

ANONYMISOINTISUUNNITELMA KVALITATIIVISELLE AINEISTOLLE: ITÄ-SUOMEN PSYKOLOGIAN OPISKELIJOIDEN KOKEMUKSIA MONIMUOTO-OPISKELUSTA 2013: OPISKELIJOIDEN HAASTATTELUT

HUOM! Esimerkki on kuvitteellinen.

SISÄLTÖ:

1. AINEISTON TIEDOT ENNEN ANONYMISOINTIA JA PALJASTUMISRISKIN ARVIOINTI
2. ANONYMISOINTITOIMENPITEET PERUSTELUINEEN

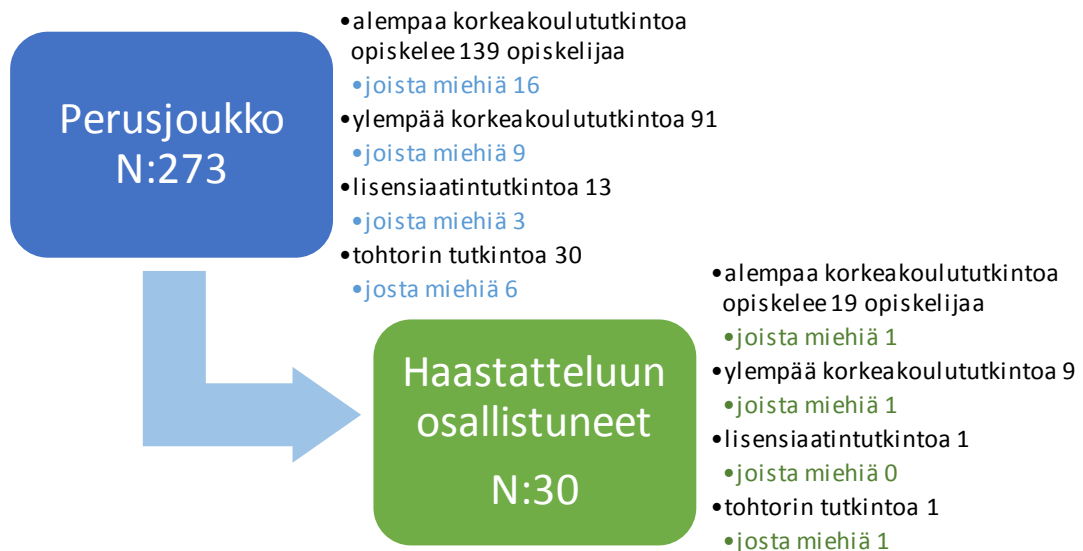
Suunnitelman tekijä(t): Annika Sallinen

Anonymisoinnin toteuttaja(t): Maija Majava, Pentti Pesukarhu

1. AINEISTON TIEDOT ENNEN ANONYMISOINTIA JA PALJASTUMISRISKIN ARVIOINTI

Alla on esitetty anonymisoinnin ratkaisujen taustalla olevat tekijät, jotka perustuvat [Tietoarkiston aineistonhallinnan käsikirjan ohjeisiin](#).

Perusjoukko ja otanta: Perusjoukkona tutkimuksessa olivat kaikki Itä-Suomen yliopiston syksyllä 2013 läsnäolevat psykologian pääaineopiskelijat (N: 273). Otanta on kokonaisotanta ja otos on itsevalikoitunut. Haastattelukutsu lähetettiin kaikille s-postilla ja halukkaat osallistuivat tutkimukseen. Haastattelut tehtiin yksilöhaastatteluina ja niihin osallistui lopulta 30 opiskelijaa. Osallistumisprosentti on 11. Perusjoukosta ja lopulta kyselyyn vastanneita kuvaavat seuraavat anonymisoinnin kannalta merkitykselliset tiedot. Perusjoukon tiedot ovat Vipusesta¹.



¹ Vipunen on opetushallinnon tilastopalvelu. Vipusesta saa tilasto- ja indikaattoritietoa eri sektoreiden koulutuksesta ja koulutuksen jälkeisestä sijoittumisesta, korkeakouluissa tehdystä tutkimuksesta sekä väestön koulutusrakenteesta ja opiskelijoiden sosioekonomisesta taustasta. Ks. vipunen.fi

Kun tarkastellaan Vipusesta perusjoukon vastaajien sukupuolta, miehet ovat vastaajissa vähemmistö, sillä 273:stä vastaajasta heitä on vain 34 eli 12,5 %. Katsottaessa ikää, harvinaisimpia ikäryhmiä perusjoukossa ovat 15-19-vuotiaat (N:3) ja kaikki ikäryhmät 40 ikävuodesta eteenpäin. Naisten määrä ikä- ja tutkintotietojen (kandi, maisteri jne.) valossa on anonymiteetin kannalta kohtuullinen, mutta sen sijaan miehet ovat harvinaisia havaintoja kaikissa ikäluokissa. Miehillä suurimman ikäluokan muodostavat kandia suorittavat 20-24-vuotiaat miehet (N: 8). Muut ikäluokat ovat yhtä suuria tai pienempiä kuin 3. Yksi harvinainen ikä- ja tutkintoyhdistelmä on 40-44-vuotiaat tohtorin tutkintoa suorittavat miehet, sillä heitä kohdejoukossa on vain yksi.

Haastatteluun osallistui lopulta kolme miestä, joista kaksi opiskelee kandin- tai maisterin tutkintoa ja yksi tohtoriksi. Vastaaja kuuluu ikäluokkaan 40-44 eli on edellisen kappaleen lopussa mainittu harvinainen tapaus.

Aineiston sisältö: Haastattelu toteutettiin haastatteluteemojen avulla, jotka käsittelivät psykologian opiskelijoiden kokemuksia monimuoto-opiskelusta Itä-Suomen yliopistossa. Haastatteluissa tiedusteltiin esim. onko ollut kursseilla, jossa hyödynnetään monimuoto-opiskelua, mitä hyvää ja huonoa monimuoto-opiskelussa on ja millä tavoin vastaaja oppii mielestään parhaiten. Haastatteluissa ei kysytty vastaajan lähipiiriin tai elämänkulkuun liittyviä kysymyksiä. Vastaajia pyydettiin haastattelun alussa, etteivät he toisi esille läheisten, opiskelukavereiden tai opettajien nimiä.

Haastatteluissa tuli esille mm. seuraavia epäsuoria tunnisteita:

Jokaiselta vastaajalta on kerätty taustatietoihin

- sukupuoli (mies/nainen/muu)
- ikä (vuoden tarkkuudella)
- opintojen vaihe (kandi/maisteri/tohtori tai lisensiaatti)

Haastatteluiden sisältöjen epäsuoria tunnisteita ovat mm.:

- millä kurssilla ollut mukana
- opintojen aloitusvuosi
- tarkkoja harjoitustehtävien kuvauksia
- edeltäviä opiskelutietoja yliopistoista ja koulutusaloista
- nykyinen opintosuuntaus
- kolmansista osapuolista kerrottiin esim. graduohjaajien ja samoilla kursseilla olleiden opiskelukavereiden nimiä

Aineisto ei sisällä tietosuojalainsäädännön mukaisiin erityisiin henkilötietoryhmiin lukeutuvia tietoja. Aineiston sensitiivisimmät tiedot liittyvät ryhmäoppimistilanteissa ilmenneisiin konflikteihin ja opiskelukavereihin tai ohjaajiin liittyviin arvosteluihin.

Aineistoa ei voi alkuperäisenä pitää anonyyminä, koska aineistossa on henkilönimien lisäksi epäsuoria tunnisteita, jotka muodostavat harvinaisia havaintoja. Henkilöt saattavat olla tunnistettavissa erityisesti seuraavia tietoja yhdistämällä:

- sukupuoli
- opintojen vaihe
- opintojen aloitusvuosi
- tarkka ikä, koulutushistoria
- graduohjaajan nimi

Aineiston ikä: Aineisto on nuori, sillä se on kerätty 2010-luvulla. Vastaajien tietoja voidaan olettaa löytyvän internetistä.

Vastaajista muualta saatavat tiedot: Vastaajista voi olla saatavilla lisätietoja muista tietolähteistä esim. sosiaalisesta mediasta ja vipunen.fi-palvelusta yliopisto- ja pääainetietojen avulla. Lisäksi mikäli urkkija epäilee henkilön opiskelevan yliopistossa, hän voi tarkistaa asian yliopiston opintotoimistosta, jos henkilö on antanut nimen luovuttamiseen luvan. Tulee ottaa myös huomioon, että pro gradu -, lisensiaatin ja tohtorin tutkielmat julkaistaan avoimesti netissä nimellä, vuosiluvulla ja pääainetiedoittain. Lisäksi tulee ottaa huomioon ihmisten tietämys toisistaan yliopistolla, sillä tulevat aineiston jatkokäyttäjät ovat mahdollisesti samassa yliopistossa työskenteleviä henkilöitä, jopa opiskelukavereita.

Ulkopuolinen henkilö voi yrittää tunnistaa yksittäisen henkilön yhdistämällä aineiston tietoja netistä löytyviin tutkielmatietoihin ja sosiaalisen median tietoihin (esim. Facebook). Yhdistettäviä tietoja voivat olla esimerkiksi opintojen pääaine, yliopisto, opintojen aloitusvuosi, sukupuoli (tässä tapauksessa erityisesti miesopiskelijat) ja tieto vastaajan edellisestä oppilaitoksesta. Helpoimmin vastaajista voi tunnistaa miesopiskelijat ja erityisesti lisensiaatin- ja tohtoriopiskelijat (miespuoleiset) henkilöt, sillä kohdejoukossa heidän määränsä on hyvin pieni. Tohtoriopiskelijoiden yhteystiedot ovat usein yliopiston nettisivuilla.

Vastaajan sosiaalisille verkostoille kuten perheenjäsenille, kavereille ja opettajille henkilöiden tunnistaminen on helpompaa riippuen, mitä tietoja henkilöllä on vastaajasta. Opiskelukaverit tai opettajat ovat myös potentiaalisia aineiston jatkokäyttäjiä, joten anonymisoinnissa pyritään estämään tunnistettavuus myös lähellä vastaajan elinpiiriä olevilta henkilöiltä.

Käytettävyys vs. anonymiteetti: Tässä tutkimuksessa ollaan kiinnostuneita monimuoto-opiskelusta erityisesti sen mukaan, kuinka pitkällä opinnot ovat ja mitkä ovat omat oppimistavat. Taustatiedoista halutaan erityisesti säilyttää tieto opintojen vaiheesta eli ts. suoritettavasta tutkinnosta, muut taustatiedot voidaan uhrata anonymisoinnille. Tarkastelu sukupuolen mukaan olisi mielenkiintoinen tutkimuksissa, mutta vähäinen miesten määrä voi este tiedon yleistettävyydelle.

2. ANONYMISOINTITOIMENPITEET

Kaikki anonymisoinnit on merkitty [hakasulkeisiin] pl. kohtia joiden anonymisointi on pakko tehdä ilman merkintöjä ks. kohta 1.

1. Mitä taustatietoja jätetään?

- ✓ Ikä luokiteltuna:
 - alle 30 vuotias
 - 30 vuotias tai yli
- ✓ Opintojen vaihe luokiteltuna:
 - kandi
 - maisteri
 - lisensiaatti/tohtoriopiskelija

Jos ikä käy ilmi haastattelun sisällöstä, se luokitellaan yllä esitettyjen luokitusten mukaisesti.

Kandi ja maisterivaiheen tietoa ei tarvitse muuttaa haastattelun sisällöstä. Mikäli henkilö mainitsee, että hän opiskelee tohtori- tai lisensiaatin opintoja, tieto tulee luokitella yllä mainitulla tavalla.

Kohdan voi poistaa myös ilman merkintöjä. Esimerkki:

Näin vanhempana opiskelijana, olen 36-vuotta, opiskelu on jo työläämpää → Näin vanhempana opiskelijana, olen [30 tai yli], opiskelu on jo työläämpää

Meillä tohtoriopintoihin ei kuulu monimuoto-opetusta muuta kuin yhdellä kurssilla → Meillä [lensiaatti/tohtoriopintoihin] ei kuulu monimuoto-opetusta muuta kuin yhdellä kurssilla

Olen tohtoriopiskelija. → Olen [lensiaatti/tohtoriopiskelija]

Perustelut: Jos aineistoon jätettäisiin täsmälliset tiedot iästä, tieto tohtori- ja lensiaattitutkinnoista ja sukupuolesta, taustatiedot voisivat mahdollistaa henkilön tunnistamisen. Etenkin miehet ja tohtori- tai lensiaattitutkinnon suorittajat voidaan tunnistaa yhdistämällä tiedot muihin tietoihin. Taustatiedoista päätettiin jättää pois kokonaan sukupuoli, sillä aineistossa on vain kolme miestä. Sukupuolten välinen vertailu pienellä aineistolla voisi jäädä tutkimuksellisesti epäluotettavaksi, vaikka sukupuolella voikin olla merkitystä monimuoto-opiskelussa. Yhdistettäessä tohtori- ja lensiaatin opiskelijat yhteen muodostuu 43 henkilön ryhmä, mikä voidaan katsoa tarpeeksi isoksi joukoksi tunnistettävyyden kannalta.

2. Haastatteluista poistetaan vastaajia ja kolmansia henkilöitä koskevat suorat tunnisteet, kuten vastaajan s-postiosoite, jossa on henkilön koko nimi.

Perustelut: Suorat ja vahvat suorat tunnisteet poistetaan aina, koska henkilön voi tunnistaa helposti niiden avulla.

3. Henkilönimen tilalle annetaan keksitty nimi. Henkilönimissä poistetaan sukunimi ja keksitty nimi annetaan vain etunimelle. Huom! Jos kyse on yliopiston henkilökunnasta, heidät kategorisoidaan nimen antamisen sijaan. Jos samassa haastattelussa mainitaan useampi opettaja tai ohjaaja, nimille voi antaa kirjain tai numerotunnisteet, jotta heidät voidaan erottaa. Esimerkit:

Helvi Mutala → [Maria]

Matti Tukkonen (tiedetään että on graduohjaaja) → [graduohjaaja 1] tai [opettaja 1]

Perustelut: Henkilönimi on usein suora tunniste ja se pitää poistaa. Tässä nimen tilalle vaihdetaan keksitty nimi, jotta henkilöt voidaan yhdistää saman haastattelun sisällä. Opettajat ja ohjaajat kategorisoidaan, sillä he voivat olla tunnistettavissa jo yksin sukupuolen perusteella, jos ohjaajien tai henkilökunnan sukupuolijakauma on epätasainen. Jos graduohjaaja paljastuu, helpottuu myös opiskelijan tunnistaminen.

4. Poistetaan haastattelujen sisällöistä vastaajasta ja kolmansista henkilöistä seuraavia tietoja:

- a. kurssien nimet, joiden osallistujamäärä on ollut pieni esim. alle 20 ja tiedot kurssien sisällöstä, jos kurssi on tunnistettavissa niiden perusteella
- b. tieto, josta ilmenee, että haastateltava on mies. Tieto tulee poistaa ja merkitä (tai jättää poisto merkitsemättä) niin, ettei miessukupuolisuutta voi päätellä kontekstista.
- c. paikkakunnat pl. Joensuu (jos eivät ole yleisiä ilmauksia, kuten ”Helsingissä on paljon työpaikkoja”)
- d. koulutushistoria (kuten ”olen opiskellut aikaisemmin biologiaa Oulun yliopistossa”)
- e. harvinaiset asemat esim. ainejärjestön hallituksessa oleminen
- f. poista myös kolmansia osapuolia, kuten kavereita tai graduohjaajia koskevien henkilöiden ominaisuudet, joiden perusteella heidät voi tunnistaa.

Perustelut: Edellä mainitut tiedot ja muut sen kaltaiset tiedot tulee poistaa, jotta henkilö ei muodostu harvinaiseksi tai ainutlaatuiseksi havainnoksi. Poistettavien tietojen kohdalla tulee arvioida, voiko tietoa saada ulkopuolisista lähteistä tai paikallistietämyksellä selville. Sukupuolitieto poistetaan miesten kohdalla, koska heitä on eri tutkintoryhmissä/opintojen vaiheissa liian vähän. Koulutustaustat, asuin- ja syntymäpaikkakunnat, asemat tai tittelit voivat käydä ilmi internetistä.

5. Luokitellaan, kategorisoidaan tai sanallistetaan yleisemmälle tasolle muut epäsuorat tunnisteet, joiden perusteella henkilöt voisivat olla tunnistettavissa. Näin säilytetään aineiston sisällöstä tietoa edes yleisemmällä tasolla sen sijaan, että tieto poistetaan kokonaan. Alla on esimerkit luokittelusta, kategorisoinnista ja sanallistamisesta.

- a. Luokittelu: Opiskeluvuosien määrä tai tarkka opintojen aloitusvuosi luokitellaan kahden tai kolmen vuoden välein.
- b. Kategorisointi: Verkossa opiskelu ja työskentely ovat tuttuja töistäni sosiaalitoimesta. → Verkossa opiskelu ja työskentely ovat tuttuja töistäni [työpaikasta].
- c. Sanallistaminen (ts. asioiden esittämistä karkeammalla tasolla, mutta voi myös sisältää tietojen tiputtamista pois): Energiaa ei ole riittänyt verkko-opiskeluun kotona, koska tyttäreemme kumpurajalka on vaatinut viime aikoina lisäleikkauksia. → Energiaa ei ole riittänyt verkko-opiskeluun kotona, koska [lapsiin liittyvän syyn vuoksi].

Perustelut: Samat kuin kohdassa 4.