

Yhteiskuntatieteellisen datan hallinta

TTA: Kerro meille datastasi, Tampere 15.5.2013

Mari Kleemola
Tietopalvelupäällikkö
mari.kleemola@uta.fi



YHTEISKUNTATIETEELLINEN TIETOARKISTO
FINLANDS SAMHÄLLSVETENSKAPLIGA DATAARKIV
FINNISH SOCIAL SCIENCE DATA ARCHIVE

Yhteiskuntatieteellinen tietoaarkisto

- > OKM:n rahoittama valtakunnallinen palveluresurssi
 - erillinen laitos TaY:n yhteydessä
 - perustettu v. 1999 alussa SA:n aloitteesta
- > Tehtävä: sähköisten tutkimusaineistojen pitkäaikaissäilytys ja jatkokäyttö
- > Toiminnan perustana asiantunteva, nopea, tasapuolinen ja maksuton palvelu
 - autamme tutkimusaineistojen haussa ja käytössä
 - neuvomme aineistojen tietosuojakysymyksissä
 - annamme ohjeita aineistonhallintaan ja pitkäaikaissäilytykseen
- > CESSDA-ERIC:n kansallinen palveluntuottaja

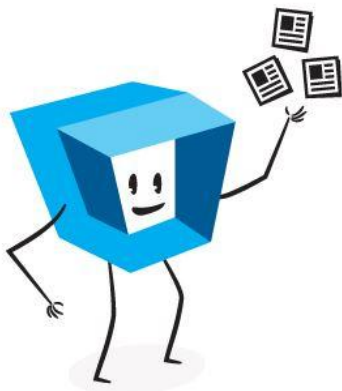
Arkistoidut aineistot

- > Kvantitatiivisia aineistoja (havaintomatriiseja) n. 1200
- > Kvalitatiivisia (tekstiaineistoja) n. 150
- > Aineistovarannon kasvu n. 100 aineistoa/vuosi (n. 1 Gt)
- > 2012: jatkokäyttöön 688 aineistoa
 - 30% kasvu vuodesta 2011
- > Tulossa: aineistoportaali AILA



Aineistot tieteenaloittain (20 yleisintä); lisäksi:

- rauhantutkimus, kulttuurintutkimus, perhetutkimus, oikeustiede, naistutkimus, aluetiede, kehityspsykologia, organisaatiotutkimus, kasvatopsykologia, lapsitutkimus, informaatiotutkimus, kriminologia, sosiaaliantropologia, tieteentutkimus, teologia, työelämän tutkimus, kehitystutkimus, liikuntatiede, maaseutututkimus, lääketiede, uskontotiede, kaupunkitutkimus, sukupuoliututkimus, miestutkimus, taloussosiologia, homo- ja lesbotutkimus, johtamistiede, kansatiede, gerontologia, kielitiede, tulevaisuudentutkimus, etnologia, folkloristiikka, historia, homotutkimus, kirjallisuudentutkimus, lesbotutkimus, queer-tutkimus



Miksi tiedonhallintaa ja jatkokäyttöä?

Tutkimusaineistojen tiedonhallinnalla tarkoitetaan sitä, että tutkimusaineistot (data) ja niihin liittyvä kuvaileva tieto (metatieto, metadata) on luotu, tallennettu ja järjestetty siten, että aineiston säilyy käyttökuntoisena ja luotettavana ja että tietosuoja ja tietoturva on varmistettu aineiston koko elinkaaren ajan.

Yhteiskuntatieteellinen tietoarkisto on edistänyt yhteiskuntatutkimuksen aineistojen avoimuutta.

> Lue lisää



1 Miksi tiedonhallintaa ja jatkokäyttöä?

- Osa hyvää tieteellistä käytäntöä
- Tutkimusaineiston elinkaari
- Jatkokäytön ja avoimen saatavuuden hyödyt

2 Tiedonhallinnan suunnittelu

- Aineistonhallintasuunnitelma
- Aineiston keruun suunnittelu
- Tutkimusaineiston tekijänoikeudet
- Tutkimushankkeen jälkeen
- Kaikkea ei kannata arkistoida

3 Tutkittavien informointi

- Informoinnin osa-alueet
- Tarkistuslista
- Arkaluonteiset tiedot

4 Tiedostoformaattit ja ohjelmistot

- Tiedostojen konvertointi
- Tutkimusaineiston digitointi
- Havaintomatriisit ja tilasto-ohjelmistot
- Tekstiaineistot
- Kuva-aineistot
- Ääni- ja äänikuvatallenteet

5 Kvantitatiivisen datan käsittely

- Havaintomatriisin tallennus
- Muuttujien nimet ja selitteet
- Muuttujan arvot ja niiden selitteet
- Muuttujien muunnokset
- Puuttuva tieto
- Painomuuttujat
- Syntaksin käyttö

6 Kvalitatiivisen datan käsittely

- Tiedostonimet ja datalistaus
- Litterointi
- Yksikkökohtaisten tietojen kirjaaminen

7 Tunnisteellisuus ja anonymisointi

- Tunnisteellisuus tutkimusaineistoissa
- Anonymisoinnin lähtökohdat
- Kvantitatiivisen aineiston anonymisointi

8 Aineiston kuvailu ja metadata

- Metadatan tallennus
- Tutkimuksen toteutuksen kuvailu
- Aineistonkeruunstrumentti
- Tiedostojen kuvailu

9 Fyysinen säilytys

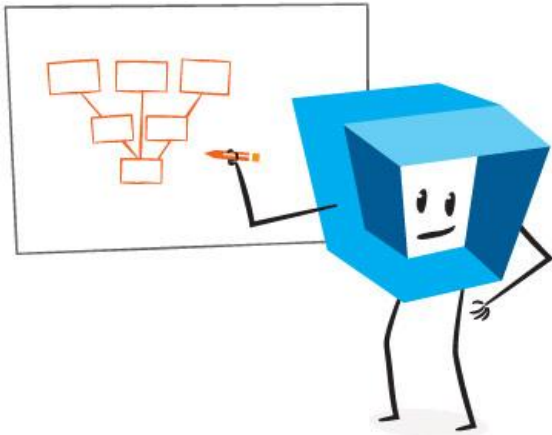
- Tiedostojen nimeäminen ja hallinta
- Varmuuskopiointi
- Migraatio ja virkistäminen
- Tallennusmediat

Datan hallinta projektin aikana: Tutkimus-aineistojen tiedonhallinnan käsikirja

<http://www.fsd.uta.fi/tiedonhallinta/>

Tiedonhallinnan suunnittelu

- Aineistonhallintasuunnitelma
- Aineiston keruun suunnittelu
- Tutkimusaineiston tekijänoikeudet
- Tutkimushankkeen jälkeen
- Kaikkea ei kannata arkistoida



Aineiston kuvailu ja metadata

- Metadatan tallennus
- Tutkimuksen toteutuksen kuvailu
- Aineistonkeruulinstrumentti
- Tiedostojen kuvailu
- Muuttujien kuvailu
- Kontekstitiedot ja paradata
- Metadataformaattit

Datan hallinta projektin jälkeen

- > Aineiston arkistointi ja jatkokäyttö ajoittuu usein (mutta ei aina!) projektin jälkeiseen aikaan
- > Vastuu aineiston ”hoivaamisesta” siirtyy pois projektilta
 - fyysinen säilytys, käyttökunnon säilytys
 - aineiston löydettävyys ja saatavuus
 - käyttöehdoista huolehtiminen,...
- > Arkistoinnin ja uudelleen hyödyntämisen onnistuminen edellyttää aineiston hyvää hallintaa elinkaaren alusta asti.



Kiitos!

*Tietoarkisto turvaa aineistosi
sekä tukee tutkimusdatan
avointa saatavuutta ja
informoitua jatkokäyttöä.*