



10.07.2020

JulkICT
Antti Helin

Yhteenveto julkisen sektorin tietojen uudelleenkäytön nykytilaa koskevasta kyselystä

1 Johdanto

Avoimesta datasta ja julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä annetun direktiivin (EU) 2019/1024 täytäntöönpanon valmistelun yhteydessä laadittiin internet-kysely, jonka avulla haluttiin kartoittaa julkisen sektorin tietojen uudelleenkäytön nykytilaa ja epämuodollisesti kerätä näkemyksiä liittyen direktiivissä käytettäviin käsitteisiin ja direktiivin edellyttämiin muutoksiin sekä niiden sovittamiseen kansalliseen lainsäädäntöön.

Kysely lähetettiin 3. kesäkuuta 2020 kaikille ministeriöille niiden työryhmissä olevien edustajien välityksellä. Lisäksi kysely lähetettiin erikseen 8. kesäkuuta 2020 opetus- ja kulttuuriministeriön edustajan suosituksesta korkeakoulujen rehtorineuvostoille Unifin ja Arenen kautta sekä CSC:lle. Kyselyyn oli aikaa vastata 17. kesäkuuta 2020 mennessä, mutta kysely suljettiin lopulta vasta 8. heinäkuuta 2020.

Vastaukset pyydettiin koordinoitusti joko koko hallinnonalalta ao. ministeriön kautta yhteenvedettynä tai organisaatiokohtaisesti yksi vastaus per organisaatio.

Julkisia yrityksiä ei lähestytty kyselyllä, koska kyselylle varatussa aikataulussa ei ollut mahdollista selvittää sopivaa jakelua, eikä kyselyn lähettäminen kaikille julkisille yrityksille ollut realistista.

2 Vastaukset ja vastaajat

Kyselyyn saatiin 29 vastausta. Vastaajat edustivat seuraavia organisaatioita:

- Kuopion kaupunki
- Suomen ympäristökeskus (2 vastausta)
- Hämeenlinnan kaupunki
- Valtioneuvoston kanslia, omistajaohjausosasto
- Ilmatieteen laitos
- Taiteen edistämiskeskus (2 vastausta)
- Tilastokeskus
- Puolustusministeriö
- Museovirasto

- Energiavirasto
- Maanmittauslaitos
- Turvallisuus- ja kemikaalivirasto Tukes
- Kuusamon kaupunki
- Kuntaliitto
- Ympäristöministeriö
- Finnvera
- Celia - näkövammaisten kirjasto
- Sisäministeriö
- Traffic Management Finland Oy
- CSC
- Kilpailu- ja kuluttajavirasto
- Liikenne- ja viestintävirasto Traficom
- Helsingin kaupunki
- Työ- ja elinkeinoministeriö
- Väylävirasto
- Business Finland
- Maa- ja metsätalousministeriö

Ottaen huomioon aihepiirin laajuuden, ei tätä voida pitää edustavana otantana Suomen julkisen sektorin toimijoista, joilla on avattua tai avattavaa tietoa hallussaan. Tiukan aikataulun sekä rajallisten valmistelumahdollisuuksien vuoksi tätä ei odotettukaan.

Esimerkiksi kaikilta ministeriöiltä ei saatu vastauksia ja tutkimustiedon osalta selkeä puute on se, että korkeakouluilta ei saatu lainkaan vastauksia.

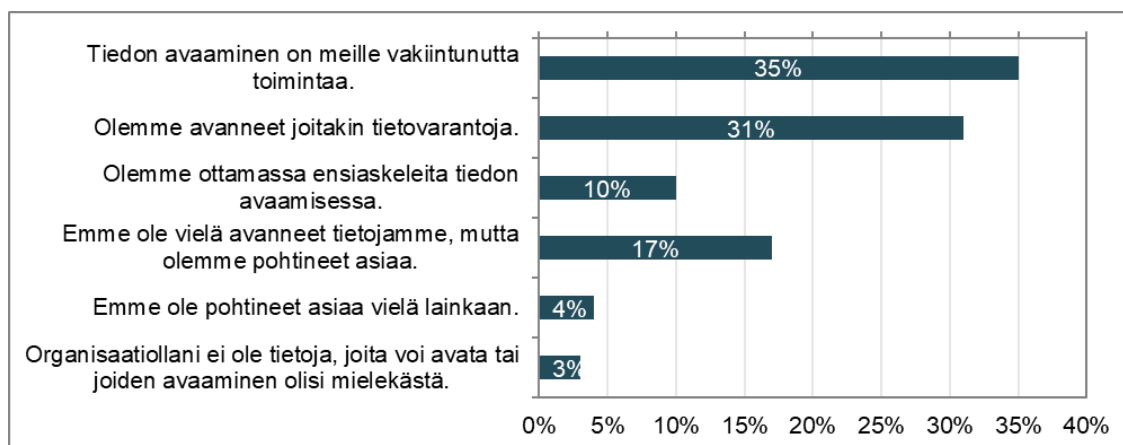
Tästä syystä tarkempia tuloksia on käsitelty muistiossa varsin yleisluontoisesti ja niitä on pidettävä korkeintaan suuntaa-antavina. Kysely ei korvaa vaikutustenarvioinnissa tehtäviä tarkempia haastatteluita tai hallituksen esityksestä järjestettävää julkista kuulemistä, tai lausuntokierrosta.

3 Organisaation tietojen avaamisen nykytila

Vastaajien organisaatioista 22 (76 %) on jo avannut tietovarantojaan, kun taas 7 ei ole.

Näistä 22 organisaatiosta 13 (59 %) on vienyt ainakin osan tiedoistaan avoindata.fi-palveluun, 6 ei ole vienyt, ja 3 ei osaa sanoa.

Kun vastaajia pyydettiin kuvailemaan omaa organisaatiotaan suhteessa tiedon avaamiseen, vastaukset olivat seuraavanlaiset:



4 Julkisen sektorin tietojen uudelleenkäytön esteet ja uudelleenkäytön edistäminen

4.1 Lainsäädäntöön liittyvät esteet

Vastaajista 15 (52 %) tunnisti joitakin lainsäädännössä olevia esteitä julkisen sektorin tietojen uudelleenkäytölle. Niitä pyydettiin tarkentamaan avovastauksentässä.

Esteiksi nähtiin:

- liian tiukat tai liian laajat laissa olevat salassapitoperusteet
- EU:n tietosuojasetus, jota pidettiin liian tiukkana tai jonka tulkintaa pidettiin liian tiukkana
- epäselvyys siitä, milloin tieto on henkilötietoa tai sisältää henkilötietoa, esimerkiksi paikkatiedot
- tekijänoikeudet, varsinkin tutkimustiedon¹ osalta
- toisiolaki (552/2019) sallii tietojen käytön tutkimukseen, mutta rajoittaa julkista ja avointa jakamista

Myös ylipäätään julkisuuslakia ja tietosuojasetusta sellaisenaan pidettiin esteenä. Tässä on huomioitava se, että vastaaja ei välttämättä ole pitänyt sitä kielteisenä, että esimerkiksi henkilötietojen laaja avaaminen ei lainsäädännön mukaan ole mahdollista.

Yksi vastaaja piti tietojen saatavuutta ja varsinkin henkilötietoja koskevan sääntelyn hajautumista eri lakeihin ongelmallisena. Tällöin vaikka tietojen saaminen tai avaaminen olisikin mahdollista, on sen selvittäminen kallista ja aikaa vievää: ”Toisin sanoen, käytännön toiminta on järjestetty siten, että käyttöön saaminen kestää pitkään, vaikka lainmukainen käyttötarkoitus ja tarve aineistojen käytölle olisi olemassa.”

Tutkimustiedon kohdalla kyse on myös julkisuuslain ja korkeakouluja koskevien erityislakien välisestä suhteesta. ”Tutkimustieto, joka on asetettu saataville maksullisen tietovarannon kautta, pitäisi olla myös saatavilla ilmaiseksi julkisuuslain nojalla sähköpostitse niin kauan kuin kyse on julkisesta asiakirjasta, joka on tallennettu sähköisesti. Yliopistolaisissa tai ammattikorkeakouluissa ei ole kuitenkaan otettu huomioon tilannetta, jossa sinänsä julkinen tutkimustieto on asetettu ilmaisen tai maksullisen tietovarannon kautta saataville.”

¹ Tässä kohtaa kyselyä direktiivin tutkimustiedon määritelmää ei ole vielä käsitelty, joten vastaaja on luultavasti tarkoittanut tieteellisiä julkaisuja.

Nykyisen lainsäädännön tulkitsemisessa 41 prosenttia (n=12) koki olevan ongelmia tai väärinkäsityksiä, jotka haittaavat julkisen sektorin hallussa olevien tietojen uudelleenkäyttöä. Kaikki vastaukset eivät kuitenkaan koskeneet nimenomaan lainsäädännön tulkitsemista vaan yleensä ongelmia tiedon uudelleenkäytössä.

Vastauksissa mainittuja ongelmia ja väärinkäsityksiä ovat:

- lainsäädännön erilainen tulkinta eri organisaatioissa
- tietojen yhdisteleminen, jota kautta voi paljastua salassa pidettäviä asioita
- tietosuoja-asetuksen tai muun lainsäädännön liian tiukka tulkitseminen varmuuden vuoksi, ettei erehdyksessä luovuteta tietoa jota ei pitäisi luovuttaa
- tietosuoja-asetuksen ja kansallisen lainsäädännön tulkitseminen yhdessä, koska käsitteet eivät ole yhdenmukaisia
- epäselvää, mikä lähdekoodin asema on julkisuuslaissa
- tekijänoikeudet; esimerkiksi viranomaiselle saapuvan asiakirjan liitteenä voi olla myös asiakirjoja, jotka ovat tekijänoikeuden alaisia
- se, että yleensä tiedosta peritään maksua.

4.2 Viranomaisille asetettavat uudet velvollisuudet

Avoimella kysymyksellä kysyttiin, mitä uusia velvollisuuksia viranomaisille pitäisi asettaa tiedon uudelleenkäytön edistämiseksi. Vastauksen sai jättää tyhjäksi, jos ei kokenut, että uudet velvollisuudet ovat tarpeen. 13 vastaajaa jätti kokonaan vastaamatta, ja lisäksi kaksi vastaajaa totesi itse vastauksessaan, että uudet velvollisuudet eivät ole tarpeen ja että asiaa olisi parempi lähestyä tuen ja kannustamisen ennemminkin kuin velvoittamisen kautta. Noin puolet vastaajista ei siis pitänyt uusia velvollisuuksia lainkaan tarpeellisina.

Uudenlaisia velvollisuuksia voisivat olla:

- Tiedon hyödyntäminen ja avaaminen (TiHA) -hankkeessa tuotettavien laatu-kriteerien noudattaminen
- Tietojen toimittaminen sähköisessä muodossa keskitettyihin kansallisiin tietovarantoihin
- Asiakaslähtöisen tiedon tuottaminen viranomaiskeskeisyyden sijaan
- ”Tietotilinpäätös”, jossa on säännöllisesti lueteltava suljetut aineistot ja perusteltava, miksi niitä ei ole avattu
- CC 4.0 BY –lisenssin käyttäminen
- Metatietojen käytön ja muotojen vakioiminen, avatun datan tarkka kuvaileminen
- Avointen rajapintojen käyttäminen, mikä kuitenkin edellyttää selkeitä pelisääntöjä ja sitoutumista jatkuvuuden takaamiseksi
- Tilastotietojen koostamisen kehittäminen niin että niitä ei tarvitse luoda käsin, niin kuin paikoin edelleen tehdään.

4.3 Avoimen tiedon käyttäjien oikeuksien vahvistaminen

Vastakohtana aiemmalle kysymykselle kysyttiin, miten tiedon käyttäjien oikeuksia tarvitsisi vahvistaa tai lisätä. Vastaajista 18 jätti vastaamatta tai totesi, ettei oikeuksia tarvitse vahvistaa. Vastauksissa tämän ja edellisen kysymyksen kohdalla oli myös jonkin verran päällekkäisyyttä, eikä päällekkäisiä vastauksia ole lueteltu uudelleen.

Uusia tai vahvistettavia oikeuksia voisivat olla:

- kaiken julkisella rahalla tuotetun tiedon, jonka käyttöä ei ole rajoitettu, tekeminen avoimeksi

- rajapintojen pitäisi olla selvästi dokumentoituja, yhdessä käyttäjien kanssa sovittuja ja niiden olisi oltava pysyvästi käytettävissä
- hyödyntäjällä pitäisi olla oikeus riittävän kattaviin metatietoihin, myös kone-luettavassa mudoossa
- Avoimen tiedon palvelulupaus (SLA); käyttäjän pitäisi voida luottaa siihen, että hänellä on täysi oikeus käyttää tietoa toivomallaan tavalla, ja että tieto on ajantasaista, sekä käytettävissä pysyvästi avoimesti ja maksutta,
- OmaData-periaate pitäisi saada käyttöön käytännössä.

4.4 Olemassa olevat käytännön järjestelyt tiedon uudelleenkäytön helpottamiseksi ja uudelleen-käytettävien asiakirjojen hakemisen helpottamiseksi

Vastaajat kertoivat seuraavista käytännön järjestelyistä:

- Kaupunginarkiston tietopalvelu
- Avoin Hämeenlinna -julistus
- Avoindata.fi (mainittiin useaan kertaan)
- Paikkatietoikkuna
- AWS Public Dataset -ohjelma
- Pirkanmaan ELY-keskuksen tarjoama ympäristöneuvonta (ympäristötieto-direktiivin toimeenpano)
- Suomi.fi -palveluväylä
- TEM:n Asiakastietovaranto (julkiseksi luokitellun datan osalta)
- Celia on rakentanut datamart-järjestelmän, joka kerää valituista tietoläh-teistä dataa keskitettyyn tietokantaan. Datasta voidaan näin jalostaa tarvit-tava informaatiopaketti.
- Tiedonhallintalaki ja sen velvoittamat yhteentoimivuus sekä tiedonhallinta-kartat
- Kirjastojen, arkistojen ja museoiden aineistopalvelu Finna.fi
- Muut erilaisten organisaatioiden omat portaalit ja aineistot, kuten Tilasto-keskus, SYKE ja MML.

Käytännön järjestelyt ovat siis lähinnä erilaisia portaaleja ja verkkosivuja, jo-den kautta avoimet tietoaineistot sekä niiden metatiedot ovat saatavilla. Näistä järjestelyistä ovat vastuussa portaalin tai verkkosivun ylläpitäjä, tai avattujen tietojen omistaja.

Muina järjestelyinä mainittiin Avoin Hämeenlinna –julistus, Celian datamart-järjestelmä, ja tiedonhallintalaki.

5 Direktiivin täytäntöönpanoon liittyvät tarkentavat kysymykset

5.1 Hyödynnetäänkö Suomessa standardilisenssejä ja pitäisikö niiden käytöstä säätää laissa?

12 vastaajan (41 %) mukaan Suomessa käytetään standardilisenssiä. Vastaa-jat nimesivät rauenneen JHS-suosituksen mukaisen lisenssin Creative Com-mons 4.0 BY tai Creative Commons 0, eivät muita lisenssejä.

Suosituksen taso nähtiin hyvänä eikä yksikään vastaaja pitänyt lisenssin käy-tön säätämistä lakiin tarpeellisenä.

5.2 Esimerkkejä organisaatio- ja tutkimusalakohtaisista tietovarastoista

Esimerkkejä tällaisista tietovarastoista ovat:

- digitraffic.fi

- digitransit.fi
- tiedejatutkimus.fi
- etsin.fi
- metax.fi
- Helsingin kaupungin keskitetty tietovarasto HETA ja HRI-palvelu
- Syken ja Luken portaalit.

5.3 Tutkimustiedon saatavuutta tai maksullisuutta koskeva, jo olemassa oleva sääntely

21 vastaajaa (78 %) ei tunnistanut olemassa olevaa sääntelyä, joka koskisi tutkimustiedon saatavuutta tai maksullisuutta. Muut vastaajat nimesivät julkisuuslain, valtion maksuperustelain ja toisiolain sekä ”sektorikohtaiset erityislait” niitä kuitenkaan tarkemmin nimeämättä.

5.4 Kansalliset avoimen saatavuuden toimintapolitiikat

Tutkimustiedon avoimen saatavuuden osalta vastaajat tunnistivat seuraavat toimintapolitiikat:

- tiedonhallintalaki ja siitä tulevat suositukset
- laki paikkatietoinfrastruktuurista
- hallituksen tietopoliittinen selonteko
- tutkimuslaitosten omat linjaukset, kuten Syken tutkimusdatapolitiikka
- Avoimen tieteen julistus 2020-2025, Avoimen tieteen linjaukset (avoim-tiede.fi)
- valtioneuvoston periaatepäätös aiheesta vuodelta 2010
- ministeriöiden asetukset eri laitosten ja viranomaisten maksullisista suoritteista
- LVM:n hallinnonalan tietokartta 2017 ja tarkastuskertomus 12/2019.

Valmisteilla ovat Tiedon hyödyntämisen ja avaamisen hanke (TIHA) sekä kuntasektorin omia datastrategioita.

Lopuksi kysyttiin vielä, mitä avoimen saatavuuden toimintapolitiikkoja pitäisi ottaa käyttöön:

- toimintapolitiikka, jolla varmistetaan, että erityisesti uudet aineistot tuotetaan avoimesti saataville heti alusta alkaen.
- DORA-julistus, johon myös Suomen Akatemia on sitoutunut.
- Kaikki toimialat kattava yleispolitiikka ja sen selkeä käytännön tulkinta, jota tukisi lainsäädännöllinen yhteentoimivuus eri toimialojen välillä, erityisesti tutkimustiedon, henkilötiedon ja arkaluonteisen tiedon tutkimuskäytön osalta.

Yhden vastaajan mukaan ”aluksi lienee syytä tehdä avoimen dataa edellyttävät linjaukset ja politiikka, jotka ovat edellytyksiä avoimen saatavuuden toimintapolitiikan muodostamiselle”.

5.5 Dynaaminen data

Vastaajista 18 (62 %) piti direktiivin dynaamisen datan määritelmää riittävän selkeänä ja yksiselitteisenä. Määritelmän parantamisen osalta todettiin, että päivittymisen tiheyteen olisi saatava jokin tarkempi rajausta, kuinka tiheästi tietoa on päivitettävä, jotta kyse olisi dynaamisesta datasta. Lisäksi toivottiin dynaamisen datan tarkempaa määrittelemistä tapauskohtaisesti ja havainnollisten esimerkkien käyttämisestä. Asiakirja-käsitettä pidettiin hämäävänä, ja että parempaa olisi puhua tietoaineistoista eikä asiakirjoista.

Vastaajista 11 (38 %) totesi, että heidän organisaatiollaan on hallussa dynaamista dataa. 10 vastasi kieltävästi ja 8 vastasi, ettei osaa sanoa.

Niistä 11:stä, jotka vastasivat, että heidän organisaatiollaan on hallussa dynaamista dataa, 9 vastasi myös, että heidän organisaationsa on asettanut dynaamista dataa reaaliaikaisesti saataville verkkopalvelun tai rajapinnan kautta.

Vastaajilta kysyttiin myös dynaamisen datan saataville asettamisesta sääntelemistä yksityiskohtaisemmin. Koska suunnitellun lakimuutoksen mukaan dynaaminen data on asetettava saataville rajapinnan kautta, pyynnöstä, julkisuuslaissa määrätyn aikarajan puitteissa, kysyttiin vastaajilta, tulisiko julkisen sektorin elimiä ja julkisia yrityksiä lainsäädännöllä velvoittaa asettamaan dynaaminen data saataville ilman pyyntöä, ja pitäisikö niitä velvoittaa erikseen varautumaan erilaisin teknisin järjestelyin siihen, että tiedot voidaan toimittaa määräajassa rajapinnan kautta.

Vastaajista 6 oli sitä mieltä, että julkisen sektorin elimet ja julkiset yritykset pitäisi velvoittaa avaamaan dynaaminen data rajapinnan kautta aina heti keräämisen jälkeen, ilman erillistä pyyntöä.

Vastaajista 7 oli sitä mieltä, että ne pitäisi velvoittaa toteuttamaan tarvittavat rajapinnat ja muut tekniset ratkaisut ennakoivasti, jotta dynaaminen data voidaan luovuttaa heti tai viimeistään kuukauden kuluessa, ja säännöllisesti sen jälkeen.

Vastaajista 3 oli sitä mieltä, että dynaamisen datan avaaminen voi tapahtua vasta tiedon luovutusta koskevan pyynnön seurauksena, eikä ole tarpeen velvoittaa julkisen sektorin elimiä ja julkisia yrityksiä varautumaan pyyntöihin esim. toteuttamalla teknisiä ratkaisuja.

13 vastaajaa ilmoitti, ettei osaa vastata tai ettei vastaajalla ole mielipidettä asi-
aan.

Lopuksi dynaamisesta datasta sai antaa vapaasti mahdollisia huomioita.

Dynaaminen data olisi tärkeää avata ennakoivasti ilman pyyntöä, koska yritykset eivät välttämättä pyydä tai osaa pyytää tiettyä tietoaineistoa saataville. Ilmatieteen laitoksen tuottaman avoimen datan vaikutusarvion kyselyssä 54 yrityksen edustajaa ilmoitti IL:n datojen tuottaneen heille uutta liiketoimintaa viimeisen 3 vuoden aikana. Nämä samat yritykset kuitenkin eivät tyypillisesti kysy niiden aineistojen perään, joita laitos ei ole avannut.

Dynaamisen datan tarjoamisessa olisi myös huomioitava julkisten yritysten toimintakyky sekä dynaamisen datan tarjoamisesta aiheutuvat kustannukset.

Lisäksi nostettiin esille se, että on selvennettävä, kohdistuuko velvoite avata tieto rajapinnan kautta palveluntuottajana toimivaan julkiseen yritykseen, jos tiedon varsinainen omistaja onkin jokin toinen taho, esim. viranomaisen.

Datan huomautettiin vaativan yleensä käsittelyä, jotta se on käyttökelpoista. Tässä mielessä vaatimus reaaliaikaisesta saatavuudesta nähtiin erikoisena. Jos kyse onkin ns. raakadatasta, tiedon tuottajia pitäisi velvoittaa ja ohjata ennakolta ilmoittamaan ja kuvaamaan, missä aikataulussa kerättävä data on saatavilla. Tästä ei kuitenkaan saisi syntyä merkittäviä lisäkustannuksia, tai sitten nämä kustannukset olisi korvattava jotenkin tiedon hyödyntäjien toimesta.

5.6 Muita yleisiä huomioita direktiivin täytäntöönpanoon liittyen

Useampi vastaaja korosti julkisomisteisten yritysten merkitystä ja halusi ne myös julkisuuslain soveltamisalan piiriin.

Virastojen, varsinkin pienten virastojen, resursseista pitäisi pitää huolta, kun esitetään niille uudenlaisia velvoitteita, jotka vaativat esimerkiksi merkittäviä muutoksia tietojärjestelmiin. Useat vastaajat kiinnittävät riittävään resursointiin huomiota. Lisäksi pitäisi panostaa yhteisiin ja keskitettyihin toimenpiteisiin täytäntöönpanon osalta, ja täytäntöönpanon aikataulun pitäisi olla järkevä.

Mahdollinen päällekkäinen lainsäädäntö kuten INSPIRE-direktiivi ja tiedonhallintalaki tulisi ottaa huomioon.

Lainsäädännöllisten esteiden lisäksi datan avaamisen esteenä on ennen kaikkea se, että haluttu aineisto ei välttämättä ole vielä digitaalisessa muodossa. Avaaminen millä tahansa tavalla vaatisi aluksi panostusta digitoimiseen. Samalla olisi kuitenkin myös selvitettävä tekijänoikeudelliset kysymykset.

Uudelleenkäyttö ei saisi muodostua kilpailevaksi käyttötarkoitukseksi alkuperäiseen tarkoitukseen nähden. Kysymyksessä pitäisi olla aito uudelleenkäyttö.

Tiedon laatuun, yhteisiin toimintatapoihin ja standardien käyttämiseen tulee kiinnittää huomiota. Tietoturvallisuus on myös varmistettava.

6 Yhteenveto

Kyselyssä ei noussut esiin erityisiä tarpeita uudelle lainsäädännölle direktiivin täytäntöönpanemiseksi, joita ei olisi jo käsitelty valmistelun yhteydessä. Tiedon hyödyntämisen ja avaamisen toimet ovat, kuten asiantuntijatyöryhmän väliraportissakin on todettu, lähinnä sellaisia, jotka eivät vaadi uutta sääntelyä. Erityisesti valtiovarainministeriön TiHA-hankkeen merkitys direktiivin täytäntöönpanossa korostui vastauksissa ja muutenkin tiedon avaamisen edistämässä.

Vastaajat pääsääntöisesti ymmärsivät direktiivin sisällön kyselyssä annettujen ohjeiden ja tiivistysten pohjalta, mutta kysely osoitti, että on oltava huolellisia, kun viestitään siitä, mitä direktiivin tavoitteet ovat ja mitä sen minimisääntelyn täytäntöönpano kansallisessa asiayhteydessä itse asiassa tarkoittaa ja mikä hallituksen esityksen suhde muun muassa julkisiin yrityksiin ja tutkimustietoon on.

Lisäksi kyselyn avulla on saatu selville useita sektorikohtaisia esimerkkejä erilaisista käytännön järjestelyistä ja avoimuutta edistävästä käytössä olevista toimista, joita voidaan hyödyntää komission notifikaation yhteydessä ja joilla voidaan rikastaa myös varsinaista hallituksen esitystä.