



Euroopan unionin  
neuvosto

Bryssel, 20. marraskuuta 2025  
(OR. en)

15712/25

TELECOM 420  
COMPET 1209  
CYBER 339

## SAATE

---

Lähettäjä: Euroopan komission pääsihteeri, allekirjoittajana johtaja Martine  
DEPREZ

Saapunut: 19. marraskuuta 2025

Vastaanottaja: Thérèse BLANCHET, Euroopan unionin neuvoston pääsihteeri

---

Kom:n asiak. nro: COM(2025) 835 final

---

Asia: KOMISSION TIEDONANTO EUROOPAN PARLAMENTILLE JA  
NEUVOSTOLLE  
DATAUNIONISTRATEGIA –  
DATAA TEKOÄLYN KÄYTTÖÖN

---

Valtuuskunnille toimitetaan oheisena asiakirja COM(2025) 835 final.

Liite: COM(2025) 835 final



Bryssel 19.11.2025  
COM(2025) 835 final

**KOMISSION TIEDONANTO EUROOPAN PARLAMENTILLE JA NEUVOSTOLLE**

**DATAUNIONISTRATEGIA – DATAA TEKOÄLYN KÄYTTÖÖN**

## 1. Johdanto – Dataa tekoälyn käyttöön

Tekoälyn käyttö muuttaa maailmantaloutta. EU:ssa tarvitaan nyt suuria määriä korkealaatuaista dataa, jotta se voi kilpailla ja edistää innovointia. Ilman dataa EU ei voi kehittää vahvoja tekoälymalleja, optimoida terveydenhuolto- tai energiajärjestelmää eikä ylläpitää teollisuuden johtoasemaa. Erityisesti pienten ja keskisuurten yritysten laajentumisen ja kilpailukyvyn säilyttämisen kannalta on ratkaisevan tärkeää, että datan saatavuutta parannetaan.

EU on rakentanut vahvan perustan turvallisille ja yhteentoimiville datan sisämarkkinoille datasäädöksen<sup>1</sup> ja muiden keskeisten säädösten sekä yhteisiin eurooppalaisiin data-avaruuksiin<sup>2</sup> tehtyjen investointien avulla. Samaan aikaan tekoälyn maanosa -toimintasuunnitelmalla<sup>3</sup> ja tekoälyn soveltamisstrategialla<sup>4</sup> on luotu edellytykset sille, että EU voi saavuttaa johtoaseman tekoälyn kehittämisessä ja käyttöönotossa.

EU:n haasteina ovat tekoälyn kehittämiseen tarvittavan datan niukkuus ja kasvava geopoliittinen kilpailu, jossa data nähdään enenevässä määrin strategisena voimavarana. Suuria määriä arvokasta dataa siiloutuu tai jää hyödyntämättä, mikä johtuu myös datasääntöjen kirjavuudesta. Samaan aikaan maailmanlaajuiset kilpailijat pääsevät hyödyntämään dataa nopeammin ja tällä tavoin saavat teknologista ja teollista etua.

Sääntöjen noudattamisen helpottamiseksi ja ennustettavuuden parantamiseksi laaditun digitaalialan koontiasetuksen tavoitteena on yksinkertaistaa datan sääntely-ympäristöä **yhdistämällä neljä säädöstä yhdeksi yhtenäiseksi datakehyykseksi**. Lisäksi strategiaan liittyy **datasäädöksen mukainen kattava tukipaketti**, jolla tuetaan yrityksiä ja helpotetaan sääntöjen noudattamista. Mallisopimusehdoilla, pilvipalvelusopimusten vakiosopimuslausekkeilla ja erityisellä neuvontapalvelulla vähennetään oikeudellista monimutkaisuutta ja tuetaan etenkin pk-yrityksiä, jotta ne selviävät velvoitteistaan ja voivat kohdentaa voimavaransa innovointiin. Mallilausekkeet tulevat käyttöön sekä yritysten ja viranomaisten välisissä että yritysten välisissä sopimussuhteissa, ja niillä on tarkoitus edistää datan luomista ja jakamista sekä sopimusten yksinkertaistamista.<sup>5</sup>

Eurooppalaisessa dataunionistrategiassa keskitytään sääntöjen sijasta tuloksiin. Niiden saavuttamiseksi EU ryhtyy toimiin kolmella keskeisellä osa-alueella:

- **parannetaan datan saatavuutta tekoälyä varten esimerkiksi datalaboratorioilla, jotka tarjoavat luotettavia pseudonymisointipalveluja ja kokoavat yhteen dataresursseja**

---

<sup>1</sup> Euroopan parlamentin ja neuvoston asetukset (EU) 2023/2854, annettu 13 päivänä joulukuuta 2023, datan oikeudenmukaista saatavuutta ja käyttöä koskevista yhdenmukaisista säännöistä ja asetuksen (EU) 2017/2394 ja direktiivin (EU) 2020/1828 muuttamisesta.

<sup>2</sup> Euroopan komissio, komission yksiköiden valmisteluasiakirja yhteisistä eurooppalaisista data-avaruuksista, SWD(2024) 21 final, 24.1.2024.

<sup>3</sup> Euroopan komissio (2025). Tekoälyn maanosa -toimintasuunnitelma. Komission tiedonanto Euroopan parlamentille, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle. COM(2025) 165 final. Bryssel: Euroopan komissio.

<sup>4</sup> Euroopan komissio, Tekoälyn soveltamisstrategia, COM(2025) 723 final, Bryssel, 8.10.2025.

<sup>5</sup> [Päivitetyt tekoälyä koskevat EU:n mallisopimuslausekkeet | Julkisten ostajien yhteisö.](#)

julkisilta ja yksityisiltä toimijoilta laadukkaiden data-aineistojen tuottamiseksi yritysten ja tutkijoiden käyttöön

- **virtaviivaistetaan datasääntöjä**, jotta yritysten ja tutkijoiden olisi helpompi jakaa dataa, myös **uudistamalla evästeitä koskevaa suostumusta** evästeväsymyksen vähentämiseksi, kuitenkin niin, että samalla suojellaan oikeuksia
- **vahvistetaan EU:n maailmanlaajuista asemaa kansainvälisissä datavirroissa** puuttamalla perusteettomiin kaupan esteisiin, jotta EU:n yrityksillä olisi tasapuoliset toimintaedellytykset maailmanlaajuisessa kilpailussa.

## 2. Pohjana Euroopan datastrategia (2020–2025)

EU loi vuoden 2020 Euroopan datastrategialla<sup>6</sup> oikeudellisen ja institutionaalisen pohjan turvallisille ja oikeudenmukaisille datan sisämarkkinoille. Sen tavoitteena oli mahdollistaa datan koko potentiaalin käyttöönotto innovoinnissa ja kasvussa, kuitenkin siten, että samalla suojellaan oikeuksia. Generatiivisen tekoälyn ja kasvavan geopoliittisen kilpailun myötä on kuitenkin selvää, että EU:n on mentävä tomissaan rakentamiaan perustoja pidemmälle.

Euroopan datastrategia edisti keskeistä lainsäädäntöä, jolla rakennettiin luottamusta, edistettiin datan jakamista ja selvennettiin sääntöjä koko datan arvoketjussa. **Datanhallinta-asetuksella** luotiin mekanismeja datan luotettavaa jakamista varten, säänneltiin datan välityspalveluita, otettiin käyttöön yrityksiä koskeva kehys datan vapaaehtoiselle jakamiselle yleisen edun mukaisia tarkoituksia varten (vapaaehtoisuuteen perustuva data-altruismi) ja helpotettiin tiettyjen suojattujen julkisen sektorin data-aineistojen saatavuutta. **Datasäädöksellä** vapautetaan dataa selventämällä verkkoon liitetyistä tuotteista ja niihin liittyvistä palveluista saatavan datan saatavuutta ja käyttöä koskevia oikeuksia. Lisäksi **avointa dataa koskevassa direktiivissä** ja siihen liittyvässä arvokkaita tietoaineistoja koskevassa täytäntöönpanosäädöksessä (jota on sovellettu kesäkuusta 2024 alkaen) säädetään, että tietyt julkisen sektorin hallussa olevat data-aineistot on asetettava vapaasti ja avoimesti saataville koneellisesti luettavassa muodossa. Nykyiseen lainsäädäntökehykseen liittyy kuitenkin yhä haasteita ja ratkaistavia kysymyksiä, kuten epä johdonmukainen kansallinen täytäntöönpano ja liikesalaisuuksiin liittyvä epävarmuus.

Euroopan datastrategian puitteissa käyttöön otettuja tukitoimenpiteitä ovat muun muassa Euroopan datainnovaatiolautakunnan kanssa tehtävä yhteistyö jäsenvaltioiden toimien koordinoimiseksi ja standardointitoimeksianto luotettavaa dataa koskevan EU:n kehityksen luomiseksi.<sup>7</sup>

---

<sup>6</sup> Euroopan datastrategia – Euroopan digitaalista tulevaisuutta rakentamassa, julkaisutoimisto, 2020, <https://data.europa.eu/doi/10.2775/645928>.

<sup>7</sup> Euroopan komissio, Commission Implementing Decision C(2025) 4135 of 1 July 2025 on a standardisation request to the European standardisation organisations as regards a European Trusted Data Framework in support of Regulation (EU) 2023/2854 of the European Parliament and of the Council (ei saatavana suomeksi), saatavilla osoitteessa: [https://ec.europa.eu/growth/tools-databases/enorm/mandate/614\\_en](https://ec.europa.eu/growth/tools-databases/enorm/mandate/614_en) (avattu 27.10.2025).

Edistääkseen datan sisämarkkinoiden toteutumista komissio täydensi kansallisen ja yksityisen sektorin toimia investoimalla vuosina 2021–2024 kaikkiaan 336 miljoonaa euroa 14 strategiseen eurooppalaiseen data-avaruuteen, jotka kattavat keskeiset talouden alat ja yleisen edun kannalta merkittävät alat. Data-avaruudet tarjoavat turvallisen infrastruktuurin ja hallintokehykset vapaaehtoisuuteen ja sovittuihin ehtoihin perustuvaa datan jakamista varten. Tällä hetkellä suurin haaste on se, kuinka näitä toimia voidaan laajentaa, jotta saadaan aikaan EU:n laajuista vaikutusta.

---

*Eurooppalainen syövän kuvantamista koskeva data-avaruus (Cancer Image Data Space) kattaa anonymisoidut kuvat ja annotoinnin. Siihen on tarkoitus kerätä yli 60 miljoonaa syöpäkuvaa vuoteen 2027 mennessä.*

---

### 3. Kolme haastetta, joihin EU:n on nyt vastattava

Tekoälyteknologian ja -palveluiden muuttaessa globaalia toimintaympäristöä EU:n on vastattava kiireellisesti kolmeen uuteen strategiseen haasteeseen, jotka ovat datan niukkuus, sääntelyn monimutkaisuus ja kasvava maailmanlaajuinen kilpailu.

#### Datan niukkuus: Innovoinnin rakenteellinen pullonkaula

Generatiivisen tekoälyn, suurten kielimallien ja agenttisen tekoälyn<sup>8</sup> yleistyessä pääsystä laajoihin, laadukkaisiin, uusiin ja alakohtaisiin data-aineistoihin on tullut ratkaiseva tekijä maailmanlaajuisen kilpailukyvyn kannalta. Epoch AI -tutkimusinstituutin mukaan suurten kielimallien kouluttamiseen käytettävien data-aineistojen koko kaksinkertaistuu noin kuuden kuukauden välein.<sup>9</sup>

Suurten kielimallien ja muun kaltaisten perusmallien koulutusta varten tarvitaan valtavia ja monipuolisia data-aineistoja. Tutkimukset viittaavat siihen, että jos suuntaukset jatkuvat nykyisenlaisina, julkisesti saatavilla oleva koulutusdata saatetaan käyttää loppuun vuosien 2026 ja 2032 välisenä aikana.<sup>10</sup>

EU pyrkii vastaamaan kahteen haasteeseen: i) laadukkaiden data-aineistojen, myös alakohtaisten data-aineistojen, tuominen laajemmin saataville ja ii) sen varmistaminen, että näiden data-aineistojen käsittelyyn tarvittavaa laskentainfrastruktuuria on saatavilla suuressa mittakaavassa. Monilla eurooppalaisilla yrityksillä, erityisesti pk-yrityksillä ja startup-yrityksillä, ei ole käytettävissään riittävää määrää eikä riittävän monipuolista dataa eikä

---

<sup>8</sup> Agenttisella tekoälyllä tarkoitetaan tekoälyjärjestelmiä, jotka pystyvät tekemään päätöksiä ja toimimaan itsenäisesti. Tämän ansiosta agentit pystyvät ymmärtämään kieltä, tekemään tehtäviin liittyviä päätelmiä, toimimaan itsenäisesti ennalta määriteltyjen tavoitteiden saavuttamiseksi, olemaan vuorovaikutuksessa ympäröivän maailman kanssa ja organisoimaan vuorovaikutusta myös ihmisten kanssa.

<sup>9</sup> Robi Rahman ja David Owen (2024), "The size of datasets used to train language models doubles approximately every six months", julkaistu Epoch AI:n verkkosivustolla osoitteessa <https://epoch.ai/data-insights/dataset-size-trend>.

<sup>10</sup> Villalobos, P., Ho, A., Sevilla, J., Besiroglu, T., Heim, L., ja Hobbhahn, M. (2024), *Position: Will we run out of data? Limits of LLM scaling based on human-generated data*, julkaisussa K. Chaudhuri, S. Jegelka, L. Song, D. L. Silver ja Y. Ermon (toim.), "Proceedings of the 41st International Conference on Machine Learning" (osa 235, s. 42085–42101), PMLR, <https://proceedings.mlr.press/v235/villalobos24a.html>.

mahdollisuutta käyttää EU:n tason laskentakapasiteettia, joita tarvitaan kilpailukykyisten tekoälyratkaisujen kehittämiseen. Ilman kiireellisiä toimia EU on vaarassa jäädä jälkeen.

### **Sääntelyn monimutkaisuus: hajanaisuus hidastaa laajentamista**

EU otti vuoden 2020 eurooppalaisen datastrategian myötä käyttöön olemassa oleviin sääntöihin perustuvia merkittäviä säädöksiä, kuten datanhallinta-asetuksen<sup>11</sup>, datasäädökseen ja erilaisia alakohtaisia säädöksiä, kuten eurooppalaista terveystietoaluetta koskevan asetuksen<sup>12</sup>. Kussakin näistä hankkeissa keskityttiin rajattuihin kysymyksiin, kuten datan jakamisen mekanismeihin, arvon oikeudenmukaiseen jakautumiseen ja datan säilytyspaikkaa koskeviin kuormittaviin vaatimuksiin. Yleisen tietosuoja-asetuksen<sup>13</sup> ja alakohtaisen lainsäädännön monimutkainen vuorovaikutus ja epätasainen täytäntöönpano eri jäsenvaltioissa ovat kuitenkin luoneet hajanaisen sääntely-ympäristön ja oikeudellista epävarmuutta, myös viranomaisille, ja lisänneet säännösten noudattamisesta aiheutuvia kustannuksia, mikä on vaikuttanut erityisesti startup- ja pk-yrityksiin.

Esimerkiksi datan välityspalvelujen tarjoajiin, joiden toimiala on edelleen uusi, sovelletaan rajoittavia oikeudellisia velvoitteita, jotka rajaavat niiden kasvumahdollisuuksia. On välttävää rasittamasta varhaisvaiheen ekosysteemejä suhteettomilla vaatimuksilla, jotka haittaavat datan jakamisen mallien ja data-avaruuksien käyttöönottoa. Jotta innovointeja saadaan vauhditettua, EU:n on yksinkertaistettava datan saatavuutta ja käyttöä koskevia sääntöjä.

### **Maailmanlaajuinen kilpailu: data strategisena voimavarana**

Tekoälykilpailussa korkea lisäarvoa tuottavan datan saatavuus on keskeinen strateginen etu. Maailmanlaajuisesti datasta on tullut geopoliittinen voimavara, ja datan saatavuutta, säilytyspaikkaa ja valvontaa käytetään enenevässä määrin vallan välineinä. Samalla kun EU pyrkii lisäämään datavirtojen avoimuutta, tietoturvallisuutta, oikeudenmukaisuutta ja luotettavuutta, muut lainkäyttöalueet pitäytyvät ehdottomuuteen tai protektionismiin perustuvissa strategioissa. EU:n ulkopuolella käytössä olevat järjestelyt, jotka koskevat datan säilytyspaikkaa ja joilla rajoitetaan datan käyttöä, rajoittavat EU:n pääsyä maailmanlaajuisiin resursseihin ja altistavat EU:n yritykset taloudellisille ja turvallisuusriskeille. Jotta voidaan hyödyntää eurooppalaisen tekoälyn koko potentiaali, unionin on käsiteltävä dataa keskeisenä strategisena resurssina ja investoitava turvallisiin, laadukkaisiin, yhteentoimiviin ja eurooppalaisten arvojen ja standardien mukaisiin data-aineistoihin. On sekä talouden että turvallisuuden kannalta välttämätöntä, että Euroopan kykyä kerätä, kuratoida ja käyttää omaa dataansa vahvistetaan. EU:n on turvattava hyödylliset datavirrat, suojattava arkaluonteiset muut kuin henkilötiedot EU:ssa ja tuettava digitaalista suvereniteettia teknologisen kilpailun lisääntyessä.

---

<sup>11</sup> Euroopan parlamentin ja neuvoston asetukset (EU) 2022/868, annettu 30 päivänä toukokuuta 2022, eurooppalaisen datan hallinnoinnista ja asetuksen (EU) 2018/1724 muuttamisesta (datanhallinta-asetus) (EUVL L 152, 3.6.2022, s. 1).

<sup>12</sup> Euroopan parlamentin ja neuvoston asetukset (EU) 2025/327, annettu 11 päivänä helmikuuta 2025, eurooppalaisesta terveystietoalueesta sekä direktiivin 2011/24/EU ja asetuksen (EU) 2024/2847 muuttamisesta.

<sup>13</sup> Euroopan parlamentin ja neuvoston asetukset (EU) 2016/679, annettu 27 päivänä huhtikuuta 2016, luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta.

## 4. Eurooppalaisen dataunionistrategian kolme pilaria

### Pilari I: Parannetaan laadukkaan datan saatavuutta tekoälyä ja innovointia varten

EU:n kilpailukyky tekoälyn ja digitaalisen innovoinnin alalla riippuu laadukkaan datan saatavuudesta ja infrastruktuurista, joka mahdollistaa datan turvallisen ja laajamittaisen jakamisen. EU on jo rakentanut vahvan perustan luomalla yhteiset eurooppalaiset data-avaruudet ja hallintokehykset sekä tekemällä merkittäviä investointeja pilviteknologiaan ja laskentaan. Seuraavana haasteena on siirtyä pilottihankkeista ja hajanaisista aloitteista saumattomaan, yhteentoimivaan ja kestäväan dataekosysteemiin, joka tukee läpimurtoinnovaatioita ja vahvistaa EU:n digitaalista suvereniteettiä.

Tämän saavuttamiseksi komissiolla on kaksi toisiaan täydentävää toimintalinjaa. Ensinnäkin se käynnistää lippulaivahankkeita, joilla puututaan EU:n välittömmimpiin pullonkauloihin, joita ovat kriittisten data-aineistojen rajallinen saatavuus, tekoälyn laajamittaista kehittämistä varten tarvittavan infrastruktuurin riittämättömyys ja luotettavien ympäristöjen tarve, mukaan lukien datalaboratoriot, jotka toimivat linkkinä data-avaruuksien ja tekoälyn kehittäjien välillä. Datalaboratoriot toimivat erikoistuneina palvelupisteinä, jotka tarjoavat turvallisia ympäristöjä, käytännön välineitä ja asiantuntijatukea datan yhdistämiseen, kuratointiin, pseudonymisointiin ja anonymisointiin. Ne auttavat yrityksiä, erityisesti pk-yrityksiä, muuttamaan datan resurssiksi, jota voidaan käyttää tekoälyn kouluttamiseen, ja samalla varmistavat datan hallinnan. Nämä toimet tukevat tekoälyn soveltamisstrategiaa auttamalla varmistamaan, että datan saatavuus tukee suoraan tekoälyn käyttöönottoa ja innovointia eri teollisuudenaloilla ja julkisella sektorilla. Toiseksi komissio tukee näitä toimia vahvistamalla horisontaalisia edellytyksiä, joihin kuuluvat datan yhdistämisen oikeudellinen selkeys, datan laatua koskevia standardit ja investoinnit synteettistä

---

### Data-avaruudet ja datalaboratoriot: Euroopan tekoälyekosysteemin rakenneosat

**Yhteiset eurooppalaiset data-avaruudet** ovat datan jakamisen ekosysteemejä. Ne perustuvat pilvi-infrastruktuuriin ja selkeisiin hallintosääntöihin, joissa määritellään, kenellä on pääsy dataan ja kuka voi käyttää ja jakaa sitä. Ne yhdistävät julkiset ja yksityiset toimijat luotettuihin tiedonvaihtomekanismeihin eri aloilla ja niiden välillä.

**Datalaboratoriot** toimivat datapalvelujen tarjoajina, jotka yhdistävät data-avaruudet tekoälyekosysteemiin. Ne tarjoavat yrityksille ja tutkijoille turvallisen ja käytännönläheisen pääsyn laadukkaisiin tietoaisteisiin ja tukea EU:n sääntöjen noudattamisessa. Lisäksi ne tarjoavat välineitä, ohjausta ja luotettavia ympäristöjä datan yhdistämistä, kuratointia, tunnisteiden lisäämistä ja pseudonymisointia varten.

Data-avaruudet tarjoavat jäsenneiltyjä luotettavan datan lähteitä, ja datalaboratoriot muuttavat datan hyödynnettävissä olevaksi resurssiksi innovointia ja tekoälyn kehittämistä varten. Tällä tavoin voidaan varmistaa saumaton siirtyminen datan saatavuudesta sen soveltamiseen.

---

dataa<sup>14</sup> koskeviin valmiuksiin, joiden avulla varmistetaan laajuus, luottamus ja pitkän aikavälin kestävyys kaikilla aloilla.

### **i. Laajennetaan yhteisiä eurooppalaisia data-avaruuksia**

**Yhteiset eurooppalaiset data-avaruudet** ovat keskeisessä asemassa datan sisämarkkinoiden rakentamisessa. Seuraavassa vaiheessa niitä laajennetaan ja ne yhdistetään tekoälyinfrastruktuuriin datalaboratorioiden ja tekoälytehtaiden kautta, jotka muuttavat EU:n data-aineistot luotettavaksi tekoälyn käyttövarannoksi. Näillä sekä tekoälyn soveltamisstrategiaan kuuluvilla toimilla varmistetaan, että data-avaruuksista on mahdollista siirtyä suoraan tekoälyn kehittämiseen ja käyttöönottoon eri aloilla.

**Pilvipohjainen Simpl-välitysohjelmistokokonaisuus<sup>15</sup>** mahdollistaa eri hankkeiden ja ohjelmien yhteentoimivuuden avoimeen lähdekoodiin perustuvien, modulaaristen ja tietoturvallisten komponenttien avulla. Tämä vähentää pk-yritysten esteitä ja luo nopeampia yhteyksiä ekosysteemien välille. Data-avaruuksien tukikeskus edistää käyttöönottoa erityisesti pk-yrityksissä jakamalla tietoa ja antamalla käytännön ohjausta.

Yhteisiä eurooppalaisia data-avaruuksia koskevassa tulevassa EU:n rahoituksessa asetetaan etusijalle yleisen edun kannalta merkittävät alat, kuten terveys, liikkuvuus, energia, julkishallinto ja ympäristö, kun taas kehittyneillä aloilla, kuten valmistusteollisuudessa ja rahoituksessa, siirrytään markkinavetoisiin malleihin. Komissio tukee tätä siirtymää edistämällä standardointia, yhteentoimivuutta ja yhteisinvestointikehyksiä. Keskeisinä tavoitteina ovat edelleen loppukäyttäjien integrointi, tekoälyvalmius ja rahoituksen kestävyys.

Osana tekoälyn soveltamisstrategian lippulaivatoimia EU lisää **yhteisten eurooppalaisten data-avaruuksien** hyödyntämistä tekoälyn käyttöönoton nopeuttamiseksi keskeisillä aloilla ja tukee etulinjan tekoälymallien kehittämistä Frontier AI Initiative -aloitteen kautta. Nämä toimet liittyvät läheisesti muihin

---

#### *Eurooppalaisen terveysdata-avaruuden seuraavat vaiheet:*

*Eurooppalainen terveysdata-avaruus toimii keskeisenä siltana terveysdataekosysteemien ja tekoälyn kehittämisen välillä ja mahdollistaa sen, että datalaboratoriot ja tekoälytehtaat voivat hyödyntää anonymisoituja ja synteettisiä data-aineistoja luotettavissa käsittely-ympäristöissä.*

*Maaliskuusta 2029 alkaen potilaskertomusten ja sähköisten lääkemääräysten vaihto on mahdollista kaikkien jäsenvaltioiden välillä sähköisten terveystietojen toissijaisen käytön ohella. Maaliskuuhun 2031 mennessä tämä laajennetaan kattamaan lääketieteelliset kuvat, laboratoriotulokset ja loppulausunnot. Lisäksi mahdollistetaan genomi- ja muiden tietojen toissijainen käyttö.*

---

<sup>14</sup> Synteettinen data on keinotekoisesti tuotettua dataa, jota ei kerätä todellisista tapahtumista vaan joka on laadittu tilastollisesti jäljittelemään todellisen data-aineiston ominaisuuksia, malleja ja suhteita.

<sup>15</sup> Simpl on avoimeen lähdekoodiin perustuva tietoturvallinen väliohjelmistokokonaisuus, joka mahdollistaa datan saatavuuden eurooppalaisia datahankkeita ja ohjelmia varten sekä tukee eurooppalaisten datahankkeiden yhteentoimivuutta. Se tarjoaa useita yhteensopivia komponentteja, jotka ovat vapaasti käytettävissä ja joiden osalta noudatetaan datan laatua ja yhteiskäyttöä koskevia yhteisiä standardeja. Ks. <https://simpl-programme.ec.europa.eu/>.

tekoölyn soveltamisen lippulaivahankkeisiin, joihin kuuluvat muun muassa teollisuuden yleiskäyttöisiä tekoölymalleja koskeva hanke (Foundational Models for Industry), tekoölyyn perustuvaa lääketieteellistä tutkimusta koskeva hanke (AI-powered Pharma Discovery) ja autonomista ajamista edistäviä kaupunkeja koskeva hanke (Autonomous Drive Ambition Cities) ja joissa kaikissa hyödynnetään yhteisten eurooppalaisten data-avaruuksien kautta saatua alakohtaista dataa. Useat käytännön sovellukset pohjautuvat yhteisiin eurooppalaisiin data-avaruuksiin: terveydenhuollon tekoölypohjaiset seulontakeskukset hyödyntävät eurooppalaista terveystiedon avaruutta<sup>16</sup> diagnosoimisen välineiden validoinnissa, valmistusteollisuuden data-avaruutta käytetään valmistusta koskevan luotettavan datan yhdistämisessä erikoistuneiden ja etulinjan tekoölymallien kouluttamiseksi ja yhteisen eurooppalaisen maatalouden data-avaruutta hyödynnetään maatalous- ja elintarvikealaa koskevassa tekoölyalustassa, jolla tuetaan tekoölypohjaisten viljelyvälineiden käyttöönnottoa.

Vuodesta 2026 alkaen data-avaruuksien käyttöönnottoa jatketaan ensisijaisilla aloilla. EU tukee toimia meneillään olevilla noin 100 miljoonan euron investoinneilla, jotka mahdollistavat luotettavan ja laajamittaisen datan käytön tekoölysovelluksia varten. **Eurooppalainen terveystiedon avaruus** tukee tekoölypohjaista diagnostiikkaa ja yksilöllistettyä hoitoa. Se toimii keskeisenä siltana terveystiedon ekosysteemien ja tekoölyn kehittämisen välillä ja mahdollistaa sen, että datalaboratoriot ja tekoölytehtaat voivat hyödyntää anonymisoituja ja synteettisiä data-aineistoja luotettavissa käsittely-ympäristöissä. Yhteinen eurooppalainen **liikkumistiedon avaruus** puolestaan mahdollistaa ajoneuvojen, infrastruktuurin ja logistiikan verkottumisen turvallisemman ja ympäristöystävällisemmän liikenteen mahdollistamiseksi. Energian data-avaruus tukee älykkäitä ja joustavia energiapalveluja, ja **mediatiedon avaruus** vauhdittaa luovia aloja edistämällä tekoölypohjaisia kulttuuri-innovaatioita. Datalaboratoriot toimivat data-avaruuksien käytännöllisinä yhteyspisteinä ja auttavat organisaatioita saamaan, käsittelemään ja käyttämään dataa tehokkaasti tekoölyä varten. Näissä puitteissa **eurooppalainen oikeudellinen data-avaruus** laajentaa oikeudellisten ja tuomioistuintietojen saatavuutta oikeuskäytännön ja lainsäädännön yhteisten tunnusten ja metadatan avulla ja näin mahdollistaa sen, että näitä tietoja voidaan käyttää oikeusteknologiassa (LegalTech). Tässä yhteydessä selvitetään, onko tarpeen kehittää sopimusehtojen datapooli automatisoitua sopimuksentekoa varten.

Komissio aikoo nopeuttaa ympäristöalan digitalisointia **Euroopan vihreän kehityksen ohjelman data-avaruuden** avulla. Näin DigitalGreenTech-yhteisö pääsee hyödyntämään uudelleenkäytettäviä komponentteja ja laadukkaita data-aineistoja monialaisten ratkaisujen laajentamiseksi. Ensisijaisiin toimiin kuuluvat muun muassa eurooppalaista vesiresilienssistrategiaa edistävät datapohjaiset palvelut, lupamenettelyjen digitalisointi, tekstiilien jäljitettävyyttä ja luontohyvytyksiä koskevat pilottihankkeet sekä kehittynyt metsien seuranta avoimeen ja luottamukselliseen dataan perustuvan koneoppimisen avulla.

---

<sup>16</sup> Toimien pohjana ovat myös EU:n syöväntorjuntasuunnitelma, eurooppalainen elämäntieteiden strategia ja sydän- ja verisuoniterveyttä edistävä EU:n suunnitelma.

**Eurooppalainen puolustusdata-avaruus** luo luotettavan ympäristön operatiivisen, teollisen ja tutkimusdatan yhdistämistä varten, jotta voidaan kehittää seuraavan sukupolven puolustusjärjestelmiä, lisätä teollisia valmiuksia ja vähentää riippuvuutta kolmansien maiden toimittajista, mikä vahvistaa EU:n teknologista suvereniteettia. Komissio tutkii yhteistyö- ja tiedonvaihtomahdollisuuksia Ukrainan datapohjaisesta puolustuksesta saatujen tietojen perusteella. Hanketta kehitetään yhdessä jäsenvaltioiden ja asiaankuuluvien sidosryhmien, kuten yritysten, kanssa.<sup>17</sup>

## ii. Datalaboratoriot

Kuten tekoälyn maanosa -toimintasuunnitelmassa kuvataan, datalaboratoriot ovat erikoistuneita tiloja, jotka toimivat linkkinä datan haltijoiden, yhteisten eurooppalaisten data-avaruuksien, alakohtaisten dataekosysteemien ja EU:n tekoälyekosysteemin välillä. Datalaboratoriot<sup>18</sup> tarjoavat käytännön palveluja, kuten datan yhdistämistä<sup>19</sup>, kuratointia<sup>20</sup>, tunnisteen lisäämistä ja pseudonymisointia<sup>21</sup>. Niiden tarkoituksena on auttaa toimijoita, erityisesti startup- ja scale-up-yrityksiä, jakamaan ja käyttämään dataa turvallisesti, helpottaa yhteistyöhön perustuvaa tekoälyn kouluttamista ja tukea tekoälymallien kehittämistä keskeisillä aloilla, kattaen erilaiset hallinto- ja lisensointimallit. Datalaboratoriot tukevat tekoälyn soveltamisstrategiaa auttamalla siirtymään laadukkaan datan saatavuudesta konkreettiseen tekoälyn käyttöön. Tällä tavoin ne toimivat käytännön mahdollistajina ja nopeuttavat kokeilua, käyttöönottoa ja laajentamista. Niitä voidaan käyttää myös data-avaruuksien ja muiden datainfrastruktuurien puolesta tehtävissä, jotka edellyttävät kehittyneitä tekoälyresursseja, esimerkiksi synteettisen datan tuottamisessa tai yksityisyyden ja liikesalaisuuksien turvaamisessa, ja tällä tavoin auttaa organisaatioita jakamaan ja käyttämään dataa turvallisesti.

Yhdistämällä julkisia ja yksityisiä resursseja datalaboratoriot auttavat korjaamaan markkinoiden keskeistä toimintapuutetta eli monipuolisen ja laadukkaan datan rajallista saatavuutta ja haluttomuutta jakaa yksityisomistuksessa olevaa dataa tekoälyn kouluttamista varten. Ne toimivat olemassa olevien kanavien ja kehysten kautta eivätkä vaadi suoraa datasiirtoa. Näin data-avaruudet pysyvät luotettavina infrastruktuureina, joissa dataa hallinnoidaan ja tuodaan saataville, kun taas datalaboratoriot voivat toimia operatiivisena

---

<sup>17</sup> Hanketta ohjaa Euroopan puolustusviraston toteutettavuustutkimus, jonka on määrä valmistua vuoden 2025 loppuun mennessä.

<sup>18</sup> Joissakin yhteyksissä samankaltaisia välineitä, jotka mahdollistavat jäsenllyyn, tietoturvallisen ja luotettavan datan käytön eri ympäristöissä, kutsutaan datasäiliöiksi. Yhdessä laajemman datan kontittamisen kanssa ne toimivat täydentävänä tietojenvaihdon järjestämisen ja hallinnoinnin tapana ja edistävät EU:n tekoälyekosysteemien yhteentoimivuutta ja johdonmukaisuutta.

<sup>19</sup> Yhdistämisellä tarkoitetaan useista lähteistä peräisin olevan datan yhdistämistä ja jakamista yhteen keskitettyyn arkistoon tai jaettuun ympäristöön.

<sup>20</sup> Kuratoinnilla tarkoitetaan datan ja sen tunnisteen organisointia, integrointia, validointia ja ylläpitoa saatavuuden ja käytön parantamiseksi.

<sup>21</sup> Asetuksen (EU) 2016/679 4 artiklan 5 kohta: ”[pseudonymisoinnilla tarkoitetaan] henkilötietojen käsittelemistä siten, että henkilötietoja ei voida enää yhdistää tiettyyn rekisteröityyn käyttämättä lisätietoja, edellyttäen että tällaiset lisätiedot säilytetään erillään ja niihin sovelletaan teknisiä ja organisatorisia toimenpiteitä, joilla varmistetaan, ettei henkilötietojen yhdistämistä tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön tapahdu”.

rajapintana, joka mahdollistaa datan tietoturvallisen ja lisäarvoa tuottavan käytön tekoälyä varten.

Osallistuminen on vapaaehtoista, ja datan haltijat päättävät, miten ja milloin dataa käytetään ja kuka sitä voi käyttää. Tietoja ei siirretä ilman nimenomaista suostumusta. Kaikki toimet suojataan luottamuksellisuutta turvaavilla tiukoilla suojatoimilla, ja niitä tuetaan yksityisyyden suojan varmistavilla ja hajautetuilla tekniikoilla, joihin kuuluvat muun muassa federoitu oppiminen, homomorfinen salaus ja monenkeskinen laskenta. Dataa voidaan käsitellä paikallisesti tai solmupisteiden välillä ilman sulauttamista yhteen datavarastoon. Näin voidaan varmistaa, että data pysyy alkuperäisen haltijan valvonnassa. Tämä malli, josta erityisesti pk-yritykset hyötyvät, tukee EU:n tietosuojasääntöjen noudattamista, turvaa luottamuksellisuutta ja auttaa rakentamaan luottamusta samalla kun datan käyttöä laajennetaan tekoälyn kehittämistä varten.

EU:n laskentakapasiteetti on kehittynyt EuroHPC-yhteisyrityksen puitteissa toteutetusta tieteeseen suuntautuneesta suurteholaskennasta tekoälytehtaisiin, jotka laajentavat tätä toiminta-ajatusta tekoälyn kehittämisen edistämiseksi yhdistämällä laskentainfrastruktuurin datan saatavuuteen ja kokeiluihin. Tulevat tekoälyn gigatehtaat laajentavat tekoälylaskentalaitteistoja edelleen.

Osana näitä toimia perustetaan ensimmäiset datalaboratoriot tekoälytehdasaloitteen ja EuroHPC:n puitteissa. Ne tarjoavat tietoturvallisia ympäristöjä ja datapalveluja, joiden avulla tekoälyn kehittäjät voivat käyttää yhteisiä eurooppalaisia data-avaruuksia esimerkiksi terveydenhuollon, valmistuksen ja ilmaston kuin myös kielten, kyberturvallisuuden ja kulttuuriperinnön aloilla. Sen varmistamiseksi, että datalaboratorioiden palvelut tavoittavat yritykset ja julkishallinnot, ne tekevät tiivistä yhteistyötä eurooppalaisten digitaali-innovointikeskittymien kanssa, jotka toimivat käyttäjille suunnattuina yhteyspisteinä ja auttavat liittämään datatarpeet konkreettisiin sovelluksiin.

Datalaboratorioita perustetaan lisää itsenäisesti myös muille aloille, jotta voidaan vastata alakohtaisiin, esimerkiksi energia-alaa koskeviin, erityis- tai tutkimustarpeisiin. Tulevat tekoälyn gigatehtaat toimivat laajennettuina tekoälylaskentalaitoksina, ja niiden avulla valmistellaan datalaboratoriomallin kaupallista käyttöönottoa kaikkialla EU:ssa muuttamalla se omavaraiseksi palveluekosysteemiksi, joka yhdistää laskenta-, data- ja tekoälyinnovaatiot.

**Datalaboratoriot** tarjoavat palveluja erityisesti yhdeksällä keskeisellä osa-alueella:

- **Data-avaruuksien ja tekoälyekosysteemien välinen silta:** käytännön yhteydet, joiden avulla yritykset voivat saada laadukasta ja yhteentoimivaa dataa ja jotka tuovat yhteiset eurooppalaiset data-avaruuksien tekoälyn kehittäjien, infrastruktuurien ja alakohtaisten ekosysteemien saataville.
- **Tekninen infrastruktuuri ja välineet:** datasäiliöt datan tehokkaaseen tallentamiseen ja järjestämiseen ja niitä täydentävät tietoturvalliset ympäristöt, jotka mahdollistavat arkaluonteisten tietojen käsittelyn datalähteellä, käyttövalmiit välineet datan esikäsittelyyn sekä yksityisyyttä suojaavat tekniikat anonymisointia ja synteettisen

datan tuottamista varten. Käyttökelpoisuutta, nopeutta ja laajennettavuutta koskevilla tiukoilla vaatimuksilla varmistetaan, että välineet ovat yksinkertaisia, luotettavia ja helppokäyttöisiä.

- **Datan yhdistäminen:** yrityksiä tuetaan julkisista ja rajoitetuista lähteistä peräisin olevan datan – erityisesti innovatiivisiin tarkoituksiin käytettävän datan – yhdistämisessä yhteisiin eurooppalaisiin data-avaruuksiin kuuluvien luotettavien datanjakamismekanismien avulla. Datalaboratoriot auttavat yrityksiä varmistamaan, että datan vaihtaminen tai yhdistäminen tapahtuu EU:n kilpailulainsäädännön mukaisesti. Komissio tukee datalaboratorioita tässä roolissa hyödyntämällä ja täydentämällä yrityksille kohdennettuja horisontaalisia suuntaviivoja, jotka sisältävät yhteistyötä ja yhteisiä resursseja koskevia käytännön ohjeita, ja antamalla erityisohjeita parhaista käytännöistä tietojen vaihdossa ja yhdistämisessä. Lisäksi yksittäiset datalaboratoriot voivat pyytää räätälöityjä lisäohjeita epävirallisia ohjeita koskevan tiedonannon mukaisella pyynnöllä.
- **Pseudonymisointi- ja anonymisointipalvelut:** kehittyneiden välineiden ja asiantuntemuksen tarjoaminen henkilötunnisteiden poistamiseen tai peittämiseen. Näihin palveluihin sisältyvät muun muassa pseudonymisointia, anonymisointia ja differentiaalista yksityisyyttä koskevat tekniikat, jotka mahdollistavat datan tietoturvallisen uudelleen käytön ja samalla säilyttävät sen analyttisen käytettävyyden.
- **Synteettisen tiedon tuottaminen:** tuetaan todellisen datan tilastollisia ominaisuuksia jäljittelevien laadukkaiden synteettisten data-aineistojen luomista samalla kun varmistetaan arkaluonteisten tai luottamuksellisten tietojen tietosuoja. Datalaboratoriot tarjoavat välineitä ja asiantuntemusta synteettisen datan tuottamiseksi, validoimiseksi ja vertailemiseksi tekoälymallien kouluttamista ja testausta varten, täydentävät anonymisointia ja parantavat datan saatavuutta arkaluonteisilla aloilla.
- **Datan kuratointi, tunnisteiden lisääminen ja vektorointi:** kattava tuki data-aineistojen puhdistamista, tunnisteiden lisäämistä, annotointia, rikastamista ja vektorointia varten, jotta saadaan luotettavaa, edustavaa ja käyttökelpoista dataa tekoälyn kouluttamista varten. Tähän sisältyvät laadunvarmistusprosessit, avoin dokumentointi ja alakohtaisia tunnisteita koskeva yhteistyö asiantuntijayhteisöjen kanssa.
- **Sääntelyyn liittyvä ohjaus ja koulutus:** räätälöity neuvonta, jolla autetaan yrityksiä noudattamaan EU:n lainsäädäntöä, sekä tekoälyn kehittäjille kohdennettu koulutus datan käytöstä ja oikeudellisista velvoitteista, kuten tekoälyä koskevista säännöksistä, tekijänoikeuksista, liikesalaisuuksista ja kilpailulainsäädännöstä.
- **Data-avaruuksien ja tekoälyekosysteemien välinen silta:** käytännön yhteydet, joiden avulla yritykset voivat saada laadukasta ja yhteentoimivaa dataa ja jotka yhdistävät yhteiset eurooppalaiset data-avaruudet tekoälyn kehittäjiin, infrastruktuureihin ja alakohtaisiin ekosysteemeihin.

**Datan saatavuuden helpottaminen:** kysyntälähtöinen palvelu, jossa startup- ja pk-yritykset voivat viestiä datatarpeistaan ja jossa datalaboratoriot auttavat niitä löytämään niiden tarpeisiin vastaavia data-aineistoja ja voittamaan markkina-, oikeudelliset tai hallinnolliset esteet.

---

### *Miten datalaboratorio toimisi käytännössä?*

*Jäsenvaltiossa X toimiva yritys kehittää tekoälyyn perustuvia ennakoivia huoltojärjestelmiä sähköajoneuvoja varten, mutta sillä on vaikeuksia saada riittävästi laadukasta anturidataa eri automalleista ja latausinfrastruktuureista. Yksittäiset valmistajat epäröivät jakaa näitä tietoja liikesalaisuuksiin, yksityisyyden suojaan ja kilpailusääntöihin liittyvien näkökohtien vuoksi. Tekoälytehtaat tarjoavat tarvittavat laskentaresurssit ja integroitujen datalaboratorioidensa kautta datanhallintapalvelut, jotka auttavat näiden esteiden poistamisessa.*

*Datalaboratorion kautta yritys saisi käyttöönsä luotettavia, anonymisoituja ja yhdistettyjä data-aineistoja, jotka ovat peräisin eri lähteistä, kuten julkisilta latausoperaattoreilta ja osallistuvilta alkuperäisiltä laitevalmistajilta, sekä muuta liikkumisdata-avaruuden kautta löydettyä dataa.*

*Tekoälytehtaan osana toimiva datalaboratorio tarjoaisi*

- turvalliset ympäristöt reaaliaikaisten anturitietojen analysoimiseksi federoidun oppimisen avulla ilman, että tiedot lähtevät alkuperäisen laitevalmistajan järjestelmistä*
- anonymisointipalvelut, joilla varmistetaan, että kuljettajaan ja ajoneuvoon liittyvän datan käyttö on yksityisyyden suoja koskevien vaatimusten mukaista*
- sääntelyyn liittyvät ohjeet dataan pääsyä koskevien datasäädöksen säännösten soveltamisesta ja liikesalaisuuksien suojaamisesta*
- datan kuratointivälineet, joilla yhdenmukaistetaan erilaisia anturimuotoja ja laatustandardeja.*

*Tällä tavoin laboratorio toimisi siltana liikkumisdata-avaruuden ja tekoälyekosysteemin välillä, tarjoaisi yritykselle mahdollisuuden kouluttaa vakaita tekoälymalleja ja turvaisi samalla valmistajien liiketoimintaa suojaavan luottamuksellisuuden.*

---

### **iii. Pilvipalveluja ja tekoälyä edistävä säädös**

Kestävä datakeskuskapasiteetti sekä itsenäiset pilvi- ja tekoälypalvelut ovat edellytys sille, että EU voi saavuttaa tässä strategiassa asetetut tavoitteet. Koska dataa tuotetaan yhä enemmän, myös tarve kerätä, tallentaa, yhdistää ja käsitellä dataa kasvaa. Jotta voidaan minimoida viiveet<sup>22</sup> ja vähentää riippuvuutta muualla maailmassa sijaitsevasta infrastruktuurista, EU:ssa on oltava riittävä datakeskuskapasiteetti.

Jotta voidaan varmistaa kestävä datakeskusinfrastruktuuri sekä itsenäiset pilvi- ja tekoälypalvelut EU:n yrityksille ja julkishallinnoille, komissio aikoo ehdottaa vuoden 2026

---

<sup>22</sup> Viiveellä tarkoitetaan aikaa, joka kuluu, kun data siirtyy verkon pisteestä toiseen.

ensimmäisellä neljänneksellä **pilvipalveluja ja tekoälyä edistävää säädöstä**. Tällä aloitteella tuetaan innovointia pilvipalvelujen ja tekoälyn koko arvoketjussa huippuprosessorien integroinnista kestäviin jäähdytysteknologioihin sekä tekoälylaitteisiin ja -ohjelmistoihin. Sillä myös nopeutetaan kestävä datakeskuskapasiteetin käyttöönottoa ja varmistetaan, että EU:lla on tarvittava infrastruktuuri tietoturvallisia ja itsenäisiä pilvi- ja tekoälypalveluja varten.

#### **iv. Strategiset data-aineistot: julkisen sektorin, tieteelliset, kulttuuriset ja kielelliset resurssit**

EU:n kilpailukyky tekoälyn alalla riippuu laadukkaan, jäsennellyn ja luotettavan datan saatavuudesta. Tieteelliset, kulttuuriset ja kielelliset data-aineistot ovat ratkaisevan tärkeitä tekijöitä, jotka mahdollistavat vakaat tekoälymallit, tieteelliset läpimurrot ja teknologisen itsenäisyyden.

Avointa dataa koskevan direktiivin mukaisia julkisen sektorin hallussa olevia viitetietoaineistoja on tarkoitus laajentaa. Arvokkaat tietoaineistot<sup>23</sup> on asetettava saataville maksutta ohjelmointirajapintojen kautta koneellisesti luettavassa ja tarvittaessa massaladattavassa muodossa. Komissio ehdottaa vuonna 2026 arvokkaiden tietoaineistojen luettelon laajentamista kattamaan muun muassa oikeudelliset, tuomioistuini- ja hallinnolliset tiedot. Tästä hyötyvät erityisesti startup- ja pk-yritykset. Komissio seuraa myös, onko uusien data-aineistojen lisäämiselle tarvetta.

Tieteelliset tiedot ovat jo osoittautuneet muutosvoimaisiksi, kuten AlphaFold-järjestelmän<sup>24</sup> perusteella voidaan todeta. Hyvin jäsennellyt tietokannat vähentävät tutkimus- ja kehityskustannuksia, nopeuttavat innovointia ja avaavat uusia mahdollisuuksia materiaali-, lääke-, energia- ja bioteknologia-aloille. Tämän pohjalta komissio jatkaa nykyisten tietokantojen kartoittamista, asettaa prioriteetteja asiantuntijoiden kanssa, turvaa käyttöoikeudet ja rahoittaa uusia digitaalisia infrastruktuureja tutkimus- ja teknologiainfrastruktuuria koskevan eurooppalaisen strategian mukaisesti. Tältä osin **eurooppalaiset avoimen tieteen pilvipalvelut**, jotka muodostavat yhteisen eurooppalaisen tutkimuksen ja kehityksen data-avaruuden, ovat kehittämässä datavarastoa luotettavalle alustalle, jonka avulla voidaan jakaa ja uudelleenkäyttää laadukasta, löydettävissä olevaa, saatavilla olevaa, yhteentoimivaa ja uudelleenkäytettävissä olevaa (findable, accessible, interoperable and reusable, FAIR) tutkimusdataa, -välineitä ja -palveluja yli tieteenalojen ja rajojen Euroopassa. Tämä tukee tekoälyyn liittyvää tieteellistä toimintaa eurooppalaisen tekoälyn tutkimusneuvoston (RAISE) puitteissa.<sup>25</sup> Samaan aikaan tuleva ehdotus

---

<sup>23</sup>Avointa dataa koskevan direktiivin liitteen I mukaisesti arvokkaat tietoaineistot ovat peräisin seuraavista luokista: paikkatiedot, maan havainnointi ja ympäristö, säätiedot, tilastot, yritys- ja yritysten omistustiedot ja liikkuvuustiedot. Uusia luokkia voidaan lisätä.

<sup>24</sup> AlphaFold on Deep Mindin kehittämä tekoälyjärjestelmä, joka ennustaa proteiinirakenteita syväoppimisen ja suurien datamäärien avulla. Tämä auttaa nopeuttamaan läpimurto tutkimusta monilla biologian aloilla.

<sup>25</sup> Euroopan komissio (2025), komission tiedonanto Euroopan parlamentille ja neuvostolle, Eurooppalainen tekoälystrategia tieteessä: Edellytysten luominen eurooppalaiselle tekoälyn tutkimusneuvostolle (RAISE), Bryssel, 8.10.2025, COM(2025) 724 final.

eurooppalaista tutkimusaluetta (ERA) koskevaksi säädökseksi<sup>26</sup> vahvistaa tutkimuskäyttöön tarkoitettujen julkisrahoitteisten tutkimustulosten, julkaisujen ja data-aineistojen jakamisen, saatavuuden ja uudelleenkäytön oikeudellista kehystä.

Myös EU:n kulttuuri- ja kieliresursseja laajennetaan. Europeana-aloitteen puitteissa Euroopan kulttuurilaitoksilta saatua yli 30:tä miljoonaa digitoitua teosta tullaan hyödyntämään tekoälyn kehittämisessä.<sup>27</sup> Komissio myös tutkii tapoja vahvistaa yleisradioyhtiöiden ja tekoälyn tarjoajien välistä yhteistyötä ja edistää lisensointia, jotta yleisradioyhtiöiden audiovisuaalisia arkistoja voitaisiin käyttää tekoälyn kouluttamiseen, ottaen huomioon oikeudenhaltijoille maksettavat korvaukset.

**Eurooppalaisen kielidata-avaruuden ja kieliteknologian allianssin (ALT-EDIC)** pilottihankkeissa joukkoistetaan alakohtaisten data-aineistojen hankintaa, myös pienemmistä kielistä. Saadut aineistot täydentävät jo saatavilla olevaa 477 miljardin tokenin datavarantoa, joka vastaa suuruudeltaan merkittävimpiä tekoälyn koulutusdata-aineistoja. Tämä auttaa varmistamaan, että myös harvinaisia kieliä käytetään suurten kielimallien kehittämisessä, mikä edistää tekoälyjärjestelmien näillä kielillä tuottamien tulosten laatua.

#### **v. Horisontaaliset tekijät: synteettinen data, datan yhdistäminen ja standardit**

Lippulaivahankkeiden ohella EU tarvitsee myös horisontaalisia toimenpiteitä, jotka ulottuvat eri aloille ja auttavat laajentamaan koko datataloutta.

*Synteettinen data auttaa saavuttamaan johtoaseman tekoälyn alalla*

Synteettinen data<sup>28</sup> voi mahdollistaa tekoälyn kouluttamisen aloilla, joilla dataa on niukasti saatavilla tai se on arkaluonteista. Tällaisia aloja ovat muun muassa harvinaisten sairauksien tutkimus, robotiikka ja autonomisen ajamisen reunatapaukset. Synteettinen data mahdollistaa tekoälymallien kehittämisen ilman, että henkilö- tai omistusoikeuden suojaamat tiedot paljastuvat, ja tällä tavoin edistää sekä kilpailukykyä että yksityisyyden suojan säilyttävää innovointia.

Tämän potentiaalin hyödyntämiseksi komissio laatii ohjeita ja standardeja synteettisen datan luotettavaa käyttöä varten, tarkastelee asiaan liittyviä oikeudellisia kysymyksiä, konsultoi vapaaehtoista eurooppalaista sertifiointijärjestelmää ja tutkii, voitaisiinko perustaa ”synteettinen datatehdas”, joka mahdollistaisi laajamittaisia data-aineistoja tuottavan suurteholaskennan. Synteettisen datan tuotantotekniikoita koskevaa huipputason tutkimus- ja kehitystyötä tuetaan myös Horisontti Eurooppa -ohjelmasta saatavalla rahoituksella.

---

<sup>26</sup> Euroopan komissio, tulossa oleva ehdotus eurooppalaista tutkimusaluetta (ERA) koskevaksi säädökseksi, ilmoitettu komission työohjelmassa 2025, Bryssel, 11.2.2025, saatavilla osoitteessa <https://commission.europa.eu/strategy-and-policy/strategy-documents/commission-work-programme/commission-work-programme-2025>.

<sup>27</sup> Kulttuuriperinnön digitaalinen alusta Europeana: <https://www.europeana.eu/fi> (avattu 27.10.2025).

<sup>28</sup> Ks. edellä oleva määritelmä.

## *Strategisen datan yhdistämisen edellytysten selvittäminen*

Monilla yrityksillä esimerkiksi terveydenhuollon, liikkuvuuden, energian, maatalouden ja valmistusteollisuuden aloilla ei ole mahdollisuutta käyttää suuria ja monipuolisia data-aineistoja, joita tarvitaan kehittyneiden tekoälymallien kouluttamiseen. Yhdistämällä tuotteiden ja palvelujen tuotantosyklin alkuvaiheisiin liittyvää dataa voitaisiin tuottaa yhteisiä etuja, mutta oikeudellinen epävarmuus ja pelko kilpailulainsäädännön rikkomisesta jarruttavat yhteistyötä.

Komissio jatkaa toimia, jotta yrityksille voidaan tarjota selkeät oikeudelliset puitteet ja jotta voidaan siirtyä säännöistä tuloksiin, kuten Euroopan kilpailukyvyyn tulevaisuutta koskevassa Draghin raportissa kehotetaan. Kilpailijoiden välisiä yhteistyösopimuksia koskevissa vuoden 2023 horisontaalisissa suuntaviivoissa on jo selitetty, milloin datan yhdistäminen on EU:n kilpailulainsäädännön mukaista. Niissä myös annetaan käytännön esimerkkejä ja tietoa suojatoimien soveltamisesta.

---

*Draghin raportti: ”EU:n suurten data-aineistojen puutetta voitaisiin paikata siten, että mallien koulutuksessa hyödynnettäisiin dataa, jota useat yksittäisen sektorin EU-yritykset luovuttavat vapaasti yhteiseen käyttöön. Järjestelyä tukisivat avoimen lähdekoodin kehitykset, ja se suojattaisiin kilpailuviranomaisten kilpailuoikeudelliselta valvonnalta.”*

---

Helpottaakseen edelleen datalaboratorioiden kautta toteutettavaa laillista ja toimivaa datayhteistyötä komissio antaa erityisiä ohjeita parhaista käytännöistä datan vaihtamisessa ja yhdistämisessä.

Lisäksi komissio voi epävirallisia ohjeita koskevan tiedonannon mukaisesti antaa pyynnöstä kilpailulainsäädäntöä koskevia ohjeita erityisiin dataan liittyviin monikansallisiin hankkeisiin ja aloitteisiin, jotka edistävät rajat ylittävää innovointia, teollisuuden häiriönsietokykyä ja tekoälyn kehittämistä. Tekemällä datan yhdistämisestä luotettavan ja oikeudellisesti turvallisen vaihtoehdon EU voi saada aikaan tehokkuusetuja ja nopeuttaa läpimurtoja keskeisillä aloilla.

### *Korkeammat standardit datan laadulle ja keräämiselle*

Ilman luotettavia standardeja kaikkein kunnianhimoisimpiakin datan yhteiskäyttöaloitteita uhkaa sirpaloitumisen ja vähäisen käyttönoton riski. Luotettavaa dataa koskevassa EU:n kehityksessä<sup>29</sup> on jo vahvistettu säännöt yhteiskäytölle, metadatalle ja hallinnolle, mutta lisätoimet ovat tarpeen, jotta esiin nousevat ongelmat voidaan ratkaista.

Komissio antaa standardointitoimeksiannon, jotta voidaan laatia **datan eurooppalaiset laatustandardit**, jotka kattavat täydellisuuden, johdonmukaisuuden, alkuperän, semanttisen selkeyden ja hallinnon. Standardit antavat yrityksille, sääntelyviranomaisille ja tutkijoille yhteiset vertailuarvot luotettavia data-aineistoja varten. Näillä toimilla täydennetään käynnissä olevia tekoälysäädöksen mukaisia datan laatua ja dokumentaatiota koskevia

---

<sup>29</sup> Ks. Euroopan komissio, Implementing Decision C(2025) 4135 on the European Trusted Data Framework (ei saatavana suomeksi).

standardointitoimia ja varmistetaan johdonmukaisuus datanhallintaa ja tekoälyn kehittämistä koskevien vaatimusten välillä.

Lisäksi käynnistetään aloite, jonka tarkoituksena on mahdollistaa **annotointi- ja tunnistekäytäntöjen** standardointi ja tällä tavoin helpottaa datan löytämistä, yhdistämistä ja uudelleenkäyttöä ja samalla varmistaa luottamus datan alkuperään ja käyttöolosuhteisiin, mikä on ratkaisevan tärkeää tekoälyn kouluttamisen laajentamisen ja monialaisen uudelleenkäytön kannalta. Sidosryhmien välisessä työpajassa tarkastellaan myös standardeja, jotka koskevat datan keräämistä verkkoon liitetystä tuotteista, antureista ja kameroista – mukaan lukien otanta, metadata, aikaleimaus, kalibrointi ja eheys – ja puututaan tietojen tehokkaan yhdistämisen ja uudelleenkäytön keskeiseen esteeseen.

### **Lippulaivat**

- Käynnistetään ensimmäiset **datalaboratoriot**, joiden tehtävänä on laajentaa datan saatavuutta ja jotka yhdistetään tekoälyekosysteemeihin (vuoden 2025 viimeinen neljännes). Ne tarjoavat myös luotettavia pseudonymisointipalveluja.
- **Käynnistetään tekoälyä varten tarvittavaa laadukasta dataa koskeva aloite:** laajennetaan arvokkaita tietoaaineistoja avointa dataa koskevan direktiivin mukaisesti (vuoden 2026 viimeinen neljännes), perustetaan sidosryhmäfoorumi yleisradioyhtiöiden ja tekoälyn kehittäjien kanssa (vuoden 2026 toinen neljännes), asetetaan saataville 30 miljoonaa digitoitua kulttuurisinetä tekoälyn kouluttamista varten (vuoden 2026 viimeinen neljännes) ja käynnistetään joukkoistamista koskeva aloite alakohtaisen datan ja pienempien eurooppalaisten kielten kielidatan hankintaa varten (vuoden 2026 toinen neljännes).

## **Pilari II: Virtaviivaistetaan datasääntöjä**

EU:n datakehityksen on oltava selkeä, käytännöllinen ja innovoinnille suotuisa. Hallinnollisten rasitteiden keventämiseksi ja kilpailukyvyyn parantamiseksi komissio esittää **digitaalialan koontiasetuksen** eli lainsäädäntöehdotuksen, jolla pyritään muun muassa nykyaikaistamaan ja lujittamaan EU:n horisontaalista dataa koskevaa säännöstöä. Lisäksi komissio ilmoittaa toimista, jotka koskevat vaatimustenmukaisuuden tarkastamisessa käytettävää ns. yhden klikkauksen periaatetta (*one-click compliance*) eli automaattisoitua lakisäätteistä raportointia, sekä toimista, jotka koskevat datasäädöksen tukipakettia, joka sisältää vakiosopimuksia, vakiolausekkeita, korvauksia ja liikesalaisuuksia koskevia ohjeita, sekä pk-yritysten oikeudellista neuvontapalvelua.

### **i. Yksinkertaistetaan EU:n datasäännöstöä**

EU:n datasäännöstö on kasvanut nopeasti. Sen puitteissa on luotu uusia oikeuksia, mutta samalla myös monimutkaisuus ja pirstaleisuus ovat lisääntyneet. Yksinkertaistaminen on tarpeen, jotta voidaan vähentää sääntöjen noudattamisesta aiheutuvia kustannuksia, helpottaa sääntöjen soveltamista ja tukea innovointia entistä paremmin.

Tätä varten komissio esittää edellä mainitun digitaalialan koontiasetuksen. Sillä päivitetään säännöstöä ja poistetaan tarpeettomat rasitteet samalla kun turvataan EU:n datatalouden keskeiset periaatteet. Koontiasetuksessa keskitytään seuraaviin ensisijaisiin uudistuksiin:

- **Poistetaan vanhentuneet säännöt.** Koontiasetuksella kumotaan muiden kuin henkilötietojen vapaata liikkuvuutta koskeva asetus<sup>30</sup>, koska sen tehtävät kuuluvat jo datasäädöksen soveltamisalaan, samalla kun nimenomaisesti säilytetään muiden kuin henkilötietojen vapaan liikkuvuuden periaate ja perusteetonta datan säilytyspaikkaa koskeva kiello.
- **Virtaviivaistetaan datan jakamista koskevia sääntöjä.** Koontiasetuksella kumotaan datanhallinta-asetus ja siirretään sen keskeiset säännökset datasäädökseen. Datan välityspalveluiden tarjoajia koskevista velvoitteista tehdään selkeämpiä, kevyempiä ja vapaaehtoisia toteuttamiskelpoisten mallien ja laajemman käyttöönnoton edistämiseksi.
- **Lujitetaan datan jakamista koskevaa kehystä julkisen sektorin osalta.** Säännöt, jotka jakautuvat tällä hetkellä datanhallinta-asetuksen ja avointa dataa koskevan direktiivin kesken, säilytetään ja yhdistetään yhdeksi datasäädöksen luvuksi. Tällä tavoin voidaan yksinkertaistaa velvoitteita ja samalla säilyttää avoimuus, läpinäkyvyys ja oikeudenmukainen saatavuus. Uudella kehyksellä puututaan myös vallan epätasapainoon datan jakamisessa ja varmistetaan oikeudenmukaiset olosuhteet ja konkreettiset hyödyt pk-yrityksille. Datalaboratoriot ilmoittavat lupaavista uusista julkisen sektorin data-aineistoista, joita säännöt eivät vielä kata.
- **Nyky aikaistetaan evästeitä ja samankaltaisia teknologioita koskevia sääntöjä.** Koontiasetuksella uudistetaan evästeitä koskevia sääntöjä, jotka sisältyvät tällä hetkellä sähköisen viestinnän tietosuoja-direktiiviin, ja tuodaan ne yleisen tietosuoja-asetuksen kehykseen. Siinä ehdotetaan käytännöllisiä ratkaisuja, kuten sitä, että tiettyihin vähäriskisiin tarkoituksiin käytetyt evästeet ja samankaltaiset teknologiat olisi katsottava laillisiksi, kun taas muunlaisissa tarkoituksissa käytön olisi perustuttava johonkin yleisen tietosuoja-asetuksen mukaisista oikeusperustoista. Lisäksi evästeistä kertoviin ilmoituksiin liittyviä sääntöjä yksinkertaistetaan ottamalla käyttöön yhden klikkauksen vaihtoehto. Se velvoittaa verkkosivustot ottamaan käyttäjän tekemän valinnan huomioon myös selainten kautta. Digitaalisen koontiasetuksen lisäksi sähköisen viestinnän tietosujakehystä uudistetaan sen varmistamiseksi, että nykyiset säännöt vastaavat nykypäivän tarpeita ja mahdollistavat ihmisten ja yritysten tehokkaan

---

*Dataunionia koskevan julkisen kuulemisen osanottaja: ”Sähköisen viestinnän tietosujasäännöt on saatettava pikaisesti ajan tasalle. Voimassa olevat säännöt on laadittu täysin erilaisessa teknologisessa kontekstissa eivätkä vastaa markkinoiden nykyisiä tarpeita.”*

---

<sup>30</sup> Euroopan parlamentin ja neuvoston asetus (EU) 2018/1807, annettu 14 päivänä marraskuuta 2018, muiden kuin henkilötietojen vapaan liikkuvuuden kehyksestä Euroopan unionissa.

suojelun. Samalla varmistetaan, etteivät perusoikeudet vaarannu, ja säilytetään journalismin riippumattomuus. Keskeiset säännökset sisällytetään muihin säädöksiin, jolloin direktiivi voidaan peruuttaa.

- **Luodaan innovointia edistävä tietosuojakehys. Yleisen tietosuoja-asetukseen** tehtävillä kohdennetuilla muutoksilla selvennetään erityisesti henkilötietojen määritelmää, yhdenmukaistetaan EU:n tasolla, milloin tietosuoja koskevia vaikutustenarviointeja olisi tehtävä, yksinkertaistetaan valvontaviranomaisille tehtäviä ilmoituksia tietoturvaloukkauksista, virtaviivaistetaan tietoturvaloukkauksista ilmoittamista ottamalla käyttöön yksi yhteispiste EU:ssa ja yksinkertaistetaan ilmoitusvelvoitteita tapauksissa, joissa on perusteltu syy olettaa, että yksityishenkilöillä on jo tiedot, ja rekisteröidylle aiheutuva riski on vähäinen. Lisäksi selvennetään, että oikeutettu etu voi olla oikeusperusta tekoälyn kouluttamiselle, kattaen erityisten datakategorioiden satunnaisen käsittelyn, ja selkeytetään automatisoituja yksittäispäätöksiä koskevia säännöksiä.

Yksi keskeinen muutos koskee luotettavan anonymisoinnin käyttämistä tekoälyä varten tarvittavan datan vapauttamiseksi. Henkilötietojen riittävää anonymisointia koskeva epävarmuus on usein pääasiallinen syy, joka estää datan jakamisen. Yrityksillä on erityisesti vaikeuksia määrittää, milloin pseudonymisoitua dataa ei enää pidetä henkilötietoina tiettyjen yhteisöjen osalta. Tämä epävarmuus tekee tietojen jakamisesta monimutkaisempaa silloin, kun yleisen tietosuoja-asetuksen vaatimuksia noudatetaan varotoimenpiteenä. Komissio tukee yrityksiä määrittämällä keinot ja kriteerit sen määrittämiseksi, onko pseudonymisoinnista saatu data henkilötietoja tiettyjen yhteisöjen osalta.

Tässä yhteydessä arvioidaan käytettävissä olevien tekniikoiden viimeisintä kehitystä ja laaditaan kriteerit uudelleentunnistamisen riskin arvioimiseksi. Vaikka yritykset ovat edelleen täysin vastuussa yleisen tietosuoja-asetuksen noudattamisesta, ne voivat käyttää näitä keinoja ja kriteerejä osoittaakseen, että tiedot eivät voi johtaa uudelleentunnistamiseen. Muutoksilla helpotetaan myös tekoälymallien kouluttamista. Samalla sovelletaan asianmukaisia suojatoimia. Näiden muutosten tavoitteena on selkeyttää tekoälyn kehittämisen oikeudellista kehystä, myös arkaluonteisten tietojen satunnaisen käsittelyn osalta tapauksissa, joissa kehittäjät ovat aidosti pyrkineet poistamaan tällaista dataa, ja samalla suojella yksilöiden oikeuksia ja yritysten kilpailukykyä.

- **Hiotaan datasäädöstä käytännön täytäntöönpanoa varten. Datasäädöksen** keskeisin sisältö pysyy muuttumattomana. Yritysten ja viranomaisten välinen datan yhteiskäyttö rajataan hätätilanteisiin. Tällä tavoin kevennetään hallinnollista taakkaa ja samalla turvataan kriisitoimet. Kohdennetuilla lisämuutoksilla estetään datavuodot EU:n ulkopuolelle, otetaan käyttöön räätälöityjä järjestelmiä pilvipalveluja varten ja poistetaan älysovimuksia koskevat säännökset.
- **Kevennetään yritysten laajentumiseen liittyviä rasitteita.** Pk-yrityksiä koskeviin datasäädöksen ja avointa dataa koskevan direktiivin säännöksiin lisätään uusi pienten

midcap-yritysten (250–749 työntekijää) luokka ja integroidaan datanhallinta-asetuksen määräykset.

## **ii. Luodaan tulevaisuudenkestävä datakehys**

Osana digisääntelyn toimivuustarkastusta komissio jatkaa EU:n datasäännösten tarkistamista, jotta se säilyisi johdonmukaisena, oikeasuhtaisena ja innovointia edistävänä. Siinä kiinnitetään erityistä huomiota pk-yrityksiin ja tunnistetaan päällekkäisyydet, puutteet ja epäselvät yhteisvaikutukset, myös alakohtaisten datasäännösten kanssa, jotta voidaan luoda ennakoitavammat monialaiset puitteet.

Lisäksi nykyaikaistetaan digialan lainsäädäntöä ja tietosuojaa.<sup>31</sup> Kohdennettujen mukautusten tarkoituksena on helpottaa vaatimusten noudattamista ja vahvistaa noudattamisen valvontaa ja tällä tavoin tukea elinvoimaisten ja luotettavien innovaatioiden kehittämistä.

Datan välitystoiminnasta on tullut kasvava huolenaihe. Tietyt yritykset keräävät ja kokoavat henkilötietoja ja käyvät kauppaa niillä ilman, että yksityishenkilöt ovat tästä tietoisia tai antaneet tälle merkityksellisen suostumuksensa, tai ilman valvontaa. Tällaiset läpinäkymättömät käytännöt heikentävät tietosuojalainsäädännön ja yksityisyyden peruseriaatteita, vääristävät kilpailua ja heikentävät kansalaisten luottamusta digitaalimarkkinoihin. Nykyisten sääntöjen noudattamisen valvontaa onkin tehostettava. Komissio arvioi, tarvitaanko lisätakeita, jotta voidaan hillitä näitä käytäntöjä, lisätä datakaupan avoimuutta ja varmistaa, että yksityishenkilöt ja yritykset voivat luottaa dataan pääsyä ja sen vaihtoa koskeviin käytäntöihin EU:ssa.

## **iii. Yhden klikkauksen periaate**

Yrityksiltä kuluu merkittävä määrä aikaa ja rahaa vaatimustenmukaisuuden todentamiseen. Myös digitaalisessa muodossa olevaa dataa on usein muotoiltava uudelleen ja toimitettava uudelleen useille viranomaisille, jotka tarkistavat sen manuaalisesti. Tämä päällekkäisyys tuo valvontaan sirpaleisuutta ja vie resursseja pois innovoinnista.

Sääntöjen yksinkertaistamisen lisäksi EU investoi teknologioihin, joilla vaatimustenmukaisuuden tarkastamista voidaan automatisoida. Komissio tukee yhteisiä datamalleja, yhteentoimivuuseriaatteita ja automatisoituja analyyseja Horisontti Eurooppa - puiteohjelmasta ja Digitaalinen Eurooppa -ohjelmasta saatavilla varoilla. Pilottihankkeissa on jo osoitettu, millä tavoin reaaliaikaiset automaattiset vaatimustenmukaisuuden tarkastukset toimisivat käytännössä. Varhainen esimerkki tästä toimintatavasta on tuotelainsäädännön puitteissa käyttöön otettava digitaalinen tuotepassi.

Näiden kokemuksiin perustuva ”yhden klikkauksen periaate” mahdollistaisi vaatimustenmukaisuuden koneellisen todentamisen ja yritysten tietojen muuttamisen standardoiduiksi digitaalisiksi vaatimustenmukaisuustodistuksiksi. Näin se toimisi samaan tapaan kuin digitaalinen tuotepassi, joka mahdollistaa tuotteiden vaatimustenmukaisuuden automaattisen todentamisen.

---

<sup>31</sup> Komission vuotuinen työohjelma [EUR-Lex - 52025DC0870 – FI – EUR-Lex](#).

Yhden klikkauksen periaate voisi olla erityisen arvokas kyberturvallisuuden kaltaisilla aloilla, joilla yrityksiin sovelletaan NIS 2 -direktiivin<sup>32</sup>, kyberkestävyyssäädöksen<sup>33</sup> ja muiden kehysten mukaisia vaatimuksia.

**Eurooppalainen yrityslompakkoasetus** tukee keskeisellä tavalla tätä lähestymistapaa. Se tarjoaa luotettavan ja yhteentoimivan digitaalisen ympäristön todennettavissa olevien valtuustietojen, myös vaatimustenmukaisuustodistusten, tallentamiselle, hallinnoinnille ja jakamiselle. Yritykset voisivat käyttää **eurooppalaisia yrityslompakoita** digitaaliseen tunnistautumiseen, ekosysteemin käyttäjien tunnistamiseen ja validoimiseen sekä vaatimustenmukaisuustodistusten toimittamiseen eri EU:n sääntöjen noudattamisen osoittamiseksi. Julkisen sektorin sääntelyelimille puolestaan tarjotaan tietoturvallinen ja välitön pääsy varmennettuihin tietoihin. Ajan myötä **eurooppalainen yrityslompakko** vakiintuu infrastruktuuriksi, jolla tuetaan hallinnollisia prosesseja, kuten lisensointia, julkisia hankintoja ja rahoituksen saantia ja joka mahdollistaa saumattoman digitaalisen vuorovaikutuksen yritysten ja viranomaisten välillä kaikkialla sisämarkkinoilla.

Luottamuksen ja oikeusvarmuuden takaamiseksi on olennaisen tärkeää määrittää, kuka on vastuussa mahdollisista virheistä, väärinkäytöksistä tai järjestelmän toimintahäiriöistä – riippumatta siitä, onko kyse yhtiöstä, sertifioijasta vai sääntelyviranomaisesta. Sen vuoksi komissio tarkastelee näitä kysymyksiä tulevassa julkisessa kuulemisessa ja arvioi sekä mahdollisuuksia että suojatoimia, joita tarvitaan, jotta voidaan luoda luotettava ja vastuullinen automatisoitu ekosysteemi, joka edistää vaatimusten noudattamista.

Tällainen järjestelmä paitsi vähentäisi pk- ja midcap-yritysten kustannuksia myös antaisi päättäjille tietoa siitä, miten säännöt toimivat käytännössä, ja tällä tavoin edistäisi näyttöön perustuvaa sääntelyä. Yhden klikkauksen periaatteesta voisi tulla EU:n digialan yksinkertaistamisohjelman kulmakivi, jossa kilpailukyky sovitetaan yhteen luottamuksen ja vastuuvollisuuden kanssa.

#### **iv. Autetaan yrityksiä noudattamaan datasäädöstä**

Datan käyttöä ja jakamista koskevat keskeiset säännöt on annettu **datasäädöksessä**. Varmistaakseen, että yritykset, erityisesti pk-yritykset ja pienet midcap-yritykset, pääsevät hyödyntämään koko potentiaaliaan ja keskittymään innovointiin byrokratian sijaan, komissio on jo julkaissut koonnin usein kysytyistä kysymyksistä<sup>34</sup> ja ajoneuvoissa olevia tietoja koskevat ohjeet<sup>35</sup>. Näitä toimia täydennetään myöhemmin laajemmalla tukipaketilla.

---

<sup>32</sup> Euroopan parlamentin ja neuvoston direktiivi (EU) 2022/2555, annettu 14 päivänä joulukuuta 2022, toimenpiteistä kyberturvallisuuden yhteisen korkean tason varmistamiseksi kaikkialla unionissa, asetuksen (EU) N:o 910/2014 ja direktiivin (EU) 2018/1972 muuttamisesta sekä direktiivin (EU) 2016/1148 kumoamisesta (NIS 2 -direktiivi), EUVL L 333, 27.12.2022, s. 80–152.

<sup>33</sup> Euroopan parlamentin ja neuvoston asetukset (EU) 2024/2847, annettu 23 päivänä lokakuuta 2024, digitaalisia elementtejä sisältävien tuotteiden horisontaalisista kyberturvallisuusvaatimuksista ja asetusten (EU) N:o 168/2013 ja (EU) 2019/1020 ja direktiivin (EU) 2020/1828 muuttamisesta (kyberkestävyyssäädös), EUVL L [2847], 20.11.2024.

<sup>34</sup> Euroopan komissio, Frequently Asked Questions – Data Act, versio 1.3, Bryssel, 12.9.2025, saatavilla osoitteessa <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-frequently-asked-questions-about-data-act> (avattu 27.10.2025).

<sup>35</sup> Euroopan komissio, Ohjeistusta ajoneuvodatasta liittyen asetukseen 2023/2854 (datasäädös), C(2025) 6119 final, Bryssel, 12.9.2025.

Välittömästi toteutettavat toimenpiteet:

- datan jakamista koskevat mallisopimusehdot, jotta voidaan vähentää oikeudellista monimutkaisuutta ja kaupankäyntikuluja ja lisätä yritysten luottamusta uusiin kumppanuuksiin
- pilvipalveluja koskevat vakiosopimuslausekkeet, jotta voidaan helpottaa palveluntarjoajan vaihtamista ja tehdä sopimuksista oikeudenmukaisempia ja tällä tavoin tukea kilpailua ja innovointia Euroopan pilvipalvelumarkkinoilla.

Vaiheittain toteutettavia muita toimenpiteitä:

- ohjeet kohtuullisesta korvauksesta sen selventämiseksi, mitä datan jakamisesta voidaan veloittaa, mikä lisää oikeusvarmuutta sekä datan haltijoiden että datan vastaanottajien kannalta (vuoden 2026 ensimmäinen neljännes)
- uudet ohjeet tietyistä datasäädöksen määritelmistä (vuoden 2026 ensimmäinen neljännes)
- datasäädöstä koskeva oikeudellinen tukipalvelu, joka tarjoaa yrityksille suoraa apua uusien sääntöjen konkreettiseen soveltamiseen liittyvissä kysymyksissä ja jossa pk-yrityksille annetaan etusija, jotta niiden tiedusteluihin voidaan vastata nopeasti ja kohdennetusti (vuoden 2025 viimeinen neljännes).

Yhdessä nämä toimenpiteet helpottavat datasäädöksen vaatimusten noudattamista, vähentävät tarpeettomia kustannuksia ja tarjoavat yrityksille selkeyttä ja lisäävät niiden luottamusta, jota ne tarvitsevat voidakseen tarttua uusiin mahdollisuuksiin EU:n datataloudessa. Komissio seuraa tiiviisti sopimusvälineiden, erityisesti mallisopimusehtojen ja vakiosopimuslausekkeiden, käyttöönottoa ja tarvittaessa tarkastelee, täydentää tai muuttaa niitä datan jakamisen kansainväliseen kehitykseen mukautumiseksi.

Komissio pyrkii saamaan aikaan synergioita julkisten ostajien yhteisön ja eurooppalaisten data-avaruuksien välillä julkisen sektorin tehokkuuden parantamiseksi hyödyntäen eurooppalaisen terveysdata-avaruuden ja terveydenhuollon tehokkuutta käsittelevän suurten ostajien työryhmän (Big Buyers Working Group) välillä laadittua suunnitelmaa.<sup>36</sup>

#### **Lippulaivatoimet**

- ehdotus **datalainsäädännön konsolidoimiseksi** (vuoden 2025 viimeinen neljännes)
- ehdotus evästeitä ja samankaltaisia teknologioita koskevien **sähköisen viestinnän tietosuojasääntöjen päivittämisestä** (vuoden 2025 viimeinen neljännes)
- ehdotus **yleisen tietosuoja-asetuksen kohdennetuiksi mukautuksiksi** (vuoden 2025 viimeinen neljännes)

<sup>36</sup> [Can the European Health Data Space enable better procurement? – Big Buyers are investigating | Julkisten ostajien yhteisö.](#)

- vaatimustenmukaisuuden tarkastamisessa käytettävää yhden klikkauksen periaatetta (one-click compliance) koskevan aloitteen käynnistäminen (vuoden 2025 viimeisestä neljänneksestä alkaen)
- datasäädöksen täytäntöönpanoa tukevien toimenpiteiden käyttöönotto (vuoden 2025 viimeisestä neljänneksestä alkaen)

### **Pilari III: Euroopan datasuvereniteetin säilyttäminen kansainväliseen datapolitiikkaan liittyvillä strategisilla toimenpiteillä**

Datasuvereniteetti on EU:n digitaalisen tulevaisuuden ytimessä. Tämä tarkoittaa, että EU:n on edelleen valvottava dataan pääsyä ja sitä, miten sitä käytetään ja suojataan EU:n alueella ja sen ulkopuolella. Suvereniteetti edellyttää avoimuutta luotettavien kumppanien kanssa, mukaan lukien rajat ylittävä tiedonvaihto, mutta oikeudenmukaisin, tietoturvallinen ja EU:n arvojen ja etujen mukaisin ehdoin. Tilannetta, jossa ulkomaisilla toimijoilla on rajaton pääsy EU:n markkinoille, kun taas EU:n yritykset kohtaavat perusteettomia esteitä EU:n ulkopuolella, ei voida enää hyväksyä.

Suvereniteetin turvaaminen tarkoittaa myös EU:n selviytymiskyvyn suojelua. Kyberhyökkäykset, teknologiavuodot, valvonta ja pakottamiseen perustuvat riippuvuudet vaarantavat kriittisen datan. EU:n on varmistettava arkaluonteisten data-aineistojen saatavuus, eheys ja tietoturvasuus, jotta estetään niiden väärinkäyttö tai hyväksikäyttö, etenkin kun kyse on EU:n ulkopuolisista toimijoista.

Tätä varten komissio toteuttaa strategiaa, joka perustuu sekä avoimuuteen että vahvuuteen. Sen puitteissa datan saatavuutta ja rajat ylittäviä siirtoja koskevista oikeudenmukaisista edellytyksistä tehdään digitaalisen kaupankäynnin pilari, EU:sta peräisin olevia arkaluonteisia muita kuin henkilötietoja suojellaan selkeillä suojatoimilla ja yhteistyötä luotettavien kumppanien kanssa syvennetään. Lisäksi komissio pyrkii luomaan maailmanlaajuisia hallintomalleja, joissa otetaan huomioon EU:n edut ja arvot ja estetään pirstoutuminen kilpaileville aloille. Tämä strategia täydentää EU:n tietosuojasäännösten puitteissa kehitettyä pitkäkestoista lähestymistapaa henkilötietojen tietoturvasiin siirtoihin.

---

*Sidosryhmäkyselyssä 75 prosenttia osallistujista kannatti sitä, että EU ottaa määrätietoisan lähestymistavan muiden kuin henkilötietojen kansainvälisiin siirtoihin.*

---

Vaikka EU on luonut vankan oikeudellisen kehyksen ja edistänyt datan vapaata ja turvallista liikkuvuutta kansainvälisesti, uudet perusteettomat datan säilytyspaikkaa koskevat vaatimukset, vientivalvonta ja syrjivät säännöt EU:n ulkopuolella uhkaavat heikentää suvereniteettia. Sen vuoksi komissio aikoo jatkossa puolustaa EU:n etuja ja sääntelyä

riippumattomuutta määrätietoisemmin ja toteuttaa oikeasuhtaisia toimenpiteitä, jos avoimuutta käytetään väärin tai haavoittuvuuksia käytetään aseena.

### **i. Oikeudenmukaiset rajat ylittävät datansiirrot ja EU:n arkaluonteisia muita kuin henkilötietoja koskevat suojatoimet**

Komissio ottaa kansainvälisessä digitaalisessa kaupankäynnissä käyttöön oikeudenmukaiset edellytykset ja tehokkaat valvontatoimet rajat ylittäviä datansiirtoja varten. Nykyiseen epätasapainoiseen tilanteeseen, jossa EU:n dataa siirretään EU:n ulkopuolelle ilman asianmukaisia suojatoimia, puututaan jäsennellyllä tietojenvaihdolla, jota toteutetaan muun muassa EU:n digitaalisten kumppanuuksien ja vuoropuhelujen puitteissa.

Jos puutteita esiintyy edelleen, komissio toteuttaa objektiivisiin kriteereihin perustuvia oikeasuhtaisia toimia noudattaen kaikilta osin unionin kansainvälisiä sitoumuksia. Komissio antaa vuoden 2026 toisella neljänneksellä ohjeet sen arvioimiseksi, miten EU:n yhteisöjä kohdellaan kolmansissa maissa. Lisäksi se kehittää vuoden 2026 ensimmäisellä neljänneksellä tietovuotojen torjuntaa varten välineistön, jonka avulla voidaan puuttua datan säilytyspaikkaa koskeviin vaatimuksiin, markkinoilta sulkemiseen, riittämättömiin suojatoimiin tai muuhun perusteettomaan kohteluun. Välineistö voi perustua esimerkiksi kauppasopimusten valvontaa koskevaan asetukseen<sup>37</sup> tai pakottamistoimien vastaiseen välineeseen<sup>38</sup> tai muihin niiden kaltaisiin välineisiin sekä soveltuvien osin taloudelliseen turvallisuuteen liittyviin näkökohtiin. Välineistössä keskitytään teknologioihin ja parhaisiin käytäntöihin, joilla voidaan vahvistaa EU:n häiriönsietokykyä. Jos rakenteellisiin vääristymiin tai jatkuviin syrjiviin käytäntöihin ei puututa, komissio harkitsee tarvittaessa lisätoimenpiteitä, jotta oikeudenmukaiset edellytykset datan saatavuudelle ja käytölle voidaan varmistaa.

Samalla komissio pyrkii suojaamaan EU:n arkaluonteisia muita kuin henkilötietoja aiempaa paremmin, mikä täydentää henkilötietojen suojaa, joka on taattu yleisellä tietosuojasetuksella ja tietosuojan riittävyttä koskevilla päätöksillä. Komissio hyväksyy sidosryhmien kanssa tehtyyn yhteistyöhön ja perusteellisten riskinarviointien tuloksiin perustuvan ensimmäisen kohdennettujen toimenpiteiden paketin vuoden 2026 kolmanteen neljänneksen mennessä.

### **ii. Yhdistetään EU:n datan jakamisen ekosysteemit samanmielisten kolmansien maiden ekosysteemeihin**

Tietosuojaa, kyberturvallisuutta, valvontaa ja oikeussuojakeinoja koskeva EU:n oikeudellinen kehys on luotettava toimintaperusta EU:n ulkopuolisille datan haltijoille. Komissio edistää turvallisia, yhtenäisiä ja yhteentoimivia yhteyksiä EU:n dataekosysteemien ja samanmielisten kumppanien välillä houkutellakseen lisää datavirtoja EU:hun.

---

<sup>37</sup> Euroopan parlamentin ja neuvoston asetus (EU) N:o 654/2014, annettu 15 päivänä toukokuuta 2014, unionin oikeuksien harjoittamisesta kansainvälisten kauppasääntöjen soveltamista ja täytäntöönpanoa varten ja neuvoston asetuksen (EY) N:o 3286/94 muuttamisesta, EUVL L 189, 27.6.2014, s. 50–58.

<sup>38</sup> Euroopan parlamentin ja neuvoston asetus (EU) 2023/2675, annettu 22 päivänä marraskuuta 2023, unionin ja sen jäsenvaltioiden suojelemisesta kolmansien maiden harjoittamalta taloudelliselta pakottamiselta (pakottamistoimien vastainen väline), EUVL L 322, 27.11.2023.

Suunniteltuihin toimenpiteisiin kuuluvat seuraavat: i) tuetaan palveluja ja infrastruktuuria, kuten yhteisiä eurooppalaisia data-avaruuksia, saumattomien rajat ylittävien siirtojen mahdollistamiseksi, ii) tarjotaan vakiosopimuslausekkeiden kaltaisia välineitä, joilla autetaan yrityksiä varmistamaan laillinen vaihto ja iii) sisällytetään rajat ylittävää datan siirtoa koskevat sitoumukset kahden- ja useammankeskisiin kansainvälisiin sopimuksiin.

Komissio pyrkii lisäämään lähentymistä ja yhteentoimivuutta edistämällä luotettavaa dataa koskevaa EU:n kehystä kansainvälisissä vuoropuheluissa ja digitaalisen kumppanuuden verkostossa. Se tutkii myös, voidaanko luoda luotettavuusmerkki, joka mahdollisesti liittyisi data-avaruuksien kypsyysmalliin. Kyseessä olisi standardoitu kehys, joka auttaa arvioimaan data-avaruushankkeiden valmiuksia ja tukee yhteistyötä ulkomailla toimivien hallitusten ja yritysten kanssa.

### **iii. Vahvistetaan EU:n ääntä maailmanlaajuisessa datanhallinnassa**

Kilpailevat datanhallintamallit pirstaloivat maailmanlaajuisesta toimintaympäristöä. Komissio aikoo tehostaa EU:n lähestymistapojen edistämistä kansainvälisellä tasolla, erityisesti uusissa puitteissa, ja vahvistaa liittoumia samanmielisten kumppaneiden kanssa.

Vuoteen 2026 mennessä komissio ja Euroopan ulkosuhdehallinto syventävät ja solmivat kansainvälisen digitaalistrategian<sup>39</sup> mukaisesti datanhallintaan liittyviä kumppanuuksia digitaali-alalla, tekevät yhteistyötä niiden kumppaneiden kanssa, joilla on yhteisiä tavoitteita, ja jatkavat digitaalista kauppaa koskevien sopimusten ja perinteisiin kauppasopimuksiin sisältyvien digiasioita koskevien lukujen kehittämistä. Komissio tekee edelleen aktiivista yhteistyötä G7- ja G20-ryhmien, OECD:n ja YK:n kaltaisilla foorumeilla soveltamalla muun muassa OECD:n julistusta viranomaisten pääsystä henkilötietoihin.

Erityistä huomiota kiinnitetään EU:n lähestymistapojen edistämiseen ja molempia osapuolia hyödyttävään yhteistyöhön ehdokasmaiden, mahdollisten ehdokasmaiden ja lähimpien naapurimaiden kanssa. EU aikoo myös tutkia yhteistyössä kumppaneiden kanssa mahdollisuutta perustaa yhteinen alusta valikoidulle korkean lisäarvon julkiselle tiedolle (esim. kulttuuriperintöä koskeva tieto) ja jatkaa luotettavia järjestelyjä arkaluonteisen datan siirtojen, viranomaisten pääsyn ja alakohtaisten sääntöjen osalta.

#### **Lippulaivatoimet**

- ohjeet sen arvioimiseksi, käsitelläänkö EU:n dataa oikeudenmukaisesti EU:n ulkopuolella (vuoden 2026 toinen neljännes)
- välineistön luominen perusteettomien säilytyspaikkaa koskevien vaatimusten, markkinoilta sulkemisen, riittämättömien suojatoimien ja tietovuotojen torjumiseksi (vuoden 2026 toinen neljännes) ja toimenpiteiden hyväksyminen arkaluonteisten muiden kuin henkilötietojen suojaamiseksi (vuoden 2026 kolmas neljännes)

<sup>39</sup> Euroopan komissio ja unionin ulkoasioiden ja turvallisuuspolitiikan korkea edustaja, yhteinen tiedonanto Euroopan parlamentille ja neuvostolle – Euroopan unionin kansainvälinen digitaalistrategia, JOIN(2025) 140 final, Bryssel, 5.6.2025.

## **5. Eurooppalainen dataunionistrategia: datan hyödyntäminen tekoälyä varten**

Jotta voidaan varmistaa kilpailukyky tekoälyn aikakaudella, eurooppalainen dataunionistrategia siirtää painopisteen sääntöjen asettamisesta tulosten saavuttamiseen. Strategialla puututaan datan niukkuuteen, sääntelyn monimutkaisuuteen ja maailmanlaajuiseen kilpailuun vuonna 2020 luotujen puitteiden pohjalta.

Keskeisenä hallintofoorumina toimii edelleen Euroopan datainnovaatiolautakunta, jota on uudistettu, jotta jäsenvaltioiden ja teollisuudenalojen kanssa voidaan käydä syvällisempää teknistä ja strategista vuoropuhelua. Samaan aikaan tekoälyn soveltamisallianssista tulee tärkein kanava alakohtaiselle palautteelle. Sen avulla varmistetaan, että yritykset, tutkijat ja julkiset toimijat pääsevät muovaamaan strategian toteutusta. Tekoälyn seurantakeskus seuraa uusia suuntauksia ja tuottaa tietoa EU:n toimintapolitiikasta.

Kohdennetuilla toimilla laajennetaan laadukkaan datan käyttöä, yksinkertaistetaan sääntely-ympäristöä ja vahvistetaan EU:n roolia maailmanlaajuisissa datavirroissa. Pk-yritysten ja innovoijien osalta tämä tarkoittaa sääntöjen noudattamisesta aiheutuvien kustannusten vähenemistä, helpompaa tietojen saantia ja suotuisampaa kansainvälistä ympäristöä.

Vain mitattavat tavoitteet tuovat tuloksia. Tämän vuoksi komissio on ilmoittanut sisämarkkinoiden etenemissuunnitelmasta, jolla pyritään vauhdittamaan etenemistä ja nopeuttamaan prosesseja. Dataunionistrategiaa voidaan tarvittaessa käyttää etenemissuunnitelman edistämiseen, jotta päättäjiä ja alan toimijoita, erityisesti pk-yrityksiä, voidaan auttaa esteiden poistamisessa ja datan sisämarkkinoiden toteuttamisessa.

Yhdessä tekoälyn soveltamisstrategian kanssa dataunionistrategialla varmistetaan, että EU:n datavaroilla voidaan edistää suoraan tekoälyn kehittämistä, hyödyntämistä ja käyttöönottoa kaikilla aloilla.

Pitkän aikavälin visio on selkeä: itsenäinen eurooppalainen datatalous, jossa data siirtyy tietoturvallisesti ja vastuullisesti ja tukee tekoälyä, vauhdittaa innovointia ja vahvistaa kilpailukykyä.