



EUROOPAN KOMISSIO

Bryssel 12.12.2011
KOM(2011) 882 lopullinen

**KOMISSION TIEDONANTO EUROOPAN PARLAMENTILLE, NEUVOSTOLLE,
EUROOPAN TALOUS- JA SOSIAALIKOMITEALLE SEKÄ ALUEIDEN
KOMITEALLE**

**Avoim data
Innovoinnin, kasvun ja läpinäkyvän hallinnon moottori**

KOMISSION TIEDONANTO EUROOPAN PARLAMENTILLE, NEUVOSTOLLE, EUROOPAN TALOUS- JA SOSIAALIKOMITEALLE SEKÄ ALUEIDEN KOMITEALLE

Avoim data Innovoinnin, kasvun ja läpinäkyvän hallinnon moottori

1. JOHDANTO

EU 2020 -strategian keskeisenä päämääränä on viedä Euroopan taloudet korkealle ja vakaalle kasvu-uralle. Tätä varten Euroopan on vahvistettava innovointimahdollisuuksiaan ja hyödynnettävä resurssinsa mahdollisimman hyvin.

Yksi tällainen resurssi on julkinen data¹, eli kaikki se tieto, jonka julkiset elimet Euroopan unionissa ovat tuottaneet, keränneet tai hankkineet. Esimerkkejä tästä ovat paikkatieto, tilastotieto, säätiedot, julkisrahoitteisissa tutkimushankkeissa tuotettu tieto ja kirjastojen digitoidut kirjat. Tällainen tieto tarjoaa merkittäviä – toistaiseksi hyödyntämättömiä – mahdollisuuksia uudelleenkäyttöön uusissa tuotteissa ja palveluissa sekä kustannussäästöihin hallinnoissa. Jos tämä voimavara vapautettaisiin avoimesti käyttöön, voitaisiin EU:ssa saavuttaa vuosittain 40 miljardin euron taloudelliset hyödyt. Julkisen datan avaaminen edistäisi myös kansalaisten osallistumista poliittiseen ja yhteiskunnalliseen toimintaan ja tukisi politiikan eri osa-alueita, kuten ympäristöpolitiikkaa.

Tässä tiedonannossa esitellään Euroopan digitaalistrategian puitteissa toimenpidekokonaisuus, jolla pyritään poistamaan nykyiset esteet ja hajanaisuus EU:ssa. Se koostuu kolmesta toisiaan vahvistavasta toimintalinjasta:

- Datan uudelleenkäytön oikeudellisen kehyksen mukauttaminen. Tämän tiedonannon ohessa annetaan ehdotukset tarkistetuksi direktiiviksi julkisen sektorin hallussa olevien tietojen uudelleenkäytöstä sekä tarkistetuksi komission päätökseksi komission oman tiedon uudelleenkäytöstä.
- Rahoitusvälineiden käyttö avoimen datan tueksi sekä käyttöönoton tukeminen esimerkiksi luomalla yleiseurooppalaisia dataportaaleja.
- Jäsenvaltioiden välisen koordinoinnin ja kokemustenvaihdon helpottaminen.

Toimenpiteissä keskitytään aloihin, jotka liittyvät sisämarkkinoiden toimintaan ja joilla yhteisillä standardeilla ja lähestymistavoilla voidaan luoda uusia parempia palveluja ja tietotuotteita eurooppalaisille kuluttajille. Ne tukeutuvat kansallisiin tiedon saatavuutta koskeviin periaatteisiin, mutta eivät vaikuta niihin.

¹ Käsitteitä ”data” ja ”tieto” käytetään tässä tiedonannossa toistensa synonyymeinä, ja niillä viitataan kaikenlaiseen sisältöön tallennusvälineestä riippumatta.

2. MITÄ AVOIN DATA VOI TARJOTA EUROOPALLE?

2.1. Käyttämättömät taloudelliset ja liiketoimintamahdollisuudet

Vuoden 2009 *Digital Britain* -raportissa² dataa kuvattiin ”*innovoinnin raaka-aineeksi*” ja ”*osaamistalouden elinehdoksi*”. Suuren osan tästä innovoinnin raaka-aineesta tuottavat, keräävät ja ostavat hallinnot eri puolilla EU:ta. Se on keskeinen materiaali monenlaisissa uusissa tietotuotteissa ja palveluissa, jotka hyödyntävät uusia mahdollisuuksia analysoida ja visualisoida eri lähteistä peräisin olevaa dataa.

Tällaiset tuotteet ovat esimerkiksi yritysten päätöksenteon tukijärjestelmiä, sijaintiperusteisia palveluja tai ajoneuvojen navigointijärjestelmiä tai vaikkapa älypuhelimille suunniteltuja sääpalveluja ja muita sovelluksia.

Paikkatietosektorin markkinoiden koko ja kasvu kuvastaa julkisen datan potentiaalia työpaikkojen synnyttäjänä. Saksassa paikkatietomarkkinoiden arvoksi arvioitiin 1,4 miljardia euroa vuonna 2007. Kasvua vuodesta 2000 oli peräti 50 prosenttia³. Alankomaissa paikkatietosektori työllisti 15 000 kokopäiväistä työntekijää vuonna 2008. Myös muilla aloilla, kuten meteorologinen data, juridinen tieto ja liiketoimintatieto, markkinat ovat kasvaneet tasaisesti.

Hiljattain tehdyssä katsauksessa⁴ julkisen sektorin tiedon kokonaismarkkinoiksi EU:ssa arvioidaan 28 miljardia euroa (2008). Saman lähteen mukaan EU27-alueella voitaisiin vuosittain saavuttaa noin 40 miljardin euron taloudelliset hyödyt, jos julkisen sektorin tietoa avattaisiin edelleen mahdollistamalla sen helppo uudelleenkäyttö. Julkisen sektorin tiedon käytöstä ja sovelluksista vuosittain saatavat suorat ja välilliset hyödyt EU27-taloudessa olisivat yhteensä 140 miljardin euron luokkaa.

2.2. Vastauksia yhteiskunnallisiin haasteisiin

Kyky käsitellä dataa älykkäästi on olennainen edellytys yhteiskunnallisiin haasteisiin vastaamiseksi. Dataa ymmärtämällä voidaan esimerkiksi parantaa kansallisten terveydenhuoltojärjestelmien kestävyyttä. Vuoden 2011 McKinsey-raportissa⁵ arvioidaan, että datan tehokkaalla hyödyntämisellä voitaisiin luoda vuosittain 300 miljardin dollarin lisäarvo.

Edistyminen genomiikassa, lääkekeksinnöissä ja vakavien sairauksien, kuten syövän ja sydänsairauksien, diagnosoinnissa ja hoidossa riippuu yhä enemmän datan kehittyneistä keruu- ja analysointitekniikoista.

Datan hallinta on olennaista myös *ympäristöhaasteisiin* vastaamiseksi. Esimerkkejä tästä ovat energiankulutustapojen tutkiminen energiatehokkuuden parantamiseksi tai ilmansaastepitoisuuksien analysointi liikenteen hallinnassa. Tietoon perustuva päätöksenteko liikenteen, maankäytön ja ilmastonmuutoksen aloilla on enenevässä määrin riippuvaista käytettävissä olevan datan analysoinnista.

² <http://www.official-documents.gov.uk/document/cm76/7650/7650.pdf>.

³ *Assessment of the Re-use of Public Sector Information in the Geographical Information, Meteorological Information and Legal Information sectors*, MICUS, joulukuu 2008.

⁴ *Review of recent studies on PSI re-use and related market developments*, G. Vickery, elokuu 2011.

⁵ *Mc Kinsey report on 'Big Data, The next frontier for innovation, competition and productivity'*, toukokuu 2011.

2.3. Tieteen edistymisestä nopeampaa

Tieteellinen työ tapahtuu yhä enemmän maailmanlaajuisena yhteistyönä internetissä käyttäen erittäin suuria datakokoelmia, valtavia laskentaresursseja ja hyvin suorituskykyisiä visualisointitekniikoita. Tietotekniikan mahdollistama tutkimus (ns. **eScience**) on olennainen tekijä pyrittäessä vastaamaan 2000-luvun haasteisiin tieteen edistämiseksi ja oppimisessa. Tarvittava data tulee simulaatioista, digitaalisista kojeista, anturiverkostoista ja seurantakeskuksista.

Vuonna 2010 julkistetussa raportissa *Riding the Wave*⁶ korostettiin datan keskeistä merkitystä tieteelle sekä sen mahdollisuuksia muuttaa koko tieteellisen prosessin luonnetta.

Tieteellisen datan laaja saatavuus esimerkiksi auttaa eri alojen tutkijoita työskentelemään yhdessä samalla data-aineistolla, luomaan täysin uusia tieteellisen tutkimuksen muotoja ja etsimään korrelaatioita tutkimustulosten välillä.

Tietotekniikkaa hyödyntävän tieteen tuoma muutos tieteellisessä prosessissa lisää tutkimuksen tuottavuutta ja auttaa löytämään uusia ja odottamattomia ratkaisuja yhteiskunnallisiin haasteisiin. Lisäksi julkisrahoitteisen tutkimuksen ja kaupallisen sektorin ristipölytys internetiä hyödyntävällä eurooppalaisella tutkimusalueella lisää innovoinnin nopeutta ja vaikuttavuutta.

2.4. Toimia tarvitaan kaikilla tasoilla: paikallisesti, alueellisesti, kansallisesti ja EU-tasolla

Julkista dataa tuotetaan kaikilla hallinnon tasoilla. Se, millä ehdoin dataa asetetaan saataville kaupallista tai ei-kaupallista uudelleenkäyttöä varten, vaikuttaa kilpailuun ja kilpailukykyyn.

Monet julkisen sektorin tietoon perustuvat tuotteet ja palvelut ovat luonteeltaan rajat ylittäviä. Mobiilipalvelun ei haluta katkeavan rajalla, ja koko Eurooppaa koskevat yritys- ja yhteisötietopalvelut, joissa on aukkoja yhden tai useamman maan kohdalla, eivät herätä kovin paljon kiinnostusta. Jotta voitaisiin hyödyntää sisämarkkinoiden kokoa ja auttaa pk-yrityksiä kasvamaan yli kansallisten rajojen, olisi helpotettava palvelujen toteuttamista koko EU:n kattavasti.

Näin ollen koko Euroopan unionissa olisi tarjottava samat perustoimintaedellytykset. Näin luotaisiin tasavertainen toimintaympäristö ja saataisiin aikaan julkiseen dataan perustuville innovatiivisille tuotteille ja palveluille kukoistavat markkinat.

3. NYKYTILANNE JA TÄHÄNASTISET SAAVUTUKSET

3.1. Nykytilanne

Pyrkimys kohti avointa dataa on saamassa pontta useissa jäsenvaltioissa. Niissä ajatus on omaksuttu syistä, jotka liittyvät läpinäkyvyyteen, hallinnon tehostamiseen ja uudelleenkäytön taloudelliseen potentiaaliin. Ne tukevat avointa hallintoa lainsäädännöllä ja käytännön toimenpiteillä, kuten tuottamalla dataa koneellisesti luettavassa muodossa ja luomalla dataportaaleja.

⁶ *Final report of the High Level Expert Group on Scientific Data*, lokakuu 2010.

Yhdistyneessä kuningaskunnassa on perustettu data.gov.uk-portaali, joka kokoaa dataa hallinnon kaikilta tasoilta. Vastaavia portaaleja on muissakin jäsenvaltioissa, esim. ETALAB Ranskassa. Dataportaaleja on myös alueellisella tasolla, kuten dadesobertes.gencat.cat Kataloniassa ja dati.piemonte.it Piemontessa Italiassa.

Jäsenvaltioiden välillä on kuitenkin eroja avoimeen dataan liittyvässä aloitteellisuudessa ja tietoisuudessa. Riskinä on, että Eurooppa hukkaa avoimen datan tarjoamat mahdollisuudet ja jää jälkeen alueista, joilla noudatetaan vakiintuneita avoimen datan periaatteita.

3.2. Tähänastiset saavutukset EU-tasolla

3.2.1. Lainsäädäntökehys ja poliittiset aloitteet

(1) Direktiivi julkisen sektorin tiedon uudelleenkäytöstä

Vuoden 2003 direktiivissä julkisen sektorin tiedon uudelleenkäytöstä⁷ säädetään Euroopan tasoiset yleiset lainsäädäntöpuitteet ja yhdenmukainen vähimmäistaso. Direktiivin vuoden 2009 uudelleentarkastelussa kävi ilmi, että vuoteen 2003 verrattuna tapahtuneesta edistymisestä huolimatta julkisen sektorin tiedon käytölle rajojen yli oli edelleen esteitä. Osa näistä esteistä voidaan poistaa nykylainsäädännöllä, osaa ei.

(2) Poliittiset aloitteet

Yleistä uudelleenkäyttöpolitiikkaa täydennetään yksittäisten alojen lainsäädäntö- ja politiikka-aloitteilla. Näistä ovat esimerkkejä:

- ympäristötiedon saatavuutta koskeva direktiivi ja INSPIRE-direktiivi, joilla pyritään **ympäristötiedon** mahdollisimman laajaan levittämiseen ja keskeisten datakokonaisuuksien yhdenmukaistamiseen⁸,
- komission tiedonanto *Meriosaaminen 2020*⁹, jolla pyritään muun muassa tekemään **meridatan** käytöstä helpompaa ja halvempaa,
- vuonna 2008 hyväksytyyn **älykkäiden liikennejärjestelmien (ITS) käyttöönottoa koskevaan toimitasuunnitelmaan**¹⁰ sisältyvät aloitteet, joilla pyritään muun muassa tuomaan matkustustiedot ja reaaliaikaiset liikennetiedot yksityisten palveluntarjoajien saataville,
- **tieteellisen tiedon** avointa saatavuutta koskeva komission politiikka¹¹, mihin sisältyy EU-rahoitteisissa hankkeissa tuotettavien julkaisujen vapaata saatavuutta koskeva pilottihanke sekä yleiseurooppalainen helpon osallistumisen mahdollistava **avoimia tietokantoja yhdistävä verkostoinfrastrukturi** – tässä yhteydessä tärkeä tekijä on myös Yhteisen tutkimuskeskuksen JRC:n julkaisutietokanta,

⁷ EUVL L 345, 31.12.2003, s. 90.

⁸ Direktiivit 2003/4/EY, EUVL L 41, s. 26, 14.2.2003 ja 2007/2/EY, EUVL L 108, s. 1, 14.3.2007.

⁹ KOM(2010) 461 lopullinen.

¹⁰ KOM(2008) 886 lopullinen/2.

¹¹ Tiedonanto *Tieteellinen tieto digitaaliaikana*, KOM(2007) 56, tiedonanto *Tieto- ja viestintätekniikka tieteen palveluksessa*, KOM(2009) 108 ja niihin liittyvät neuvoston päätelmät.

- **kulttuuriperinnön** digitointia koskevat periaatteet ja Europeanan (Euroopan digitaalisen kirjaston, arkiston ja museon) kehittäminen, joilla pyritään varmistamaan digitoitujen kulttuuriaineistojen ja niihin liittyvien metatietojen mahdollisimman laaja käyttö.

Julkisen sektorin datan uudelleenkäytössä niin kaupallisiin kuin ei-kaupallisiin tarkoituksiin olisi kaikilta osin noudatettava **yksityisyyden suojaa koskevaa kansallista ja EU-lainsäädäntöä**. Hallinnon avoimen datan lisäämistä ja henkilötietojen suojaa koskevat tavoitteet voivat tukea toisiaan, jos ne perustuvat julkisen sektorin harjoittamaan aktiiviseen ja tietoiseen tietohallintoon.

Avoimen datan periaatteiden käytössä olisi lisäksi kunnioitettava tinkimättä kolmansien osapuolten **teollis- ja tekijänoikeuksia** ja noudatettava henkistä omaisuutta koskevista kansainvälisistä yleissopimuksista johtuvia Euroopan unionin velvoitteita.

3.2.2. Tutkimuksen, kehityksen ja innovoinnin yhteisrahoitus

Komissio on tukenut avointa dataa rahoitusohjelmissaan, erityisesti tutkimuksen ja kehittämisen puiteohjelmissa, kilpailukyvyn ja innovoinnin puiteohjelmassa ja ISA-ohjelmassa. Hankkeet kattavat monenlaisia tutkimus- ja sovellusaloja ja niihin osallistuu monentyyppisiä organisaatioita.

***Linked Open Data (LOD2)** -hanke¹² käynnistyi syyskuussa 2010 ja kestää 4 vuotta. Siinä tutkitaan web-tekniikan käyttöä alustana datan ja tiedon yhdistämisessä ja pyritään semanttisten teknologioiden avulla parantamaan julkisen sektorin tiedon käytettävyyttä.*

***OpenAIRE**-hanke¹³ käynnistyi joulukuussa 2009 ja siinä on mukana osallistujia 25 EU-maasta ja useista assosioituneista maista. Hankkeessa pyritään luomaan osallistava infrastruktuuri tutkimustiedon avointa saatavuutta koskevalle komission Open Access Pilot -hankkeelle.*

ISA-ohjelmaan kuuluvassa, semanttista yhteentoimivuutta koskevassa SEMIC.EU-hankkeessa¹⁴ edistetään ajatusta hallinnon metatietojen avoimuudesta ensiaskelena kohti metatietojen yhteensopivuutta sekä kansallisella että eurooppalaisella tasolla.

3.2.3. Komission oman tiedon avoimuus

Vuonna 2006 komissio otti käyttöön omien tietoresurssiensa avointa uudelleenkäytettävyyttä koskevat periaatteet antamalla päätöksen komission tietojen uudelleenkäytöstä. Päätöksen myötä kaikki yleisesti saatavilla olevat komission tiedot ovat käytettävissä kaupallisiin ja ei-kaupallisiin tarkoituksiin. Tällaista tietoa ovat esimerkiksi Eurostatin tilastotiedot ja EU-toimielinten käännoismuistit. Tieto on saatavilla yleensä ilmaiseksi; poikkeustapauksissa voidaan periä jakelusta aiheutuvat lisäkustannukset.

¹² <http://lod2.eu/>.

¹³ <http://www.openaire.eu/>.

¹⁴ <http://www.semic.eu/>.

4. HAASTEET JA MAHDOLLISUUDET

4.1. Uusia mahdollisuuksia teknologian edistymisen myötä

Useat kehityssuuntaukset luovat uusia mahdollisuuksia datan – julkisen sektorin data mukaan luettuna – uudelleenkäyttöön uusissa tietotuotteissa ja -palveluissa.

Ensinnäkin *datan määrä* maailmassa kasvaa räjähdysmäisesti. Yhdysvaltain kongressin kirjasto keräsi yksin vuoden 2011 huhtikuussa 235 teratavua dataa¹⁵. Maailmanlaajuisesti tuotetun datan määrän vuotuiseksi kasvuksi arvioidaan 40 prosenttia.

Osa tästä kasvusta tulee *uusista datatyypeistä*. Kuljetusalalla, ajoneuvoliikenteessä, energianjakelussa ja muilla aloilla on käytössä jo nyt yli 30 miljoonaa verkotettua anturiyksikköä¹⁶.

Samanaikaisesti on nähtävissä vallankumous *teknologioissa*, joita käytetään datan analysoinnissa, hyödyntämisessä ja käsittelyssä. Esimerkkejä tästä ovat monikieliset hakutekniikat ja merkityssisältöjen automaattinen louhinta anturiverkostoista.

4.2. Haasteet ja jäljellä olevat esteet

Viime vuosina julkisen datan avaamisessa on jossain määrin edistytty, mutta asiaan liittyy vielä erilaisia esteitä.

Oikeudellinen kehys

Julkisen sektorin tiedon uudelleenkäytöstä vuonna 2003 annetulla direktiivillä aikaansaadusta vähimmäistason yhdenmukaistamisesta huolimatta **kansallisissa säännöissä ja käytännöissä on edelleen merkittäviä eroja**. Tämä aiheuttaa hajanaisuutta tiedon sisämarkkinoilla ja vaikeuttaa rajat ylittävien tietopalvelujen luomista.

Selkeimmin erot näkyvät hinnoittelussa, kun joissain tapauksissa veloitetaan aiheutuneet kustannukset ja joissain tapauksissa uudelleenkäyttö on ilmaista tai lähes ilmaista. Hiljattain tehtyjen selvitysten analyysi osoittaa, että avoimen datan malli johtaa makrotaloudellisesta näkökulmasta parempiin kokonaistuloksiin¹⁷. Kaikkien kustannusten kattamisesta lisäkustannusmalliin siirtyneitä julkisen sektorin elimiä koskevat tapaustutkimukset osoittavat, että siirtyminen lisäsi uudelleenkäyttöä, mutta hyödytti myös kyseisiä julkisen sektorin elimiä itseään¹⁸. Lisäksi avoimen datan malli poistaa mahdolliset monopolistiset taipumukset, jos data on saatavilla vain yhdestä lähteestä.

Julkisten organisaatioiden ja yksityisen sektorin tietoisuus asiasta

Toinen keskeinen tekijä on, että **julkiset organisaatiot eivät ole riittävän tietoisia** avoimen datan potentiaalista. Hallinnan menettämistä pelätään edelleen laajalti. Osa huolenaiheista on perusteltuja, kuten yksityisyyden suojaan, kansalliseen turvallisuuteen ja kolmansien

¹⁵ McKinsey, edellä mainitussa tutkimuksessa.

¹⁶ Ks. ed. alaviite.

¹⁷ G. Vickery, edellä mainitussa tutkimuksessa.

¹⁸ *Pricing of Public Sector Information*, Deloitte consulting ja muut, kesäkuu 2011.

osapuolten immateriaalioikeuksien suojaan liittyvät näkökohdat. Osa pelonaiheista vaikuttaisi kuitenkin pikemminkin tekosyiltä olla toimimatta¹⁹.

Hallintojen suhtautumisen muuttaminen edellyttää korkeimman tason vahvaa poliittista sitoutumista ja dynaamista vuoropuhelua eri intressitahojen – muun muassa hallintojen ja julkisen datan haltijoiden, yritysten ja akateemisen maailman – välillä. Pilotti- ja testihankkeet, hyviä käytänteitä koskeva tiedonvaihto ja edistämiskampanjat (käyttäen esimerkiksi avoimeen dataan liittyviä kilpailuja) voivat auttaa julkista sektoria omaksumaan avoimen datan kulttuurin. Näin voidaan lisätä myös yritysten tietoisuutta datan saatavuudesta ja sen tarjoamista mahdollisuuksista.

Tekniset ja käytännön hidasteet

Samoin on vielä ratkaistavana **tekniisiä ja käytännön** kysymyksiä. Kielimuurit ja yhteentoimivuusnäkökohdat on hoidettava niin, että eri organisaatioiden ja eri maiden tietoresurssit voidaan yhdistää. Tiedon saatavuus koneellisesti luettavassa muodossa sekä rajattu yhteisesti sovittu metatietokokonaisuus auttaisi datan ristiviittauksia ja yhteentoimivuutta ja parantaisi näin huomattavasti sen uudelleenkäyttöarvoa. Lisäksi tarvitaan tekninen infrastruktuuri, jolla varmistetaan tiedon saatavuus myös pitkällä aikavälillä. Myös datan analysointi- ja visualisointityökaluihin liittyvää tutkimusta, kehitystä ja innovointia on tuettava enemmän.

4.3. Johtopäätös

Julkisen datan hyödyntäminen tarjoaa valtavia mahdollisuuksia EU:n talouden ja kuluttajien hyvinvoinnin näkökulmasta. Nykyiset sääntelyvälineet ja niiden toteutus, hallintojen ja yritysten tietoisuuden puute ja innovatiivisten teknologioiden hidas omaksuminen pidättelevät kuitenkin julkisen datan todellisten uudelleenkäyttömarkkinoiden kehittymistä, eivätkä anna mahdollisuutta saada maksimaalisia hyötyjä datan ja kehittyvien teknologioiden tarjoamista uusista mahdollisuuksista.

Näiden näkökohtien vuoksi komissio on päättänyt tarkistaa ja lujittaa julkista dataa koskevaa strategiaansa kohdentamalla oikeudellista kehystä ja käytettävissä olevia tukijärjestelmiä uudelleen.

5. TOIMET EUROOPAN TASOLLA: TARKISTETTU STRATEGIA JULKISTA AVOINTA DATAA VARTEN

Komission tarkistettu strategia rakentuu kolmelle toisiaan täydentävälle toimintalinjalle:

- datan uudelleenkäytön puitteiden mukauttaminen, mukaan luettuina lainsäädäntö, eisitovat säädökset ja toimintapolitiikka,
- rahoitusvälineiden hyödyntäminen asettamalla avoin data etusijalle tutkimus-, kehitys-, innovointi- ja infrastruktuuriohjelmissa,
- jäsenvaltioiden välisen koordinoinnin ja kokemustenvaihdon helpottaminen.

¹⁹ Muutamia esimerkkejä osoitteessa
http://www.dr0i.de/lib/2011/07/04/a_sample_of_data_hugging_excuses.html.

5.1. Avoimuutta ja uudelleenkäyttöä suosiva tarkistettu lainsäädäntökehys

Tämän tiedonannon ohessa annetaan ehdotus julkisen sektorin tietoa koskevan direktiivin tarkistamiseksi. Tarkistuksilla

- otetaan käyttöön periaate, jonka mukaan kaikki julkinen tieto, johon ei nimenomaisesti sovelleta jotakin poikkeuksista, on uudelleenkäytettävissä kaupallisiin ja ei-kaupallisiin tarkoituksiin,
- rajoitetaan määrä, joka julkisen sektorin tiedosta voidaan veloittaa, enintään jakelusta aiheutuviin lisäkustannuksiin; poikkeustapauksissa on edelleen mahdollista periä kaikki tiedon tuottamisesta ja jakelusta aiheutuvat kustannukset,
- laajennetaan direktiivin soveltamisalaa kirjastoihin, arkistoihin, museoihin ja korkeakoulujen kirjastoihin niin, että tästä aiheutuvat rahoitusvaikutukset jäävät mahdollisimman pieniksi, eikä näille laitoksille aiheudu suurta hallinnollista rasitetta.

Direktiivin muutoksista aiheutuu myös muille julkisen sektorin elimille korkeintaan vähäisiä rahoitus- ja hallintorasitteita, joita vastaan saadaan huomattavasti merkittävämpiä hyötyjä koko yhteiskunnalle.

Samanaikaisesti komissio on päivittämässä omien tietoresurssiensa uudelleenkäyttöä koskevia sääntöjä. Tärkeimmät muutokset koskevat Yhteisen tutkimuskeskuksen JRC:n tuottaman tutkimustiedon saattamista uudelleenkäytettävyyden piiriin, päätöksen täytäntöönpanoa parantavia toimia ja tiedon tarjontaa koneellisesti luettavassa muodossa. Vuoden 2012 aikana komissio selvittää, miten vastaavia sääntöjä voitaisiin ottaa käyttöön muissa EU-toimielimissä ja keskeisissä EU:n erillisvirastoissa.

Komissio ottaa tulevissa lainsäädäntö- ja politiikka-aloitteissaan esimerkiksi liikenteen, ympäristön, meripolitiikan ja avaruuspolitiikan alalla kaikilta osin huomioon avoimen datan tarjoamat mahdollisuudet. Tarvittaessa se pyrkii julkisen sektorin tiedon uudelleenkäytöstä annettua direktiivissä asetettuja vähimmäisehtoja kunnianhimoisempaan lopputulokseen, jotta dataresurssit saataisiin mahdollisimman laajaan käyttöön.

5.2. Ei-sitovia säädöksiä tutkimustiedon avoimuudesta

Tutkimusdatan erityispiirteiden vuoksi komissio tulee esittelemään *tiede- ja tutkimusdataa* ja niihin liittyviä infrastruktuureja koskevan strategiansa yksityiskohdat erillisissä asiakirjoissa. Se aikoo vuonna 2012 antaa tiedonannon ja suosituksen tieteellisen tiedon saatavuudesta ja säilyttämisestä. Komissio pyrkii yhdessä jäsenvaltioiden kanssa saamaan jäsenvaltiot vauhdittamaan konkreettisen toimenpidekokonaisuuden pohjalta toimiaan tieteellisen tiedon tuomiseksi avoimesti saataville. Samassa yhteydessä komissio määrittelee, miten Euroopan unionin rahoittaman tutkimustoiminnan tuloksia käsitellään.

5.3. Rahoitus- ja tukitoimenpiteet: Avoimeen dataan liittyvän tutkimuksen, kehityksen ja innovoinnin edistäminen

Komissio jatkaa rahoitusohjelmissaan tukea hallintojen datan avaamiseen tähtääville toimille. Se pyrkii eri välineiden kautta stimuloimaan markkinoita, testaamaan innovatiivisia ratkaisuja ja edistämään niiden kehitystä sekä varmistamaan avoin data -mallin mahdollisimman laajan omaksumisen.

(1) Tutkimuksen, kehityksen ja innovoinnin tukeminen

Komissio jatkaa tukeaan datankäsittelyteknologian, esim. datan louhinnan, analysoinnin ja visualisoinnin, *tutkimukselle ja kehitykselle*. Kaudella 2011–2013 komissio käyttää noin 100 miljoonaa euroa näiden alojen t&k:hon. Tiedonhallinta on yksi tieto- ja viestintäteknikan kaavailuista painopistealoista myös Horisontti 2020 -kokonaisuudessa, joka kattaa EU:n tuen tutkimukselle ja innovoinnille kaudella 2014–2020.

Komissio tukee *teknologista innovointia ja teknologian käyttöönottoa* pilottitoimin testaamalla ja esittelemällä innovatiivisia sovelluksia, kuten paikkatietojärjestelmiä ja sijaintisidonnaisia palveluja sekä luovan sisällön sovelluksia koulutuksen, kulttuurin ja muotiteollisuuden alalla. Näitä tuetaan kaudella 2012–2013 kilpailukyvyyn ja innovoinnin puiteohjelman tieto- ja viestintäteknikkaosioista. Sen jälkeen tukea on määrääntänyt Horisontti 2020 -ohjelmasta.

Lisäksi komissio auttaa järjestämään *avoimeen dataan liittyviä kilpailuja* uusien tietopalvelujen kehittämisen edistämiseksi ja tekee aloitteita, joilla parannetaan julkisen sektorin dataan perustuvia uusia tietopalveluja rakentavien yrittäjien mahdollisuuksia saada pääomarahoitusta.

(2) Datainfrastruktuurien tukeminen – yleiseurooppalaiset dataportaalit

Helpottaakseen sellaisten tietotuotteiden ja -palvelujen kehittämistä, joissa yhdistellään dataa eri puolilta Euroopan unionia, komissio pyrkii luomaan kaksi toisiinsa yhteydessä olevaa dataportaalia.

Vuonna 2012 se ryhtyy ylläpitämään portaalia, jolla tuodaan komission omat ja muiden EU-toimielinten ja EU:n erillisvirastojen dataresurssit helposti saataville ja käytettäväksi. Samanaikaisesti määritellään metadatakokonaisuus, jolla on suuri uudelleenkäytettävyysspotentiaali, ja pyritään edistämään sen käyttöä. Eurooppalaiselle datalle pyritään varmistamaan mahdollisimman laaja käyttö innovatiivisissa tietotuotteissa komission tiedon uudelleenkäytöstä annetun komission päätöksen mukaisesti.

Komissio pyrkii yhdessä jäsenvaltioiden, julkisen sektorin elinten ja alueellisten tiedonkokoajien kanssa luomaan yleiseurooppalaisen dataportaalin, joka käynnistäisi toimintansa vuonna 2013 ja tarjoaisi suoran pääsyn monenlaisiin datakokonaisuuksiin eri puolilta EU:ta (komission portaalin tarjoama data mukaan luettuna). Tässä hyödynnetään mahdollisuuksien mukaan olemassa olevia rakenteita ja saavutuksia ja portaalin sisältöä kasvatetaan asteittain keskeisiin datakokonaisuuksiin kaikista jäsenvaltioista.

Perustamisvaiheessa (2011–2013) tukea annetaan kilpailukyvyyn ja innovoinnin puiteohjelmasta. Kaudella 2014–2020 julkisen datan eurooppalaisen sähköisen palveluinfrastruktuurin rahoitus tulee Verkkojen Eurooppa -välineestä²⁰.

Komissio jatkaa myös rahoitustaan digitointitoiminnalle ja Europeana-alustalle, jotka ovat tärkeitä datan uudelleenkäyttöön liittyviä osa-alueita.

(3) Tutkimusalan datainfrastruktuurien tukeminen

²⁰ Ks. välinettä koskevat komission ehdotukset, KOM(2011) 665 ja KOM(2011) 657/3.

Komissio tukee seitsemännessä puiteohjelmassa ja kaavailee tukevansa jatkossa Horisontti 2020 -ohjelmassa pysyvän ja vankkatoimisen palveluinfrastruktuurin kehittämistä Euroopassa tieteellistä dataa varten. Infrastruktuurin on määrä vastata dataintensiivisen tieteen ja tutkimuksen tarpeita 2020-luvulla, ja kehitystyössä noudatetaan *Riding the Wave* -raportin suosituksia. Se tuo saataville ja käytettäväksi tiedon koko kirjon jalostamattomasta havainto- ja koedatasta julkaisuihin kaikilla tieteenaloilla.

Tähän infrastruktuuriin liittyy teknisiä, organisatorisia ja sääntelyllisiä näkökohtia, jotka vaativat mittavaa koordinoitua jäsenvaltioiden kanssa, mutta myös kolmansien maiden ja kansainvälisten järjestöjen kanssa, jotta voidaan varmistaa maailmanlaajuinen yhteentoimivuus ja vastavuoroiset käyttöoikeudet. Komissio kehittää yhdessä kansainvälisten kumppanimme kanssa standardeja datan maailmanlaajuisista saatavuutta ja yhteentoimivuutta varten.

5.4. Koordinoitimet jäsenvaltio- ja EU-tasolla

Komissio helpottaa edelleen jäsenvaltioiden välistä koordinoitua ja kokemustenvaihtoa erityisesti seuraavien toimintamuotojen kautta:

- PSI-ryhmä, joka on jäsenvaltioiden asiantuntijaryhmä hyviä käytänteitä ja julkisen sektorin tiedon uudelleenkäyttöä tukevia aloitteita koskevaa tiedonvaihtoa varten,
- julkisen sektorin tietoa koskeva verkkoportaali, joka sisältää uutisia Euroopassa tapahtuvasta kehityksestä, hyvistä käytänteistä, esimerkkejä uusista tuotteista ja palveluista sekä oikeustapauksia, joissa on kyse julkisen sektorin tiedon uudelleenkäytöstä,
- LAPSI-verkosto, joka analysoi julkisen sektorin tietoon liittyviä oikeudellisia kysymyksiä ja edistää keskustelua tutkijoiden ja sidosryhmien keskuudessa, ja jonka puitteissa tullaan laatimaan saatavuus- ja uudelleenkäyttöperiaatteita ja -käytäntöjä koskeva ohjeisto,
- ISA-ohjelman toimet semanttisen yhteentoimivuuden alalla.

Komissio tukee jatkossakin politiikan neuvoo-antavia ryhmiä – joiden toimintaan se myös osallistuu – , kuten sähköisiä infrastruktuureja käsittelevä *e-Infrastructures Policy Forum* ja tieteen e-infrastruktuurin kehittämisen mietintäryhmä *e-Infrastructures Reflection Group*, jotka ovat tärkeitä jäsenvaltioiden välisessä koordinoinnissa tieteellisiin datainfrastruktuureihin liittyvissä kysymyksissä.

6. PÄÄTELMÄT

Julkisten organisaatioiden eri puolilla Euroopan unionia tuottama, keräämä tai ostama tieto on keskeinen voimavara tietotaloudessa. Tällä hetkellä läheskään sen koko potentiaalia ei hyödynnetä. Tässä tiedonannossa komissio ehdottaa Euroopan julkisen sektorin tietoresurssien potentiaalinen vapauttamiseksi käytännön toimia, jotka ulottuvat julkisen sektorin tiedon uudelleenkäytöstä annetun direktiivin tarkistamisesta yleiseurooppalaisen dataportaalin perustamiseen.

Jäsenvaltiot voivat edesauttaa kehitystä kohti avointa dataa viemällä julkisen sektorin uudelleenkäytöstä annettavan tarkistetun direktiivin nopeasti kansalliseen lainsäädäntöönsä ja panemalla sen viipymättä täytäntöön. Näin luodaan tarvittavat edellytykset avoimeen dataan perustuvalla taloudellisella toiminnalla ja edistetään rajat ylittäviä sovelluksia.

Lisäksi jäsenvaltioiden olisi laadittava ja otettava käyttöön avointa dataa koskevat toimintaperiaatteet, jotka perustuisivat eri puolilta EU:ta saatuihin kokemuksiin hyvistä toimintamalleista. Tukea olisi annettava esimerkiksi avoimeen dataan liittyville pilottihankkeille ja kilpailuille ja erityisesti sellaiselle toiminnalle, jossa on kyse rajat ylittävien tuotteiden ja palvelujen kehittämisestä.

Lopuksi komissio pyytää jäsenvaltioita omalta osaltaan osallistumaan yleiseurooppalaisen dataportaalien kehittämiseen. Komissio käynnistää jäsenvaltioiden asiantuntijoiden kanssa keskustelut vakaan perustan varmistamiseksi portaalille. Käyttöönottoaiheessa jäsenvaltioiden on asetettava portaalien kautta saaville laaja valikoima datakokonaisuuksia, jotta siitä tulisi menestyvä palvelu ja perusta taloudelliselle toiminnalle.

Komissio pyytää Euroopan parlamenttia ja neuvostoa luomaan vastuualueidensa puitteissa tarvittavat edellytykset julkisen sektorin tiedon uudelleenkäytölle Euroopan unionissa ja tukemaan hankkeita ja infrastruktuureja, joilla Euroopan julkisesta datasta voidaan tehdä innovaatioiden, kasvun ja läpinäkyvyyden moottori.

Avoimen datan strategia, avaintoimenpiteet

Lainsäädännölliset puitteet

- Ehdotus tarkistetuksi direktiiviksi julkisen sektorin tiedon uudelleenkäytöstä, joulukuu 2011,
- tarkistettu komission päätös komission tiedon uudelleenkäytöstä, joulukuu 2011, pyrkimykset laajentaa järjestelmää muihin EU-toimielimiin ja EU:n erillisvirastoihin, 2012,
- avoimen datan periaatteet otetaan huomioon alakohtaisissa lainsäädäntö- ja politiikka-aloitteissa.

Eurooppalaiset avoimen datan portaalit ja alustat

- Portaalit komission ja muiden EU-toimielinten ja -virastojen dataan, kevät 2012,
- yleiseurooppalainen dataportaalit eri puolilla EU:ta sijaitseviin datakokonaisuuksiin, kevät 2013, pohjautuen vuonna 2011 jäsenvaltioiden kanssa tehtyyn valmistelutyöhön,
- eurooppalaisen avoimen datan sähköisen palveluinfrastruktuurin yhteisrahoitus Verkkojen Eurooppa -välineestä 2014–2020.

Tieteen avoin data

- Tiedonanto ja suositus jäsenvaltioille tieteellisestä tiedosta, 2012 alku,
- tieteellisten julkaisujen avointa saatavuutta koskevan pilottihankkeen laajentaminen koko Horisontti 2020 -ohjelmaan sekä tutkimusdatan avointa saatavuutta koskeva pilottihanke.

Tutkimus ja innovointi

- Avoimen datan kannalta merkitykselliset tutkimus- ja innovointihankkeet seitsemännessä puiteohjelmassa, kilpailukyvyyn ja innovoinnin puiteohjelmassa ja Horisontti 2020 -ohjelmassa sekä rahoitus tutkimusinfrastruktuureille, jotka tukevat tutkimusartikkelien ja tutkimusdatan avointa saatavuutta,
- avoimeen dataan liittyvät kilpailut (2012–2013) ja pääomarahoituksen saatavuuden parantaminen alan yrittäjille.

Aikataulutavoitteet

Komissio sitoutuu

- käynnistämään komission avoimen datan portaalin vuonna 2012,
- käynnistämään vuonna 2013 monikielisellä käyttöliittymällä ja hakutoiminnoilla pilottivaiheen portaalin, joka tarjoaa datakokonaisuuksia eri puolilta Eurooppaa.

Jäsenvaltioiden kanssa tehtävässä työssä tarkoituksena on

- laatia ja ottaa käyttöön avoimen julkisen datan periaatteet kaikissa jäsenvaltioissa vuoden 2013 alkuun mennessä,

- tuoda kolmasosa kaikesta saatavilla olevasta rakenteisesta hallintojen datasta jäsenvaltioissa yleiseurooppalaisen dataporttaalin hakutoiminnon piiriin vuoteen 2015 mennessä.

Odotettu kokonaisvaikutus:

- vuoteen 2017 mennessä (kolme vuotta ajankohdasta, johon mennessä julkisen sektorin tiedon uudelleenkäyttöä koskeva tarkistettu direktiivi on saatettava osaksi kansallista lainsäädäntöä) julkisen sektorin tiedon uudelleenkäytön kokonaishyödyt EU:ssa vuodessa 100 miljardia, sisältää uuden liiketoiminnan ja julkisen sektorin palveluissa saavutetut tehokkuushyödyt.