

Tilastoista ja niiden tulkinnasta

Sinikka Luokkanen

Tilastoitavien tietojen määrittäminen, niiden rajaaminen ja tulkinta ovat kirjastojen tilastointityön ja tilastojen luotettavuuden kannalta ydinasioita. Näitä tekijöitä joudutaan punnitsemaan tilastoinnin perusteita rakennettaessa, mutta myös tilastointityössä sekä tilastoitavan kohteen muuttuessa. Kirjoituksessa pohditaan tieteellisten kirjastojen nykyisten kirjastotilastojen tuottamisen ongelmakohtia ja kehittämistarpeita. Yhteistilastojen kehittämiseen kaivataan tiiviimmin myös kirjastojen panosta mukaan.

Tieteellisten kirjastojen toimintaa ja voimavaroja kuvaavilla tilastoilla on pitkät perinteet Suomessa. 1990-luvulla tilaston kokoamisvastuu siirtyi Helsingin yliopiston kirjastolle, nykyiselle Kansalliskirjastolle. Suurin osa tilastoitavista asioista on kansainvälisen ISO 2789 standardin mukaisia tietoja.

Ensimmäinen kansainvälinen ISO standardi, ISO 2789 *International library statistics*, vahvistettiin vuonna 1974. Standardin sivumäärä oli 4. Se kumottiin vuonna 1991, jolloin vahvistettiin 9-sivuinen ISO 2789:1991 standardi. Nykyinen voimassaoleva 52-sivuinen versio on vahvistettu vuonna 2003. Suomi on ollut mukana kehittämissä standardeja.

Kansallista työskentelyä ja työryhmiä on kuvannut mm. Maarit Olander Tietolinjassa ilmestyneessä artikkelissaan (Olander, 2000). Samassa Tietolinjan numerossa ovat myös Tuulikki Nurminen, Marja Anttosen ja Ritva Hännisen artikkelit, joissa he kuvaavat edustamiensa kirjastojen ja ao. sektoreiden tilastointiin liittyviä odotuksia ja ongelmia.

On samalla kertaa lohdullista ja turhauttavaa huomata, että nykyisiä ongelmia standardin mukaisessa tilastoinnissa on ollut jo sen käyttöön-

ottovaiheessa – lohdullista, kun ei ole yksin ja turhauttavaa, kun mikään ei ole vuosien saatossa muuttunut.

Tilastoinnin ongelmiin on vaikea löytää yksinkertaista kuvausta, koska kaikki näyttävät liittyvän toisiinsa. Yhtäältä on tulkintaeroja siinä, mitä halutaan kuvata/mitata ja toisaalta mitä yksittäiset määrittelyt tarkoittavat.

Mitä tilastoilla kuvataan?

Eräs tärkeistä peruskysymyksistä on, tilastoidaanko kirjaston toimintaa vai organisaation ”kirjastotoimintaa”? Jos tilastoidaan kirjaston toimintaa, niin yksi tyypillisimmistä ongelmista on, että standardissa ei ole mukana kaikkia niitä tehtäväkokonaisuuksia, joita kirjastossa nykyisin tehdään. Nurminen luettelee esimerkkeinä julkaisu-toiminnan, verkkopalveluiden tuottamisen, mikro-tukineuvonnan, elektronisen julkaisu-toiminnan ja kehysorganisaation julkaisujen vaihdon ja myynnin (Nurminen, 2000). Viimeisinä vuosina uusien tehtäväkokonaisuuksien kirjo lienee vain laajentunut.

Tilastoitaessa organisaation kirjastotoimintaa mukaan tilastointiin pitäisi ottaa myös muu kuin kirjaston hankkima ja/tai rahoittama tietoaineis-

to kokonaisuudessaan. Osa kirjastoista ei missään tapauksessa halua tätä monistakin syistä. Kaksi tärkeintä syytä lienee tietojen keräämisen työläys ja haluttomuus ottaa mukaan aineistoja, jotka on hankittu vain hyvin rajoitetun henkilöstön käyttöön.

Toisaalta on organisaatioita, jossa on linjattu, että kirjaston tietohuollon mukaisesti tehtäviin kuuluu myös rajattuun käyttöön tarvittavan aiheiston hankinta ja kokoelmanhoito laajasti ymmärrettyinä, vaikka asianomaisten aineistojen hankintaa ei maksettaisikaan kirjastolle osoitetuista toimintamäärärahoista.

Nykyisen tieteellisten kirjastojen tilastotietokannan valmiiksi lasketuista 24 tunnusluvusta erilainen tulkinta oheisesta kysymyksestä vaikuttaa suoraan ainakin neljään tunnuslukuun kokoelmien koon, kartunnan ja hankintakulujen kautta. Molemmat näkökannat ovat mielestäni perusteltuja ja siten oikeutettuja. Tulkinterot johtavat vääjäämättä siihen, että kun perustilastotiedot eivät ole vertailukelpoisia, niin mitkään niiden perusteella lasketut tunnusluvut eivät ole luotettavia.

Mitä tilastoidaan?

Nykyiseen KITT-tilastoon kysytään yli kolme sataa yksittäistä tietoa. Suurin osa kysymyksistä koskee kokoelmia. Kokoelmatiedot jaotellaan yksityiskohtaisesti. Onko jaottelu jo niin yksityiskohtainen, että alan ulkopuolisen henkilön on vaikea nähdä metsä puilta?

Saapuvissa kausijulkaisuissa tilastoidaan sekä tilausten että nimekkeiden määrää. Yhteistilasto-ohjeen mukaan kohtaan ilmoitetaan kaikkien saapuvien kausijulkaisujen määrä (Yhteistilasto-ohje, 2003). Jos kirjasto tilaa yhtä nimekettä kolmena kappaleena, niin tilausten määrä on kolme, mutta nimekkeiden yksi.

Vuoden 2004 tilastoissa yliopistokirjastoilla on kausijulkaisunimekkeitä enemmän kuin tilauksia. Kun ”tilausten määrän” tulisi kuvata kaikki saapuvia kausijulkaisuja – kukin nimeke vähintään kertaalleen laskettuna, niin on vaikea ym-

märtää mistä johtuu, että tilauksia on vähemmän kuin nimekkeitä. Jos ”tilausten määrä” ymmärretään koskevan vain ostettavia tilauksia, niin silloin tilastosta ei mistään saada saapuvien kausijulkaisujen kokonaismäärää.

Kokoelmaosuudessa on myös luettelointi, jossa KITTissä eritellään myös uutuusluettelointi, johon ohjeen mukaisesti tilastoidaan julkaisut, jotka on julkaistu enintään kolme vuotta sitten. Näin määriteltynä tilaston uutuusluettelointi kuvaa, kuinka paljon luetteloidusta kartunnasta on ilmestynyt viimeisen kolmen vuoden. Määrittely on ollut aikana, jolloin tehtiin paljon takautuvaa luettelointia, ihan käyttökelpoinen. Mutta onko se vielä relevantti ja jos on, niin olisiko parempi puhua uutuushankinnasta kuin uutuusluetteloinnista?

Edellä kuvaamani saapuvien nimekkeiden tilastointiin liittyvän epäselvyyden kanssa samankaltainen tilanne ymmärtääkseni on tiloissa. Yhteistilasto-ohjeen mukaan tiloihin neliömetreinä ilmoitetaan kirjaston käytössä olevat hyötyneliömetrit ja kohtaan kirjastorakennuksen kokonaispinta-ala vähintäänkin edellä mainitut hyötyneliömetri. Tässä kohdassa kun tavoitteena on saada tilastoihin mukaan myös aulat, portaikat ja muut yleiset tilat, joita hyötyneliöihin ei lasketa. Vuoden 2004 tilastoissa sekä ammattikorkeakoulu- että yliopistokirjastojen yhteissummissa hyötyneliömäärät ovat kuitenkin suuremmat kuin kokonaisneliömäärät.

Mikä on kehysorganisaatio?

Tänään Suomessa on useita kirjastoja, jotka palvelevat useampaa organisaatiota. Marja Anttonen kuvasi hyvin ammattikorkeakoulusektorin tilannetta, missä useat ammattikorkeakoulun kirjastot palvelevat samalla myös toisen asteen opetusta (Anttonen, 2000).

Toiminta voi pohjautua joko samaan ylläpitoorganisaatioon tai palvelusopimuksiin. Molemissa tapauksissa ainakin kohdeväestön kokoa, oman organisaation opiskelija ja henkilökuntamäärä käyttävät tunnusluvut ovat vinoutuneita,

jos toista palveltavaa kokonaisuutta ei huomioida. Talouslukujen osalta tilanne on vielä moniulotteisempi. Kuitenkin standardi mielestäni yksiselitteisesti toteaa, että vain äärimmäisissä tapauksissa (”in extreme cases”) useampaa kuin yhtä organisaatiota palvelevien kirjastojen pitää jaksaa toimintansa ja tilastoida ne erikseen. (ISO 2789:2003, 3.1.5 Library, Note 2).

Voimassaoleva yhteistilasto-ohjeen mukaan

”Jos kirjastolla on useampia kuin yksi tehtävä (esim. yleinen kirjasto ja ammattikorkeakoulukirjasto tai ammattikorkeakoulukirjasto ja toisen asteen kirjasto) ilmoitetaan tiedot siinä ryhmässä, jota kirjasto pääasiallisesti palvelee sekä kehysorganisaation tiedot, lainausaste, talouden ja henkilökunnan osalta tämän toiminnan osuutena.”

Mitä ohjeessa tarkoitetaan lainausasteella? Kun tilastossa kysytään kirjaston kokonaislinalukuja eikä asiakasryhmittäisiä lainauslukuja, niin yksittäisenä syötettävänä tietona lainausaste jää epäselväksi.

Toisaalta, kun kirjasto on avoin ja palvelee myös muita kuin tilastoinnin kehysorganisaatiota, niin eikö ”toisen tehtävän” osuus tulisi tilastoida kokonaisuudessaan myös talous- ja henkilöstöosuuksissa vähintään maksullisen palvelun osuuksissa? Jos palvelu perustuu palvelusopimukseen, tilanne lienee kiistaton. Jos sitten vain saman ylläpito-organisaation tapauksissa toisen palvelutavan organisaation tiedot jäävät kokonaan tilastoinnin ulkopuolelle, niin tilastot antaisivat todella virheellisen kuvan tilanteesta.

Oma tulkinta standardista on toinen. Kun kirjastolla on pääasiallinen tehtävä, niin tiedot ilmoitetaan siinä ryhmässä ilman, että kirjaston toimintaan kuvaavia lukuja jaettaisiin ensi- ja toissijaisten tehtävien perusteella. Vaikka tilastoitaisi jaettaisiin nykyisen ohjeen mukaisesti talous- ja henkilöstöluvut ”ryhmän / toiminnan osuuteen”, niin edelleen vinoutuneiksi tunnusluvuiksi jäisivät mm. kohdeväestöön suhteutetut kokoelmia,

kartuntaa ja tiloja kuvaavat luvut. Kun tilastointiin asioihin on talouspuolelle lisätty mm. yhteisprojektien osuudet, niin on mielestäni entistä selvempää, että toissijaisten organisaation osuudet on tilastoitava jossakin kohdassa. Yksimielisyys olisi hyvä löytää siitä, että missä ja miten.

Suomessa on jo usean organisaation yhteisiä kirjastoja ja suunnitteilla lienee liuta lisää. Yhteiskirjastojen kohdalla kehysorganisaation omien opiskelijoiden ja henkilökunnan määrien sekä taloustietojen osalta voidaan joutua sopimaan eri tilastoitavien yksiköiden kesken toimintatavoista, jotta samoja henkilöitä ja euroja ei tilastoitaisi useampaan kertaan.

Kansallisten tilastojen kerääminen

Standardin liitteessä C ohjaa kansallisten ja muiden yhteenlaskuun perustuvien koottujen tilastojen tuottamista. Standardin mukaan puuttuvat osuudet on arvioitava (”grossing up”) sen antamien laskentakaavojen mukaisesti. Standardin mukaan julkistetussa tilastossa pitää ilmoittaa se, jos yhteenlasketussa tilastossa on käytetty laskennallista täydentämistä sekä mitä laskentatapaa on käytetty. Yksittäisen organisaation tietojen täydentämisestä liitteen C laskentakaavoilla ei siis ole kyse.

Ainakaan perustamisvaiheessa KITTin rakennetta ei tehty sellaiseksi, että puuttuvien tietojen täydentäminen kansallisen tason tilastoissa olisi tehty standardin vaatimusten mukaisesti. Tiedosani ei ole, kuinka täydennys tehdään, jos/kun sitä KITTissä tehdään.

AMKOTA, KITT ja KOTA

AMKOTA tilasto on talouden osalta suorastaan niukka. Ammattikorkeakoulujen taloutta kuvataan vain opiskelijakohtaisella yksikköhinnalla. KOTA tilastoissa yliopistojen budjettirahoitus jaotellaan palkkoihin, tilakustannuksiin, rakensinvestointeihin ja muihin kuluihin. Budjettirahoituksen lisäksi tarkastellaan ulkopuolista rahoitusta, joka jaotellaan seitsemään alaryhmään. Kustannukset KOTA tilastoi tulosalueittain: kou-

lutus, tutkimus ja taiteellinen toiminta sekä yhteiskunnalliset palvelut.

KITTTin taloustiedot ovat erittäin yksityiskohtaiset. Mitä yksityiskohtaisemmin kysytään, sitä suurempi on virheiden todennäköisyys sekä tulokintojen että puhtaiden syöttövirheiden osalta.

AMKOTA ja KOTA tilastot tiivistävät organisaatioiden toiminnan kuvausta muutamiiin keskeisiin tilastoihin. Niiden perusteella ulkopuolisenkin on mahdollista saada hyvä yleiskuva tilanteesta. KITTT sitä vastoin kuvaa yksityiskohtaisia asioita.

Kirjasto- ja tietopalvelualan ammattilaisella on paikka paikoin vaikeuksia ymmärtää, mitä tilastoilla kuvataan. Kuinka alan ulkopuolinen päättäjä/vaikuttaja näkeekään KITTTin satoine tilastokohtineen? Tulisiko päättäjille selkeämpi kuva tieteellisestä kirjastotoiminnasta Suomessa, jos keskeisimmät KITTTin tilasto- ja/tai tunnusluvut koostettaisiin paljon suppeammaksi tilastoyhteenvedoksi. Koko KITTT olisi toki alan käytössä julkisesti verkossa, mutta sitä ei markkinoitaisi alan ulkopuolisille käyttäjille.

Yhteistilastojen kehittämiseen yhteistä pohjaa 2000-luvun alussa nykyistä tietokantapohjasta KITTT tehtäessä Kansalliskirjasto nimesi hankkeelle ohjausryhmän ja projektiryhmän, jossa osallistujasektorit olivat edustettuina. Viime aikoina KITTTiin on tehty muutoksia, josta ymmärtääkseni ei ole neuvoteltu kaikkien tilastoinnissa mukana olleiden tahojen kanssa. Yhteistilaston luoneen mukaista olisi sopia asioista yhteisesti, joten olisiko aiheellista Kansalliskirjaston perustaa pysyvä KITTT-ohjausryhmä?

Yhteistilaston yhteydessä on käytetty slogania ”KITTT on kirjaston peili”. Peilinkin, antaakseen värittymätöntä ja vinoutumatonta kuvaan maailmasta, täytyy olla laadultaan hyvä ja ominaisväril-

tään neutraali tasopeili. Muutoin vika ei ole aina katsojassa, vaikka kuva ei miellyttäisikään. ☞

Lähteet:

Ammattikorkeakoulut 2004: Taulukoita AMKOTA tietokannasta (2005) Viitattu 30.9.2006 http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2005/liitteet/opm_312_opm23.pdf?lang=fi

Anttonen, Marja (2000) Ammattikorkeakoulukirjastojen tilastoinnin näkökulmia ja ongelmia, Tietolinja 4/99 Viitattu 30.9.2006 <http://www.lib.helsinki.fi/tietolinja/0499/tietolinja0499.pdf>

ISO 2798:2003 Information and documentation - International library statistics

Nurminen, Tuulikki (2000) Kuvaavatko tilastot riittävän relevantisti kirjastojen tilannetta?, Tietolinja 4/99 Viitattu 30.9.2006 <http://www.lib.helsinki.fi/tietolinja/0499/tietolinja0499.pdf>

Olander, Maarit (2000) Kirjastotilasto uudistuvat, Tietolinja 4/99 Viitattu 30.9.2006 <http://www.lib.helsinki.fi/tietolinja/0499/tietolinja0499.pdf>

Tieteellisten kirjastojen tilastotietokanta. Viitattu 30.9.2006 <http://yhteistilasto.lib.helsinki.fi/>

Yhteistilasto-ohje 2003. Tulostettu 30.9.2006. Saatavana Yhteistilastoon kirjautumisen jälkeen Yhteistilaston ajankohtaista -sivulta.

Yliopistotilastot 2004: Taulukoita KOTA-tietokannasta (2005) Viitattu 30.9.2006 http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2005/liitteet/opm_309_opm26.pdf?lang=fi

Tietoa kirjoittajista:

Luokkanen, Sinikka, tietopalvelupäällikkö
Hämeen ammattikorkeakoulu
email. sinikka.luokkanen@hamk.fi