data? Esimerkiksi oma aineistoni on Kaisassa kenen tahansa luettavissa.
 - MO: DMP on silloin huomattavasti helpompi. Tämän voi mainita suunnitelmassa.
 - AT: Eli se tosiaan riittää?
 - MO: Se riittää, mutta DMP pitää palauttaa.
- OneDriveen liittyen: onko se riittävä tallennuspaikka, jos ei puhuta sensitiivisestä datasta?
 - MO: Jos tera riittää, niin kyllä. Helsingin yliopiston kautta käytetty OneDrive takaa käyttäjälle ja tämän datalle paremman turvan kuin kaupalliset palvelut.
 - Tutkija: Mutta tulevaisuus ei välttämättä ole HY:ssä, data kannattaa laittaa myös omaan taskuun.

Ryhmä 5: AVAAMINEN

- Kysymys Tietoarkistosta nousi keskustelussa esiin. Onko se soveltuva paikka historian aineistojen säilyttämiseen? Anu Lahtinen tuo esiin, että Tietoarkiston käytettävyys voi olla historian tutkijan kannalta heikkoa, jos vain osa hänen aineistostaan sopii Tietoarkistoon. Kannattaako tällaista pikkuruista ja työtä vaativaa datasettiä tehdä Tietoarkistoa varten, vai kannattaako data laittaa kokonaisuudessaan jonnekin muualle?