



YHTEISKUNTATIEEELLINEN  
TIETOARKISTO



HANNELE KECKMAN-KOIVUNIEMI & MARI KLEEMOLA

# OPAS KVANTITATIIVISTEN TUTKIMUSAINEISTOJEN JATKOKÄYTTÖÖN



HANNELE KECKMAN-KOIVUNIEMI & MARI KLEEMOLA

**OPAS KVANTITATIIVISTEN TUTKIMUSAINEISTOJEN  
JATKOKÄYTTÖÖN**

Yhteiskuntatieteellisen tietoarkiston julkaisija 2  
Hannele Keckman-Koivuniemi & Mari Kleemola (2006)  
Opas kvantitatiivisten tutkimusaineistojen jatkokäyttöön

Julkaisija: Yhteiskuntatieteellinen tietoarkisto (FSD)  
Käyntiosoite: Åkerlundinkatu 2 A, 5 krs, Tampere  
Postiosoite: Yhteiskuntatieteellinen tietoarkisto  
33014 Tampereen yliopisto  
Puhelin: (03) 3551 8519  
Faksi: (03) 3551 8520  
fsd@uta.fi  
Kotisivu: <http://www.fsd.uta.fi>

Myynti: Tiedekirjakauppa TAJU  
PL 617, 33014 Tampereen yliopisto  
Puhelin: (03) 3551 6055  
Faksi: (03) 3551 7685  
taju@uta.fi  
<http://granum.uta.fi>

ISSN 1459-8906  
ISBN 951-44-6805-8 (nid.)  
ISBN 978-951-44-7507-8 (pdf)

Graafinen suunnittelu  
Vinjetti Ky

Taitto  
Marita Alanko

Tampereen yliopistopaino Oy – Juvenes Print  
Tampere 2006

## Johdanto

Arkistoitua aineistoa hyödyntävä tutkija kohtaa analyysissä pääosin samanlaisia mahdollisuuksia ja ongelmia kuin itse keräämäänsä, uutta aineistoa käyttävä tutkijakin. Arkistoitujen aineistojen käyttöön liittyy kuitenkin kysymyksiä, joita esimerkiksi suomalaisessa yhteiskuntatieteellisessä menetelmäkirjallisuudessa on sivuttu toistaiseksi vain lyhyesti. Tämä selittyy ainakin osittain sillä, että Suomessa ei ennen Yhteiskuntatieteellisen tietoarkiston perustamista vuonna 1999 ollut keskitettyä tutkimusaineistoarkistoa.

Arkistoitujen tutkimusaineistojen käyttömahdollisuudet ovat moninaiset. Aineistoja voidaan käyttää opetuksessa ja opiskelussa harjoitteluun. Niitä voi käyttää oman tutkimuksen vertailuaineistoina. Niiden avulla voidaan tutkia sekä nykyisyyttä että menneisyyttä, vertailla ilmiöitä sekä toistaa tai laajentaa olemassa olevia tutkimuksia. Useita tietolähteitä yhdistämällä saa myös monipuolisemman kuvan tutkimuskohteesta kuin yhtä aineistoa käyttämällä. Aineistojen jatkokäyttö tehostaa monin eri tavoin resurssien käyttöä. Parhaimmillaan opiskelijoilta, tutkijoilta ja tutkimusprojekteilta säästyy sekä aikaa, rahaa että työtä.

Tämä teos antaa perusohjeita yhteiskuntatieteellisten, erityisesti kvantitatiivisten tutkimusaineistojen löytämiseen ja jatkokäyttöön. Oppaamme alkaa jatkokäytön määrittelyllä, jonka jälkeen kerromme muun muassa jatkokäytön eduista. Oman lukunsa saavat myös aineistojen tilaaminen, jatkokäytön haasteet ja jatkokäyttäjän velvollisuudet. Yksittäiset luvut ovat melko itsenäisiä kokonaisuuksia, joten lukija voi halutessaan hypätä suoraan itseään kiinnostavaan asiaan. Varsinaiset tutkimusmenetelmät jäävät aihepiirimme ulkopuolelle.<sup>1</sup>

Opas on suunnattu ensisijaisesti perustutkinto-opiskelijoille, mutta myös muille arkistoitujen aineistojen käytöstä kiinnostuneille. Vaikka keskitymme tietoarkiston kautta saatavilla oleviin kvantitatiivisiin kyselyaineistoihin, teksti soveltuu joiltain osin ohjenuoraksi myös kvalitatiivisten aineistojen jatkokäyttäjille.

Kiitokset Sami Borgille, Marja Alastalolle ja Matti Vuorensyrjälle, jotka ovat lukeneet käsikirjoituksemme ja antaneet arvokkaita kommentteja ja vinkkejä. Kiitos myös Taina Jääskeläiselle, joka valaisi meitä kansainvälisten vertailuaineistojen kääntämiseen liittyvissä kysymyksissä, ja Lotta Alholle, Sanni Haveriselle ja Anna Lappalaiselle, jotka kommentoivat tekstiämme opiskelijan näkökulmasta. Kaikki tekstiin mahdollisesti jääneet puutteet ovat tietysti meidän vastuullamme.

Tampereella marraskuussa 2006

*Hannele Keckman-Koivuniemi*

*Mari Kleemola*

<sup>1</sup> Menetelmistä lisää esim. Menetelmäopetuksen tietovaranto MOTV: <http://www.fsd.uta.fi/menetelmaopetus/>



# Sisällys

Johdanto	3
1 TUTKIMUSAINEISTOJEN JATKOKÄYTTÖ JA SEN EDUT	7
1.1 Aineistojen jatkokäytön määritelmiä	7
1.2 Jatkokäytön etuja	8
2 ARKISTOITUJEN AINEISTOJEN KÄYTTÖMAHDOLLISUUDET	10
2.1 Uusi teema ja näkökulma	10
2.2 Tausta-aineisto	11
2.3 Vertaileva tutkimus	11
2.4 Uuden tutkimuksen suunnittelu ja menetelmien kehittäminen	12
2.5 Tulosten tarkistaminen ja todentaminen	12
2.6 Opetus ja opiskelu	13
3 ETSI JA LÖYDÄ SOPIVA AINEISTO	15
3.1 Tutkimusongelma, teoria ja aineiston valinta	15
3.2 Aineiston alustava arviointi	16
3.3 Aineiston etsiminen Yhteiskuntatieteellisestä tietoaarkistosta	19
3.4 Kyselyaineistoja muualta maailmasta	21
3.5 Muita aineistoja muualta: esimerkkejä tilasto- ja rekisteriaineistoista	22
4 AINEISTON TILAAMINEN JA KÄSITTELY	23
4.1 Aineiston tilaaminen tietoaarkistosta	23
4.2 Datatiedoston avaaminen ja käsittely	23
5 TUTKIMUSAINEISTOJEN JATKOKÄYTÖN HAASTEITA	25
5.1 Jatkokäytön hyödyt suuremmat kuin haitat	25
5.2 Operationalisoinnit, kysymysten järjestys ja asiayhteys	26
5.3 Otantaongelmat ja kato	27
5.4 Erilaiset tiedonkeruumenetelmät	28
5.5 Mittarien vertailtavuus	29
5.6 Kansainvälisten vertailuaineistojen käytön erityispiirteet	31
6 AINEISTOJEN JATKOKÄYTTÖ JA HYVÄT TIETEELLISET KÄYTÄNNÖT	34
6.1 Jatkokäyttö osana aineistojen elinkaarta	34
6.2 Tutkimusaineistoon viittaaminen	34
6.3 Julkaisuista tiedottaminen	35
6.4 Tietosuoja tutkimusaineiston säilyttämisessä ja käytössä	35
7 LOPUKSI	36
Sanasto	37
Lähteet	39
Hyödyllisiä verkkosivuja	43
Liitteet	44
Liite 1. Käyttölupahakemus	44
Liite 2. Käyttöehtositoumus	45



# 1 Tutkimusaineistojen jatkokäyttö ja sen edut

## 1.1 Aineistojen jatkokäytön määritelmiä

Olemassa olevien kvantitatiivisten aineistojen hyödyntäminen yhteiskuntatieteellisessä tutkimuksessa ja opetuksessa on vakiintunut kansainvälinen käytäntö. Etenkin yhdysvaltalainen ja brittiläinen alan lähdekirjallisuus on monipuolista ja kattavaa. Englanniksi olemassa olevien aineistojen käyttöön viitataan useimmiten termeillä ”secondary analysis”, ”secondary research”, ”re-use”, ”re-work” ja ”reanalysis”. Tietoarkistossa olemme päätyneet käyttämään termiä ”jatkokäyttö”.

Alan uranuurtajan Herbert Hymanin mukaan survey-aineistojen<sup>2</sup> jatkokäytössä tuotetaan uutta tietoa aihealueista, jotka eivät olleet alkuperäisten tutkimusten painopisteitä (Hyman 1972, 1). Jatkokäyttö voi myös sisältää tulkintoja, päätelmiä tai lisätietoja, joita ei esitetty ensimmäisessä tutkimusraportissa ja sen päätuloksissa (Hakim 1982, 1). Kautta linjan jatkokäytön tunnuspiirteenä mainitaan, että aineisto on kerätty jo aiemmin ja että aineistoa käyttää joku muu kuin alkuperäinen tutkija. (Dale, Arber & Procter 1988, 3–4; Uusitalo 1999, 94; Frankfurt-Nachmias & Nachmias 1996, 304; Kiecolt & Nathan 1985, 9–10; Stewart 1984, 11; Hakim 2000, 24.)

Suomalaisen yhteiskuntatieteellisen menetelmäkirjallisuuden yleisoppaissa olemassa olevien tutkimusaineistojen käyttö on yleensä mainittu lyhyesti yhtenä mahdollisuutena tutkimuksen teossa. Koska Yhteiskuntatieteellinen tietoarkisto aloitti toimintansa vuonna 1999, viittauksia tietoarkiston palveluihin on vasta uusimmissa oppikirjoissa.

Rolf Alameri ja Pentti Pöyhönen (1954, 14) tarkoittivat 1950-luvulla ”valmiista aineistoista” kirjoittaessaan tilastoaineistoja tai painettuja taulukkoja. Itse kerättyjen aineistojen käsittelystä puhuessaan he tosin kannustivat esimerkiksi neulakorttien käyttämiseen, koska työn säästön lisäksi ”jää kortisto palvelemaan mahdollisia jatkotutkimuksia” (emt., 30). Seuraavalla vuosikymmenellä Antti Eskola (1966, 71) mainitsi yhtenä ”sekundäärianalyysin” muotona asianomaisen tutkijan luvalla käyttöön saadun perusaineiston, esim. reikäkorttien, ”uudelleenritytely”. Reikäkortit ovat vuosikymmenten saatossa vaihtuneet nopeasti käsiteltäviksi datatiedostoiksi, mutta periaate ei ole vanhentunut.

Hannu Uusitalo (1991, 94) kutsuu olemassa oleviin aineistoihin perustuvaa tutkimusta sekundaarianalyysiksi ja korostaa, että tutustuminen omasta ongelmasta aikaisemmin tehtyyn tutkimukseen on välttämätöntä. Hän toteaa myös, että ”[e]rityisesti surveytutkimuksissa aineisto muodostuu helposti niin runsaak-

2 Survey-käsitteen määrittelystä lisää ks. Alastalo 2005, 14–17.



si, että sitä voidaan käyttää myös muuhun kuin vain alunperin ajateltuun tarkoitukseen.” Tapani Alkula, Seppo Pöntinen ja Pekka Ylöstalo (1995, 54–55) tuovat esiin, että ”[a]ikaisemmin kerättyjen tutkimusten aineistoja voidaan myös analysoida uudelleen. Tällaisessa jälkianalysissä aineisto analysoidaan uudesta näkökulmasta ja sen pohjalta pyritään esittämään uusia tulkintoja.”

Myös Sirkka Hirsjärvi (2004, 175) on tähdentänyt, että ”[t]utkimusaineistojen keruussa olisi syytä pyrkiä ekonomiseen ja tarkoituksenmukaiseen ratkaisuun.” Jokaisen tutkimusongelman ratkaisemiseen ei tarvita itse kerättyä aineistoa. Aina-kin ongelman joihinkin osiin voi saada vastauksen jo muiden keräämien aineistojen eli sekundaariaineistojen pohjalta.

Tutkimusaineistojen jatkokäytölle on siten ominaista toisten kokoaman, yleensä johonkin muuhun tarkoitukseen keräämän aineiston hyödyntäminen. Usein jatkokäyttäjä käyttää erilaisia menetelmiä kuin mitä alkuperäiset tutkijat ovat käyttäneet. Jatkokäyttö voi koostua yhtä lailla vain yhden tutkimusaineiston kuin useiden eri aineistolähteiden käytöstä ja näiden sisältämien tietojen yhdistelystä.

## 1.2 Jatkokäytön etuja

Tutkimusaineistojen jatkokäyttöä käsittelevässä kirjallisuudessa luetellaan monia jatkokäytön etuja<sup>3</sup>. Useimmiten korostetaan tieteen avoimuutta, tutkimustulosten kontrolloitavuutta ja tiedon kumuloitumista sekä käytännöllisiä, sosiaalisia ja taloudellisia näkökohtia.

Tieteen avoimuuteen kuuluu mahdollisuus muiden esittämien tutkimustulosten arviointiin. Tieteellinen tieto on luonteeltaan kumuloituvaa: uudet ideat ja tutkimustulokset täydentävät aikaisempia tuloksia. Kumuloituminen onnistuu parhaiten silloin, kun tieteellisen työn tulokset ja myös tutkimusaineistot ovat tiedeyhteisön vapaasti saatavilla. Tutkijan on perinteisesti oletettu tutustuvan ja viittaavan aiheestaan aiemmin julkaistuihin tutkimuksiin, mutta hyvä olisi perehtyä myös aiempiin empiirisiin tutkimusaineistoihin.

Tilastollisesti edustavien tutkimusaineistojen suunnittelu, kerääminen ja tallentaminen edellyttävät paljon resursseja. Opiskelijoille ja yksittäisille tutkijoille uuden oman tutkimusaineiston kerääminen saattaa olla taloudellisesti mahdotonta. Aineistojen jatkokäyttö tehostaa sekä tutkimusta että tutkimusrahoituksen käyttöä: kun jokaisen tutkimusongelman ratkaisemiseksi ei tarvitse kerätä uutta aineistoa, säästyy rahaa, aikaa ja työtä. Erityisen kalliita ovat vertailevan ja aika-

3 Ks. mm. Alkula, Pöntinen & Ylöstalo 1995, 54–55; Dale, Arber & Procter 1988, 45–55; Van Deth 2003, 292–294; Fienberg, Martin & Straf 1985, 9–15; Frankfort-Nachmias & Nachmias. 1996, 304; Heaton 2004, 26–28; Hedrick 1985, 124–132; Hirsjärvi 2004, 175; Hyman 1972, 6–24; Law 2005, 9–14; Procter 1993, 256–257; Sieber 1991, 12–14; Stewart 1984, 14; Szabo & Strang 1997, 66–74; Uusitalo 1999, 94; Mochmann, Ekkehard & de Guchteneire 1998.

sarjatutkimuksen vaatimat valtakunnalliset ja kansainväliset aineistot, jotka ovat kuitenkin ensiarvoisen tärkeitä sosiaalisten ilmiöiden selittämiseksi ja muutoksen ymmärtämiseksi. Tutkimuksessa voi olla myös tarpeellista käyttää erityyppisiä aineistoja, esimerkiksi kvalitatiivisia ja kvantitatiivisia aineistoja rinnakkain.

Lisäksi arkistoitujen aineistojen kohdalla useat eettiset kysymykset on ratkaistu jo ennen jatkokäyttöä. Aineistot ovat esimerkiksi yleensä valmiiksi anonymisoituja. Käyttämällä arkistoituja aineistoja vältetään myös vastaajien toistuva vaivaaminen.



## 2 Arkistoitujen aineistojen käyttömahdollisuudet

Arkistoituja aineistoja voidaan käyttää useisiin eri tarkoituksiin (Corti & Thompson 2004, 331–334; Seale 2004, 360). Tässä luvussa tarkastelemme lähemmin seuraavia eri käyttömahdollisuuksia:

1. uusi teema ja näkökulma
2. tausta-aineisto
3. vertaileva tutkimus
4. tutkimuksen suunnittelu ja menetelmien kehittäminen
5. tulosten tarkistaminen ja todentaminen
6. opetus ja opiskelu.

### 2.1 Uusi teema ja näkökulma

Tutkijat eivät läheskään aina analysoi koko tutkimusaineistoaan, joten arkistoidut aineistot sisältävät paljon täysin käyttämätöntä materiaalia. Tietoarkistoon on tallennettu runsaasti alun perin lähinnä hallinnollisiin tarkoituksiin kerättyjä aineistoja, joista ei ole välttämättä julkaistu muuta kuin päällimmäisiä tuloksia. Esimerkiksi alun perin asukkaiden kuntapalveluja ja hallintoa koskevien mielipiteiden selvittämistä varten kerätyillä kaupunkipalvelututkimuksilla on myöhemmin tutkittu ulkona kävelyn koettua turvattomuutta (Myllyniemi 2006), perusterveydenhuollon kustannus-vaikuttavuutta (Hämäläinen 2005) ja yhteyttä taloudellisen resurssoinnin ja kuntalaisten palveluita koskevien mielipiteiden välillä (Uitus-Mäntylä 2003). Kaikki mainitut kolme tutkimusta ovat pro gradu -töitä.

Joskus menneeseen aikaan liittyviin tutkimusongelmiin ei edes ole saatavilla muuta kuin toisten aikaisemmin keräämiä aineistoja. Yhteiskuntatieteellisen tietoarkiston vanhimmat aineistot ovat 1970-luvulta. Uusia menetelmiä käyttäen tai uuden teorian näkökulmasta vanhakin aineisto tuoreutuu.

Tietoarkistosta saa myös tuoreita aineistoja, joita ei ole analysoitu aikaisemmin. Vuosittain heti aineistonkeruun jälkeen tulee tietoarkiston kautta saataville International Social Survey Programme -hankkeen eli ISSP:n Suomen aineisto. Lisäksi joka toinen vuosi kerättävä European Social Survey (ESS) -aineisto tulee ensikäyttöön tietoarkistojen kautta. Vuonna 2002 kerätyllä Suomen ESS-aineistolla on tutkittu muun muassa, korostuvatko samat selittävät muuttujat analysoitaessa kahta subjektiivisen elämänlaadun ja hyvinvoinnin mittaria eli onnellisuutta ja tyytyväisyyttä elämään (Kouvo & Räsänen 2005), vertailtu poliittista osallistumista ja asenteita Pohjoismaissa (Borg 2005) ja selvitetty suomalaisten luottamusta yhteiskunnallisiin instituutioihin (Mattila & Sänkiaho 2005).

## 2.2 Tausta-aineisto

Arkistoituja tutkimusaineistoja voi käyttää uuden tutkimuksen tausta-aineistoina. Niiden avulla voi kuvailla tutkimuksen yhteiskunnallisia, sosiaalisia, historiallisia ja kansainvälisiä konteksteja. Eurobarometrien<sup>4</sup> avulla voidaan tarkastella vaikkapa eurooppalaisen identiteetin kehittymistä, eurooppalaisten yleistä tyytyväisyyttä elämään tai tyytyväisyyttä demokratian toimivuuteen. Esimerkiksi Oikeusolot 2004 -raportissa on eurobarometrejä ja World Values Survey<sup>5</sup> -aineistoja käyttäen luotu yleiskuvaus Suomen oikeuslaitoksen nauttimaan luottamukseen sekä tarkasteltu eurooppalaisten luottamusta Euroopan oikeusasiamieheen ja EY-tuomioistuimeen (Litmala 2004).

## 2.3 Vertaileva tutkimus

Arkistoidut tutkimusaineistot mahdollistavat ajalliset ja kansainväliset vertailut. Lisäksi tutkija voi halutessaan vertailla oman uuden tutkimusaineistonsa perusteella saamia tuloksia muiden aiemmin saamiin tuloksiin. Pro gradu -tutkielmassa vertailuaineiston käyttö voi tuoda kaivattua lisäarvoa opiskelijan omalle työlle.

Arkistoitujen aineistojen mahdollistamasta ajallisesta pitkittäisvertailusta käy esimerkiksi Kimmo Grönlundin, Heikki Paloheimon ja Hanna Wassin (2005, 124–125) artikkeli, jossa käsitellään suomalaisten äänestysaktiivisuutta eri väestöryhmissä 1975–2003. Vaalitulokset eivät aiheesta kerro, joten aineistoina tutkimuksessa on käytetty muun muassa puolueiden ajankohtaistutkimuksia ja EVAn kansallisia asennetutkimuksia. Niissä on säännöllisesti kysytty, mitä puoluetta vastaaja äänesti edellisissä eduskuntavaaleissa.

Eri maiden vertailuun tietoarkistosta löytyy useita kansainvälisiä aineistosarjoja. Lisäksi arkistojen kansainvälinen yhteistyö tuo tarjolle paljon yksittäisiä tutkimusaineistoja, jotka mahdollistavat kahden tai useamman maan vertailun.

Arkistoitujen aineistojen avulla voi hyvin tehdä vertailevaa tutkimusta myös ajankohtaisista aiheista. Tietoarkiston toimittamien datojen avulla on esimerkiksi tarkasteltu digitaalista kuilua ja vertailtu Internetin käyttöä ja tietoyhteiskuntakehityksen vaihetta Euroopan eri maissa (Räsänen 2006).

4 Eurobarometrien trendit:  
[http://www.gesis.org/en/data\\_service/eurobarometer/standard\\_eb\\_trend/](http://www.gesis.org/en/data_service/eurobarometer/standard_eb_trend/)

5 World Values Survey -hankkeen kotisivu: <http://www.worldvaluessurvey.org/>

## 2.4 Uuden tutkimuksen suunnittelu ja menetelmien kehittäminen

Arkistoituja aineistoja voi käyttää uusien tutkimusten suunnitteluun, teorioiden rakentamiseen ja tutkimusmenetelmien kehittämiseen. Niillä voi testata mittareita ja analyysimenetelmiä. Esimerkiksi Timo Toivonen (2005) on Eurobarometri 56.0 -aineiston avulla vertaillut moniulotteisten ristiintaulukoiden ja monimuuttujamallien mahdollisia hyviä ja huonoja puolia.

Arkistoidut aineistot antavat vinkkejä aineistonkeruuseen ja kyselylomakkeen laatimiseen. Tutustumalla aiemmin käytettyihin koti- ja ulkomaisiin lomakkeisiin saa käsityksen siitä, mitä tietoa on jo tarjolla ja millaisia kysymyksiä omassa uudessa tutkimuksessa on mahdollista kysyä. Aineistojen datatiedostoihin (havaintomatriiseihin) ja aineistosta tuotettuun kirjallisuuteen perehtymällä voi saada vastauksia myös siihen, miten erilaiset operationalisoinnit ovat käytännössä toimineet. Tietoarkiston Internet-sivujen käyttötilastossa onkin suosituimpien sivujen joukossa useita yksittäisiä kyselylomakkeita.

## 2.5 Tulosten tarkistaminen ja todentaminen

Tieteellinen tieto rakentuu vanhan pohjalle, joten tieteen kehittyminen edellyttää tieteen avoimuutta ja julkisuutta. Julkisuuden vaatimus ei koske vain tutkimuksen tuloksia, vaan se ulottuu myös tutkimuksessa käytettyihin aineistoihin (Niiniluoto 2001).

Olennainen osa tiedettä on mahdollisuus tutkimusten toistamiseen, tulosten testaamiseen ja todentamiseen. Samaa aineistoa samoilla tutkimusmenetelmillä analysoitaessa olisi päädyttävä samoihin tuloksiin. Koko tiedeyhteisöllä olisi oltava mahdollisuus arvioida käytettyjä aineistoja ja analyysimenetelmiä. Aineistojen avoin saatavuus parantaakin tutkimuksen laatua.

Esimerkiksi Pertti Töttö (2004, 162–192) on tarkastellut kriittisesti Suvi Ronkaisen (1999) tulkintoja sukupuolten vastaamistyylien eroista. Tietoarkistoon tallennettujen kysely- ja haastatteluaineistojen avulla Töttö on testannut, missä määrin Ronkaisen tulkinnat ovat sidoksissa Ronkaisen käyttämään aineistoon ja päästäänkö muita aineistoja analysoimalla Ronkaisen kanssa yhteneviin tuloksiin.



Kuva 1. Menetelmäopetuksen tietovarannon etusivu.

## 2.6 Opetus ja opiskelu

Tietoariston aineistoja hyödynnetään paljon opetuksessa ja opiskelussa. Vuonna 2005 aineistojen käyttöluvista 42 prosenttia myönnettiin opiskelu- ja opetuskäyttöön ja 30 prosenttia opinnäytetöihin. Myös Menetelmäopetuksen tietovaranto, MOTV, on opettajien ja opiskelijoiden suosima. MOTV on sekä kvantitatiivisten että kvalitatiivisten tutkimusmenetelmien verkko-oppimisympäristö. Se käsittelee analyysimenetelmien ohella myös tutkimuksen teon keskeisiä kysymyksiä. Tietovarannossa on myös vapaasti ladattavia harjoitusaineistoja.

Arkistoiduilla aineistoilla voi harjoitella tutkimuksen tekoa: tutkimuksen suunnittelua, tiedonhankintaa, hypoteesien asettamista ja testausta, tilasto-ohjelmien ja analyysimenetelmien käyttöä sekä raportointia tieteenalan viittauskäytöineen. Aineistojen avulla opiskelijat voivat käytännössä tutustua tieteenalansa tutkimusaineistoihin ja arvioida niissä tehtyjä valintoja.

Valmiin aineiston käyttö säästää aikaa. Aika ei mene tietojen keräämiseen ja tallentamiseen, vaan analyysimenetelmiä voidaan tarvittaessa havainnollistaa ja harjoitella heti. Aitojen aineistojen käyttö tekee harjoittelusta mielekkäämpää kuin kuvitteellisten demoaineistojen käyttö. Hyötyä on siitäkin, että menetelmäkurssin harjoitusaineistona tutuksi tullutta tutkimusaineistoa voi halutessaan myöhemmin käyttää myös kandidaatintutkielmassa, teemaseminaareissa ja pro gradu -työssä.

Tietoarkiston aineistot mahdollistavat opiskelijoillekin kattavan kvantitatiivisen tutkimuksen tekemisen. Esimerkiksi Harri Myllyniemi päätyi käyttämään pro gradussaan arkistoitua aineistoa sen vuoksi, että hän halusi laajan, mielellään valtakunnallisen tutkimusaineiston, jossa olisi runsaasti käyttökelpoisia taustamuuttujia (Myllyniemi 2004). Ohjaajan kanssa kannattaa neuvotella oman oppiaineen käytännöistä, esimerkiksi teoriaosuuden ja empiirisen osuuden suhteesta valmiita aineistoja käytettäessä.<sup>6</sup>

---

6 Ks. myös Juha Kääriäisen haastattelu "Itsenäisiä tulkintoja valmiista aineistoista". Tietoarkisto 16 (1/2005). Saatavissa: <http://www.fsd.uta.fi/tietoarkistolehti/16/kaariainen.html>

## 3 Etsi ja löydä sopiva aineisto

### 3.1 Tutkimusongelma, teoria ja aineiston valinta

Aineiston käyttötarkoitus vaikuttaa siihen, mikä aineisto valitaan käyttöön. Esimerkiksi tutkimusmenetelmien harjoituskurssilla saattaa olla perusteltua käyttää aineistoa, joka on helposti ja nopeasti saatavilla, ja valita sitten tutkimusongelma aineiston mukaan. Jos tarkoituksena on harjoitella tilasto-ohjelmiston käyttöä, valitaan aineisto, jonka muuttujat sisältävät tarvittavia mitta-asteikot.

Silloin kun aineistoa käytetään tutkimukseen tai tutkimuksen teon harjoitteluun, valinnan tulisi kuitenkin aina perustua tutkimusongelmaan. Teoreettinen viitekehys ja kysymyksenasettelu ohjaavat sekä aineiston että tutkimusmenetelmien valintaa (Toivonen 1999, 13; Alkula, Pöntinen & Ylöstalo 1995, 15). Tämä periaate koskee sekä uusien että olemassa olevien aineistojen hankintaa ja käyttöä. Kuten Uusitalo (1991, 98) muistuttaa, tutkimuksen ei ”pitäisi olla aineistovetoista, vaan ongelmat ratkaisevat, millaisia aineistoja on tarpeen hankkia.”

Tietoa olemassa olevista aineistoista saadaan tietoaarkiston lisäksi esimerkiksi samoista tutkimuskysymyksistä kiinnostuneilta tutkijoilta, ohjaajilta, tieteellisistä artikkeleista ja tutkimusorganisaatioista. Tässä keskitymme arkistoitujen aineistojen etsimiseen ja löytämiseen.

Ideaalitulanteessa siis arkistoidutkin aineistot valitaan käyttöön tutkimusongelman ja teoreettisen viitekehysten ehdoilla. Käytännössä tilanne on usein toinen. Saatavilla olevat arkistoidut tutkimusaineistot on aina kerätty johonkin toiseen tarkoitukseen, joten omiin tutkimuskysymyksiin soveltuvien aineistojen löytäminen voi olla hankalaa. Ja vaikka aihepiiriltään sopiva aineisto löytyisikin, käytetyt käsitteet ja mittarit eivät ehkä ole juuri toivottuja tai muuttujien luokitukset eivät salli tutkimusongelman ratkaisemista.

Arkistoitujen aineistojen kohdalla haasteena onkin useimmiten löytää tutkimuskysymystä parhaiten vastaava aineisto. Joskus aineiston käyttökelpoisuutta voidaan lisätä esimerkiksi luokittelemalla uudelleen tai yhdistämällä muuttujia. Toisinaan tutkimusongelma joudutaan sovittamaan saatavilla olevan aineiston mukaan. Äärimmäisessä tapauksessa havaitaan, että haluttuun kysymykseen ei saada vastausta käytettävissä olevasta tutkimusaineistosta. Silloin saattaa olla tarpeen luopua aineiston jatkokäyttösuunnitelmista. (Elder, Pavalko & Clipp 1993, 5.) Joka tapauksessa monipuolinen menetelmien hallinta antaa liikkumavapautta aineistojen valinnassa.



## 3.2 Aineiston alustava arviointi

Jatkokäyttäjät välttävät monta aikaa vievää ja työteliästä tutkimusprosessin vaihetta hyödyntäessään toisten keräämää aineistoa. Käytettävä aineisto pitää kuitenkin valita huolellisesti. Lähdekritiikki on arkistoitujen tutkimusaineistojen kohdalla aivan yhtä tärkeää kuin muidenkin lähteiden kohdalla.

Arkistoidun aineiston kohdalla lähdekritiikin ensimmäinen askel on tutustuminen aineiston kuvailutietoihin jo ennen aineiston tilaamista. Esimerkiksi tietoariston verkkosivuilla voi tutustua lähes kaikkien arkistoitujen aineistojen koodikirjoihin. Koodikirja sisältää metadatan eli aineistoa kuvailevaa tietoa. Tyypillisesti koodikirjasta löytyvät aineiston tekijätiedot, sisällön kuvailu, otantatiedot, aineistonkeruun kuvailu, muuttujien suorat jakaumat ja kyselylomake.

The screenshot shows a web browser window with the URL <http://www.fsd.uta.fi/aineistot/luettelo/FSD2133/>. The page title is "YHTEISKUNTATITEELLINEN Tietoarkisto". The main heading is "FSD2133 ISSP 2005 : työorientaatiot III : Suomen aineisto".

**Asiasanat:** ajankäyttö, ammattitaito, ansiotyö, työolot, työsuhde

Tutkimuksessa selvitetään suomalaisten suhtautumista työhön. Kaikilta vastaajilta kysyttiin, mihin elämän eri osa-alueisiin he haluaisivat käyttää enemmän ja mihin vähemmän aikaa, kuinka tärkeinä vastaajat pitävät eri asioita työssään, millaisessa ja minkä kokoisessa yrityksessä vastaajat haluaisivat työskennellä ja millainen työtilanne olisi vastaajille mieluisin. Lisäksi esitettiin erilaisia töitä koskevia väittämiä.

Ansiotyössä olevilta tiedusteltiin lisäksi heidän haluaan pidentää tai lyhentää työaikaansa, työuupumuksesta, mahdollisuudesta vaikuttaa työn alkamis- ja päättymisaikoihin sekä sitä, miten vastaaja voi vaikuttaa päivittäisten töidensä järjestelyyn, tuntuuko työpaikka turvatulta, ovatko tulot ja etenemismahdollisuudet hyvät, kuinka mielenkiintoiselta työ tuntuu, onko työ itsenäistä ja yhteiskunnallisesti hyödyllistä. Lisäksi haluttiin tietää, kuinka sitoutuneeksi työpaikkaansa vastaaja tuntee itsensä, kuinka helppoa tai vaikeaa olisi löytää uusi työpaikka ja kuinka helppoa tai vaikeaa työnantajan olisi korvata vastaaja uudella työntekijällä. Tiedusteltiin myös, kuinka todennäköisesti vastaaja yrittää löytää uuden työpaikan toisesta organisaatiosta ja onko vastaaja huolissaan työpaikkansa menettämisestä. Lisäksi tiedusteltiin ATK-laitteiden käyttöön kuluvaa aikaa työpäivän aikana ja sitä, tekeekö vastaaja päätöksen lisäksi jotakin muuta työtä.

Vastaajilta, jotka eivät ole ansiotyössä, kysyttiin ovatko he koskaan olleet ansiotyössä vuoden tai pidempään, milloin viimeinen ansiotyö päättyi ja mikä oli päättymisen pääasiallinen syy, haluaako vastaaja päästä ansiotyöhön ja kuinka todennäköistä uuden työpaikan löytäminen olisi ja etsiikö vastaaja tällä hetkellä työtä. Lisäksi tiedusteltiin, mitä keinoja vastaaja on käyttänyt työpaikan löytämiseksi, onko hän osallistunut ammattitaitoa kehittävään koulutukseen ja mistä vastaaja saa pääasiallisen toimentulonsa.

Taustatietoina kysyttiin mm. sukupuoli, syntymävuosi, kotitalouden koko, koulutus, työelämään osallistuminen, ammatti, elinkeino tai toimiala, säännöllinen viikoittainen työaika, ammattiasema, esimiesasema, työnantaja (yksityinen/julkinen sektori), ammattiliiton jäsenyys, äänestyskäyttäytyminen, uskonnollisuus, tulot sekä vastaajan asuinalueetta kuvaavia tietoja.

»Pitkä aineistokuvaus

YHTEYSTIEDOT | [fsd@uta.fi](mailto:fsd@uta.fi)

Kuva 2. Yksittäisen aineiston kuvailu

Koodikirjasta löytyy myös lista julkaisuista, joissa aineistoa on jo käytetty. Nämä julkaisulistat eivät aina ole täydellisiä, mutta ne auttavat hahmottamaan,

**[K38] Ajatelkaa pääasiallisen työnne työtunteja ja ansaitsemianne tuloja (mukaan lukien säännölliset ylityöt). Jos voisitte valita vain yhden seuraavista kolmesta vaihtoehdosta, minkä valitsisitte.**

#### Kysymyksen esiteksti

Mikäli olette nykyisin ansiotyössä, vastatkaa kysymyksiin 38 - 59 pääasiallisen työnne perusteella. Mikäli ette ole ansiotyössä, siirtykää kysymykseen 60.

#### Kysymysteksti

*Ajatelkaa pääasiallisen työnne työtunteja ja ansaitsemianne tuloja (mukaan lukien säännölliset ylityöt). Jos voisitte valita vain yhden seuraavista kolmesta vaihtoehdosta, minkä valitsisitte?*

#### Frekvenssit

selite	arvo	n	%	v. %
Pidentäisin työaikaani ja ansaitsisin enemmän	1	112	8.3	14.7
Pitäisin työaikaani ennallaan ja ansaitsisin saman verran	2	516	38.4	67.6
Lyhentäisin työaikaani ja ansaitsisin vähemmän	3	105	7.8	13.8
En osaa sanoa	8	30	2.2	3.9
tieto puuttuu (SYSMIS)	.	582	43.3	—
		1345	100.0	100.0

Kuva 3. Muuttujan kuvaus koodikirjassa (FSD2133).

mihin aineistoa on jo käytetty.<sup>7</sup> Arkistojen tarjoaman dokumentaation lisäksi kannattaa tutustua aineiston keränneen tutkimushankkeen pääraportteihin, joissa on usein kerrottu hyvinkin tarkasti aineiston keruusta ja arvioitu muun muassa aineiston vinoumia.

Olemassa olevan aineiston laatua ja sopivuutta omaan tutkimusongelmaan voi arvioida etsimällä vastauksia seuraavassa esitettyihin kysymyksiin (ks. mm. Stewart 1984, 23–31; Dale, Arber & Procter 1988, 20–31; Seale 2004, 361; van Deth 2003, 297). Käytännössä vastaukset näihin kysymyksiin löytyvät tai niiden ainakin pitäisi löytyä koodikirjoista tai tutkimushankkeiden pääraporteista. On kuitenkin tärkeää huomata, että aineisto saattaa olla hyvinkin käyttökelpoinen, vaikka kaikkia vastauksia ei löytyisikään. Käyttötarkoitus ja -tapa ratkaisevat aineiston sopivuuden.

<sup>7</sup> Julkaisulista ei sen sijaan kerro, mitä on mahdollisesti aiottu tutkia, mutta jätetty syystä tai toisesta tutkimatta. Saattaa olla, että aineistoa jo käyttäneillä tutkijoilla on siis tietoa siitä, mihin aineistoa ei kannata käyttää tai mihin aineistosta vielä löytyisi vastauksia.

#### ALKUPERÄISTÄ TUTKIMUSTA KOSKEVAT KYSYMYKSET

- Mikä oli tutkimuksen tarkoitus?
- Oliko kyse kuvailevasta vai selittävästä tutkimuksesta?
- Mikä oli tutkimuksen teoreettinen viitekehys?
- Kuka tutkimuksen rahoitti?

#### OTANTAA KOSKEVAT KYSYMYKSET

- Mikä on perusjoukko ja miten se valittiin?
- Minkälainen otanta oli?
- Mitä populaatiota aineisto edustaa?
- Voidaanko aineiston pohjalta tehdä yleistyksiä?
- Onko vastanneissa tai vastaamatta jättäneissä vinoumaa?

#### AINEISTONKERUUTA KOSKEVAT KYSYMYKSET

- Milloin aineisto kerättiin?
- Miten aineisto kerättiin: postitse, puhelimitse, Internetissä, kasvokkain?
- Kuka aineiston keräsi?
- Miten keruun laatu varmistettiin?
- Mikä oli vastausprosentti?

#### AINEISTONKERUUISTRUMENTTIA KOSKEVAT KYSYMYKSET

- Miten tutkimusongelmat oli operationalisoitu?
- Mitkä olivat kysymysten ja vastausvaihtoehtojen tarkat sanamuodot?
- Mikä oli kysymysten järjestys?

#### AINEISTOA (DATATIEDOSTOA) KOSKEVAT KYSYMYKSET

- Sisältääkö aineisto tarvittavat taustamuuttujat?
- Onko avovastaukset koodattu?
- Ovatko koodausohjeet saatavilla?
- Kuka aineiston koodasi?
- Minkälaisia uudelleenkodeauksia tehtiin?
- Miten muuttujat on luokiteltu?
- Vastaavatko aineiston jakaumat alkuperäisissä tutkimusraporteissa esitettyjä jakaumia?

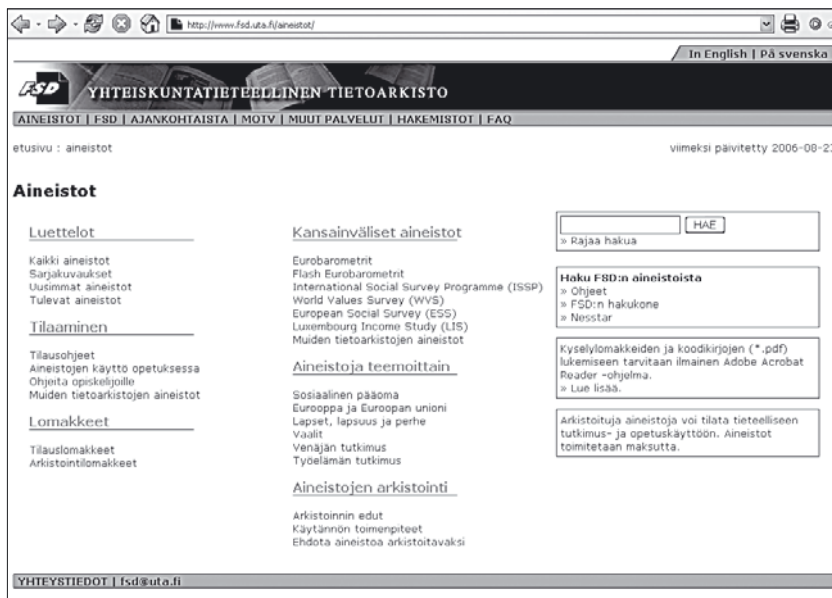
#### OMAAN TUTKIMUSONGELMAAN LIITTYVÄT KYSYMYKSET

- Mihin kysymyksiin etsitään vastauksia?
- Mitä hypoteesejä mahdollisesti testataan?
- Mihin teoreettiseen viitekehukseen aineiston analysointi perustuu?
- Ovatko aineiston muuttujat valmiiksi tarvittavassa muodossa tai riittävän helposti muunnettavissa tarvittavaan muotoon?

Lisäksi kannattaa pohtia, miten keruuajankohdan historialliset, poliittiset, taloudelliset tai kulttuuriset olosuhteet ovat mahdollisesti vaikuttaneet kyselyn vastauksiin.

### 3.3 Aineiston etsiminen Yhteiskuntatieteellisestä tietoaarkistosta

Yhteiskuntatieteellisen tietoaarkiston aineistovarannossa on tätä kirjoitettaessa noin 700 tutkimusaineistoa. Aineistot on kuvailtu hyvin yksityiskohtaisesti tietoaarkiston verkkosivuilla.



Kuva 4. Tietoaarkiston aineistosivu

Arkistoitu kvantitatiivinen aineisto on tyypillisesti kirje-, puhelin- tai käyntikyselyaineisto, jossa havaintoyksikkönä on henkilö ja jossa vastaukset on tallennettu havaintomatriiseiksi. Aineistojen analysoimiseen tarvitaan tilastomatematiikka ohjelmisto (esimerkiksi SPSS, SAS tai Stata) ja vähintään perustiedot kvantitatiivisista menetelmistä. Tietoaarkistossa on myös jonkin verran kvalitatiivisia, tekstimuotoisia tutkimusaineistoja.

Tutkimusaineiston etsiminen tietoaarkistosta kannattaa aloittaa pohtimalla seuraavia kysymyksiä:

- Mitä tutkitaan? Mikä on tutkimusongelma?
- Ketä tutkitaan? Millaista havaintoyksikköä tutkimusongelma edellyttää?
- Millaisia mittareita tai muuttujia tarvitaan? Mitä tutkittavilta on mahdollisesti kysytty?
- Tarvitaanko suomalaisen aineiston lisäksi vertailuaineistoa jostain muusta maasta?
- Mitä ajankohtaa tai ajanjaksoa tutkitaan?

Myös tutkimuksen tai harjoitustyön aikataulu saattaa vaikuttaa aineiston valintaan. Useimmat arkistoidut aineistot saa tietoarkistosta parissa päivässä. Joissain tapauksissa arkistoitu aineisto on kuitenkin vielä alkuperäisellä tutkijalla aktiivikäytössä, jolloin aineisto tai osa sen muuttujista saattaa olla määräaikaisessa käyttökiellossa (embargossa). Lisäksi joidenkin aineistojen toimittamiseen tarvitaan ensin alkuperäisen tutkijan lupa. Tiedot mahdollisista käyttörajoituksista löytyvät aineiston kuvailusta.

Tietoarkisto tarjoaa useita erilaisia hakuvälineitä sen mukaan, minkälaista aineistoa ja minkä tyyppistä tietoa haetaan. Kotimaisia aineistoja voi selailla vaikka tieteenaloitain tai asiansanoittain. Hakuja voi kohdistaa kysymysteksteihin tai vastausvaihtoehtoihin. Tarkemmat hakuohjeet löytyvät tietoarkiston aineistohakusivulta<sup>8</sup>.

	fsd_no	fsd_vr	fsd_pro	fsd_id	nro	k1	k2	k3	k4	k4b	k5	k6	k7	k8	k9_1	k9_2
1	2133	1,0	1	1	1001	2	1940	3	1	1960	2	.	.	.	.	.
2	2133	1,0	1	2	1002	1	1951	2	1	1972	4	3	1	.	1990	.
3	2133	1,0	1	3	1003	1	1954	2	1	1978	2	.	.	.	.	.
4	2133	1,0	1	4	1004	1	1976	2	1	2002	2	2	.	.	.	.
5	2133	1,0	1	5	1005	2	1980	3	1	2001	2	.	.	.	.	.
6	2133	1,0	1	6	1006	1	1989	1	2	.	4	3	1	.	1989	.
7	2133	1,0	1	7	1007	2	1942	5	2	.	1	.	.	.	.	.
8	2133	1,0	1	8	1008	1	1939	2	1	1968	2	2	.	.	.	.
9	2133	1,0	1	9	1009	1	1951	2	1	.	.	.	1	.	1988	.
10	2133	1,0	1	10	1010	2	1959	1	2	.	.	1	1	.	.	.
11	2133	1,0	1	11	1011	1	1977	1	2	.	1	.	.	.	.	.
12	2133	1,0	1	12	1012	1	1986	1	2	.	5	3	2	.	1988	1995
13	2133	1,0	1	13	1013	2	1953	2	1	1971	2	2	.	.	.	.
14	2133	1,0	1	14	1014	2	1951	5	2	.	1	1	.	.	.	.

Kuva 5. Havaintomatriisi (SPSS Data View, FSD2133)

8 [http://www.fsd.uta.fi/aineistot/aineistot\\_hakeminen.html](http://www.fsd.uta.fi/aineistot/aineistot_hakeminen.html)

### 3.4 Kyselyaineistoja muualta maailmasta

Eri maiden tietoarkistojen verkoston välityksellä suomalaistenkin tutkijoiden ja opiskelijoiden käytettävissä on tuhansia kyselyaineistoja eri puolilta maailmaa. Lähes kaikissa Euroopan maissa toimii yhteiskuntatieteellinen tietoarkisto<sup>9</sup>. Lisäksi tietoarkistoja on muun muassa Yhdysvalloissa, Australiassa ja Afrikassa. Niiden arkistoimat aineistot ovat suomalaisten tutkijoiden ja opiskelijoiden käytettävissä joko suoraan tai Yhteiskuntatieteellisen tietoarkiston kautta. Aineistot saa käyttöönsä yleensä maksutta, mutta käytännöt vaihtelevat maittain. Muiden arkistojen aineistoista kiinnostuneiden kannattaakin ottaa yhteyttä Yhteiskuntatieteelliseen tietoarkistoon.

Laajoihin kansainvälisiin sosiaalitieteellisiin vertailututkimusaineistoihin tutustuminen kannattaa myös aloittaa Yhteiskuntatieteellisen tietoarkiston aineistosivuilta. Sivuilla esitellään muiden muassa World Values Survey, European Social Survey, Luxembourg Income Study, International Social Survey Programme ja Eurobarometrit.

Yhteiskuntatieteellinen tietoarkisto ja sen myötä suomalaiset yliopistot ovat jäseniä alan suurimmassa arkistossa, ICPSR:ssä, joka on arkistoinut kymmeniä tuhansia aineistoja ympäri maailmaa. ICPSR:n tietokantaan voi tehdä monipuolisia hakuja, ja aineistojen kuvailutiedot ovat rekisteröityneiden käyttäjien vapaasti saatavilla.<sup>10</sup> Jos ICPSR:n valikoimista löytyy sopiva tutkimusaineisto, sen voi tilata käyttöönsä Yhteiskuntatieteellisestä tietoarkistosta. Myös ICPSR:n aineistot toimitetaan maksutta.

Kaikille tietoarkistoille yhteisenä tehtävänä on alusta asti ollut tutkimusaineistojen arkistointi, aineistojen jakelu jatkokäyttöön sekä menetelmäopetuksen tukeminen.<sup>11</sup> Lisäksi arkistot ovat usein mukana koordinoimassa suuria kansainvälisiä tutkimushankkeita. Esimerkiksi Saksan tietoarkisto hoitaa Eurobarometri- ja International Social Survey Programme -aineistojen arkistoinnin ja jakelun, ja Norjan tietoarkiston vastuulla on European Social Survey -aineisto<sup>12</sup>.

9 Eurooppalaisten tietoarkistojen kotisivuille pääsee alan eurooppalaisen kattojärjestön, CESSDAn, sivuilta: <http://www.cessda.org/>

10 <http://www.icpsr.umich.edu/>

11 Ensimmäinen yhteiskuntatieteellinen tietoarkisto, Roper Center, perustettiin Yhdysvaltoihin vuonna 1947. Euroopan ensimmäinen data-arkisto Zentralarchiv für Empirische Sozialforschung avattiin Saksan Kölnissä vuonna 1960. Pohjoismaista ensimmäisenä oman data-arkiston sai Norja vuonna 1971, Tanska seurasi perässä vuonna 1973 ja Ruotsi 1985.

12 Ks. ISSP: <http://www.issp.org/> ja ESS: <http://www.europeansocialsurvey.org/>

### 3.5 Muita aineistoja muualta: esimerkkejä tilasto- ja rekisteriaineistoista

Suomalaisten ja kansainvälisten tilastoaineistojen lähteille pääsee parhaiten Tilastokeskuksen sivuilta<sup>13</sup>. Hyödyllinen on muun muassa StatFin-tilastopalvelu<sup>14</sup>, joka tarjoaa Suomea koskevaa tilastotietoa veloitusetta. Tilastokirjaston WebStat-tietokannasta<sup>15</sup> voi puolestaan etsiä Internetin kautta saatavilla olevia tilastoja. Eurooppalaisia tilastotietoja löytyy runsaasti Eurostatin online-tietokannoista, jotka ovat maksutta itsepalvelukäytössä Eurostatin verkkosivujen<sup>16</sup> kautta. Tilastokirjaston EU-tilastotietopalvelu<sup>17</sup> toimii itsenäisen EU-tilastotiedonhaun tukena.

Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus Stakes julkaisee virallisia sosiaali- ja terveyshuollon tilastoja ja alan tietokantoja verkkosivuillaan<sup>18</sup>. Stakesin ja muiden sosiaali- ja terveysalan toimijoiden yhteistyönä kehitetty indikaattori-pankki SOTKANet<sup>19</sup> sisältää puolestaan kuntapohjaista hyvinvointidataa, esimerkiksi kuntakohtaisia tietoja väestön terveydestä.

Jotkut yliopistokirjastot ovat ostaneet tilastotietokantoja omien asiakkaidensa käyttöön. Esimerkiksi OECD:n tilastoja sisältävä SourceOECD, Yhdistyneiden kansakuntien Internet-tietokanta UNCDB, Kansainvälisen valuuttarahaston verkkotietokanta ja Maailmanpankin World Development Indicators -tietokanta ovat käytössä joissakin kirjastoissa. Lisäksi Kansainvälinen työjärjestö, ILO, julkaisee nykyisin tilastojaan vapaasti verkossa Laborsta-palvelussaan. (Laaksonen 2005.)

Rekisteriaineistoista kiinnostuneiden kannattaa tutustua Rekisteritutkimuksen tukikeskuksen, ReTKin, verkkosivuihin<sup>20</sup>. ReTKi edistää kansallisten rekisterien tutkimuskäyttöä erityisesti terveys- ja sosiaalitieteissä. Keskuksen sivusto sisältää muun muassa opastusta terveydenhuollon salassa pidettävien tietojen käyttöilupien hakuun sekä linkkejä muutamien suomalaisten terveysalan rekisterien kuvauksiin.

Lisää linkkejä tilastoihin, tietokantoihin ja tutkimusaineistoihin löytyy Yhteiskuntatieteellisen tietoarkiston linkkisivulta<sup>21</sup>.

13 <http://www.tilastokeskus.fi/>

14 <http://www.stat.fi/tup/statfin/>

15 <http://www.tilastokeskus.fi/tup/tilastokirjasto/>

16 <http://europa.eu.int/comm/eurostat/>

17 <http://www.tilastokeskus.fi/tup/esds/>

18 <http://www.stakes.fi/>

19 <http://www.sotkanet.fi/>

20 <http://www.rekisteritutkimus.fi/>

21 <http://www.fsd.uta.fi/linkit/>

## 4 Aineiston tilaaminen ja käsittely

### 4.1 Aineiston tilaaminen tietoaarkistosta

Tietoaarkistoon arkistoituja aineistoja annetaan tieteelliseen tutkimus-, opetus- ja opiskelukäyttöön. Aineiston voi tilata siten esimerkiksi opetusmateriaaliksi tai tieteellistä artikkelia, väitöskirjaa, pro gradua, kandidaatintutkielmaa, seminaarityötä tai menetelmäkurssin harjoitustyötä varten. Kaupallisiin tarkoituksiin aineistoja ei toimiteta. Tietoaarkiston kautta käytettävissä on myös muiden maiden tietoaarkistoihin tallennetut tuhannet tutkimusaineistot.

Aineisto tai aineistot tilataan täyttämällä, allekirjoittamalla ja lähettämällä tietoaarkistoon kaksi lomaketta: käyttöluvahakemus ja käyttöehtositoumus (liitteenä).

Käyttöluvahakemuksessa ilmoitetaan tiedot muun muassa aineiston käyttäjästä, käyttötarkoituksesta, tutkimuksen mahdollisesta rahoituksesta ja toivotusta tiedostomuodosta (formaattista). Perustutkinto-opiskelijoilta vaaditaan lisäksi ohjaajan allekirjoitus. Käyttöehtositoumukseseen on koottu aineiston käyttöä koskevat ehdot. Jokaisen käyttäjän on allekirjoitettava oma käyttöehtositoumuksensa.

Hyväksytyään käyttöluvahakemuksen tietoaarkisto toimittaa tutkimusaineiston yleensä parissa päivässä. Datatiedoston lisäksi jatkokäyttäjälle toimitetaan aineiston kuvailutiedot sisältävä koodikirja ja mahdolliset kyselylomakkeet PDF-formaatissa. Nämä perustiedostot toimitetaan maksutta.

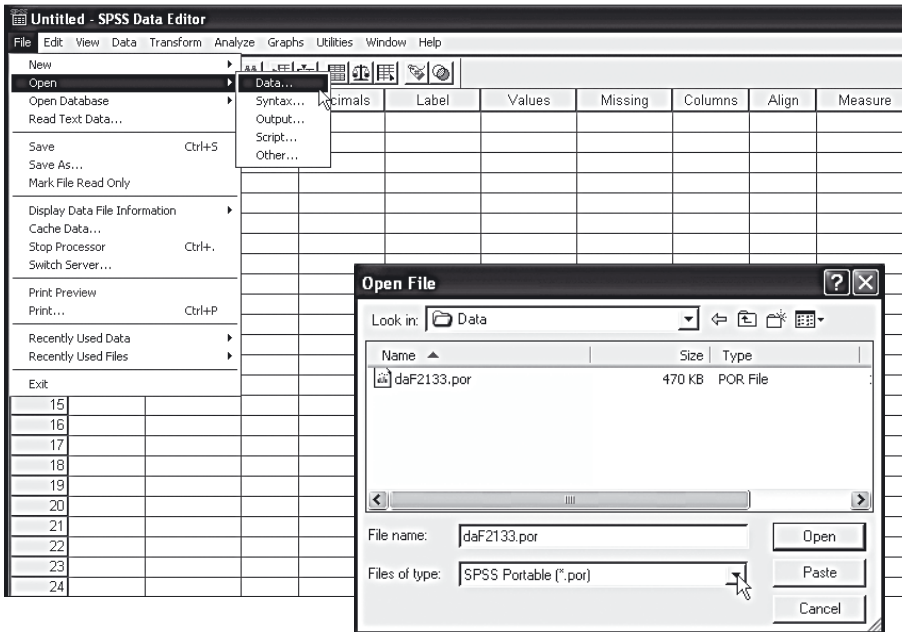
### 4.2 Datatiedoston avaaminen ja käsittely

Tietoaarkistosta saatu kvantitatiivinen aineisto on yleensä SPSS portable -muotoinen datatiedosto eli havaintomatriisi (tiedostopääte \*.por). Datatiedoston lisäksi toimitetaan koodikirja ja kyselylomake tai -lomakkeet PDF-muodossa (\*.pdf).

SPSS portable on versiosta riippumaton, siirrettävä tiedostomuoto. Portable-tiedosto voidaan tarvittaessa lukea eri ohjelmilla<sup>22</sup>, esimerkiksi SAS-ohjelmistolla tai tavallisella tekstinkäsittelyohjelmalla. SPSS-ohjelmassa portable-tiedoston saa avaamista varten näkyviin valitsemalla pudotusvalikosta File -> Open -> Data ja valitsemalla kohdasta 'Files of type' tiedostomuodoksi SPSS portable (\*.por).

22 Lista tuetuista formaateista: <http://www.fsd.uta.fi/aineistot/jatkokaytto/tiedostoformaattit.html>





Kuva 6. SPSS portable -tiedoston avaaminen.

Jos tutkimusaineisto on toimitettu sähköpostin liitteenä, liitetiedostot kannattaa ennen avaamista tallentaa oman koneen kiintolevyille. Samoin cd-rom-levyllä toimitetuista tiedostoista kannattaa ottaa työkopio omalle kiintolevyille. Aineistoa pitää aina kuitenkin säilyttää niin, että sitä voivat käyttää vain käyttöoikeuden omaavat. Aineistoa ei siis ole syytä tallentaa mikroluokan koneen kovalevyille, vaan esimerkiksi salasanalla suojattuun kotihakemistoon.

Ennen analyysyä tulee varmistaa, että havaintomatriisitiedoston muuttujien jakaumat vastaavat koodikirjassa olevia jakaumia. Samalla voi tarkistaa, miten puuttuvat tiedot on määritelty. Jos aineisto on hyvin suuri, työkopiosta kannattaa ajojen nopeuttamiseksi pudottaa pois oman tutkimuksen kannalta tarpeettomat muuttujat tai havainnot.

Tietoarkiston toimittamissa datoissa on mukana yleensä vain suoraan kyselylomakkeeseen perustuvat muuttujat. Useinkaan näitä ei voi käyttää sellaisinaan, vaan jatkokäytössä muuttujia on luokiteltava uudelleen tai on luotava omia muuttujia esimerkiksi yhdistämällä useita eri muuttujia. Oman työn sujuvuuden kannalta kaikki muuttujien yhdistämiset tai muutokset kannattaa dokumentoida huolellisesti.

## 5 Tutkimusaineistojen jatkokäytön haasteita

### 5.1 Jatkokäytön hyödyt suuremmat kuin haitat

Jatkokäyttäjä kohtaa aineiston analyysissä yleensä aivan samoja haasteita kuin uudenkin aineiston kerännyt tutkija. Jatkokäyttäjänkin pitää olla tietoinen esimerkiksi otantaan ja mittaamiseen liittyvistä yleisimmistä ongelmista ja rajoitteista. Koska arkistoituja aineistoja käytetään usein vertailuaineistoina, vertailututkimukselle tyypilliset ongelmat korostuvat.

Ongelmien esiintyminen ei suinkaan tarkoita, että tutkimusaineistoa ei voisi käyttää. Lähes kaikkiin ongelmiin on olemassa ratkaisu tai ainakin hyväksyttävä kompromissi. Aineiston käytettävyyttä voidaan parantaa esimerkiksi hyödyntämällä useampia aineistoja tai muokkaamalla aineistoa erilaisilla *ex post* -toimenpiteillä, kuten muodostamalla uusia luokitteluja tai indikaattoreita. Olemassa olevan aineiston käyttöön liittyvät edut ovat yleensä aina suurempia kuin haitat. (Elder, Pavalko & Clipp, 10–11, van Deth 2003, 306; Jacob 1994, 375–381; Kiecolt & Nathan 1985, 52–72.)

Jatkokäytön on joskus pelätty johtavan luovuuden katoamiseen. Jos tutkijat käyttävät toistuvasti samoja aineistoja, lopputuloksena saatetaan tutkia vain tiettyjä ilmiöitä tiettyjen muuttujien avulla, mikä voi uhata tieteen kehitystä. Toisaalta yhdenmukaisten mittareiden käyttäminen on edellytys vertailevalle tutkimukselle. Niin kauan kuin kyselyaineistoihin otetaan mukaan uusia mittareita vanhojen rinnalle, paikallaan polkemisen vaaraa ei kuitenkaan ole. (Kiecolt & Nathan 1985, 14.)

Kyselytutkimuksissa tutkijan ja tutkittavien välinen etäisyys on tyypillisesti suuri. Kirjekyselyissä vastaaja täyttää lomakkeen itse ja haastattelututkimuksissakin haastattelijana on yleensä joku muu kuin tutkija itse. Näin saatetaan helposti menettää tietoa esimerkiksi lomakkeessa käytettyjen luokittelujen toimivuudesta ja siitä, miten kysymykset on ymmärretty. Toisten kokoamia aineistoja käytettäessä etäisyys tutkittaviin ja koko aineiston keräämisen kontekstiin luonnollisesti kasvaa. (Alastalo 2005, 245–248, 263–264.) Etäisyyttä voi yrittää kuroa umpeen tutustumalla esimerkiksi alkuperäisen tutkimuksen tuottamiin julkaisuihin, aineiston dokumentaatioon ja aineistossa mahdollisesti oleviin avokysymysten vastauksiin.

Arkistoidun tutkimusaineiston käytön tulee aina perustua aineiston vahvuuksiin. Kaikkien analyysissä käytettävien mittareiden on luonnollisestikin oltava luotettavia (reliaabeleja) ja päteviä (valideja). Usein aineiston erilaisia luotettavuustekijöitä on jo käsitelty aineistoon pohjautuvissa aikaisemmissa julkaisuissa.

## 5.2 Operationalisoinnit, kysymysten järjestys ja asiayhteys

Yhteiskuntatieteellisissä tutkimuksissa tarkastellaan usein hyvin abstrakteja käsitteitä, kuten esimerkiksi suvaitsevaisuutta, tasa-arvoa tai onnellisuutta. Kvantitatiivinen tutkimus edellyttää näiden abstraktien käsitteiden muuntamista ensin mitattaviksi asioiksi ja sitten mittareiden luontia. Prosessia kutsutaan operationalisoinniksi. Toisistaan poikkeavat operationalisoinnit tuottavat samasta käsitteestä hyvinkin erilaisia mittareita. Lisäksi erilaisten ilmiöiden, termien ja käsitteiden merkitys riippuu aina kontekstista. Jatkokäyttäjän tehtävänä on käytettävissä olevan dokumentaation pohjalta päätellä, ovatko saatavilla olevat muuttujat valideja omien teoreettisten käsitteiden indikaattoreiksi. Kansainvälisten vertailututkimusten reliabiliteettia ja validiteettia testataan usein lisälomakkeilla, joissa samaa asiaa kysytään eri tavalla eri vastaajilta<sup>23</sup>.

Tulosten tulkinnassa on hyvä pohtia myös vastaajalle esitettyjen kysymysten järjestystä. Tästä järjestysvaikutuksesta<sup>24</sup> antavat esimerkin Schwarz ja Hippler (1995, 95–96).

Vastausjakaumiin voi vaikuttaa myös asiayhteys, jossa kysymys esitetään. Turun yliopiston sosiaalipolitiikan laitoksella tehtiin vuonna 1993 kyselytutkimus, jossa selvitettiin kansalaisten näkemyksiä julkisen sektorin leikkauksista ja sosiaalipolitiikan suunnasta (FSD1212). Aineiston avulla tutkittiin muun muassa kysymysten muotoilun vaikutusta mielipidejakaumiin (Kangas 1995, 129–136).

### ESIMERKKI 1.

#### JÄRJESTYSVAIKUTUS

Schwarz ja Hippler (1995) keräsivät tutkimusaineistonsa talvella 1991 haastatella 229 vastaajaa puhelimitse ja lähettämällä 228 vastaajalle kyselylomakkeen postitse (näistä 183 vastaajaa palautti lomakkeen). Tutkimuksen keskeisenä kysymyksenä oli, kuinka paljon rahaa vastaajat olisivat valmiita lahjoittamaan auttaakseen Venäjän kärsiviä kansalaisia. Heti ennen tätä lahjoitushalukkuutta kartoittavaa kysymystä tai heti sen jälkeen esitettiin kaksi veroihin liittyvää kysymystä, joista ensimmäisessä kysyttiin hyväksyntää tulevalle arvonlisäveron korotukselle ja toisessa haluatko vastaajat valtion mieluummin lisäävän hyvinvointimenoja vai vähentävän veroja. Vastaajien halu tukea venäläisiä oli suurin niissä puhelinhaastatteluissa, joissa verokysymykset tulivat vasta lahjoituskysymyksen jälkeen. Lahjoitusvalmius pieni huomattavasti, kun vastaajille esitettiin ensin veronkorotuksiin liittyvät kysymykset. Postikyselyssä vastaajien ilmoittamat rahasummat olivat kysymysten järjestyksestä riippumatta pienempiä kuin puhelinkyselyssä. Postikyselyssä siis myös jäljempänä tulleet verokysymykset vaikuttivat lahjoitusvalmiutta koskevaan vastaukseen.

23 Ks. esimerkiksi European Social Survey (ESS): <http://www.europeansocialsurvey.org/>

24 Dillman (2002, 89) esittelee laajemmin erilaisia järjestysvaikutuksia aiheuttavia tekijöitä ja esimerkkejä.

## ESIMERKKI 2.



## ASIAHYTEYDEN VAIKUTUS

Heikki Ervasti ja Olli Kangas (1995) tutkivat, miten tiedon lisääminen tai yhteisvastuullisuuteen ja moraaliin vetoaminen vaikuttavat kysymysten vastausjakaumiin. Aineisto kerättiin vuonna 1993 kahdella kyselylomakkeella. Vastaajista puolet vastasi ensimmäiseen ja puolet toiseen lomakkeeseen. Toisen lomakkeen joidenkin likert-asteikollisten väittämien sanamuotoa muutettiin hieman. Moraaliin vetoamisen vaikutusta tarkasteltiin kahdella asenneväittämällä. Ensimmäinen väite ”yhteiskunnassa vahvojen ryhmien tulee pitää huolta heikommassa asemassa olevista” esitettiin toisessa lomakkeessa muodossa ”inhimillisessä yhteiskunnassa vahvojen ryhmien tulee pitää huolta heikommassa asemassa olevista”. Toisen väitteen ”työssä olevilta tulisi kerätä työttömyysvero, joka käytettäisiin työttömien toimeentulon parantamiseksi” sanamuoto oli toisessa lomakkeessa ”Työttömyys aiheuttaa perheille suuria toimeentulovaikeuksia. Siksi työssä olevilta tulisi kerätä työttömyysvero, joka käytettäisiin työttömien toimeentulon parantamiseksi.” Pieni muutos vaikutti joissakin tapauksissa merkittävästi vastausjakaumiin. Ervasti ja Kangas havaitsivat, että yleisellä asenteellisella tasolla valtaosa vastaajista suhtautuu hyvin myönteisesti yhteisvastuullisuuteen: vahvojen on tuettava heikompiaan. Vastaajien tulotason ja sosioekonomisen aseman tarkastelussa havaittiin, että hyväituloiset ja ylemmät toimihenkilöt, jotka korkeampien verojensa takia ovat tukijan roolissa, suhtautuvat muita ryhmiä kriittisemmin yhteisvastuullisuuteen. Tarkemmissa kysymyksissä solidaarisuuteen suhtauduttiin varauksellisesti. Kaksi kolmesta vastaajasta ei halunnut maksaa työttömille kerättävää työttömyysveroa. Erityisveron vastustajien osuus laski noin puoleen, kun viitattiin perheiden toimeentulovaikeuksiin.

### 5.3 Otantaongelmat ja kato

Otantatutkimuksissa tehtävänä on usein arvioida äärellisen perusjoukon ominaisuuksia satunnaisotoksen perusteella. Satunnaisotoksen perusteella saadut tulokset voidaan tietysin ehdoin yleistää koskemaan koko perusjoukkoa. Puuttuvien havaintojen, mittausvirheiden tai muiden epäluotettavuustekijöiden vähäinen esiintyminen ei välttämättä haittaa tutkimusta, kunhan niihin kiinnitetään riittävästi huomiota. Toisaalta useissa otantatutkimuksissa ollaan käytännössä usein niin etäällä otantamenetelmien käyttöedellytyksistä, ettei tavanomaisia yleistyksiä voida pitää luotettavina. (Vartia 1988, 306.)

Erityisesti ongelmia aiheuttavat muut kuin satunnaisotannalla kerätyt aineistot, joiden kohdalla tulosten otantavirheen estimointi on hankalaa tai jopa mahdotonta (Jacob 1994, 341–342).<sup>25</sup> Etenkin tilastollisessa testauksessa on hyvä myös huomata, että monien tilasto-ohjelmistojen (esimerkiksi SPSS ja SAS) perusanalyysimenetelmät perustuvat oletukseen satunnaisotannasta (Sturgis 2004).

25 Otantateoriasta ja otantavirheistä kiinnostuneiden kannattaa tutustua teoksiin Pahkinen ja Lehtonen (1989): Otanta-asetelmat ja tilastollinen analyysi ja Lehtonen ja Pahkinen (1996): Practical methods for design and analysis of complex surveys sekä Groves et al. (2004): Survey Methodology, 93–131.

Tutkijan – myös jatkokäyttäjän – pitää osoittaa aineiston edustavuus suhteessa perusjoukkoon. Valitettavasti katoanalyysijä ei ole aina sisällytetty alkuperäisten tutkimushankkeiden julkaisuihin, vaikka analyysit olisi tehtykin. Kadolla tarkoitetaan joko kokonaisten otosyksiköiden (esimerkiksi henkilöiden) puuttumista tutkimusaineistosta tai puuttuvia tietoja joidenkin havaintoyksiköiden tiedoissa<sup>26</sup>. Kyselyaineistojen kohdalla puhutaan usein myös vastaaja- tai vastauskadosta.

Katoanalyysi koskee yleensä enimmäkseen vastaajakatoa, jolloin tarkastellaan, kuinka hyvin otos vastaa perusjoukkoa. Kun aineiston havaintoyksikkönä on henkilö, verrataan esimerkiksi aineiston ikä-, sukupuoli-, siviilisäätö-, koulutus-, ammatti- ja asuinpaikkajakaumia tilasto- tai rekisteritietoihin. Satunnaisesti jakautunut vähäinen kato ei yleensä ole ongelma, mutta kato voi olla myös systemaattista. Systemaattisia vinoumia voi korjata painomuuttujien avulla. (Alkula, Pöntinen & Ylöstalo 1995, 112–113; Groves, Fowler, Couper, Lepkowski, Singer & Tourangeau 2004, 321–328).

Aineiston perusjoukon suhteen kannattaa olla tarkkana. Esimerkiksi puolueiden ajankohtaistutkimuksissa siirryttiin 1990-luvun alussa käyntihaastatteluista GallupKanavan kotitietokonevastaamiseen, ja samalla vastaajakunnan ikärakenne muuttui jonkin verran. Tuloksia ei pidä myöskään yleistää koskemaan kaikkia suomalaisia, jos aineiston perusjoukko on 18–65-vuotias manner-Suomen väestö, jolloin muun muassa alaikäiset, vanhuksset ja Ahvenanmaa rajautuvat pois.

## 5.4 Erilaiset tiedonkeruumenetelmät

Aineiston keruumenetelmä vaikuttaa kysymysten muotoiluun ja siihen, mitä yleensä on mahdollista kysyä. Postitettavalla lomakkeella ei voi luotettavasti mitata vastaajien tietotasoa ja puhelinkyselyssä kysymykset ja vastausvaihtoehdot on esitettävä tiiviimmin kuin kasvokkain tapahtuvassa käyntikyselyssä. Eri aineistonkeruuorganisaatioilla on yleensä myös oma tapansa tehdä otannat ja kouluttaa haastattelijat, mikä saattaa hieman vähentää eri organisaatioiden tuottamien aineistojen vertailukelpoisuutta. Tiedonkeruutapa vaikuttaa kuitenkin enemmän kuin keruuorganisaatio.

Tiedonkeruumenetelmien tuottamia eroja tarkastelleiden – lähinnä ulkomaisien – tutkimusten mukaan sosiaalisesti etäisemmissä tilanteissa vastaajat kertovat avoimemmin asioita, jotka eivät esitä heitä myönteisessä valossa. Kirjekyselyssä vastaajalla on enemmän aikaa vastata ja arvioida tutkimuksen luotettavuutta kuin puhelin- tai käyntikyselyissä. Vaikka kirjekyselyt tuottavatkin enemmän puuttuvia vastauksia kuin puhelin- tai käyntikyselyt, niin arkaluonteisia asioita kerrotaan kuitenkin avoimemmin ja tarkemmin. Vastaajat esimerkiksi ilmoittavat tu-

26 Ks. Tilastokeskuksen verkkokoulu: <http://www.stat.fi/tup/verkkokoulu/index.html>

lojaan tarkemmin ja kertovat laittomasta toiminnastaan sekä uhrikokemuksistaan vapaammin itse täytettävällä lomakkeella. Lisäksi kirjekyselyissä paljastetaan puhelinkyselyitä helpommin ikäviä puolia itsestä ja muista. Lomakkeen täyttäminen itse vähentää myös myönteisten asioiden yliparportointia. (Kiecolt & Nathan 1985, 66–68; Ahola 1998, 262–271; Ahola 2003, 46–52; Ahola 2004, 47–52; Piispa 2006, 150–151.)

Suomessa on saatu tosin myös päinvastaisia tuloksia. Anja Ahola (2003) on todennut, että suomalaiset kertovat hyvinvoinnistaan helpommin läheisimmässä eli käyntihaastattelutilanteessa. Myös muualla erityisesti sosiaalisesti epävarmassa asemassa olevat ihmiset kertovat aroista asioista enemmän kasvotusten. Puhelinta tietokoneavusteiset kyselyt antavatkin alempiin sosiaaliryhmiin kuuluvien hyvinvointikokemuksista myönteisemmän kuvan kuin perinteiset (so. ei tietokoneavusteiset) käyntikyselyt. (Ahola 2003, 46–52.)

Puhelinkyselyjen erityisongelmana on, että niistä jäävät pois vastaajat, joilla ei ole lainkaan puhelinta tai jotka salaavat puhelinnumeron. Suomessa suurimpien aineistonkeruuorganisaatioiden, kuten Tilastokeskuksen, Taloustutkimuksen ja TNS-Gallupin, puhelinhaastattelut kohdistuvat nykyään myös niihin henkilöihin, joilla on ainoastaan matkapuhelinliittymä.

Keruumenetelmä vaikuttaa myös otantamenetelmän valintaan – ja päinvastoin. Esimerkiksi valtakunnallisissa käyntikyselyissä on vain harvoin varaa toteuttaa yksinkertaista satunnaisotantaa, jossa missä tahansa kunnassa asuva kansalainen voi tulla poimituksi otokseen. Yleensä käyntikyselykustannuksia pienennetään ryvästämällä otosta alueellisesti.

## 5.5 Mittarien vertailtavuus

Operationalisoinnin ongelmat korostuvat vertailevassa tutkimuksessa. On selvää, että mittarit eivät voi olla kaikilta osiltaan täysin identtisiä eri aikoina, eri maissa tai eri kielillä tehdyissä tutkimuksissa. Arkistoitu aineisto heijastaa aina alkupe-  
räisten tutkijoiden näkökulmia, joihin puolestaan ovat vaikuttaneet ajan tieteelliset ja kulttuuriset arvot ja käytännöt. Lisäksi kontekstuaaliset asiat, esimerkiksi kenttätö-  
n aikaiset poliittiset tapahtumat, poikkeukselliset ilmiöt tai vaikkapa luonnonmullistukset, voivat vaikuttaa ihmisten vastauksiin siten, että tietojen vertailukelpoisuus vähenee.

Tyypillistä on, että vastausvaihtoehtojen määrä vaihtelee tai vastausvaihtoehtot ovat sisällöllisesti erilaiset. Esimerkiksi asenteiden mittauksessa käytetty ”samaa mieltä–eri mieltä” -asteikko voi olla vaikkapa kolme- tai viisiportainen. Ja vaikka kysymyksen muotoilu olisi sama, muuttujien sisältö eroaa, jos kysymys on kohdistettu vain osalle vastaajista. Vaikkapa tuloja voidaan eri tutkimuksissa kysyä joko kaikilta vastaajilta tai vain kokoaikaisilta työntekijöiltä. (Kiecolt & Nathan 1985, 56–59.)

On myös tavallista, että kysymysten sanamuoto vaihtelee eri kyselyissä. Tutkijan pohdittavaksi jääkin, kuinka paljon kysymysten ja vastausvaihtoehtojen erot mahdollisesti vaikuttavat aineistojen vertailtavuuteen. Joka tapauksessa toisistaan poikkeavat kysymyksenasettelut on hyvä mainita raportissa.

### ESIMERKKI 3.

#### KYSYMYSTEN SANAMUOTO

Politiikan tutkimuksessa sisäistä kansalaispätevyyttä mitataan usein likert-asteikollisilla väittämällä, joista yksi käsittelee kansalaisten voimattomuuden tunnetta. Englanninkielisissä kyselyissä väittämä on yleensä muotoa ”people like me don’t have any say about what the government does”. Ennen 1990-lukua tehdyissä suomalaisissa kyselyissä väittämä on esitetty tyypillisesti muodossa ”minunlaisillani ihmisellä ei ole mitään sanomista siihen, mitä hallitus tekee”<sup>27</sup>. 1990-luvulla ja sen jälkeen kerätyissä aineistoissa on käytetty (oikeampaa) termiä ”sananalta” ja väittämä on useimmiten muotoiltu ”minulla ei ole mitään sananaltaa siihen, mitä maan hallitus ja eduskunta päättävät”<sup>28</sup>. Muitakin muotoiluja on käytetty, esimerkiksi ”minunkaltaisillani ihmisellä ei ole mitään sananaltaa siihen, mitä valtiovalta tekee”<sup>29</sup> ja ”minunlaisillani ihmisillä ei ole mitään sananaltaa siihen, mitä hallitus tekee”<sup>30</sup>. Myös vastausasteikko on eri vuosina ollut erilainen, esimerkiksi vaihtoehto ”en osaa sanoa” on joskus sijoitettu asteikon keskelle, joskus asteikon loppuun.

Kyselyn kohdejoukkokaan ei ole aina selväpiirteinen ja muuttumaton. Esimerkiksi ruokakunta, kotitalous, asuntokunta ja perhe on määritely eri tavoin eri aikoina, eri kulttuureissa ja eri tutkimuksissa. Kuntien yhdistyminen tai muut maantieteellisten tai hallinnollisten rajojen muutokset saattavat aiheuttaa ongelmia aineistojen käytössä ja vertailussa. Pitkittäistutkimuksessa tulee kiinnittää huomioita myös taustatekijöiden muutoksiin. Esimerkiksi väestön koulutustaso on noussut huomattavasti viime vuosikymmenten aikana. (Kiecolt & Nathan 1985, 62–67, Kurkela & Sauli 1998, 27–42.)

Koulutus- ja ammattimuuttujiin kannattaa kiinnittää erityistä huomiota. Näitä taustamuuttujia on luokiteltu eri tavalla eri aikoina ja eri tutkimusorganisaatioissa. Suomessa työhallinto käyttää pääasiassa Pohjoismaista ammattiluokitusta ja Tilastokeskus ILO:n kansainvälisen ammattiluokituksen EU-versiota

27 Esimerkiksi aineistoissa FSD1009, FSD1011, FSD1014, FSD1016.

28 Esimerkiksi aineistoissa FSD1260, FSD2017, FSD1027, FSD1021, FSD1022, FSD1023 ja FSD1247.

29 Esimerkiksi aineistossa FSD0116.

30 Esimerkiksi aineistossa FSD2039.

(ISCO 88-COM).<sup>31</sup> Tilastokeskus myös ylläpitää ja julkaisee kansallisia luokitus-suosituksia<sup>32</sup>.

Käytännössä vertailututkimuksessa on etsittävä samaa asiaa mittaavia mittareita, ei ulkoisesti identtisiä mittareita. Sopivan aineiston hakemisessa ei kannata tuijottaa vain kysymyksen muotoiluun, vaan etsiä laajemmin aineistoja, joissa samoja asioita mitataan eri mittareilla. (van Deth 2003, 302–303; Jacob 1994, 346–363.)

## 5.6 Kansainvälisten vertailuaineistojen käytön erityispiirteet

Jatkokäyttäjien on aina syytä perehtyä huolellisesti käyttämänsä tutkimusaineiston tietosisällön dokumentointiin ja tutkimusraportteihin. On olennaista selvittää, miten perusjoukon valinta ja otos, havaintoyksikkö, aineistonkerääjien ohjeet ja aineistonkeruun muut menettelytavat, vastausprosentit, mittausvirheet, aineiston vinoumat, tutkimusajankohdat, tutkimuskysymykset sekä niistä koodatut vastaukset vaikuttavat aineiston käyttökelpoisuuteen oman tutkimuksen kannalta.

Kansainvälisessä tutkimuksessa on lisäksi perehdyttävä huolellisesti sekä tutkimusten että tutkimuskohteidenkin kulttuurisiin, kansallisiin, historiallisiin ja ajankohtaisiin yhtäläisyyksiin ja eroihin. Myös vastaajien taustatiedot, esimerkiksi koulutus, ammatti, sosiaalinen ryhmä ja tulot, ovat hyvin kulttuurisidonnaisia.

Kansainvälisten aineistojen vertailtavuutta tutkittaessa on keskeistä selvittää muuttujissa mahdollisesti olevat vinoumat ja vastaavuudet. Van de Vijver (2003, 145–148; ks. myös Lehtinen 2004, 60) ryhmittelee nämä vinoumat kolmeen eri tyyppiin: käsitteellisiin vinoumiin, kysymyskohtaisiin vinoumiin ja metodisiin vinoumiin.

Käsitteellisen vinouman tunnistaminen vaatii laajaa tietämystä tarkastelun kohteena olevista kulttuureista. Ovatko vastaajat ymmärtäneet vaikkapa uskonnollisuuteen tai hyvään vanhemmuuteen liittyvät kysymykset samalla tavalla eri puolilla maailmaa?

31 Pohjoismainen ammattiluokitus perustuu ruotsalaiseen ammattiluokitukseen ”Nordisk yrkesklassificering” (NYK) vuodelta 1983. NYK vastaa kolminumerotasolle Pohjoismaisen Työmarkkinatoimikunnan (NAUT) yhteispohjoismaista luokitusta (NYK 82). NYK 82 puolestaan pohjautuu ILO:n kansainväliseen ammattiluokitukseen (International Standard Classification of Occupations ISCO). Kansainvälisen ammattiluokituksen nykyinen, kolmas versio on ISCO-88.

32 Käytössä olevat luokitukset löytyvät verkosta: <http://www.stat.fi/tk/tt/luokitukset/>



## ESIMERKKI 4.

## KÄSITTEELLINEN VINOUMA

Kulttuureissa, joissa voimakkaat mielipiteenilmaisut eivät ole sosiaalisesti hyväksytyjä, tulokset saattavat olla erilaisia riippuen siitä, lievennetäänkö vastausasteikkojen sanamuotoja vai ei. Esimerkiksi Japanin ISSP-aineistossa on *täysin samaa mieltä-samaa mieltä* -asteikko ilmaistu muodossa, joka suomeksi käännettynä on suunnilleen *samaa mieltä-taipuvainen olemaan samaa mieltä*.

Jos mitattavat käsitteet eroavat eri kulttuureissa, pitää vertailussa nostaa käsitteellistä abstraktiotasoa senkin uhalla, että menetetään yksityiskohtaisempaa tietoa.

Kysymyskohtaiset vinoumat voivat johtua esimerkiksi huonoista käännöksistä tai alkuperäisen kysymyksen epäselvyydestä (van de Vijver 2003, 146). Kansainvälisten vertailuaineistojen kohdalla kannattaakin aina pohdiskella käännösteknisiä kysymyksiä ja ongelmia<sup>33</sup>.

## ESIMERKKI 5.

## KYSYMYSKOHTAINEN VINOUMA

Sama kysymys on saatettu erikielisissä lomakkeissa esittää hieman eri muodoissa. Esimerkiksi vuoden 2001 ISSP-aineistossa, joka kerättiin lähes 30 maassa, yhtenä kysymyksenä oli ”Kuinka usein keskustellette politiikasta *ystävienne* kanssa?”. Venäjällä kysymys kuitenkin esitettiin muodossa ”Kuinka usein keskustellette politiikasta *toisten ihmisten* kanssa”<sup>34</sup>.

Metodisista vinoumista on kysymys esimerkiksi silloin, kun otokset eivät ole vertailukelpoisia tai kyselylomakkeet tai tiedonkeruumenetelmät ovat olleet erilaiset eri maissa. Metodisen vinouman tunnistaminen vaatii huolellista tutustumista aineiston kuvailutietoihin.

Aineistosarjan samaankin keruuaaltoon kuuluvat aineistot on saatettu kerätä useina eri vuosina. Myös perusjoukko eri maissa saattaa erota esimerkiksi havaintoyksikköjen ala- ja yläikärajan suhteen. Maiden vertailua hankaloittavat myös tietojen eriasteinen saatavuus ja luotettavuus sekä nopeasti muuttuvien maiden kohdalla ajantasaisten tietojen puuttuminen (Stewart 1984, 43–44).

## ESIMERKKI 6.

## AINEISTONKERUUN AJANKOHTA

Tuorein World Values Survey on kerätty vuosina 2005 ja 2006. Tutkimukseen osallistuneet maat ovat itse saaneet valita keruujankohdan 1.4.2005–31.12.2006 väliseltä ajalta. Suomen aineisto kerättiin syksyllä 2005.

33 Käännösteknisistä kysymyksistä, ongelmista ja esimerkeistä voi lukea lisää teoksesta Behling & Law (2000). Translating Questionnaires and Other Research Instruments: Problems and Solutions.

34 ISSP-aineistot ja niiden mahdolliset ongelmat on kuvailtu tarkasti Saksan tietoaarkiston sivuilla: [http://www.gesis.org/en/data\\_service/issp/data/index.htm](http://www.gesis.org/en/data_service/issp/data/index.htm)

Vertailtaessa eri maiden dataa tai eri aikoina kerättyjä dataa tarvitaan usein tietojen harmonisointia eli yhteensovittamista. Tyypillisiä harmonisoitavia asioita ovat koulutus- ja ammattiluokitukset. Silloinkin, kun aineistonkeruu on ollut koordinoitua ja tietosisältöä on pyritty yhtenäistämään jo ennalta, kulttuuriset ja ajalliset tekijät voivat heikentää muuttujien vertailukelpoisuutta. Ongelmana on usein myös se, että aineistosarjoissakaan toistuvien kysymysten vastauksia ei ole koodattu aina samalla tavalla (van Deth 2003, 298–301).

#### ESIMERKKI 7.



#### LUOKITUSTEN HARMONISOINTI

Jan van Deth (2003) antaa esimerkin Euroopan komission rahoittamista eurobarometreistä, joita on kerätty 1970-luvun alusta alkaen yleensä kaksi kertaa vuodessa. Lähes jokainen eurobarometri sisältää tiedon vastaajan ammatista. Koska eurobarometrien aineistonkeruuta koordinoidaan keskitetysti ja ammatti on perustaustamuuttuja, voisi olettaa tietojen olevan helposti vertailtavissa. Ammattien koodaamisessa käytettiin kuitenkin vuosina 1970–1996 kaiken kaikkiaan kymmentä eri koodaustapaa ja tiedot saadaan vertailukelpoisiksi vain suuritöisellä jälkikäteen tehävällä harmonisoinnilla.<sup>35</sup>

Huolellisellakaan suunnittelulla ei pystytä kontrolloimaan kaikkia kulttuuri-eroavaisuuksista aiheutuvia ongelmia. Erialaisten selittävien taustamuuttujien käyttö saattaakin usein selventää maiden välisiä eroja vastausjakauksissa. Jos on aiheutta epäillä aineistojen tietojen vertailtavuutta, kannattaa keskittyä eri maiden vertailun sijasta tutkittavan ilmiön rakenteelliseen tarkasteluun eri maissa (Lehtinen 2004, 61). Joskus eri maiden välisiltä kulttuurisilta ja rakenteellisilta eroilta näytävät tulokset voivat selittyä pelkästään tilastointitapojen eroilla (Melin 2005, 55–58; Blom & Melin 2004, 157–163).

35 Tämänkaltaisten eurooppalaisten datajien taustamuuttujista ja niiden harmonisoinnista ks. myös Hoffmeyer-Zlotnik & Wolf (2003).

## 6 Aineistojen jatkokäyttö ja hyvät tieteelliset käytännöt

### 6.1 Jatkokäyttö osana aineistojen elinkaarta

Tutkimusaineistojen koko elinkaaren suunnittelu keruuvaiheesta projektin jälkeiseen aikaan on osa hyvää tieteellistä käytäntöä<sup>36</sup>. Taloudelliselta kannalta arkistointi takaa aineistojen kustannustehokkaan käytön. Tieteellisen tiedon luotettavuuden ja kehityksen kannalta aineistojen avoimuus on välttämätöntä.<sup>37</sup> Tietoarkiston kokemusten mukaan tutkimusorganisaatiot ja tutkijat suhtautuvat aineistojen jatkokäyttöön yleensä myönteisesti. Samoin tutkittavat henkilöt suhtautuvat yleensä hyvin myönteisesti heitä koskevien tutkimusaineistojen arkistointiin ja jatkokäyttöön. Tämä kertoo luottamuksesta tutkijoiden ammatilliseen pätevyYTEEN. (Kuula 2006A, 227–230.)

Yhteiskuntatieteellisen tietoarkiston toiminta perustuu puolestaan luottamukseen aineiston alkuperäisen kerääjän, tietoarkiston ja jatkokäyttäjän välillä. Jokainen tietoarkiston aineistojen jatkokäyttäjä allekirjoittaa käyttöehtositoumuksen (ks. liite 2), jossa määritellään aineistojen käyttöön ja säilyttämiseen liittyvät periaatteet. Periaatteet ovat luonteeltaan hyvän tieteellisen käytännön mukaisia ja ne soveltuvat lähtökohdiksi kaikkien tutkimusaineistojen käsittelyyn.

### 6.2 Tutkimusaineistoon viittaaminen

Tietoarkistosta hankittu tai muuten valmiina saatu aineisto on tutkimuksen lähdeaineisto siinä missä kirjallisuus tai tilastotkin, joten se tulee mainita lähdeluettelossa. Viitteestä tulee käydä ilmi ainakin aineiston nimi, tekijät ja jakelija.

Viittaamista helpottaakseen tietoarkisto toimittaa asiakkailleen aina aineiston lisäksi myös malliviittaustiedon, joka löytyy myös tietoarkiston kotisivuilta aineiston kuvailusta, koodikirjasta ja kyselylomakkeista. Viitteen voi merkitä lähdeluetteloon sellaisenaan tai sitä voi muokata kustantajien ja julkaisujen käytäntöjen mukaiseksi. Tässä esimerkkinä vuonna 2005 kerätyn Suomen ISSP-aineiston malliviittaus:

36 Tutkimuseettisen neuvottelukunnan ohjeet: Hyvä tieteellinen käytäntö ja sen loukkausten käsitteleminen, <http://pro.tsv.fi/tenk/htkfi.pdf>

37 Kynnys omien aineistojen arkistointiin on kuitenkin suhteellisen korkea. Syitä on monia. Aineiston toimittaminen arkistoon vie aina jonkin verran aikaa, vaikka tietoarkisto hoitaaakin datan prosessoinnin ja aineiston kuvailun. Lisäksi tutkimusaineiston erityispiirteet, omistussuhteet ja jatkokäyttäjien asiantuntemuksen riittävyys voivat arveluttaa aineiston alun perin kerännyttä tutkijaa. Myös kritiikin pelko, tutkittavan suostumus, salassapitovelvollisuus tai lainsäädäntö voivat joskus estää aineiston jatkokäytön. (Ks. mm. Heaton 2004, 27–28; Sieber 1991, 141–150; Hedrick 1985, 132–143; Jacobs & Humphrey 2004, Law 2005, 6–9; Kuula 2006A, 227–230.)

ISSP 2005 : työorientaatiot III : Suomen aineisto [elektroninen aineisto]. FSD2133, versio 1.0 (2006-03-07). Blom, Raimo (Tampereen yliopisto. Sosiologian ja sosiaalipsykologian laitos) & Melin, Harri (Turun yliopisto. Sosiologian laitos) & Tanskanen, Eero (Tilastokeskus. Haastattelu- ja tutkimuspalvelut) [tekijät]. Tampere : Yhteiskuntatieteellinen tietoarasto [jakaja], 2006.

### 6.3 Julkaisuista tiedottaminen

Tietoarastosta saatuja aineistoja saa käyttää vain käyttöluvahakemuksessa ilmoitettuun tarkoitukseen. Jos käyttötarkoitus vaihtuu tai laajenee, esimerkiksi jos haluaa käyttää tutkimustarkoitukseen saatua aineistoa opetuksessa, tietoarastolle on ilmoitettava uudesta käyttötarkoituksesta.

Tietoarastolla on velvollisuus seurata aineistojen käyttöä. Tietoarastolle onkin ilmoitettava julkaisusta, jossa on käytetty arastosta saatuja aineistoja. Ilmoitettavia julkaisuja ovat esimerkiksi kirjat, väitöskirjat, kirjan osat, raportit, artikkelit ja pro gradut. Sen sijaan kandidaatintöistä ja perusopintoihin kuuluvista seminaari- ja harjoitustöistä ei tarvitse ilmoittaa. Varsinaista julkaisua ei myöskään tarvitse lähettää tietoarastoon, bibliografinen viite riittää.

Julkaisujen viitetiedot lisätään aineistojen kuvailujen yhteyteen. Näin kootaan ja välitetään tietoa siitä, mitä on jo tutkittu. Aineistojen arkistoinnin ja jatkokäytön kannalta olisi tärkeää myös, että jatkokäyttäjät ilmoittaisivat tietoarastolle aineistossa mahdollisesti havaitsemistaan puutteista.

### 6.4 Tietosuoja tutkimusaineiston säilyttämisessä ja käytössä

Tutkimusaineiston säilyttämisessä ja käytössä on huolehdittava riittävästä tietoturvallisuudesta. Datatiedostot ja cd-rom-levyt on säilytettävä niin, että niihin pääsevät käsiksi vain käyttöluvan omaavat henkilöt. Tutkimuksen valmistuttua tai kurssin päätyttyä tietoarastosta saatu aineisto on hävitettävä. Tiedostot on tuhoettava tietokoneen levyiltä ja sähköpostin liitteistä. Cd-rom-levyt voi tuhota fyysisesti tai palauttaa tietoarastoon.

Jatkokäyttäjän on otettava huomioon myös aineistojen käyttöön ja keruuseen liittyvät eettiset kysymykset. Ihmisten yksityisyyden kunnioittaminen on yksi keskeisimmistä tutkimuseettisistä normeista. Tutkimusaineistosta ei saa pyrkiä tunnistamaan vastaajia ja tutkittavien henkilöiden anonymiteetistä on huolehdittava muutoinkin. Tietosuojasta on huolehdittava myös aineistosta tuotetuissa raporteissa. Tulokset on julkaistava esimerkiksi tilastotaulukkoina siten, että yksittäisen vastaajan tunnistaminen on mahdotonta. (Kuula 2006B, 124–137.)

Aineistojen asianmukaista säilyttämistä ja hävittämistä ohjeistetaan tietoaraston Tutkimusetiikka ja lait -sivustolla: [http://www.fsd.uta.fi/laki\\_ja\\_etiikka/etiikka\\_lait.html](http://www.fsd.uta.fi/laki_ja_etiikka/etiikka_lait.html)

## 7 Lopuksi

Tutkimusaineistot eivät ole kertakäyttötavaraa, sillä uusilla menetelmillä ja uusien teorioiden valossa vanhoistakin tutkimusaineistoista löytyy uusia puolia. Lisäksi tutkimusaineistot sisältävät yleensä runsaasti täysin käyttämättömiä muuttujia. Jatkokäyttö on myös taloudellisesti järkevää, sillä uusien aineistojen suunnittelu, kerääminen ja tallentaminen vaativat paljon resursseja. Opiskelijoille arkistoidut aineistot antavat hyvän mahdollisuuden valtakunnallisten aineistojen ja monipuolisten menetelmien käyttöön.

Olemassa olevan aineiston käyttöön liittyvät edut ovat käytännössä aina suurempia kuin haitat. Mahdollisiin ongelmiin löytyy yleensä joko ratkaisu tai ainakin hyväksyttävä kompromissi. Ongelmien tunnistamiseksi jatkokäyttäjien on syytä perehtyä huolella sekä aineistoon että aineiston dokumentointiin.

Usein suurin este jatkokäytölle on sopivan olemassa olevan tutkimusaineiston löytäminen ja saaminen käyttöön. Tietoa olemassa olevista aineistoista ei ole tai tutkijakollegan datan tiedostomuoto tai tallennusväline on vanhentunut niin, että tiedostoa ei voi enää lukea. Pahimmassa tapauksessa arvokas tutkimusaineisto on käyttökelvoton tai kadonnut kokonaan. Tietoarkistoon tallennettujen aineistojen kohdalla näitä ongelmia ei ole: arkistoitujen aineistojen kuvailut löytyvät arkiston verkkosivuilta ja tietoarkisto huolehtii aineistojen teknisestä käyttökunnosta. Lisäksi tietoarkistoon tallennettujen aineistojen kohdalla useat eettiset kysymykset on jo ratkaistu, aineistot ovat esimerkiksi yleensä valmiiksi anonymisoituja.

YHTEISKUNTATIETEELLINEN TIETOARKISTO PALVELEE  
OPISKELIJOITA JA TUTKIJOITA.

TIETOARKISTOSTA LÖYTYY RUNSASTI SEKÄ KOTIMAISIA  
ETTÄ KANSAINVÄLISIÄ KYSELYAINEISTOJA, MAKSUTTA.

SOPIVA TIETO ODOTTAÄ KÄYTTÄJÄÄNSÄ!

## Sanasto

Aineisto	elektroninen tutkimusaineisto (data) ja siihen liittyvä dokumentaatio
Alkuperäinen tutkija	henkilö, yhteisö tai taho, joka on vastuussa aineiston ajatuksellisesta sisällöstä, keruusta ja ensikäytöstä
Data	elektroninen tutkimusaineisto, tietokannan kvantitatiivisten aineistojen kohdalla tavallisesti havaintomatriisi
Dokumentaatio	aineiston sisältöä ja keruuta kuvailevat sekä jatkokäytön mahdollistavat tiedot (mm. tekijästä, sisällöstä, keruusta), metatieto, metadata.
Ex post -toimenpiteet	jälkikäiteistoimenpiteet, aineistoon tehdyt muutokset ja lisäykset aineiston keruun ja tallentamisen jälkeen.
Formaatti	tiedostomuoto
Embargo	käyttökielto, yleensä osittainen ja väliaikainen
Havaintomatriisi	kvantitatiivinen tutkimusaineisto taulukkomuodossa, datatiedosto. Tavallisesti tilasto-ohjelmissa yksi sarake edustaa yhtä muuttujaa ja yksi rivi yhtä havaintoyksikköä.
Havaintoyksikkö	tutkimuskohde, josta tehdään empiirisiä havaintoja (esim. henkilö, kotitalous, kunta)
Kvantitatiivinen tutkimusaineisto	numeerinen havaintomatriisi, datatiedosto. Tyypillinen tietokantaan arkistoitu kvantitatiivinen tutkimusaineisto on kerätty kirje-, puhelin- tai käyntikyselyllä
Koodikirja	dokumentti, joka sisältää tutkimusaineiston yleiskuvailun, muuttujien kuvailun ja suorat jakaumat sekä aineistonkeruussa mahdollisesti käytetyn kyselylomakkeen. Tietokannan tuottamat koodikirjat ovat PDF-tiedostoja.
Koodausohjeet	ohjeet, joiden mukaan tutkimusaineisto luokitellaan ja tallennetaan havaintomatriisiksi
Kuvailu	ks. dokumentaatio
Jatkokäyttö	olemassa olevan, muiden keräämän tutkimusaineiston hyödyntäminen
Jatkokäyttäjät	olemassa olevan, muiden keräämän tutkimusaineiston hyödyntäjät
Metadata	ks. dokumentaatio

Mittari	apuväline, jolla mitataan tutkittavaa ominaisuutta (esim. kyselylomake, yksittäinen kysymys)
Muuttuja	mitattava/tutkittava ominaisuus, joka on usein johdettu kyselylomakkeen kysymyksestä (Kysymys: Mikä on sukupuolenne? Muuttuja: sukupuoli)
Operationalisointi	tutkittavan abstraktin käsitteen määrittäminen konkreettiseen mitattavaan muotoon
Otanta	otoksen valinta perusjoukosta tutkimusta varten
Otos	perusjoukkoa pienempi satunnaisesti valittu havaintoyksikköjoukko, johon mittaus kohdistetaan
Perusjoukko	tutkimuksen kohderyhmä, jota koskevia tietoja halutaan kerätä
Populaatio	ks. perusjoukko
Satunnaisotanta	otannan perusmenetelmä, jossa jokaisella perusjoukkoon kuuluvalla on yhtä suuri todennäköisyys tulla poimituksi otokseen

## Lähteet

- Alameri, Rolf & Pöyhönen, Pentti (1954) *Johdatusta tilastolliseen tutkimukseen*. 2. painos. Helsinki: Helsingin yliopisto.
- Ahola, Anja (1998) Haastattelutilanne surveyvastausten rakentajana, teoksessa Seppo Paananen, Anneli Juntto & Hannele Sauli (toim.) *Faktajuttu. Tilastollisen sosiaalitutkimuksen käytännöt*. Tampere: Vastapaino, s. 257–281.
- Ahola, Anja (2003) *Massahaastattelumenetelmien erot väestön hyvinvoinnin kuvaamisessa*. Hyvinvointikatsaus 2003:2, s. 46–52. Saatavissa: [http://www.stat.fi/tup/hyvinvointikatsaus/surlab2\\_2003.pdf](http://www.stat.fi/tup/hyvinvointikatsaus/surlab2_2003.pdf) [viitattu 16.11.2006].
- Ahola, Anja (2004) *Postikysely ja puhelinhaastattelu työolobarometriin vastaamisessa*. Hyvinvointikatsaus 2004:4, s. 47–52. Saatavissa: [http://www.stat.fi/tup/hyvinvointikatsaus/surlab4\\_2004.pdf](http://www.stat.fi/tup/hyvinvointikatsaus/surlab4_2004.pdf) [viitattu 16.11.2006].
- Alatalo, Marja (2005) *Metodisuhdanteiden mahti. Lomaketutkimus suomalaisessa sosiologiassa 1947–2000*. Tampere: Vastapaino.
- Alkula, Tapani, Pöntinen, Seppo & Ylöstalo, Pekka (1995) *Sosiaalitutkimuksen kvantitatiiviset menetelmät*. 1.–2. painos. Helsinki: WSOY.
- Ammattiluokituksen soveltamisohje* (2005) Työnvälitys ja työvoiman rekrytointi -tiimi. Ohje 11.2.2005 (moniste). Helsinki: Työministeriö.
- Blom, Raimo & Melin, Harri (2004) Vertaileva luokkatutkimus idässä ja lännessä, teoksessa Risto Alapuro & Ilkka Arminen (toim.) *Vertailevan tutkimuksen ulottuvuuksia*. Porvoo: WSOY, s. 153–164.
- Borg, Sami (2005) *Kansalaisena Suomessa. Kansalaisvaikuttaminen Pohjoismaissa ja European Social Survey 2002*. Oikeusministeriön julkaisuja 2005:3. Helsinki: Oikeusministeriö.
- Corti, Louise & Thompson, Paul (2004) Secondary analysis of archived data, in Clive Seale, Giampietro Gobo, Jaber Gubrium & David Silverman (eds.) *Qualitative Research Practice*. London: Sage p. 327–343.
- Dale, Angela, Arber, Sara & Procter, Michael (1988) *Doing Secondary Analysis*. Contemporary Social Research 17. London: Unwin Hyman Ltd.
- PERUSTEELLINEN KATSAUS AINEISTOJEN JATKOKÄYTÖN ETUIHIN JA KUSTANNUKSIIN SEKÄ KÄYTÄNNÖN VINKKEJÄ AINEISTOJEN ARVIOINTIIN JA TEHOKKAASEEN KÄYTTÖÖN.
- Dillman, Don A. (2002) *Mail and Internet Surveys. The Tailored Design Method*. 2nd Edition. New York: Wiley.
- Elder, Glen H. Jr., Pavalko, Eliza K. & Clipp, Elizabeth C. (1993) *Working with Archival Data. Studying Lives*. Quantitative applications in the social sciences 88. Newbury Park: Sage.
- Eskola, Antti (1975) *Sosiologian tutkimusmenetelmät* 2. 2. p., 2. muuttam. lisäp. Helsinki: WSOY.
- Fienberg, Stephen E., Martin, Margaret E. & Straf, Miron L. (eds.) (1985). *Sharing Research Data*. Washington, D.C.: National Academy Press.
- ERILAISIA NÄKÖKULMIA AINEISTOJEN JATKOKÄYTTÖÖN JA JAKAMISEEN TUTKIJOIDEN KESKEN.
- Frankfort-Nachmias, Chava & Nachmias, David (1996) *Research Methods in the Social Sciences*. 5. ed. London: Arnold.
- Groves, R., Fowler, F., Couper, M. J., Lepkowski, J., Singer, E. & Tourangeau, R. (2004) *Survey methodology*. Wiley series in survey methodology. Hoboken, N.J.: Wiley.
- Grönlund, Kimmo, Paloheimo, Heikki & Wass, Hanna (2005) Äänestysosallistuminen, teoksessa Heikki Paloheimo (toim.), *Vaalit ja demokratia Suomessa*. Helsinki: WSOY, s. 119–146.



Hakim, Catherine (1982) *Secondary Analysis in Social Research. A Guide to Data Sources and Methods with Examples*. Contemporary Social Research Series 5. London: Allen & Unwin.

• PERUSTEOS, JOSSA TARKASTEELLAAN MUUN MUASSA JATKOKÄYTTÖÖN LIITTYVIÄ TEOREETTISIA ONGELMIA. HAKIM KÄSITTELEE AIHETTA LYHYESTI MYÖS KIRJASSAAN RESEARCH DESIGN. SUCCESSFUL DESIGNS FOR SOCIAL AND ECONOMIC RESEARCH (2000).

Hakim, Catherine (2000) *Research Design. Successful Designs for Social and Economic Research*. 2nd Edition. London: Routledge.

Harkness, Janet A., van de Vijver, Fons J. R. & Mohler, Peter Ph (2003) *Cross-Cultural Survey Methods*. Wiley series in survey methodology. Hoboken, N.J.: Wiley-Interscience.

• ARTIKKELIKOKELOMA, JOSSA KÄYDÄÄN LÄPI KANSAINVÄLISEN VERTAILEVAN SOSIAALITUTKIMUKSEN ONGELMIA JA NIIDEN RATKAISUJA.

Heaton, Janet (2004) *Reworking Qualitative Data*. London: Sage.

• KVALITATIIVISTEN TUTKIMUSAINEISTOJEN JATKOKÄYTTÖN PERUSOPAS

Hedrick, Terry E. (1985) Justifications for and Obstacles to Data Sharing, in: Stephen E. Fienberg, Margaret E. Martin & Miron L. Straf (eds.) *Sharing Research Data*. Washington, D.C.: National Academy Press, p. 123–147.

Hirsjärvi, Sirkka (2004) Aineiston hankinnan suunnittelu, teoksessa Sirkka Hirsjärvi, Pirkko Remes & Paula Sajavaara *Tutki ja kirjoita*. 10., osin uudistettu laitos. Helsinki: Tammi, s. 166–179.

Hoffmeyer-Zlotnik, Jürgen H.P. & Wolf, Christof (eds.) (2003) *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables*. New York: Kluwer Academic/Plenum Publishers.

Hyman, Herbert H. (1972) *Secondary Analysis of Sample Surveys. Principles, Procedures, and Potentialities*. New York: Wiley.

• ALAN KLASSIKKO, JOKA SOPII EDELLEEN JOHDANNOKSI AIHEESEEN. KIRJASTA ON ILMESTYNYT MYÖS

UUSINTAPAINOS VUONNA 1987 NIMELLÄ SECONDARY ANALYSIS OF SAMPLE SURVEYS (WITH A NEW INTRODUCTION). MIDDLETOWN, CT : WESLEYAN UNIVERSITY PRESS.

Hämäläinen, Maria (2005) *Perusterveydenhuollon kustannus-vaikuttavuus. Espoon, Helsingin, Kemin, Kuusankosken ja Tampereen perusterveydenhuollon kustannusten sekä kuntalaisten terveyskeskuspalveluihin tyytyväisyyden kohtaaminen vuosina 1993, 1997 ja 2001*. Tampere: Tampereen yliopisto. Sosiaalipoliittikan pro gradu -tutkielma.

Jacob, Herbert (1994) Using Published Data. Errors and Remedies, in Michael S. Lewis-Beck (ed.) *Research Practice*. International handbooks of quantitative applications in the social sciences 6. London: Sage, p. 339–389.

• AINEISTOJEN JATKOKÄYTTÖÖN LIITTYVIEN ERI LAISTEN ONGELMIEN KUTEN OTANTA- JA MITTAUSVIRHEIDEN, VALIDITEETIN JA RELIABILITEETIN SEKÄ NIIDEN RATKAISUMALLIEN YKSITYSKOHTAISTA TARKASTELUA.

Jacobs, James A. & Humphrey, Charles (2004) *Preserving Research Data*. Communications of the ACM 47:9, 27–29.

Kangas, Olli (1995) *Metsä vastaa miten huudetaan. Kysymysten muotoilun vaikutus mielipiteisiin*. Poliitikka 37:2, 128–137.

Kiecolt, K. Jill & Nathan, Laura E. (1985) *Secondary Analysis of Survey Data*. Quantitative Applications in the Social Sciences 53. Newbury Park: Sage.

• TIIVIS OPAS TUTKIMUSAINEISTOJEN HANKINTAAN JA AINEISTOJEN TEHOKKAASEEN KÄYTTÖÖN.

Kouvo, Antti & Räsänen, Pekka (2005) *Sosiaalinen pääoma, elämäntilanne ja sosiodemografiset tekijät. Käyttökelpoisia elämänlaadun ja hyvinvoinnin jäsenmysterusteita?* Janus 13:1, 21–38.

Kurkela, Reijo & Sauli, Hannele (1998) Tilastolliset luokitukset ja arki, teoksessa Sepo Paananen, Anneli Juntto & Hannele Sauli (toim.) *Faktajuttu. Tilastollisen sosiaalitutkimuksen käytännöt*. Tampere: Vastapaino, s. 27–42.

Kuula, Arja (2006A) *Tutkimusetiikka. Aineiston hankinta, käyttö ja säilytys*. Tampere: Vas-tapaino.

• TUTKIMUSAINEISTOJEN ELINKAARTA TARKASTE-LEVA OPPIKIRJA. AIHEALUEINA MM. TUTKITTAVIEN INFORMOINTI, AINEISTON HANKINNAN JURIDIikka SEKÄ TUTKIMUSAINEISTON ANONYMISOINTI JA ARKISTOINTI. ERILLISESSÄ LUVUSSA TARKASTELLAAN INTERNET-TUTKIMUSTA.

Kuula, Arja (2006B) Yksityisyyden suoja tutkimuksessa, teoksessa Jaana Hallamaa, Veikko Launis, Salla Lötjönen & Irma Sorvali (toim.) *Etiikkaa ihmistieteille*. Suomalaisen Kirjallisuuden Seuran julkaisuja Tietolipas 211 & Tutkimuseettisen neuvottelukunnan julkaisuja. Helsinki: Suomalaisen Kirjallisuuden Seura, s. 124–140.

Laaksonen, Helena (2005) *Kansainvälisten järjestöjen tilastoaineistot tehokäyttöön Suomessa? Alustavia pohdintoja ja selvitystä aineistopalvelun mahdollisuuksista*. Tietoarkistolehti 16, 2005:1. Saatavissa: <http://www.fsd.uta.fi/tietoarkistolehti/16/kvtilastot.html> [viitattu 16.11.2006].

Law, Margaret (2005) *Reduce, Reuse, Recycle. Issues in the Secondary Use of Research Data*. Alberta: University of Alberta.

Lehtinen, Marjaana (2004) *Surveyn kulttuuriset erityispiirteet kansainvälisen vertailun haasteena*. Hyvinvointikatsaus 2004:3, 59–61. Saatavissa: [http://www.stat.fi/tup/hyvinvointikatsaus/surlab3\\_2004.pdf](http://www.stat.fi/tup/hyvinvointikatsaus/surlab3_2004.pdf) [viitattu 16.11.2006].

Litmala, Marjukka (2004) Luottamus oikeuslaitokseen. Trendejä Suomessa ja Euroopassa, teoksessa Marjukka Litmala (toim.) *Oikeusolot 2004. Katsaus oikeudellisten instituutioiden toimintaan ja oikeuden saatavuuteen*. Oikeuspoliittisen tutkimuslaitoksen julkaisuja 210. Helsinki: Oikeuspoliittinen tutkimuslaitos, s. 9–27. Saatavissa: <http://www.om.fi/optula/28637.htm> [viitattu 16.11.2006].

Mattila, Mikko & Sankiaho, Risto (2005) Luottamus poliittiseen järjestelmään, teoksessa Heikki Paloheimo (toim.) *Vaalit ja demokratia Suomessa*. Helsinki: WSOY, s. 73–87.

Melin, Harri (2005) Vertailevan tutkimuksen monet lähtökohdat, teoksessa: Pekka Räsänen, Anu-Hanna Anttila & Harri Melin (toim.) *Tutkimus menetelmien pyörteissä. Sosiaalitutkimuksen lähtökohdat ja valinnat*. Jyväskylä: PS-kustannus, s. 53–65.

Mochmann, Ekkehard & Guchteneire, Paul de (1998) *The Uses of Secondary Analysis in the Social Sciences*. Saatavissa: <http://www.ifdo.org/methods/index.html> [viitattu 16.11.2006]

Myllyniemi, Harri (2004) *Gradun tekoa arkistoiduilla aineistoilla*. Tietoarkistolehti 14, 2004:2. Saatavissa: <http://www.fsd.uta.fi/tietoarkistolehti/14/gradu.html> [viitattu 2.11.2006].

Myllyniemi, Harri (2006) *Rationaalisen tuolla puolen eli tutkimus ulkona kävelyn koetusta turvattomuudesta suomalaisissa kunnissa ja kaupungeissa*. Tampere: Tampereen yliopisto. Sosiologian pro gradu -tutkielma.

Niiniluoto, Ilkka (2001) *Julkisuusperiaate ja tutkimustulosten salailu*. Tieteessä tapahtuu 2001:2. Saatavissa: <http://www.tsv.fi/ttapaht/012/niiniluoto.htm> [viitattu 16.11.2006]

Piispa Minna (2006) Kvantitatiivisen tutkimuksen eettiset lähtökohdat. Esimerkkinä naisiin kohdistuvan väkivallan kyselytutkimus, teoksessa Jaana Hallamaa, Veikko Launis, Salla Lötjönen & Irma Sorvali (toim.) *Etiikkaa ihmistieteille*. Suomalaisen Kirjallisuuden Seuran julkaisuja Tietolipas 211 & Tutkimuseettisen neuvottelukunnan julkaisuja. Helsinki: Suomalaisen Kirjallisuuden Seura, s. 141–160.

Procter, Michael (1993) Analysing Other Researchers' Data, in Nigel Gilbert (ed.) *Researching Social Life*. London: Sage, p. 255–269.

Räsänen, Pekka (2006) *Information Society for All? Structural Characteristics of Internet Use in 15 European Countries*. European Societies 8:1, 59–81. Available: <http://www.journalsonline.tandf.co.uk/openurl.asp?genr>

e=article&id=doi:10.1080/14616690500491423 [cited 16.11.2006]

Ronkainen, Suvi (1999) *Ajan ja paikan merkitykset. Subjektiviteetti, tieto ja toimijuus*. Helsinki : Gaudeamus.

Schwarz, Norbert & Hippler, Hans-J. (1995) *Subsequent Questions may influence Answers to Preceding Questions in Mail Surveys*. Public Opinion Quarterly 59:1, 93–97.

Seale, Clive (2004) Using data archives for secondary analysis, in C. Seale (ed.) *Researching Society and Culture*. 2nd Edition. London: Sage, p. 355–366.

Sieber, Joan E. (ed.) (1991) *Sharing Social Science Data. Advantages and Challenges*. Newbury Park: Sage.

• KIRJAN ARTIKKELIT KÄSITTELEVÄT MM. AINEISTOJEN AVOIMUUTTA YLEENSÄ JA AINEISTOJEN KÄYTTÖÄ OPETUKSESSA.

Szabo, Vivian & Strang, Vicki (1997) *Secondary Analysis of Qualitative Data*. Advances in Nursing Science [web journal] 20:2, 66–74 [cited 16.11.2006]. Available: <http://gateway.uk.ovid.com/gw1/ovidweb.cgi>

Stewart, David W. (1984) *Secondary Research. Information, Sources and Methods*. Applied social research methods series 4. Beverly Hills: Sage.

• KÄYTÄNNÖN VINKKEJÄ JATKOTUTKIMUKSEEN.

Sturgis, Patrick (2004) *Analysing Complex Survey Data. Clustering, Stratification and Weights*. Social Research Update 43, Autumn 2004. Available: <http://www.soc.surrey.ac.uk/sru/SRU43.html> [cited 16.11.2006]

Toivonen, Timo (1999) *Empiirinen sosiaalitutkimus. Filosofia ja metodologia*. Porvoo: WSOY.

Toivonen, Timo (2005) *Elaboraatio ja interaktio eli mitä log-lineaarilla malleilla on hankala havaita, teoksessa Pekka Räsänen, Anu-Hanna Anttila & Harri Melin (toim.)*

*Tutkimus menetelmien pyörteissä. Sosiaalitutkimuksen lähtökohdat ja valinnat*. Jyväskylä: PS-kustannus, s. 123–138.

Töttö, Pertti (2004) *Syvällistä ja pinnallista. Teoria, empiria ja kausaalisuus sosiaalitutkimuksessa*. Tampere: Vastapaino.

Uitus-Mäntylä, Katja (2003) *Raha ratkaisee - vai ratkaiseeko? Palveluiden taloudellisen resurssoinnin vaikutus kuntalaisten mielipiteisiin kunnan sosiaali- ja terveyspalveluista*. Tampere: Tampereen yliopisto. Sosiaalityön pro gradu -tutkielma.

Uusitalo, Hannu (1999) *Tiede, tutkimus ja tutkielma. Johdatus tutkielman maailmaan*. 1.–6. painos. Helsinki: WSOY.

Van Deth, Jan W. (2003) Using Published Survey Data, in Janet A. Harkness, Fons J. R. van de Vijver & Peter Ph. Mohler (eds.) *Cross-Cultural Survey Methods*. Wiley series in survey methodology. Hoboken, N.J: Wiley-Interscience, p. 291–309.

Van de Vijver, Fons J. R. (2003) Bias and Equivalence. Cross-Cultural Perspectives, in Janet A. Harkness, Fons J. R. van de Vijver & Peter Ph. Mohler (eds.) *Cross-Cultural Survey Methods*. Wiley series in survey methodology. Hoboken, N.J: Wiley-Interscience, p. 143–155.

Vartia, Yrjö O (1988) *Ovatko tilastolliset tutkimustulokset vain monimutkaisesti perustetuja mielipiteitä?* Kansantaloudellinen aikakauskirja 84:3, 305–309.

Wolf, Christof & Hoffmeyer-Zlotnik, Jürgen H. P. (2003) Measuring Demographic and Socio-Economic Variables in Cross-National Research. An Overview, in Jürgen H. P. Hoffmeyer-Zlotnik & Christof Wolf (eds.) *Advances in Cross-National Comparison. A European Working Book for Demographic and Socio-Economic Variables*. New York: Kluwer Academic/Plenum Publishers, s. 1–13.

## Hyödyllisiä verkkosivuja

Yhteiskuntatieteellinen tietoaarkisto:

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

CESSDAn kotisivut:

<http://www.cesda.org/>

ICPSR:

<http://www.icpsr.umich.edu/>

Menetelmäopetuksen tietovaranto MOTV:

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

Rekisteritutkimuksen tukikeskus ReTKi:

<http://www.rekisteritutkimus.fi/>

Tilastokeskuksen verkkokoulu:

<http://www.stat.fi/>

Tilastokeskuksen SurveyLaboratorio:

<http://www.stat.fi/tup/surveylab/>

Tietosuojavaltuutetun toimisto:

<http://www.tietosuoja.fi/>

Valtion säädöstietopankki:

<http://www.finlex.fi/fi/>

## Liite 1. Käyttölupahakemus

Tyhjennä lomake



## Käyttölupahakemus

## ARKISTO TÄYTTÄÄ

Käyttölupahakemuksen numero

Saapumispvm

Käsittelijä

## HAKIJA TÄYTTÄÄ

## Hakijan yhteystiedot

Nimi	Yliopisto, tutkimuslaitos tai vastaava	Laitos, osasto tai vastaava
Postiosoite		
Puhelin	Faksi	Sähköposti

Muiden aineistoa käyttävien henkilöiden yhteystiedot<sup>1)</sup>

Nimi	Yliopisto, tutkimuslaitos tai vastaava	Laitos, osasto tai vastaava
Postiosoite		
Puhelin	Faksi	Sähköposti

Nimi	Yliopisto, tutkimuslaitos tai vastaava	Laitos, osasto tai vastaava
Postiosoite		
Puhelin	Faksi	Sähköposti

Aineiston käyttötarkoitus ja hankkeen rahoitus<sup>2)</sup>

Aineistot, joihin käyttö lupaa haetaan	
Tutkimusaineiston numero	Tutkimusaineiston nimi
Hankkeen nimi	
Hankkeen kesto kuukauden tai vuoden tarkkuudella (esim. 12/1998 – 2000)	

- 1) Tarvittaessa eri liitteellä
- 2) Kutakin tutkimushanketta varten pitää täyttää oma käyttölupahakemus. Samaa tarkoitusta varten voi kuitenkin hakea yhdellä hakemuksella lupaa useiden aineistojen käyttöön.

TAMPEREEN  
YLIOPISTO

## Yhteiskuntatieteellinen tietoaarkisto

Akerlundinkatu 2 A 5. krs, 33014 Tampereen yliopisto  
 puh. (03) 215 8519, faksi (03) 215 8520  
 Internet: <http://www.fsd.uta.fi>  
 E-mail: [fsd@uta.fi](mailto:fsd@uta.fi)

Hankkeen lyhyt kuvaus, josta ilmenee tilattavan aineiston käyttötarkoitus	
Jos aineistoa käytetään opetuksessa, ilmoitetaan kurssin nimi ja arvio opetukseen osallistuvien lukumäärästä	
Hankkeen kokonaiskustannukset, rahoittajat ja näiden rahoitusosuudet (Suomen Akatemian rahoittamista projekteista myös hankenumero)	
Kokonaiskustannukset	
Rahoittaja	Rahoitusosuus
Rahoittaja	Rahoitusosuus
Rahoittaja	Rahoitusosuus
Rahoittaja	Rahoitusosuus
Rahoittaja	Rahoitusosuus
Rahoittaja	Rahoitusosuus
Rahoittaja	Hankenumero
<input type="checkbox"/> Suomen Akatemia	

**Kenelle ja millä tavoin aineisto toivotaan luovutettavaksi?**

<input type="checkbox"/> Tämän hakemuksen allekirjoittajalle	<input type="checkbox"/> Muulle henkilölle, kenelle <sup>3)</sup>	
<b>Vastaanottajan yhteystiedot, jos eri kuin hakija</b>		
Nimi	Yliopisto, tutkimuslaitos tai vastaava	Laitos, osasto tai vastaava
Postiosoite		
Puhelin	Faksi	Sähköposti
Käyttäjän toivoma data-aineiston tiedostoformaatti (esim. SPSS, SAS, Excel...)		
Tiedostoformaatti	Jos jätetään yksilöimättä, FSD toimittaa tiedoston SPSS Portable -muodossa	
Käyttäjän toivoma lähetyksen- tai siirtotapa (rastita)		
<input type="checkbox"/> Zip-levyke (100Mb)	<input type="checkbox"/> CD-ROM	<input type="checkbox"/> SFTP-siirto tai -haku, sovitettava erikseen
<input type="checkbox"/> 3,25" levyke	<input type="checkbox"/> Sähköpostin liitetiedosto	<input type="checkbox"/> HTTP-siirto, sovitettava erikseen
<input type="checkbox"/> Muu, mikä?	Lisätietoa tiedostoformaateista <a href="http://www.fsd.uta.fi/tiedostoformaatit.html">http://www.fsd.uta.fi/tiedostoformaatit.html</a>	

Hakijan yhteystiedot saa tallettaa FSD:n asiakasrekisteriin. Aineiston tilaajista voidaan informoida aineiston alkuperäisiä tekijöitä.

Pvm	Hakijan allekirjoitus	(Jos perustutkinto-opiskelija) Ohjaajan allekirjoitus
	Nimenselvennys	Nimenselvennys

<sup>3)</sup> Jos vastaanottaja on muu kuin hakemuksen allekirjoittaja, hakijan on ilmoitettava hänen yhteystietonsa.

**Hakijan, vastaanottajan ja kaikkien muiden aineiston käyttäjien tulee lähettää tietoarkistolle erillinen henkilökohtainen käyttöehtositoumus allekirjoitettuna.**

## Liite 2. Käyttöehtositoumus



## Käyttöehtositoumus

## ARKISTO TÄYTTÄÄ

Käyttölupahakemuksen numero	
Saapumispvm	Käsittelijä

## VASTAANOTTAJA TÄYTTÄÄ

Tämä sopimus koskee Yhteiskuntatieteellisen tietoarkiston allekirjoittaneelle luovuttamien tutkimusaineistojen ja niiden oheismateriaalin, eli yhdessä *aineiston*, käyttöä. Vastaanottajana sitoudun noudattamaan aineiston käytössä seuraavia ehtoja:

**1. Käyttötarkoitus ja aineiston hävittäminen**

Käytän aineistoa ainoastaan käyttölupahakemuksessa ilmoittamani käyttötarkoituksen mukaisesti, jonka jälkeen hävitän aineiston. Jos haluan käyttää aineistoa muihin tutkimus- tai opetustarkoituksiin, toimitan Yhteiskuntatieteelliselle tietoarkistolle uuden käyttölupahakemuksen.

**2. Tietoturvasuus**

Huolehdin aineiston säilyttämisessä ja käytössä riittävästä tietoturvasuudesta. Vastaan siitä, että aineisto ei ole muiden kuin käyttölupahakemuksessa yksilöityjen henkilöiden tai tahojen saatavilla.

**3. Tietosuoja**

Huolehdin aineiston säilyttämisessä ja käytössä sekä aineistosta julkistettavissa tuloksissa niiden henkilöiden ja tahojen tietosuojasta, joita koskevia tietoja aineisto sisältää. En pyri tunnistamaan tietosuojaa rikkovalla tavalla aineiston tietoihin liittyviä henkilöitä tai tahoja. Huolehdin myös siitä, että tietosuojanäkökohdat huomioidaan riittävästi kaikissa minulle luovutetusta aineistosta julkistetuissa tuloksissa. Hakijan yhteystiedot saa tallettaa FSD:n asiakasrekisteriin. Aineiston tilaajista voidaan informoida aineiston alkuperäisiä tekijöitä.

**4. Aineiston alkuperän mainitseminen julkaisuissa**

Ilmoitan tietoarkiston aineistolle antaman lähdeviittaustekstin kaikissa julkaisuissa, joissa aineistoa käytetään vähintään osa-aineistona. Hyväksyn sen, että aineiston arkistoon luovuttajilla ja arkistolla ei ole vastuuta aineiston uudiskäytössä tuotetuista tuloksista ja tulkinnoista.

**5. Aineistoihin liittyvät virheet ja puutteet**

Hyväksyn sen, että aineiston arkistoon luovuttajat ja arkisto eivät ole vastuussa uudiskäyttöä mahdollisesti haittaavista aineiston puutteista. Informoin tietoarkistoa aineistossa havaitsemistani puutteista.

**6. Julkaisuista tiedottaminen**

Lähetän arkistolle tiedot julkaisuista, joissa minulle luovutettua aineiston kopiota tai siitä luvallisesti otettuja kopioita on hyödynnetty.

**7. Tekijänoikeus**

En kopioi aineistoa muutoin kuin kohtien 1 ja 2 sallimassa tarkoituksessa.

**Saako hakijan yhteystiedot luovuttaa tutkijoille, jotka käyttävät samoja aineistoja/samaa aineistoa?**

Kyllä  Ei

Pvm	Allekirjoitus
	Nimenseivennys



TAMPEREEN  
YLIOPISTO

**Yhteiskuntatieteellinen tietoarkisto**

Åkerlundinkatu 2 A 5. krs, 33014 Tampereen yliopisto  
puh. (03) 215 8519, faksi (03) 215 8520  
Internet: <http://www.fsd.uta.fi>  
E-mail: [fsd@uta.fi](mailto:fsd@uta.fi)



YHTEISKUNTATIEETEELLINEN  
TIETOARKISTO

Käyntiosoite: Åkerlundinkatu 2 A, 5. krs  
Tampere  
Postiosoite: Yhteiskuntatieteellinen tietoaarkisto  
33014 Tampereen yliopisto  
Puhelin: (03) 3551 8519  
Faksi: (03) 3551 8520  
Sähköposti: fsd@uta.fi  
Kotisivu: <http://www.fsd.uta.fi/>

