

Asia: VN/5386/2020

Lausuntopyyntö ehdotuksesta tiedon hyödyntämisen ja avaamisen strategiseksi tavoitteiksi

Tiedon hyödyntämisen ja avaamisen strategiset tavoitteet

Ovatko strategiset tavoitteet kokonaisuutena realistisia ja ovatko ne vaikuttavia?

Open Knowledge Finland ry (OKFI) kiittää mahdollisuudesta lausua tiedon hyödyntämisen ja avaamisen strategisista tavoitteista. On erinomaista, että tiedon avaaminen on todettu tärkeäksi ja pyritään löytämään siihen laajasti vaikuttavia keinoja. Dokumentissa on tunnistettu tärkeitä ja kannatettavia tavoitteita.

Toisaalta tekstistä saa ajoittain vaikutelman (varsinkin tavoite 1.1), että avoimen tiedon tuottaminen nähdään keskusjohtoisena prosessina, jossa etukäteen määritellään tiedon tarvisijoita ja tarpeita. Kuitenkin avoimuus on syytä nähdä oletusarvona, jonka on toteuduttava automaattisesti kaikessa viranomaistoiminnassa. Tämä vaatii toimintakulttuurin muutoksen eikä onnistu lisäämällä uusia prosesseja vanhaan käytäntöön. Toki tätä muutosta tulee edistää myös keskitetysti esimerkiksi tulossopimuksilla.

Lisäksi näyttää siltä, että tavoitteisiin ei lainkaan liity konkreettisia ja määrällisiä mittareita, joilla voitaisiin todentaa tavoitteiden toteutumista. Tämä puute on syytä korjata, koska mittarit ryhdistävät selvästi toimintaa ja parantavat lopputulosta. Tavoitteeksi asetettu tiedon hyödyntämisen tiekartta on laadittava siten, että toimenpiteiden toteutumista ja strategisten tavoitteiden saavuttamista voidaan mitata relevanteilla määrällisillä mittareilla.

OKFI on omassa toiminnassaan noudattanut jo vuosia mahdollisimman suurta avoimuutta ja voi siksi omalta osaltaan vastata myöntävästi toimeenpanoon liittyviin avoimuuskysymyksiin. Niinpä lausunnossa keksitymme tämä kysymyksen sijaan niihin täsmennyksiin, joita tavoitepaperiin on syytä tehdä, jotta se onnistuisi saavuttamaan päätavoitteensa eli avoimemman tietokulttuurin Suomessa.

Yleistason kommenttina toteamme, että tämä strategialuonnos muistuttaa enemmän vision ja toimintasuunnitelman yhdistelmää. Se kertoo suurista tavoitteista sekä ylätasolla keinoista päästä näihin tavoitteisiin. Strategiassa tulisi kertoa, mitä tehdään vähemmän tai mitä jätetään tekemättä. Esimerkiksi: "Tietoaineistojen myyntiin perustuva liiketoiminta ajetaan alas." olisi selkeä strateginen tavoite, jota ei kuitenkaan ole mainittu.

Teema 1: Ohjaus, koordinaatio ja yhteistyö

Ovatko teeman 1 tavoitteet realistisia ja toimeenpantavissa organisaatiosi näkökulmasta?

Tämän teeman tärkein ja vaikein työ on kulttuurin muutos. Nykyinen toimintakulttuuri ei tosiasiallisesti tue avoimuutta, vaikka siitä jo paljon puhutaan. Väärästä avoimuudesta saa potkut mutta väärästä sulkeutuneisuudesta ei seuraa mitään. Edellisestä on esimerkkinä eduskunnan turvallisuusjohtaja, joka luovutti turvakameran videotallenteen medialle. Jälkimmäisestä puolestaan ovat esimerkkinä THL:n koronaepidemiamallit, joita OKFI on toistuvasti pyytänyt, mutta joita on vain vajavaisesti saatu - pääministerin ohjeen vastaisesti.

Yhdysvalloissa on laki, jonka mukaan liittovaltion verovaroin tuottama tieto on automaattisesti tekijänoikeuden ulkopuolella. Tämä on tuottanut kulttuurin, jossa monet tietoaineistot tuotetaan suoraan nettiin ilman, että organisaatioiden sisäinen tarveharkinta monimutkaistaisi käytäntöjä. Lisäksi valtion maksaman tutkimuksen julkaisujen yksinoikeuksia ei ole mahdollista luovuttaa tieteellisille lehdille, vaan artikkeleita saa julkaisemisen jälkeen jakaa vapaasti. Samansisältöinen laki pitäisi säätää Suomessakin, koska nykyinen epäterve lisensointi estää vakavasti tutkimustiedon levittämistä. Tutkijat pelkäävät julkaisemisen vaarantuvan, jos he itse lähtevät taistelemaan tiedejulkaisijoita vastaan.

Tavoitteen 1.1 keskusjohtoisuus voi auttaa ottamaan nopeasti ja laajasti käytäntöön avoimuuden toimintamalleja. Vastaavasti se voi estää ketterää alakohtaista toimintaa luomalla jäykkiä rakenteita, hidastamalla kehittämistä ja etäännyttämällä vastuuta.

Tavoite 1.2, alatavoite 3 mainitsee immateriaalioikeudet mahdollisena perusteltuna syynä estää tietoaineiston avaamisen. Nykyinen eurooppalainen tekijänoikeuslaki antaa tietokannoille automaattisen lähioikeuden (eli immateriaalisuojan). Niinpä tuon lauseen muotoilu heikentää koko alatavoitteen, koska kaiken tietoaineiston avaamisen esteeksi voi antaa perusteeksi immateriaalioikeuden. Lausetta voisi tarkentaa esim. muodolla "kolmansien osapuolten immateriaalioikeudet".

Tavoitteeseen 1.3 voisi lisätä alatavoitteen tyyliin "Suomessa käytetään kansainvälisiä standardeja tietoaineistojen formaattien, rajapintakuvausten sekä niiden lisensointiin liittyen". Ilman yhteisiä standardeja yhteistyö vaikeutuu nopeasti. Lisensseistä standardoituneita ovat jo CC0 metadatalle ja

CC BY 4.0 tietoaaineistoille. Näiden käyttöä pitäisi aktiivisesti laajentaa, koska nykytilanne on riittämätön. Erityisesti tietoaaineistojen vapaa yhdistettävyyden ja hyödynnettävyys kansainvälisissä tietoportaalissa toteutuu nimenomaan selkeästi tarjoamalla ne käyttöön tekijänoikeudettomana eli public domainina (tai CC0-luopumalausekkeella). Minkä tahansa muiden lisenssien käyttäminen Suomessa käytännössä tekee kansainvälisten tietoportaalien yhteiskäytön mahdottomaksi. Nämä tulisi ehdottomasti kirjata myöhempiin ohjaaviin dokumentteihin.

Teema 2: Strategia ja toiminta

Ovatko teeman 2 tavoitteet realistisia ja toimeenpantavissa organisaatiosi näkökulmasta?

Tietoperusteinen päätöksenteko on kannatettava tavoite. Se on erityisen kehittymätön alue avoimen tiedon näkökulmasta: vain harvoin päätöksenteon tietopohja, käytetyt mallit, arvoperusteet ja optimointikriteerit päätyvät avoimeksi tiedoksi. Niinpä olisi erityisesti satsattava siihen, että päätöksenteon dataa ja vaikutusarviointeja avataan myös. Tämä edellyttää, että tällaista tietoa systemaattisesti myös kerätään.

Tavoitteen 2.1 mukaisesti tiedontarpeet ja kansalaisten tasapuolinen palvelu tulee ottaa huomioon vahvemmin. Tietoa on tuotettava ja tarjottava laajasti Suomessa puhutuilla kielillä. Suomen vähemmistökielten digitaaliset palvelut on turvattava sekä teknisiä infrastruktuureja että sisällöntuotantoa tukemalla. Kansainväliset avoimen tiedon projektit, kuten Wikipedia tai Wikidata tarjoavat mahdollisuuksia maailman eri kielillä palvelemiseen

Tiedon laadulle (2.5, myös 3.3) voidaan määritellä teknisiä kriteereitä esimerkiksi rajapintojen, metadatan ja ajantasaisuuden osalta. Kuitenkin on huomattavasti vaikeampaa määritellä kriteereitä sille, onko tieto totta vai ei. Itse asiassa paras tunnettu keino tähän on tieteen menetelmä: tieto on tuotettava varmennetuilla menetelmillä, sen on perustuttava havaintoihin, ja on varmistettava vertaisarviointi eli taattava mahdollisuus ulkopuoliseen kritiikkiin tiedon heikkouksien löytämiseksi. Tämä ei onnistu minkään tarkistuslistan avulla vaan edellyttää koko prosessin läpinäkyvyyttä ja laadunvarmistuksen joukkoistamista. Ja tämä puolestaan edellyttää isoa kulttuurin muutosta.

Tietoa avattaessa mahdolliset tiedon virheet tai puutteet tulevat näkyvämmiksi. Tämä on kynnyksen avaamiselle, mutta toisaalta avaamisen yleistymisen nostaa omia laatutavoitteita, jotta tietoaaineiston voi luottavaisin mielin avata. Tästä huolimatta virheitä tulee tapahtumaan, joten kulttuuriin on omaksuttava tiedon päivitys, virheiden korjaaminen ja anteeksi pyytäminen. Tämän kaiken taustalla on tärkeää, että itse avatun datan taustat, eli keruumenetelmät, analyysimenetelmät, yms on dokumentoitu, versioitu ja arkistoitu.

Onneksi tällaisia käytäntöjä on kehitetty, ja niitä voidaan valtionhallinnossakin ottaa käyttöön. Eräs keino on jäsentää tietoa tietokiteiksi, joissa täsmällisten kysymysten vastaamiseen etsitään

joukkoistamalla parhaat vastaukset. Nämä vastaukset sitten jaetaan avoimena datana kaikkien käyttöön.

Tietokiteet ovat myös yksi ratkaisu tavoitteeseen 2.5 disinformaation leviämisen hillitsemiseen. Disinformaation vaikutus vähenee, kun kaikki tietävät, mistä ajankohtaisin ja vertaisarvioitu tieto jostakin aiheesta löytyy. Mutta silloin tietoa pitää tuottaa juuri niistä aiheista, joista sitä tarvitaan, vaikka tarjolla ei olisikaan suoraa mittausdataa. Tällöin on tukeuduttava asiantuntijoiden arvioihin, jotka altistetaan kritiikille.

Teema 3: Tiedon hallinta

Ovatko teeman 3 tavoitteet realistisia ja toimeenpantavissa organisaatiosi näkökulmasta?

Kun organisaatiot veloitetaan tuottamaan tietoa tietovarantoihin, on syytä mitata ja tutkia, mitä tietoa tuotetaan ja mille tiedolle on kysyntää. Mahdollisia sovelluksia käyttäjät kyllä rakentavat orgaanisesti avoimen tiedon päälle, joten käyttämisen organisoinnista ei tarvitse olla niin huolissaan. Näitä tuloksia on aktiivisesti hyödynnettävä tietotyön kehittämiseen. Tämä on tärkeämpää kuin se, että varanto on keskitetty; myös hajautetut järjestelmät toimivat hyvin, kunhan rajapinnat ovat avoimia ja tieto pääsee virtaamaan.

Kansallisten palvelujen ja formaattien yhteentoimivuus kansainvälisten palvelujen ja standardien kanssa tulee varmistaa ja sitä tulee edistää. Esimerkiksi geodataa tulee tarjota myös formaateissa, jotka eivät vaadi konversioita, jotta niitä voi käyttää yleisesti saatavilla olevissa palveluissa.

Immateriaalioikeudet voivat muodostua merkittäväksi rajoitteeksi jos suunnitelman osana ei myös yhtenäistetä käytäntöjä sen suhteen, millaisilla lisensseillä toimitaan. Työntekijän immateriaalioikeuksia on julkishallinnossa joskus käytetty perusteena tiedon käytön rajoituksiin. Kuitenkin työntekijän näkökulmasta on tärkeämpää saada työstä sen ansaitsema arvostus kuin immateriaalioikeuksien tuottama teoreettinen mahdollisuus kaupalliseen hyödyntämiseen.

Tavoitteen 3.4 mukaisten ohjelmointirajapintojen lisäksi on tuettava toimijoita, joille näiden palvelujen tuottaminen ei ole käytännön syistä mahdollista. Keskittyminen vain korkeasti teknologisiin ratkaisuihin ohittaa heikommin organisoituneiden ja resursoitujen toimijoiden, kuten yhdistysten tai pienten museoiden tiedon.

Tietoaineistoja käsittelevät analyysilähdekoodit on luokiteltava dataksi ja avattavan tiedon piiriin. Nyt käytännöt ovat sulkeutuneita, koodeja ei ajatella julkisina asiakirjoina, eivätkä harvat ennakkotapaukset ole tilannetta ainakaan korjanneet.

Teema 4: Mahdollistajat

Ovatko teeman 4 tavoitteet realistisia ja toimeenpantavissa organisaatiosi näkökulmasta?

Tavoitteen 4.3 mukaista oppimista on tehtävä huomattavasti kunnianhimoisemmin kuin mitä tavoite antaa ymmärtää. Datalukutaitoa, vastuullista datakansalaisuutta ja luovaa teknologiaosaamista on vahvistettava kaikilla tasoilla. Se on yhtälailla ammatissa toimivien ketterää täydennys- ja uudelleen koulututtamista, kansalaistaitojen kartuttamista, opiskelumahdollisuuksia, sekä koululaisten ja päiväkotilasten tekniikkaa luovasti hyödyntävää oppimista.

Yhteisten tietöalustojen (4.4) kehittäminen on kannatettavaa. Strategiaan tulee lisätä maininta, että järjestelmien itsensä pitää olla avointa koodia. Suljetussa koodissa on lisenssimaksuja ja käyttörajoituksia eikä avoimen tiedon hyödyntäminen onnistu tavoitellulla tavalla. Lisäksi se aiheuttaa toimittajaloukun, josta usein ei edes ymmärretä pyrkiä eroon. Avoin koodi on muutenkin järkevä ratkaisu: edistyneimmät tietojärjestelmät ovat avointa koodia, koska suljetut järjestelmät eivät pysty houkuttelemaan riittävästi parhaita kehittäjiä.

Alustojen lisäksi (4.4) on syytä kehittää ja tuottaa synteettistä väestödataa eli dataa joka koskee kuvitteellisia henkilöitä ja joka tilastollisesti muistuttaa todellista väestöä. Synteettinen data mahdollistaa hyvin yksilölliset päätösanalyysit ja mikrosimulaatiot ilman tietosuojaongelmia (3.5). Sen avulla päätöksenteon perusteet ja käytetyt parametrit voidaan kuvata ja arvioida paljon nykyistä tarkemmin. Myös muulle synteettiselle datalle löytyy varmasti käyttökohteita, koska se mahdollistaa standardoitujen rajapintojen ja tietorakenteiden käytön. Tilastoyhteenvedot eivät tähän yhtä hyvin pysty, ja raakadatan julkaisemisessa on usein tietosuojaan liittyviä esteitä.

Toikkanen Tarmo
Open Knowledge Finland