

T.D. WILSON

## Kirjasto- ja informaatiopalvelujärjestelmien evalvointistrategiat<sup>1</sup>

Wilson, T.D., Kirjasto- ja informaatiopalvelujärjestelmien evalvointistrategiat [Evaluation strategies for library/information systems]. Kirjastotiede ja informatiikka, 3 (4): 90–94, 1984.

The ability to identify goals and objectives is an essential prerequisite of all evaluation. Only when we can think about criteria which indicate when an objective has been met, can we think of whether or not it is possible to »measure» these criteria in a quantitative sense. Starting from these arguments the author reviews the topics »What can we evaluate?», »According to what criteria?», and, »By what research strategies?». Arguments are presented for the view that a well-founded qualitative evaluation of benefits can be very persuasive if it supports philosophical or ethical or moral values held by those responsible for making decisions about resources.

The call for evaluation is the result of some problem occurring in a system, or the librarian/information worker being under pressure to justify expenditure of staff, materials, other aspects of services, or proposed changes in services. Only the librarian/information worker can decide what needs to be evaluated.

*Address: University of Sheffield, Department of Information Studies, Sheffield S10 2TN, U.K.*

### Johdanto

Viime vuosikymmeninä runsain mitoin harjoitettu tiedonhakujärjestelmien evalvointi on vaikuttanut monin tavoin kirjasto- ja informaatiopalvelu-, K & I -järjestelmien evalvointia koskeviin käsityksiin. Sieltä sellaiset termit kuin »saanti» ja »tarkkuus» ovat siirtyneet ammattisanastoon, joskaan monet niiden käyttäjät eivät aina ymmärrä termien käyttöaluetta, vaan soluttavat ne monenlaisiin asiayhteyksiin, myös sellaisiin, joihin ne eivät ollenkaan kuulu. Näillä termeillä opeoivien evalvointien vaikeudet kiteytyvät kysymyksiin siitä mitä tarkoittaa »relevanti», mitä »relevanssi»? Määriteltäköön relevanssi niin tai näin, aina sen mittayksiköiden määrittely on ongelmallista. Samanlaisia ongelmia on myös muita.

Ongelmallisuudestaan huolimatta tiedonhakujärjestelmien evalvoinnin idea on osoittautunut hedelmälliseksi, sillä moni kirjastonhoitaja on alkanut pitää evalvointia johtamisen hyödyllisenä, jopa välttämättömänä tukitoimintana. Turhan usein evalvoinnin menetelmää valittaessa mallit kuitenkin otetaan tiedonhaun evalvoinnista, mikä merkitsee, että pääpaino on muuttujien mittaamisessa ja niiden välisissä kompromisseissa (*trade-offs*). On tietenkin ideaalista, jos tällaisia evalvointimenetelmiä voidaan käyttää, mutta useimmiten ei ole edes mahdollista löytää kirjastojärjestelmistä tällaista mitattavaa.

Syynä siihen, että tiedonhaun evalvointi onnistuu kun on onnistuakseen, on siinä, että tiedonhakujärjestelmien tavoitteet voidaan usein jäsentää varsin selvästi. Voidaan esim.

<sup>1</sup> Suomentanut Marjatta Okko.

sanoa, että ideaalinen tiedonhakujärjestelmä tulostaa käyttäjille kokoelman kaikki »relevantit» dokumentit eikä yhtään »irrelevanttia» dokumenttia. Kaiken evalvoinnin perusedellytys on, että päämäärät ja tavoitteet kyetään ilmaisemaan riittävän täsmällisesti, sillä vain silloin voidaan miettiä millaiset kriteerit osoittavat onko jokin tavoite saavutettu ja vasta sitten on aika pohtia onko vai eikö ole mahdollista »mitata» kriteerejä kvantitatiivisesti.

Miksi sitten halutaan evalvoida palveluja? Evalvoinnin idea näyttää tulleen ammatilliseen tietoisuuteen aivan viime vuosina. Tosin kirjastotutkimuksen (*library survey*) idea kehiteltiin jo 1930-luvun Yhdysvalloissa ja siinä ne evalvoinnin siemenet jo olivat. Viime aikoina evalvoinnin merkitys on korostunut, koska kirjastoilta on vaadittu entistä paljon täsmällisempiä ja tyhjentävämpiä budjettiperusteita. Yleensäkin organisaatiot tarkastelevat palvelutoimintojaan siltä kannalta, ovatko ne välttämättömiä organisaation tavoitteiden kannalta. K & I -järjestelmät eivät tässä suhteessa muodosta poikkeusta. Yrittäessään selviytyä budjettisupistuksista myös koulutuslaitokset toimivat tällä tavalla, siis tarkastelevat palvelutehtäviään suhteessa laitosten tavoitteisiin. Paikallishallinnossa mm. yleisillä kirjastoilla on samanlaisia paineita.

Tästä tilanteesta on seurannut, että kustannukset on kytketty jopa ylikorostetusti evalvointiin sen sijaan, että kirjastopalveluja olisi perusteltu lähtemällä liikkeelle niiden käyttökelpoisuudesta kirjaston käyttäjän kannalta. Raha on kieltämättä kätevä kvantitatiivinen mittayksikkö, mutta sen yksipuolinen korostaminen on mielestäni johtanut siihen, että K & I -palvelujen tuottamien hyötyjen arviointiin kelvolliset muut kriteerit ovat jääneet taka-alalle. Käsitykseni mukaan huomio olisi suunnattava juuri näihin kriteereihin.

### Mitä voidaan evalvoida?

Vastaus väliotsikon kysymykseen on hyvin yksinkertainen: Mitä tahansa organisaation toimintaa voidaan evalvoida! Evalvoivaan tarkasteluun voidaan ottaa esim:

- organisaatiossa sovellettavan johtamisrakenteen toimivuus
- informaatioaineistoihin kohdistuvat sisäiset operaatiot, kuten luettelointi, luokitus, indeksointi yms.
- käyttäjille tarkoitettut K & I -palvelut
- uudet palvelumuodot<sup>1</sup>

- uuden teknologian antamat mahdollisuudet tukea palvelutoimintaa
- minkä tahansa tekemisen vaihtoehtoiset ratkaisut
- muutossuunnittelua edeltävä kokonaisjärjestelmän toiminta

Jos kirjastonhoitaja kysyy »Onko mahdollista evalvoida X:ää?», vastaus on aina myöntävä, koska periaatteessa on mahdollista evalvoida mitä tahansa. Ongelma on siinä, että aina ei tiedetä miten se tehtäisiin, tai että sen teettäminen olisi liian kallista tai että käytettävissä olevat menetelmät eivät tuota juuri sitä tietoa, mikä kirjastonhoitajalla oli mielessä esittäessään kysymyksensä.

### Minkä kriteerien perusteella?

Seuraava kysymys on: »Minkä kriteerien perusteella X:ää voidaan evalvoida?» ja tällä kysymyksellä onkin jo joukko seuraamuksia.

Seuraavat kriteerit ovat käyttökelpoisia:

- menestyksellisyys (*»success»*)
- suorituskyky (*efficiency*)<sup>2</sup>
- tehokkuus (*effectiveness*)
- hyödyt (*benefits*) ja
- kustannukset (*costs*), joita voidaan evalvoida yksinään tai yhdessä muiden kriteerien kanssa.

On syytä pitää mielessä, että listan viimeistä lukuunottamatta kriteerit saattavat käsitellä joko kvantitatiivista tai kvalitatiivista arviointia tai molempien yhdistelmiä.

### Menestyksellisyys

Menestyksellisyyden käyttäminen kriteerinä liittyy varsin selvästi uusien palvelumuotojen tai muunlaisten innovaatioiden evalvointiin. Yhtä selvää on, että käsitys jonkin hankkeen menestyksellisyydestä saattaa olla hyvinkin subjektiivinen: Yhden henkilön vaatimattomaksi arvioima menestys on toiselle silkkaa epäonnistumista. – Poliitikot ovat mestareita kääntämään surkeat epäonnistumiset vaatimattomiksi menestyksiksi. Ehkä kirjastonhoitajan tulisi olla yhtä taitava.

<sup>1</sup> Tekijä käyttää englantilaiseen tapaan ilmaisua »new programmes of service delivery». Se sisältää ajatuksen, että kukin palvelumuoto on itsenäinen kokonaisuus, joka budjetoidaan erikseen ja jonka menestymistä voidaan tarkastella muista »ohjelmista» riippumatta (suom. huom.)

<sup>2</sup> »Efficiency» voidaan ilmaista myös sanalla toimivuus (suom. huom.)

Subjektiiivisuutta voidaan vähentää, jos menestyksellisyys kriteerit voidaan määrittellä. Ne puolestaan riippuvat uudelle hankkeelle asetetuista tavoitteista. Esimerkkinä olkoon videokasettien lainauspalvelun aloittaminen yleisessä kirjastossa. Onko hankkeen tarkoituksena

- tyydyttää jokin filosofisperustainen tavoite asettaa saataville kaikenlaisia informaatioaineistoja?
- ansaita lisätuloja perimällä palvelusta maksuja?
- kiinnittää sellaisten henkilöiden huomiota kirjastopalveluihin, jotka muutoin eivät käyttäisi kirjastoa?

Menestyksellisyys kriteerit riippuvat tarkoituksista. Ensimmäisessä tapauksessa riittää, että aineistot hankitaan ja asetetaan käytettäväksi; pelkkä teko takaa menestyksellisyys. Toisessa tapauksessa menestyksellisyys voidaan mitata toiminnan tuottaman voiton avulla; jos lainauksesta saatavat tulot ovat suuremmat kuin kokoelman muodostamisen ja ylläpidon vuotuiset kustannukset, voidaan toiminnan katsoa olleen menestyksellistä. Kolmas tapaus on jo hankalampi; se vaatii yksityiskohtaista käyttäjäseurantaa ja sen selvittämistä, onko toiminnasta seurannut täysin uusien käyttäjien suhteellisen osuuden kasvua. Evalvointia ei kuitenkaan voi lopettaa tähän. Kuinka monta uutta käyttäjää on saatava, jotta hanketta voitaisiin pitää menestyneenä? Riittääkö, että uudet käyttäjät käyttävät vain videokasettien lainauspalvelua, vai edellytetäänkö, että he käyttävät myös muita palveluja? Mikäli tällaista edellytetään, on vielä kysyttävä miten suuri muun käytön pitäisi olla, jotta voitaisiin päätellä onko hanke ollut menestyksellinen vai ei.

Mitä kompleksisemmat jonkin hankkeen vireillepanon syyt ovat, sitä monipuolisemmiksi käyvät kysymykset ja sitä aikaavievemmäksi evalvointiprosessi.

## Suorituskyky

On tapana sanoa, että jokin palvelu on toimiva, kun siihen kohdistuu se minimikäyttö, joka tarvitaan, jotta palvelu ylipäänsä voisi toimia. Kun arvioidaan polttomoottorin suorituskykyä, selvitetään miten paljon polttoainetta se käyttää verrattuna muihin yhtä voimakkaisiin moottoreihin. Urheilijain suorituskykyä voidaan mitata selvittämällä heidän hapenotto kykyään ja kykyään muuntaa se energiaksi.

K & I -järjestelmien suorituskykyä on siis mitattava voimavarojen (ihmiset, aineistot, varat) kulutuksen perusteella. Kysymykset, joihin haetaan vastauksia ovat tyyppiä »Voimmeko ylläpitää samaa tai parempaa tietopalvelun tasoa käyttämällä vähemmän ihmisiä ja ottamalla käyttöön online hakupäätteitä?» Tällaiset kysymykset kuuluvat luonteenomaisesti tilanteeseen, jossa haetaan kompromissia (*trade-offs*) ihmisten ja pääomajoukkojen välillä. Tai voidaan kysyä »Onko edullisempaa ostaa luettelointipalveluja vai tehdä työ itse?». Näin kysyttäessä perusoletamus on, että vaihtoehtojen tulee tuottaa yhtä käyttökelpoisia luetteloita.

Jälkimmäinen kysymys sisältää jo kustannus-suorituskyky -kriteerin aineksia. Vielä selvemmin on siitä kyse, kun kysytään: »Voimmeko tehdä X:n halvemmalla tekemällä sen jollain toisella tavalla kuin nyt?» On päivänselvää, että vastaamiseen tarvitaan tietoa eri vaihtoehtoihin sisältyvistä kustannustekijöistä. Tämä saattaa käydä helposti päinsä selvitetessä käytössä olevaa järjestelmää, mutta on huomattavasti vaikeampaa arvioida jonkin hypoteettisen vaihtoehdon todellisia käyttökustannuksia.

## Tehokkuus

Minkä tahansa kirjasto-operaation tai palvelun tehokkuuden evalvointi merkitsee sen arvioimista, miten hyvin järjestelmä toimii suhteessa sille asetettuihin tavoitteisiin. Lancasterin (1977, s. 1) mukaan tehokkuutta on mitattava sen perusteella, miten hyvin palvelu tyydyttää käyttäjien sille asettamat vaatimukset; mutta tottahan toki myös muiden operaatioiden kuin julkisten palvelujen tehokkuus voidaan ottaa evalvoitavaksi.<sup>1</sup>

Myös tehokkuuden evalvoinnissa tavoitteiden tarkastelu on keskeisellä sijalla. Jos evalvotaan vaikkapa luettelointiosaston tehokkuutta on päätettävä millaista kriteeriä käytetään. Teknistä kriteeriä käytettäessä tarkasteltavana on aineiston käsittelyn adekvaatin suoritusasteen ylläpitäminen suhteessa käsittelemättömän aineiston (*backlog*) määrään. Toisenlaisesta kriteeristä on kyse, jos tarkastellaan miten hyvin syntyvä luettelo tyydyttää lukijoiden tarpeet. Selvästikin on

<sup>1</sup> NB. Ei Lancaster kiellä evalvoimasta operaatioita, mutta hän katsoo, että niidenkin tehokkuuden evalvointi on ulotettava käyttäjien vaatimuksiin asti. (suom. huom.)

helpompaa evalvoida tehokkuutta, jos kriteerit sallivat kvantitatiivisen mittaamisen; on helpompaa (ja halvempaa) laskea suoritteita ja käsittelemättömän aineiston määriä kuin suhteuttaa luettelon laatu käyttäjien tarpeisiin. Kvantitatiivisten tietojen (yleensä) helpon keräämisen voi sanoa estäneen kirjastonhoitajia yrittämästä päästä perille järjestelmiensä tehokkuudesta suhteessa käyttäjien tarpeisiin, mikä kuitenkin saattaa olla merkityksellisempää tietoa.

Kysymykseen »Voimmeko ylläpitää nykyistä pienemmin kustannuksin sitä tasoa, jolla järjestelmä nyt toimii?» on vastausta haettava kustannus-tehokkuus -evalvoinnin keinoin. Tässäkin tapauksessa olemassa olevan järjestelmän kustannukset voidaan määrittellä, joskaan se ei välttämättä ole helppoa. Sen sijaan on hyvinkin vaikeaa määrittellä vaihtoehtoisten menettelytapojen kustannuksia. Kaikilla organisaatioissa tehtävillä muutoksilla on ennalta arvaamattomia seurauksia; niistä yksi saattaa olla, että kustannukset nousevat odotettua korkeammiksi.

## Hyödyt

Kirjastojen hyötyjen tulee liittyä niihin yhteisöihin, joita ne palvelevat. Tässä on esillä kirjaston (tai kirjastopalvelun) *arvo* joko yhteisölle kokonaisuutena (kaupunki, kunta, koulutuslaitos, yritys) tai yksittäisille käyttäjille. Tähän kysymykseen kirjastonhoitajan on kaikkein vaikeinta löytää vastauksia, varsinkin jos tehtävään kytketään kysymys kustannuksista eli pyydetään kustannus-hyöty-analyysia.

Vaikka onkin vaikeaa liittää hyötyihin rahallisia arvoja, ei se käsitykseni mukaan riitä syyksi välttää itse hyötyjen evalvoinnin yrittämistä. Käytettäessä arvon käsitettä muissa asiayhteyksissä esillä ovat filosofiset, eettiset tai moraaliset arvot – eikä kukaan niihin yritä kytkeä rahallisia arvoja.

Sanon saman asian toisin. Väitän, että hyvin perusteltu hyötyjen kvalitatiivinen evalvointi voi olla hyvinkin vakuuttava, jos se tukee niitä arvoja (em. arvokategoriassa), jotka voimavarojen jakamisesta vastuussa ovat hyväksyvät. Niinpä arvelen, että esim. tutkimus yleisten kirjastojen hyödyistä työtömille tarjona olevina palveluina saisi Suomessa myönteellisemmän vastaanoton kuin rva Thatcherin Englannissa.

## Mitä tutkimusstrategioita käyttäen?

Jakamalla tutkimusmenetelmät kahteen kategoriaan, »kvantitatiivisiin» ja »kvalitatiivisiin» menetelmiin (ja ottaen huomioon, että jako on ylen karkea ja että termit kuvaavat metodijatkumon kahta ääripäätä) voidaan kehittää seuraava metodien typologia:

- kvantitatiiviset menetelmät
- tietojen keruu jatkuvana toimintana (*constant monitoring*)
- toimintoja tai kustannuksia koskevien tietojen keruu kertatutkimuksessa (*ad hoc surveys*)
- »kovie» tietojen keruu käyttäjiltä kertatutkimuksessa
- kvalitatiiviset menetelmät
- informaation ja mielipiteiden hankkiminen käyttäjiltä kertatutkimuksessa.

Yllä olevan typologian lähempi tarkastelu, herättää kysymyksen siitä, mitä eri termeillä tarkoitetaan, ja myös siitä, mitä taitoja tietojen keruussa tarvitaan.

## Monitoring

Tässä on kysymys tietojen keräämisestä sitä mukaan kuin asiat eteen tulevat. Kirjastot keräävät tilastotietoja yksinkertaisimmista toiminnoistaan. Esim. lainaustilasto voidaan koota helposti, varsinkin jos se tuetaan tietokoneita hyväksi käyttäen. Hankalampaa on kerätä tietoja tietopalvelukysymyksistä (*reference questions*) lähinnä siksi, että on vaikeaa täsmentää eri tyyppisille kysymyksille ja toimeksiannoille yleisesti hyväksytyjä luokituksia, joista käyttäen kerätäisiin vertailukelpoisia tietoja. Monessa kirjastossa pyritään kuitenkin pitämään lukua tiedustelujen määräästä. Toisinaan yritetään kerätä tietoja kirjastossa tapahtuvasta aineistojen käytöstä esim. laskemalla lukupöydille jätetyt kirjat, lehdet jne. Tiettyjä toimintoja tarkkaillaan keräämällä niistä tietoja tietysin aikaväleihin; esim. lukusalien käytöstä saadaan tietoja laskemalla sen käyttäjien määrä säännöllisin välein vuoden mittaan.

Tähän tietojen keruuseen tarvittavat taidot ovat jokseenkin alkeellista laskentoa, perustaidoksi riittää yleensä yhteenlasku. Yleisenä sääntönä voidaan pitää, että mitä kiintoisammasta toiminnasta on kysymys, sitä vaikeampaa on kerätä siitä tietoja vain laskemalla yhteen määriä.

## Surveys

Survey-tutkimuksia on yleisesti ottaen kahta lajia, nimittäin kysely- ja haastattelututkimukset. Aihe on aivan liian laaja käsiteltäväksi riittävän perusteellisesti tällaisessa suppeassa katsauksessa.<sup>1</sup>

Vaadittavat taidot ulottuvat otannan perusteiden hallitsemisesta, kyselylomakkeen laatimiseen ja eritoten kysymysten muotoilemiseen (mikä on helpompaa sanoa kuin tehdä), haastattelutaitoihin, koottujen tietojen käsitelyyn ja tulkintaan asti. Kun *monitoring*-keruuta voi tehdä jokainen tervejärkinen kirjastonhoitaja ja -virkailija, *survey* vaatii tämän menetelmän tuntemusta, ellei peräti alan tutkijan konsultointiapua.

## Toimintatutkimus

Toimintatutkimus (*action research*) on tutkimustapa, jolla on sijansa evalvointistrategioissa. Tämä tutkimusmenetelmä on kehitetty muutoksen aikaansaamisen välineeksi. Siinä edetään täsmentämällä ja paikantamalla ongelmia yhdessä tutkittavan organisaation toimijoiden kanssa, keräämällä tietoja toiminnan suunnittelun perustaksi, panemalla muutos alkuun, seuraamalla tuloksia ja tarvittaessa toistamalla tämä sykli kunnes muutos on saavuttanut tarkoitetun tason. Toimintatutkimusta voidaan pitää evalvoivana innovointina. Tietojen keruuseen voidaan käyttää sekä kvantitatiivisia että kvalitatiivisia menetelmiä; informaatiota ja mielipiteitä voidaan koota tietojen keruun kaikissa vaiheissa. Toimintatutkimus on paikallaan, kun tarkoituksena on *muuttaa* jotakin tilannetta eikä vain tutkia sitä.

## Johtopäätös

Johtopäätöksiä on vain yksi ja se on hyvin yksinkertainen: Vain kirjastonhoitaja tai informaattikko voi tietää, minkä evalvointia hän tarvitsee. Tavallisesti evalvointiin tartutaan, kun järjestelmässä esiintyy jokin ongelma tai kun kirjastonhoitajan/informaattikon on kyettävä perustelemaan henkilökunnasta, aineistoista tai palveluista koituvia kustannuksia tai kun hänen on saatava aikaan jokin muutos. Pelkästä akateemisesta uteliaisuudesta tehdyllä evalvoinnilla ei mielestäni ole mi-

tään annettavaa organisaatioille, jotka pyrkivät toteuttamaan arvokasta tehtävänsä sille yhteisölle, jota kukin palvelee.

Niinpä Sinun tulee päättää, *miksi* evalvointi suoritetaan ja *mitä* evalvoidaan. Miten evalvoida kirjaston palveluja tai operaatioita saattaa olla asia, jonka ratkaisemisessa asian-tuntija-apu on hyödyksi.

Evalvoinnin suorittamiseen ei ole tarjona vain yksi tapa tehdä se. Kaikki riippuu tavoitteistasi, ongelmista, jotka aiheutuvat niistä, ja siitä tiedosta, jota tarvitset päättääksesi järjestelmistä.

Saapunut 10. 10. 1984.

## Asiasta enemmän

- Cronin, B. (1982) 'Taking the measure of service'. *Aslib Proceedings*, 34, 273–294.
- De Prosopo, E.R., Altman, E. and Beasley, K.E. (1973) *Performance measures for public libraries*. Chicago: ALA.
- King, D.W. and Bryant, E.C. (1971) *The evaluation of information services and products*. Washington, DC: Information Resources Press.
- Lancaster, F.W. (1977) *The measurement and evaluation of library services*. Washington, DC: Information Resources Press.
- Lancaster, F.W. and Cleverdon, C.W. eds. (1977) *Evaluation and scientific management of library and information services*. Leyden: Noordhoff.
- Orr, R.H. (1973) 'Measuring the goodness of library services: a general framework for considering quantitative measurements'. *Journal of Documentation*, 29, 315–332.
- Wills, G. and Oldman, C. 'An examination of cost/benefit approaches to the evaluation of library and information services', *in*: Lancaster and Cleverdon (1977) 165–184.

<sup>1</sup> Survey-tutkimuksesta on tarjona hyvää kirjallisuutta, esim. M.B. Line, »Library surveys», 2nd ed., 1982. (suom. huom.)