

Tietoalustaselvitys - JOD-tietoalustan toteutustapojen arviointi

Esiselvityksen loppuraportti

26.6.2023

Sisällysluettelo

1.	Johdanto	6
1.1.	JOD-hankkeen kuvaus	6
1.2.	Esiselvityksen taustaa	6
1.3.	JOD-palvelut	7
1.4.	Tavoitetilan kuvaus	9
1.5.	JOD-tietoalustan perusvaatimukset	11
2.	Toteutusvaihtoehtojen kuvaus	13
2.1.	Taustaa	13
2.2.	Tunnistetut JOD-tietoalustan toteutusvaihtoehdot	15
2.3.	Toteutustapa A: Muiden tietoalustojen hyödyntäminen	16
2.4.	Toteutustapa B: Ostopalveluna toteutettava tietoalusta	19
2.5.	Toteutustapa C: Kokonaan oma tietoalusta	20
3.	Tietoalustojen sopivuus JOD-hankkeen tarpeisiin	21
3.1.	Tietoalustojen kuvaukset	21
3.2.	Tietoalusta Mylly	21
3.3.	Tietoalusta Antero	23
3.4.	Tietoalusta Tietojohdaja	25
3.5.	Tietoalusta Opintoallas	27
4.	Tietoalustojen vertailu	28
4.1.	Vertailuun otettavat tietoalustat	28
4.2.	Toiminnallisten vaatimusten arviointi	29
4.3.	Toiminnallisen vertailun tulokset	31
4.4.	Ei-toiminnalliset vaatimukset	36
4.5.	Ei-toiminnallisten vaatimusten vertailun tulokset	37
5.	SWOT-analyysi	39
5.1.	SWOT-analyysin kuvaus	39
5.2.	SWOT-analyysin tulokset	40
6.	Yhteenveto	44
6.1.	Tietoalustojen vertailujen yhteenveto	44
6.2.	Suositus	45
6.3.	Selvitettävät asiat	46
7.	Liitteet	46

Versiohistoria

Laatija	Versio	päivämäärä	Muutokset
Jussi Järvinen	0.1	14.5.2023	Loppuraportin runko
Atte Roine	0.2	15.5.2023	rungon täydennystä menetelmällä ja taustoilla
Jussi Järvinen	0.3	17.5.2023	Menetelmää tarkennettu
Jussi Järvinen	0.4	23.5.2023	Toteutustapaa B2 tarkennettu
Atte Roine	0.5	30.5.2023	1. Työpajojen tulokset
Atte Roine	0.6	8.6.2023	2. työpajojen tulokset (Mylly, Tietojohtaja, Antero)
Jussi Järvinen	0.7	9.6.2023	Rakenteen ja ulkoasun päivitys
Atte Roine ja Jussi Järvinen	0.8	11.6.2023	Tarkennuksia, yhteenveto ja alustavat suositukset
Atte Roine ja Jussi Järvinen	0.9	16.6.2023	Stilisointia
Konsulttiverkko	1.0	22.6.2023	Stilisointia
Konsulttiverkko	1.1	26.6.2023	Suosituksen tarkennuksia ja stilisointia

Tiivistelmä

Esiselvitys arvioi mahdollisuuksia JOD-tietoalustan toteutustavaksi siten, että niiden avulla voidaan jatkaa JOD-tietoalustan hankintaprosessia. Ensisijainen tietoalustojen arviointikriteeri oli arvioida olemassa olevien tietoalustojen soveltuvuutta JOD-palvelujen tarpeisiin. Olemassa olevia tietoalustoja on kahdentyyppisiä: JOD-ekosysteemin toimijoiden olemassa olevia tietoalustoja sekä muita valtionhallinnon sidosyksiköiden tietoalustoja ja palveluja. JOD-ekosysteemin tietoalustat ovat koulutuksen ja työllisyyden toimijoiden tarjoamia.

Esiselvitys arvioi JOD-hankkeen asettamia vaatimuksia tietoalustalle ja teki tähän pohjautuen suosituksen tarpeisiin soveltuvista tietoalustoista. Keskeiset vaatimukset ovat yksilön tiedon hallinta, sanitisointi, tekoälyn hyödyntäminen, tekstianalytiikka sekä yleiset tietoaltaan datan prosessointiin liittyvät kyvykkyydet.

Teknisestä näkökulmasta esiselvitys tunnisti useamman tarpeisiin soveltuvan tietoalustan tässä järjestyksessä:

1. Mylly tai Tietojohdaja
2. Antero
3. Opintoallas

Mikään tietoalusta ei ole suoraan valmis JOD-hankkeen kaikkiin tarpeisiin. Teknisesti tietoalustat voivat kuitenkin laajentua ja täyttää JOD-hankkeen tarpeet. Anteroon liittyy enemmän varauksia kuin kahteen muuhun mainittuun tietoalustaan. Opintoallas ei ole varsinainen tietoalusta termin puhtaassa merkityksessä ja sen hyödyntäminen on mahdollista, mutta vaatisi suuremman muutoksen.

Esiselvitys ei annetussa aikataulussa saanut tarkennettua tietoalustojen tarjoajien mahdollisuuksia aloittaa toteutus, resursseja tai pitkän tähtäimen ylläpidon kustannuksia. Keskustelujen tasolla Mylly ja Tietojohdaja ovat valmiita aloittamaan nopeasti, mutta tämä tulee varmentaa tarjouspyynnössä. Siksi toteutuksen kannalta seuraava vaihe on vuoropuhelu tietoalustoja ylläpitävien toimijoiden kanssa tarkemmalla tasolla toteutuksen aloittamisesta. JOD-hankkeella on kunnianhimoinen aikataulu, jonka toteutuksen aikataulun reunaehdot määrittävät lopullisen valinnan.

1. Johdanto

1.1. JOD-hankkeen kuvaus

JOD-tietoalustan palvelut tulevat hyödyntämään soveltuvin osin jo käytössä olevia sidosryhmien tietoalustoja ja palveluita, jotka tulee huomioida kokonaisarkkitehtuurin näkökulmasta. Sidosryhmien tuottama data on keskeisessä roolissa palvelujen tuottamisessa. JOD-hankkeen luomat palvelut yhdistävät laajasti eri toimijoiden tuottamaa koulutus- ja työelämä tietoa. Lisäksi yksilö tuottaa itse tietoa. JOD-hanke haluaa luoda laajan ekosysteemin palveluja, joissa eri lähteiden tuottamaa tietoa hyödynnetään mutta myös palautetaan jalostettuna takaisin. Näin sidosryhmät voivat hyötyä JOD-palveluista. Tietoa hyödyntävät myös koulutustoimijat ja työelämätoimijat. Käyttäjät sekä tuottavat että hyödyntävät JOD-palvelujen tietoa muodostaen näin aidon ekosysteemin.

JOD-hankkeen tarpeena on nyt tunnistaa ne sidosryhmien tietoalustat, joissa on synergiaa JOD-palvelujen toteuttamiseksi. Lisäksi on tarve tunnistaa kokonaisuudet, jotka on toteutettava omalla tietoalustalla. Tavoitteena on hyödyntää mahdollisimman paljon olemassa olevia tietoalustoja ja näiden kyvykkyksiä sekä mahdollisesti jopa välttää oman tietoalustan toteuttaminen.

Esiselvityksen lopputuloksena on arvio JOD-hankkeen tarpeisiin parhaiten sopivasta toteutustavasta.

1.2. Esiselvityksen taustaa

Esiselvityksen pyrkimyksenä oli määrittää JOD-tietoalustalle vaihtoehtoisia toteuttamistapoja sekä arvioida näiden toteuttamiskelpoisuutta. Tavoitteena on, että JOD-hanke voi aloittaa hankinnan ja kilpailutuksen JOD-tietoalustan toteuttavan kumppanin valinnasta.

Selvityksessä keskeistä on tunnistaa ne osa-alueet, jotka JOD-hankkeen tulee toteuttaa ja hankkia palveluna eroteltuna osa-alueista, jotka ovat jo olemassa sidosryhmien hallinnoimien tietoalustojen kautta. Selvityksen painopiste on tietoaltaassa, tietovarastossa, datassa ja datan hallintaan liittyvissä kyvykkyyksissä, joista JOD-palvelut rakentuvat. Erityisen tärkeää on painottaa toteutuksessa tietoaltaan elinkaarta ja maturiteettia. JOD-tietopalvelulla on vaativia järjestelmävaatimuksia ja erityispiirteitä. Selvityksen näkökulma toteutustavan valintaan on teknologialähtöinen.

Selvityksen laajuudesta on rajattu pois mm. seuraavat näkökulmat:

- JOD-tietoallasratkaisun hallintamalli

- käyttöliittymät ja käyttötapaukset loppukäyttäjän näkökulmasta
- Sidosryhmien olemassa olevien palveluiden datan hallinnan prosessit (nämä oletetaan sellaisiksi kuin ovat, ja tarkastelu tehdään ns. "as is"-tilanteessa)

Selvityksessä painotetaan tietoallasratkaisun teknisen toteuttamisen arviointia.

1.3. JOD-palvelut

JOD-palvelut ovat koulutuksen ja työllisyyden kohtaannon ratkaisemiseen suunnattuja palveluja, jotka pyrkivät jalostamaan suosituksia esimerkiksi yksilön uravalintojen tueksi tiedon ja ennakoivan analytiikan avulla. Yksilö voi olla esimerkiksi nuori aikuinen tai myöhemmin uranvaihtoa harkitseva keski-ikäinen henkilö. Koulutusta ja työelämän tarvetta ei ole rajattu, vaan koulutuksen tarpeiden osalta laajuuteen sisältyy kaikki tutkintomuotoinen opetus. Tästä syystä JOD-palvelut kattavat hyvin laajan alueen koulutuksen ja työllisyyden tarpeita.

Laaja-alaiset tarpeet huomioiden, JOD-palveluiden ominaispiirteet voidaan tiivistää kolmeen osa-alueeseen:

JOD-palvelut ovat uudenlaisia poikkihallinnollisia palveluja

JOD-palvelut ovat poikkihallinnollisia, ja ne yhdistävät monipuolisesti eri hallinnonalojen tarjoamaa tietoa sekä yksilön omaa dataa. JOD-palvelut yhdistävät tapahtumatietoa ja päivittävät analyysin pohjana olevaa tietoa reaaliaikaisesti. JOD-palvelut poikkeavat tässä suhteessa huomattavasti perinteisistä tilastolliseen analyysiin ja raportointiin tuotetuista palveluista, joissa tiedon päivittämisen sykli on tyypillisesti harvempi. JOD-palvelut kattavat lisäksi laajan joukon opetushallinnon ja työllisyyden palvelualoja, esimerkiksi kaikki julkisin varoin tuotetut koulutuspalvelut. Siksi kaikki JOD-palvelujen tarpeet täyttävää valmista tietoalustaratkaisua ei ole.

JOD-palvelut yksilölle ja ekosysteemin toimijoille

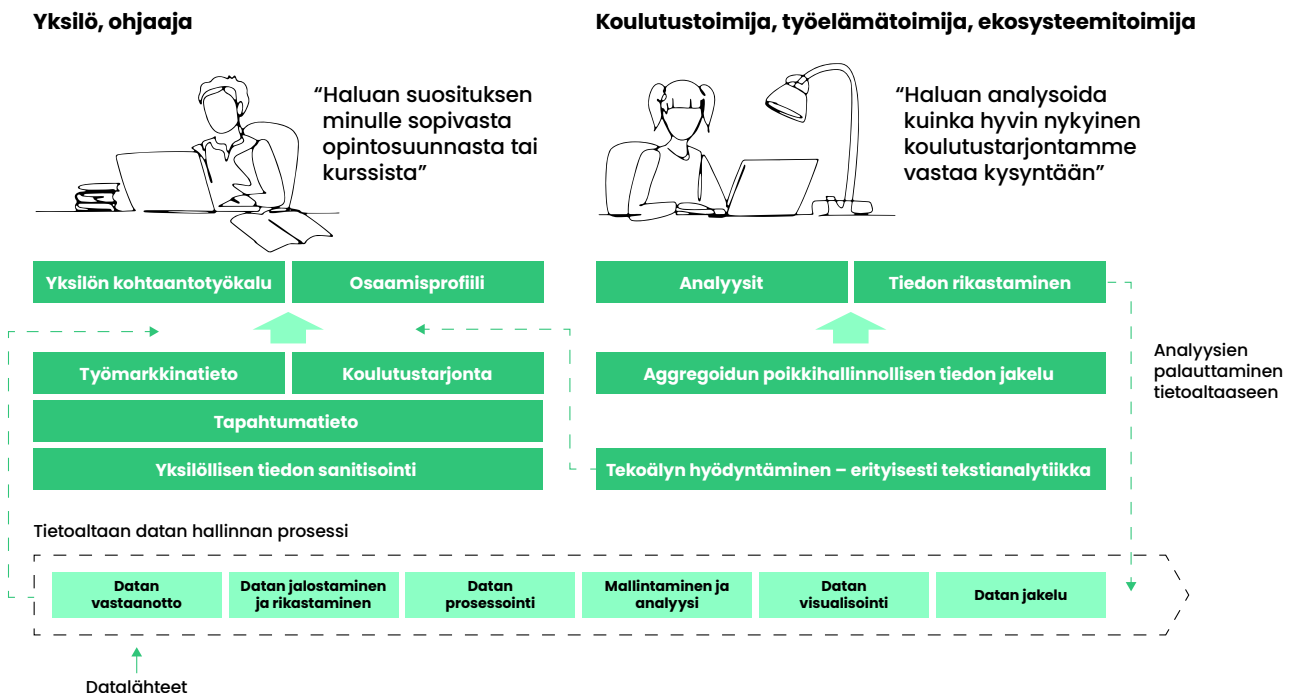
JOD-hankkeen ensisijainen tavoite on tuottaa palveluja yksilöille ja ekosysteemin toimijoille. Palvelut poikkeavat tyypillisistä hallinnonalan sisäiseen tai organisaatioiden väliseen tiedonvaihtoon liittyvistä palveluista. Kehitettävät palvelut asettavat erityisvaatimuksia esimerkiksi yksilön tietojen käsittelylle sekä tiedon sanitisoinnille ja jäljitettävyydelle. JOD-palvelut käyttävät tietoa laajasti, minkä vuoksi rekisterisuoja, rekisterin ylläpitovastuu ja mahdolliset tiedon hyödyntämiseen liittyvät rajoitukset tulee

arvioida. JOD-hanke luo tarpeita lakimuutoksille, jotta tietoa voidaan hyödyntää täysimääräisesti poikkihallinnollisesti.

Tiedon analysointi ja rikastaminen

JOD-palvelut poikkeavat perinteisestä raportointikäyttöön laaditusta tietovarastosta, jossa analyysi tehdään tyypillisesti tietovaraston ulkopuolella, eikä tuloksia suoraan palauteta tietovarastoon hyödynnettäväksi. JOD-palvelujen tavoite on jatkuvasti rikastaa tietoa sitä mukaa, kun palveluja käytetään sekä yksilön että ekosysteemitomijoiden toimesta ja tarpeisiin. Tekoäly ja tekstianalytiikka ovat keskeisessä roolissa tiedon jatkuvassa jalostamisessa.

JOD-hanke on tunnistanut ja määrittänyt 24 kappaletta JOD-palveluja, jotka on suunnattu viidelle (5) eri käyttäjäryhmälle. Käyttäjät voivat olla yksilö, ohjaaja, koulutustoimija, työelämätoimija tai ekosysteemitomija. Yksilölle JOD-palvelujen tavoite on tukea koulutus- ja uravalintojen tekemistä. Ammattimaisille käyttäjille tarjotaan laajempi joukko palveluja. JOD-hanke on tunnistanut palvelujen tietojärjestelmävaatimukset, joiden avulla on mahdollista arvioida JOD-tietoalustalle sopivin toteutustapa.

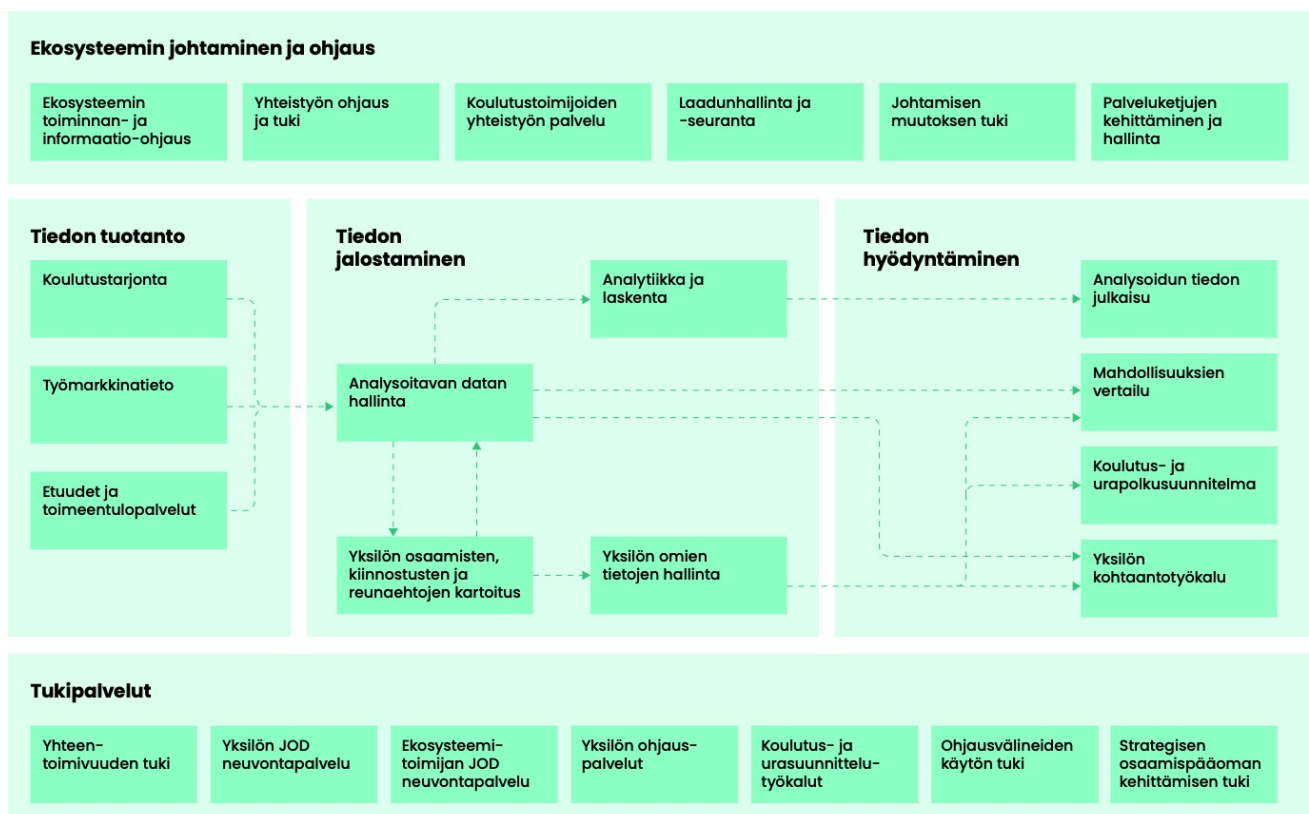


Kuva 1. Havainnekuva JOD-ekosysteemistä

Tässä esiselvityksessä ei käydä yksityiskohtaisesti läpi JOD-palvelujen vaatimuksia tietoalustalle. Esiselvitys on saanut taustamateriaalina tietoa, jossa vaatimukset on kuvattu mm. ”*RFF tavoitearkkitehtuuriraportti*” -dokumentissa. Esiselvitys on pyrkinyt hyödyntämään tätä tietoa laaja-alaisesti toteutusvaihtoehtojen kuvauksessa. Seuraavassa luvussa on tiivistelmä keskeisistä JOD-palvelujen vaatimuksista tietoalustalle.

1.4. Tavoitetilan kuvaus

JOD-hanke on määrittänyt 24 JOD-palvelua sekä näiden keskinäiset tietovirrat ja tietosisällöt. Kuvassa 2 on havainnollistettu JOD-palvelujen suhteita toisiinsa.



Kuva 2. JOD-tietoalustan palvelut (JOD-tavoitetilan arkkitehtuuri)

JOD-arkkitehtuurissa keskeistä on tiedon tuotanto, jalostaminen ja hyödyntäminen. Näiden osakokonaisuuksien avulla voidaan hahmottaa suurin osa JOD-tietoalustan tarpeista. Vaatimukset tiivistyvät ”JOD-17 Yksilön kohtaantotyökalu” -palveluun, joka toteutuessaan kokoaa tietoalustassa jalostuneet kyvykkyydet sekä yksilön tarpeista että työ- ja

koulutustarjonnan mahdollisuuksista lähtien. Tämän lisäksi keskeisinä tietolähteinä ovat koulutustarjonta ja työmarkkinatieto.

Taulukossa 1 on kuvattuna näiden kolmen keskeisen JOD-palvelun vaatimukset.

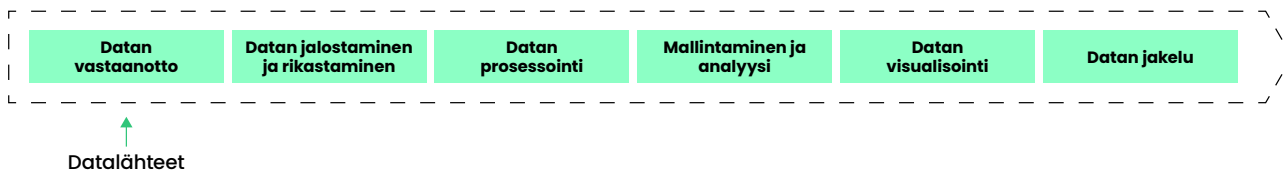
<p>Koulutustarjonta (JOD-07)</p>	<p>Koulutustarjonta on rakenteellista tietoa.</p> <p>Tavoitetilassa tieto voidaan mahdollisesti tuoda valmiissa muodossa tietoalustalle tai hyödyntää olemassa olevia tietoalustoja, joissa koulutustarjonta on saatavilla.</p>
<p>Työmarkkinatieto (JOD-08)</p>	<p>Työmarkkinatieto tuodaan useista eri tietolähteistä, jotka edustavat eri tietotyyppisiä. Kaikki tieto ei ole rakenteellisessa muodossa tai olemassa olevassa tietovarastossa, vaan sirpaloituneena. Tästä syystä JOD-tietoalustan tulee ensin kerätä tieto yhteen ennen analyysin tekemistä.</p> <p>Tiedon jalostaminen jää JOD-tietoalustan vastuulle. Jalostamisessa tarvitaan mm. tapahtumaperusteista dataa.</p>
<p>Yksilön kohtaantotyökalu (JOD-17)</p> <p>Hyödyntää mm. seuraavia muita JOD-palveluja:</p> <ul style="list-style-type: none"> ● Yksilön omien tietojen hallinta (JOD-13) ● Yksilön osaamisen...(JOD-12) ● Analytiikka ja laskenta (JOD-11) ● Analysoitavan datan hallinta (JOD-10) 	<p>Yksilön käyttäjätietojen ja tapahtumadatan yhdistäminen laskentaa varten. Hyödynnetään koulutustarjontaa ja työmarkkinatietoa.</p> <p>Yksilön osaamisprofiili siirretään JOD-tietoalustan käyttöön.</p> <p>Yksilön tiedot anonymisoidaan ennen laskentaa. Laskenta palauttaa tuloksen yksilölle käyttöliittymän avulla. Laskennan tulokset, analyysi ja tapahtumatieto rikastavat JOD-tietoalustan dataa ja sitä hyödynnetään myöhemmin, kuitenkin ilman yksilöivää tietoa.</p>

Taulukko 1. Keskeiset JOD-palvelut

1.5. JOD-tietoalustan perusvaatimukset

Palvelukohtaisten vaatimusten lisäksi tietoalustalla tulee olla perustason kyvykkyudet hallita suuria tietomääriä. Tietoalustan datan hallinnan perusprosessin vaiheet on kuvattu kuvassa 3.

Tietoalustan datan hallinnan prosessi



Kuva 3. JOD-tietoalustan datan hallinnan prosessi

Taulukossa 2 on kuvattu datan hallintaan liittyvät perusvaatimukset vaiheittain.

<p>Datan vastaanotto</p>	<p>Tietoalustan on kyettävä ottamaan dataa vastaan erilaisissa muodoissa.</p> <p>Erlaisia muotoja ovat mm. rakenteellinen data, tapahtumadata, teksti, multimedia ja muu sisältö.</p> <p>Tietoalustan ei tule rajoittaa vastaanotettavan datan muotoa.</p> <p>Dataa voidaan ottaa vastaan reaaliaikaisesti tapahtumatieloina ja kyselyinä tai eräajoina ajastetuissa erissä.</p>
<p>Datan jalostaminen ja rikastaminen</p>	<p>Tietoalustan tulee kyetä jalostamaan dataa ja rikastamaan sitä muilla tietolähteillä sekä tekemään varmistuksia datan oikeellisuudesta. Dataa voidaan myös harmonisoida tai normalisoida. Esimerkiksi tapahtumatietoa voidaan aggregoida sovituin intervalein (kyselyä tunnissa ym.)</p> <p><u>Erityishuomio tiedon anonymisoinnista:</u></p> <p>Tietoa pitää pystyä anonymisoimaan. Arvioitava siihen soveltuvat teknologiat, arvioitu teknologian elinikä, anonymisoinnin tyyppi (anonyymi/pseudonyymi) GDPR, prosessi.</p>
<p>Datan prosessointi</p>	<p>Tietoalustan tulee kyetä yhdistämään ja jalostamaan tietoa eri tietolähteistä. Yhdistelyssä tulee voida hyödyntää moderneja koneoppimisen ratkaisuja.</p> <p>Tietoalustan tulee mahdollisesti säilyttää data tai mahdollistaa datan palauttaminen.</p>

	<p>Tietoalustan tulee kyetä jalostamaan dataa erilaisiin rakenteisiin raportointia tai tietovarastoa varten.</p> <p>Tietoalustan tulee prosessoida dataa hyödyntäen koneoppimista, algoritmeja tai muita data-analyysin työkaluja.</p> <p>Tietoalustalla tulee pystyä tallentamaan ja hyödyntämään käyttäjien luomaa dataa toiminnallisuuden kehittämisessä ja uuden toiminnallisuuden luomisessa.</p> <p>Tietoalustalla on huomioitava nopeasti kehittyvän tekstianalytiikan hyödyntäminen ja sen ohjaaminen datalla ja digitaalisella ohjeistuksella.</p>
Datan visualisointi	Dataa tulee kyetä visualisoimaan käyttöliittymän avulla.
Datan jakelu	<p>Dataa on voitava jaella rajapintojen kautta, kuutioina, raporteina, raakamuodossa, Excel-muodossa.</p> <p>Tiedon omistaja voi tarjota pääsyn tietoon ja hallinnoida tiedon jakelua.</p>

Taulukko 2. JOD-tietoalustan perusvaatimukset

2. Toteutusvaihtoehtojen kuvaus

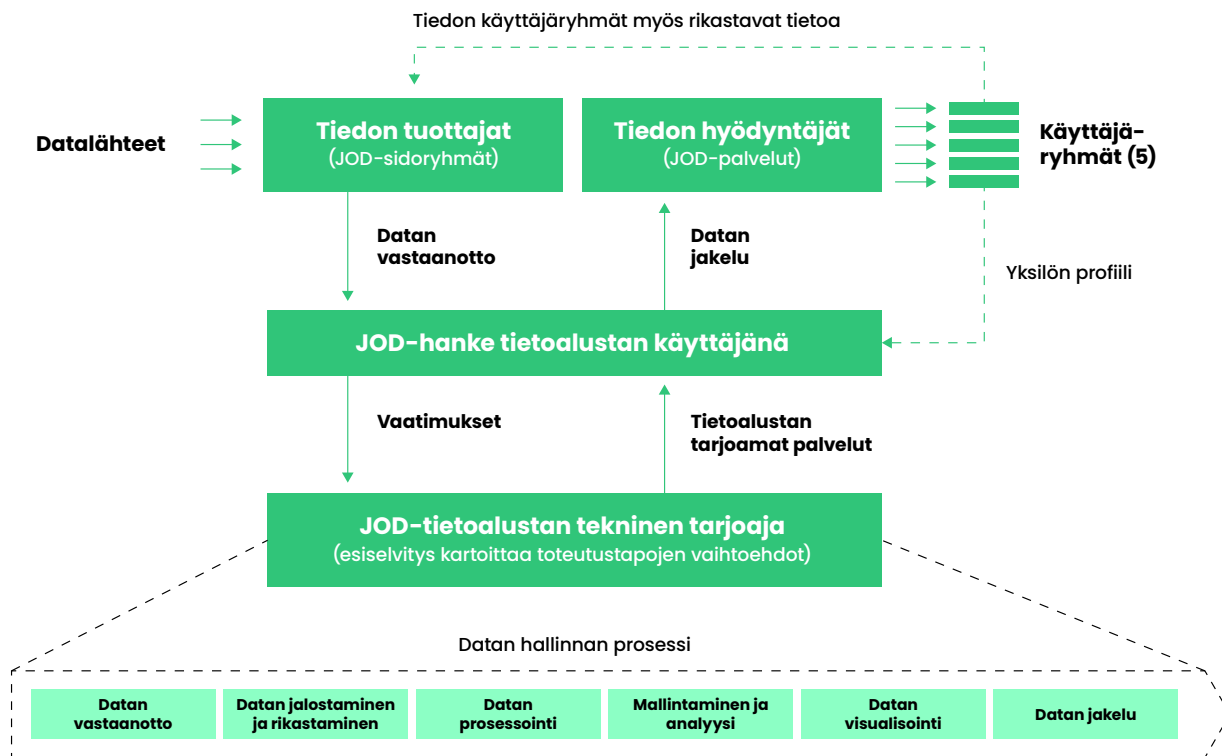
2.1. Taustaa

Esiselvitys on tunnistanut seuraavat JOD-ekosysteemin roolit suhteessa JOD-tietoalustan käyttöön:

- **Tiedon tuottaja.** Ekosysteemin toimijat, jotka toimivat datalähteenä JOD-tietoalustalle.
- **Tiedon hyödyntäjä.** Toimija, joka tarjoaa tai hyödyntää JOD-palvelujen tarjoamaa tietoa, tai mahdollisesti tarjoaa myös JOD-palvelun. Tiedon hyödyntäjät voivat olla myös palvelujen käyttäjiä.
- **Tietoalustan tekninen tarjoaja.** Tietoalustan tarjoaja JOD-palvelujen käyttöön.

Mikäli tietoalustan tarjoaja myös itse tuottaa ja hyödyntää tietoa, tämä voi olla vahvuus arvioidessa toimijaa myös tietoalustan teknisenä tarjoajana.

Kuva 4 havainnollistaa toimijoiden riippuvuuksia suhteessa tietoon, palveluihin ja tietoalustaan.



Kuva 4. Tietoalustan toimijat

Käytännön tasolla JOD-palvelujen toteuttamiseksi tarvitaan liittymiä sidosryhmien nykyisiin tai toteutuksen alla oleviin tietoalustoihin. Liittymät ovat joko dataintegraatioita tai palvelutason integraatioita, joissa JOD-tietoalusta hyödyntää muiden tietoalustojen tarjoamia palveluja. Näitä dataintegraatioita tai palveluja ovat mm.:

- Koulutuskysyntä
- Koulutustarjonta
- Yksilön osaamisprofiili
- Analytiikan työkalut
- Kohtaantotyökalun analytiikka
- Vertailutiedon tuonti
- Koulutustieto
- Urapolkutieto
- ”Minun tietoni”-dataintegraatio
- Profiilitieto
- Raportointi

Koska tavoite on välttää kokonaan oman JOD-tietöalustan toteutus, palvelujen ei tarvitse sijaita loogisesti yhdessä paikassa. Esimerkiksi mikäli jokin olemassa oleva tietöalusta jo tosiasiallisesti tarjoaa JOD-palvelujen vaatimaa dataa, tästä palvelusta voidaan jalostaa "JOD-palvelu" tai vaihtoehtoisesti tämän palvelu tieto tuodaan toiselle tietöalustalle, josta muodostuu "JOD-palvelu". Toisin sanoen, moni olemassa oleva tietöalusta voi jo osittain täyttää jonkin 24 tunnistetusta JOD-palvelusta. JOD-tietöalustan ei tarvitse luoda kaikkea itse, vaan ennemmin pyrkiä hyödyntämään jo olemassa olevien tietöalustojen palveluja ja luoda "sateenvarjo" JOD-palveluiden kokonaisuudelle.

JOD-tietöalustan toteutusvaihtoehtojen näkökulmasta onkin oleellista muodostaa käsitys olemassa olevista JOD-ekosysteemin tietöalustoista ja näiden kyvykkyyksistä.

2.2. Tunnistetut JOD-tietöalustan toteutusvaihtoehdot

Tavoitteena on lähtökohtaisesti hyödyntää mahdollisimman paljon olemassa olevia tietöalustoja. Toisaalta tiedetään, ettei mikään olemassa oleva tietöalusta täytä kokonaan kaikkia 24 tunnistetun JOD-palvelun vaatimuksia. Puutteita on esimerkiksi yksilöllisen tapahtumatiedon käsittelyssä, tiedon anonymisoinnissa ja tekoälyn hyödyntämisessä. Siksi on välttämätöntä, että JOD-hanke toteuttaa joltain osin myös omia tietöalustaratkaisuja, kuitenkin hyödyntäen mahdollisimman pitkälti sidosryhmien jo käytössä olevaa teknologiapohjaa.

JOD-tietöalustan toteutusvaihtoehtoja arvioidaan seuraavassa prioriteettijärjestyksessä olemassa olevia tietöalustoja painottaen:

1. JOD-hanke toteuttaa osittain oman tietöalustan, ja lisäksi täydentää tietöalustaa yhteisillä osilla muista tietöalustoista
2. JOD-hanke toteuttaa tietöalustan ostopalveluna hyödyntäen olemassa olevia toimijoita, joilla on tarpeisiin soveltuva tietöalusta
3. Toteutetaan kokonaan uusi JOD-tietöalusta

Ensisijaisesti hyödynnetään siis tietöalustoja, jotka ovat myös JOD-ekosysteemissä toimijoina, ja siten tulevat joka tapauksessa hyödyntämään JOD-palveluja.

Kuvassa 5 on kuvattu tunnistetut toteutusvaihtoehdot ja näiden suhde JOD-palveluihin joko palvelujen toteuttajan, tietosisältöjen tarjoajan tai pelkästään teknisen JOD-tietöalustan tarjoajana.

	Toteutustapa A: JOD-hanke toteuttaa osittain oman tietoalustan, ja lisäksi täydentää tietoalustaa yhteisillä osilla muista tietoalustoista.	Toteutustapa B: JOD-hanke toteuttaa tietoalustan ostopalveluna hyödyntäen olemassa olevia toimijoita, joilla on tarpeisiin soveltuva tietoalusta	Toteutustapa C: Toteutetaan kokonaan uusi JOD-tietoalusta
JOD-palvelut	JOD-palvelujen toteuttamiseksi hyödynnetään sidosryhmien tietoalustoja mahdollisimman paljon, jotta "oman" tietoalustan tarve jäisi pieneksi	Tietoalustan tarjoavalla toimijalla ei ole roolia JOD-palvelujen ekosysteemissä.	JOD-tietoalusta toteuttaa tarvittavat JOD-palvelut ja tarjoaa ne sidosryhmille.
Tietosisältöjen tuottajat			
Datan hallinnan prosessi	Sidosryhmien tietoalustat tarjoavat teknologiapohjan, johon JOD-palvelut toteutetaan.	Toimija tarjoaa tietoalustan ja datan hallinnan prosessit, joilla JOD-palvelut toteutetaan.	Toimija tarjoaa tietoalustan ja datan hallinnan prosessit, joilla JOD-palvelut toteutetaan.
JOD-tietoalusta			

Kuva 5. Tunnistetut JOD-toteutustavat

Seuraavissa luvuissa on kuvattu eri toteutustavat tarkemmalla tasolla.

2.3. Toteutustapa A: Muiden tietoalustojen hyödyntäminen

Tässä toteutustavassa pyritään hyödyntämään mahdollisimman pitkälti olemassa olevia tietoalustoja 24 tunnistetun JOD-palvelun toteuttamiseksi. Toteutustapana tämä voi tarkoittaa esimerkiksi "putkea" olemassa olevaan tietoalustaan, jossa olemassa oleva tietoalustapalvelu avaa kehityspotken JOD-hankkeelle. Näin JOD-hanke hyödyntää suoraan tämän tietoalustan data- ja analytiikkapalveluja. Eri tietoalustoihin voidaan avata useita putkia, joista tarvittavat JOD-palvelut kokonaisuutena muodostuvat. Näin yhden JOD-tietoalustan sijaan muodostetaan virtualisoitu kokonaisuus olemassa olevia tietoalustoja hyödyntäen.

Tämän toteutustavan tarkoituksena on hakea synergiaetuja, jotta JOD-tietoalustaa voidaan rakentaa ja ylläpitää kustannustehokkaammin yhteisiä resursseja hyödyntäen. Samalla voidaan tukea jotain elinkaarensa loppupäässä olevaa alustaa teknologian modernisoinnissa (optio).

Yhteiset tietoalustat voivat sisältää esimerkiksi seuraavia komponentteja:

- Yhteinen data, joka jaetaan näkymillä JOD-palvelujen käyttöön. Datamassoja ei siis kopioida palvelimelta toiselle, vaan tietoa ylläpidetään yhdessä paikassa ja JOD-palvelut käyttävät näkymädataa omiin palveluihinsa. Esimerkiksi Anteron ja

Opintoaltaan koulutusdata voisi olla tällaista valmiiksi jalostettua ja yhdistettyä dataa, jota käytettäisiin yhdessä.

- Yhteiset ohjelmistoratkaisut, joissa käytetään toisen data-alustan kanssa yhteisiä ETL-työkaluja, tietokantaa, mallintamistyökaluja jne.
- Yhteinen tietoallas (*data lake*) / tietovarastokerros, jonka sisällä datat ovat osin yhteisiä ja osin omia tietovarastojaan. Pilvipalveluissa on helppo tallentaa eri ratkaisujen ja palveluiden dataa samaan paikkaan joko raakamuodossa tai jalostettuna. Tämä helpottaa sen jatkokäsittelyä etenkin, jos samaa dataa käytetään hieman eri tarkoituksiin.
- Osittain yhteinen organisaatio ja prosessit esim. datan ja hallinnan osalta. Työtehtävät ovat yleensä samanlaisia, vaikka itse data vaihtelee. On varmistettava, että tiedonsiirtoa seurataan ja jos ongelmia tai poikkeuksia esiintyy, ne täytyy ratkaista. Synergiaetuja syntyy helposti, jos samat henkilöt valvovat useampaa data-alustaa.
- Replikoitavat palvelut tai ratkaisut esimerkiksi anonymisoinnissa tai tietojen yhdistelyssä ja yhtenäistämässä. Esimerkiksi henkilötietojen anonymisointi on monimutkainen ongelma, mutta kun se on jo kertaalleen ratkaistu, ratkaisusta kertynyttä oppia kannattaa hyödyntää vähintäänkin vaatimusten tai jopa valmiin koodin tasolla.

On myös mahdollista, että jokin olemassa oleva JOD-ekosysteemin tietoalusta sulautuu osaksi uutta tietoalustaa niin, että sen kaikki kyvykkyydet ja resurssit sulautuvat JOD-tietoalustaksi. Toisin sanoen olemassa olevasta tietoalustasta ”kasvaa” JOD-tietoalusta.

Esiselvitys kävi läpi sidosryhmien hallinnassa ja käytössä olevat tietoalustat. Näistä tietoalustoista kolme tietoalustaa osoittautui mahdollisiksi tässä toteutustavassa JOD-palvelujen näkökulmasta.

Taulukossa 3 on kuvattu tunnistetut sidosryhmien tietoalustat ja arvio näiden soveltuvuudesta JOD-palvelujen tietoalustaksi.

Toimija	Tietoalusta	Arvio soveltuvuudesta
Keha	Mylly	Soveltuu jatkoselvitettäväksi
Vipunen	Antero	Soveltuu jatkoselvitettäväksi

Digivisio	Opintoallas	Soveltuvuus selvittävä
DigiFinland	Tietojohtaja	Ei sovellu (ei sisällä JOD-palvelujen tietoja, mutta soveltuu vaihtoehtoon B)
OPH	Ennakoinnin tietoalusta	Ei voida hyödyntää, koska ei vielä olemassa
OPH	Koski	Ei sovellu, spesifi käyttö. Ei laaja-alainen dataalusta
OPH	Virta	Ei sovellu, spesifi käyttö. Ei laaja-alainen dataalusta
OPH	Lampi	Koski-tietojen datalähde, ei sellaisenaan riittävä JOD-tietoalustaksi
OPH	ODW	ODW on luotu haku- ja valintatietojen hakemiseen. Ei sovellu
Jotpa	Jotpa tietoalusta	Kesken, spesifi käyttö. Ei laaja-alainen dataalusta
CSC	Ostopalvelu	Ei tunnistettu alustaa, jota voisi hyödyntää

Taulukko 3. Tunnistetut tietoalustat (A)

Alustavan arvion perusteella potentiaalisiksi vaihtoehtoisiksi tässä toteutustavassa nousivat **Mylly** ja **Antero**.

Opintoaltaan osalta kyvykkyyksiin tulee suhtautua varauksella koska se ole varsinainen tietoalusta vaan virtuaalisesti luotu palvelujen kokonaisuus, joka nojautuu vahvasti datalähteisiin. Siksi Opintoaltaan osalta arviona on soveltuvuus käyttöön tietoaltaana varauksella.

Tietojohtajan päälle ei voida rakentaa suoraan vaihtoehto A:n mukaisesti, mutta se soveltuu vaihtoehtoon B.

Koski-rekisteri (opiskeluoikeuksien ja suoritusten keskitetty integraatiopalvelu) ja Virta-opintotietopalvelu (korkeakoulujen opiskelijavalintojen kehittäminen) on rakennettu tiettyjä käyttötapauksia palvelemaan, eivätkä ne näin sovellu laajoiksi data-alustoiksi.

Jotpan ja Ennakoinnin data-alustat ovat vasta kehitteillä ja ne on tarkoitettu toimijoiden spesifeihin käyttötarkoituksiin, ei laaja-alaisiksi tietoalustoiksi. Edustajat eivät ilmaisseet mahdollisuutta osallistua yhteiseen kehitykseen.

ODW on luotu haku- ja valintatietojen hakemiseen, eli se ei sovellu käyttötarkoitukseen. ODW:n arvioidaan olevan elinkaarensa loppupuolella eikä sen päälle haluta uutta kehitystä. Lampi ja Koski ovat keskittyneet opintosuoritusten hallintaan ja niiden tarkoitus halutaan pitää sellaisena. Lampi voisi hyvin toimia datalähteenä kehitettävälle JOD-tietoalustalle ja sen moderni arkkitehtuuri mahdollistaisi datan tarjoamisen virtuaalisena näkymänä. CSC:llä ei ole suoraan tähän tarpeeseen valmista alustaa hyödynnettäväksi.

Haastattelujen ja olemassa olevien materiaalien avulla tarkempaan jatkoselvitykseen vaihtoehtoiksi valittiin:

- Antero-tietoallas
- Keha Mylly
- Digivisio Opintoallas (varauksin - selvitettävä)

2.4. Toteutustapa B: Ostopalveluna toteutettava tietoalusta

Tässä vaihtoehdossa strategiana on ostaa JOD-tietoalustan toteutus palveluna joltain valtionhallinnon sidosyksiköltä. Toimija on tietoalustapalvelun tuottaja ja JOD-hanke toimii tilaajana. JOD-hankkeen tavoiteaikataulun takia vain sidosyksiköiden käyttö on mahdollista, koska avoin kilpailutusprosessi vaatisi liikaa aikaa. Palvelun tarjoava taho kehittäisi tässä tapauksessa tietoalustan "avaimet käteen"-periaatteella ja vastaisi palvelun operoinnista koko sopimuskauden ajan.

Taulukossa 4 on kuvattu esiselvityksen tunnistamat sidosyksiköiden tietoalustapalvelut.

Tarjoaja	Tietoalusta	Arvio soveltuvuudesta
DigiFinland	Tietojohtaja	Soveltuu jatkoselvitettäväksi

Keha	Mylly	Voisi soveltua (jatkoselvityksessä todettiin, ettei ole heidän suosimansa malli)
Valtiokonttori	Tietokiri	Selvitettävä
CSC	Anteron tai vastaavan tietoalustan hyödyntäminen	Selvitettävä

Taulukko 4. Tunnistetut tietoalustat (B)

Esiselvityksen perusteella sopivimmaksi vaihtoehdoksi valikoituu DigiFinlandin Tietojohtaja. Myös Keha voisi tarjota palvelun Mylly-alustaan perustuen, mutta he kokivat, että malli A istuu paremmin heidän nykyiseen hallintomalliinsa (toimiva monivirastomalli, jossa tuotteita). Arviomme on, että toteutustapa A on Myllyn tapauksessa luontevampi vaihtoehto.

Valtiokonttorin Tietokiri on saman tyyppinen tietoalusta kuin Tietojohtaja, mutta on tällä hetkellä keskittynyt talouden tunnuslukujen raportointiin. Sote-hallinnonalalle suuntautunut Tietojohtaja on hieman lähempänä JOD-hankkeen tarpeita. Valtiokonttori on kuitenkin mahdollinen vaihtoehto, joka tulisi selvittää tarkemmin, mikäli in house -kilpailutus järjestetään. Sama koskee CSC:ia toimijana.

2.5. Toteutustapa C: Kokonaan oma tietoalusta

Tässä vaihtoehdossa JOD-hanke toteuttaa kokonaan oman tietoalustansa kilpailuttamalla suoraan haluamansa toteutuskumppanin. Tämä vaihtoehto on realistinen, mikäli todetaan, ettei mikään muu olemassa oleva tietoalusta voisi soveltua JOD-hankkeen tarpeisiin edes pitkällä tähtäimellä. Huomioitavaa on, että joka tapauksessa sidosryhmien tarjoamat tietoalustat toimivat tiedon lähteinä ja hyödyntäjinä JOD-ekosysteemissä, eli oma tietoalusta ei olisi erillinen saarekkeensa, vaan kiinteä osa koulutuksen ja työllisyyden tiedolla johtamisen ekosysteemiä.

Perusteena omalle tietoalustalle on mm. helpompi hallittavuus. Heikkoutena on vastaavasti se, ettei synergiaetuja saada, ja todennäköisesti valtiokonsernitasolla joudutaan tekemään päällekkäistä työtä. Tällöinkin on hyvä arvioida, voidaanko synergiahyötyjä saavuttaa organisoinnilla ja teknologiavalinnoilla.

Havainnot Digivision kehityksestä haastatteluihin perustuen:

Kokemukset Digivision Opintoallas-tietoalustan toteuttamisesta antavat viitteitä, että kokonaan oma tietoalusta ei ole mahdollinen annetussa JOD-palvelujen toteuttamisen aikataulussa. Digivisio oli käyttänyt vuoden pelkästään toteutuksen kilpailuttamiseen. Kilpailutus on monimuotoinen prosessi, joka vaatii useita eri kilpailutuksia toteutettaville osakokonaisuuksille, kuten kapasiteettipalvelulle, kehityspalveluille, ylläpitopalveluille ym. JOD-palvelujen toteutus vuoden 2025 loppuun mennessä ei siis ole tässä vaihtoehdossa realistinen.

CSC olisi mahdollinen kumppani oman tietoalustan toteutukselle. CSC ei ole muissa vaihtoehtoissa järin relevantti toimija. Siksi suositus on, ettei vaihtoehtoa C tarkastella pidemmälle.

3. Tietoalustojen sopivuus JOD-hankkeen tarpeisiin

3.1. Tietoalustojen kuvaukset

Seuraavissa luvuissa kuvataan tarkemmalla tasolla toteutusarvioon otettavat tietoalustat JOD-tietoalustan toteutuksen näkökulmasta.

3.2. Tietoalusta Mylly

Tietoalusta	Mylly
Kuvaus käytössä olevasta ratkaisusta	<p>Myllyn palvelukonsepti koostuu useammasta kokonaisuudesta.</p> <p>Mylly kokoaa tietoa mm. taloudesta, henkilöstöstä ja palvelutoiminnasta.</p> <p>Järjestelmä sisältää sekä historiaan että ennakkointiin perustuvaa vertailutietoa toimintaympäristöstä, asiakkaista ja oman toiminnan kehittämisestä.</p>

<p>Kehityshistoria</p>	<p>Aloitettu n. 5 vuotta sitten ja tuotannossa 3-4 vuotta. Massiivinen kokonaisuus, jota kehitetään koko ajan.</p> <p>Seuraavat askeleet tekoälyn puolella ja rajapintojen hyödyntämisessä. Myllyn tiedolla johtamisen tiekartta 2023-2024.</p> <p>Kehittämishankkeet ja projektit:</p> <ul style="list-style-type: none"> • ELY-Mylly • ASKE • AJA • HANKO • TE-KOTO DIGI • KV-rekrytointi • Taloushallinto • Henkilöstöhallinto • Maksatus • TEM • AVI • KKV • Ylläpito ja pienkehittäminen <p>Teknologia-alustan ja infrastruktuurin kehittäminen</p>
<p>Datalähteet</p>	<p>Mylly kokoaa tietoa mm. taloudesta, henkilöstöstä ja palvelutoiminnasta.</p> <p>Järjestelmä sisältää sekä historiaan että ennakkointiin perustuvaa vertailutietoa toimintaympäristöstä, asiakkaista sekä oman toiminnan kehittämisestä.</p>
<p>Nykyiset palvelun käyttäjät ja sidosryhmät</p>	<p>KEHA-keskuksen tuottama ELY/TET/KEHA/TEM sekä kuntien että muiden toimijoiden käyttöön</p>

<p>Tekninen ratkaisu</p>	<p>Azuren päälle rakennettu pilvipohjainen tiedolla johtamisen tietoalusta. Ratkaisun keskiössä on tietovarasto tiedon yhdistämiseen ja mallintamiseen sekä integraatioalusta tiedon kokoamiseen eri tietolähteistä.</p> <p>Data Vault 2.0</p> <p>Integraatiodata tiedon kokoamiseen ja jakamiseen</p> <p>PowerBI-alusta tiedon raportointiin ja analysointiin</p> <p>Tietokannaksi vaihtumassa 2023 aikana vaiheittain Snowflake.</p>
<p>Toteutettavuus ja kehitysmahdollisuudet</p>	<p>Mylly on erittäin varteenotettava vaihtoehto. Heillä on valmis arkkitehtuuri, joka on toteutettu alusta-ajattelulla. Heillä on myös valmis monivirastomalli ja ratkaisutuotantokonsepti, johon uudet projektit istuvat valmiina. Konseptiin kuuluvat myös valmiit puitesopimukset, joten kilpailutuksen läpivienti on mahdollisesti kevyt.</p>
<p>Alustavan kartoituksen tunnistamat mahdollisuudet</p>	<p>Toteutustapa A: Muiden tietoalustojen hyödyntäminen</p> <p>Mahdollisesti Toteutustapa B: Ostopalveluna toteutettava tietoalusta (varauksin)</p>
<p>Alustavan kartoituksen tunnistamat huomiot ja rajoitukset</p>	<p>Tekoälyn kehitys vielä käynnistämättä, paljon kehityshankkeita, resursoinnissa saattaa ilmentyä pullonkauloja.</p>

3.3. Tietoalusta Antero

Tietoalusta	Antero
<p>Kuvaus käytössä olevasta ratkaisusta</p>	<p>Antero on tietoalusta, joka tukee Vipusen palvelua. Se on kaikki koulutusasteet ja -sektorit kattava opetushallinnon tilastopalvelu, jonka sisällöllisestä ylläpidosta vastaavat yhdessä opetus- ja kulttuuriministeriö ja Opetushallitus (60/40).</p>

Kehityshistoria	<p>Vipunen on ollut tuotannossa jo pitkään, ja se on toteutettu sekä avoimen lähdekoodin että Microsoftin OnPremise-ratkaisuilla.</p> <p>Vipusen/Anteron osalta on tunnistettu lyhyen tähtäimen kehitystarpeet.</p> <p>Nykyisen palvelun uudistamista versioon 2.0 on pohdittu. Tämä sisältäisi mahdollisesti sekä teknologiapäivityksen että uuden toimintamallin. Vipunen 2.0:n osalta ei ole tehty selvitystä eikä määrittelyä.</p>
Datalähteet	<p>Koski, Koski-raporttikanta, Varda, varhaiskasvatuksen luovutuspalvelu, ODW, eHOKS, Valpas, Valos, Ylioppilastutkintorekisteri, Organisaatiopalvelu, Koodistopalvelu, Kouta, ePerusteet, KOTA, Virta-opintotietopalvelu, Arvo, TK tilastojärjestelmä, YTJ, Suomi.fi-koodistot, openstreetmap.org</p>
Nykyiset palvelun käyttäjät ja sidosryhmät	<p>Vipunen palvelee koulutus- ja tiedepolitiikan suunnittelua ja ohjausta sekä toimii kansalaisten ja sidosryhmien avoimena tietolähteenä.</p>
Tekninen ratkaisu	<p>Vipusen tilastopalvelun muodostavan tietovaraston ja raportointiportaalin teknisestä toteutuksesta vastaa CSC - Tieteen tietotekniikan keskus oy.</p> <p>Anterossa on Microsoftin ETL-ajot ja MS SQL -tietokannat sekä tiedonrikastus ja -yhdistäminen. Niistä muodostetaan MS Tabular view, josta tarjotaan Vipusen näkymiä</p>
Toteutettavuus ja kehitysmahdollisuudet	<p>Arkkitehtuuriselvityksessä nousi esiin joitakin varauksia. Omistajuus ja kokonaisvastuurooli ovat hieman epäselviä. Merkittävä osa resursseista käytetään ylläpitoon kehittämisen sijaan, erillään muun muassa Opetushallituksen kansallisista varannoista. Ilmiöiden esiin saamiseksi joudutaan tekemään paljon käsityötä, ja käyttö on koettu vaikeaksi.</p>
Alustavan kartoituksen tunnistamat mahdollisuudet	<p>Toteutustapa A: Muiden tietoalustojen hyödyntäminen</p>

Alustavan kartoituksen tunnistamat huomiot ja rajoitukset	Jatkokehitys ratkaistava ennen etenemistä. Voisi tarjota palveluna Koski-datat valmiiksi jalostettuna joko suoraan tai Lampi-järjestelmän kautta.
---	---

3.4. Tietoalusta Tietojohdaja

Tietoalusta	Tietojohdaja (DigiFinland)
Kuvaus käytössä olevasta ratkaisusta	<p>DigiFinlandin hallinnoima tiedolla johtamisen palvelu, suunnattu tällä hetkellä sote-palvelujen tarpeisiin. Tietoalustassa on paljon saman tyyppisiä kyvykkyksiä, ja siten toteutus JOD-hankkeen tarpeisiin on mahdollista. Referenssinä Omaolo-palvelu yksilölle.</p> <p>Tietojohdaja on tietoturvallinen ja kustannustehokas julkisen sektorin tietojohdamisen palvelu. Tietojohdaja-palvelu koostuu tietojohdamisen alustasta, sen päälle toteutettavasta tietojohdamisen ratkaisusta, palvelun käyttöönottoesta sekä sovittavan laajuisesta jatkuvan palvelun kokonaisuudesta.</p> <p>Palveluratkaisu räätälöidään aina asiakastarpeen mukaisesti ottaen huomioon asiakasorganisaation omat resurssit ja kyvykkyudet.</p>
Kehityshistoria	<p>Omaolo-analytiikkaa on tuotettu Tietojohdaja-palvelun kautta kesäkuusta 2022 lähtien. Seuraava käyttötapaus (sote-tarve) meni pilotointiin lokakuussa 2022 ja palvelut on otettu tuotantoon 1.1.2023.</p> <p>Asiakkaina ovat tällä hetkellä käytännössä kaikki hyvinvointialueet sekä Helsinki & HUS.</p> <p>Keskusteluja on käyty, ja palvelua on tavoitteena laajentaa erityisesti muille toimialoille, mutta myös sote-toimialan sisällä. Päätöksentekoprosessit ovat hitaita ja toisaalta hallituksen vaihtuminen aiheuttaa tiettyä epävarmuutta, odotuksena on saavuttaa uusia asiakkuuksia tämän vuoden aikana.</p>

<p>Datalähteet</p>	<p>Tietosisällöt käsittävät väestöön ja väestörakenteeseen, elinoloihin ja elinympäristöön, sosiaali- ja terveyspalvelujen käyttöön ja sairastavuuteen liittyviä indikaattoreita sekä näistä muodostettuja mittareita.</p>
<p>Nykyiset palvelun käyttäjät ja sidosryhmät</p>	<p>Sote-tarve: Palvelu perustuu DigiFinlandin Virta-hankkeessa tuotettuun sosiaali- ja terveydenhuollon tarve- ja vähimmäistiedon sisältökokonaisuuteen, rajoittuen kuitenkin avoimen datan tietolähteisiin. Palvelu on suunnattu kaikille hyvinvointialueille (mukaan lukien Helsinki ja HUS) ja se tarjotaan sopimusperusteisena jatkuvana palveluna 1.1.2023 alkaen.</p> <p>Omaolo-analytiikka: Osana DigiFinlandin Omaolo-palvelua tuotetaan analytiikkaa Omaolo-palvelun käyttäjäorganisaatioille muun muassa oirearvioiden täyttömääristä, työjonolle ohjautumisesta ja kiireellisyydestä.</p>
<p>Tekninen ratkaisu</p>	<p>Toteutuksessa hyödynnetään Virta-viitearkkitehtuuria, jossa lähtökohtana on varmistaa tietoturvallinen ja tietosuojan takaava tekninen ympäristö tietojohdamisen kokonaisratkaisun mahdollistamiseksi.</p> <p>Azuren päälle rakennettu ratkaisu.</p> <ul style="list-style-type: none"> - Snowflake DW - Wherescape ETL - Power BI -visualisointi - Ellie-käsitelmämallinnus
<p>Toteutettavuus ja kehitysmahdollisuudet</p>	<p>Tietojohdaja on palveluna erittäin vartenotettava vaihtoehto. Heillä on valmis arkkitehtuuri ja prosessit sekä valmis palvelukonsepti. Palvelun voi ostaa kilpailuttamatta, ja siihen voi saada myös VM-lisärahoitusta.</p> <p>Digijohdaja on tilaajalle kuitenkin uusi järjestelmä eikä siellä ole valmiina koulutus- tai työelämätoimijoiden tietoa.</p>

Alustavan kartoituksen tunnistamat mahdollisuudet	Toteutustapa B: Ostopalveluna toteutettava tietoalusta
Alustavan kartoituksen tunnistamat huomiot ja rajoitukset	Sote-puolen tietoalustan kyvykkyysien yhdistäminen koulutus- ja työelämädataan saattaa avata myöhemmin merkittäviä synergioita. Ostopalvelu saattaa luoda toimittajariippuvuuden. Palvelua hankittaessa kannattaa sopia huolellisesti, mitä tapahtuu, jos palvelusopimus halutaan joskus purkaa tai toimittajaa vaihtaa.

3.5. Tietoalusta Opintoallas

Tietoalusta	Opintoallas
Kuvaus käytössä olevasta ratkaisusta	Digivisio kehittää Opintoallas-ratkaisua, joka tukee Digivisio 2030 -hanketta. Digivisio 2030 on kaikkien suomalaisten korkeakoulujen yhteinen hanke, jonka tavoitteena on luoda oppimiselle tulevaisuus, joka hyödyttää niin korkeakouluja, opiskelijoita kuin yhteiskuntaa.
Kehityshistoria	Kehitys aloitettu Q3/2022. Tarkoitus tuottaa kolme versiota, joista palvelun 1. versio on nyt tuotannossa. Loput palvelut julkaistaan 12/2024 mennessä.
Datalähteet	Tietovarannot eivät ole fyysisiä tietovarantoja vaan loogisia tai virtuaalisia. Tietovarannot: Henkilötietovaranto, profiilitietovaranto (osaaminen, suoritukset, opinto-oikeudet, suostumukset, historia) tarjontatietovaranto (korkeakoulujen tarjonta, maksutiedot, ilmoittautumistiedot), Opintoallas (korkeakoulujen data opintotietojärjestelmistä), organisaatietietovaranto (organisaation tiedot, käyttäjät ja roolit), koodistotietovaranto, analytiikka, kaikki koulutustarjonta (sekä formaali että informaali).

Nykyiset palvelun käyttäjät ja sidosryhmät	Kaikki Suomen yliopistot ja korkeakoulut (38 kpl)
Tekninen ratkaisu	Digivisio koostuu tarjontapalveluista, joihin mikä tahansa palvelu voi kytkeytyä integraatioalustan kautta. Tietoalusta koostuu mikropalvelukerroksista, joita kutsutaan tietoalustaksi. Teknisesti palvelu sijaitsee AWS:ssä ja se on Kubernetes-pohjainen ratkaisu, joka hyödyntää PostgreSQL-tietokannan hallintajärjestelmää ja MySQL-tietokantaa.
Toteutettavuus ja kehitysmahdollisuudet	Opintoallas ei ole aito tietoalusta. Se voi olla hallinnollisesti ja aikataulullisesti haastava yhdistää, mutta mikropalvelurakenne mahdollistaisi, että palvelu voisi toimia loogisena tietoalustana korkeakouludatalle tai tarjota dataa virtuaalisesti integraatorajapintansa kautta.
Alustavan kartoituksen tunnistamat mahdollisuudet	Mahdollisesti toteutustapa A: Muiden tietoalustojen hyödyntäminen (varauksin).
Alustavan kartoituksen tunnistamat huomiot ja rajoitukset	Resurssit ja organisointi kohdistettu Digivision tavoitteisiin.

4. Tietoalustojen vertailu

4.1. Vertailuun otettavat tietoalustat

Esiselvitys on valinnut seuraavat tietoalustojen tarjoajat tarkempaan toiminnalliseen ja ei-toiminnalliseen vertailuun. Taulukossa 5 on esitetty JOD-palvelujen toteutettavuus vertailuun otettavia tietoalustoja hyödyntäen.

Palvelu	Kuvaus	vaihtoehto
---------	--------	------------

Mylly	KEHA-keskuksen tuottama pilvipohjainen tiedolla johtamisen tietoaalusta.	Mylly on ensisijainen vaihtoehto toteutustavassa A.
Antero	Antero on tietoaalusta Vipusen tilastopalvelu. Sen sisällöstä vastaavat yhdessä OKM ja OPH.	Antero on varauksin soveltuva vaihtoehto toteutustavassa A. Se soveltuu paremmin koulutustietojen lähteeksi kuin JOD-tietoaalustaksi.
Opintoallas	Digivisio-hankkeen kehitteillä oleva korkeakoulujen yhteinen data-alusta.	Opintoallas soveltuu paremmin korkeakoulujen tarjoomatietojen lähteeksi kuin JOD-tietoaalustaksi.
Tietojohtaja	DigiFinlandin hallinnoima tiedolla johtamisen palvelu, joka on suunnattu tällä hetkellä sote-palvelujen tarpeisiin, mutta hakee laajennusta muihin palveluihin.	Tietojohtaja on ensisijainen vaihtoehto toteutustavassa B.

Taulukko 5.

4.2. Toiminnallisten vaatimusten arviointi

Toiminnallisten vaatimusten arviointi pitää sisällään JOD-tietoaalustalta vaadittavat keskeiset toiminnallisuudet tiedon käsittelyn ja hallinnan osalta. Lisäksi tietoaalustan on täytettävä JOD-tavoitearkkitehtuurin linjaukset ja vastaavuus yleisiin vaatimuksiin kuten saavutettavuuteen. Arkkitehtuurilinjaukset ovat löydettävissä dokumentista ”*RFF-tavoitearkkitehtuuriraportti*”.

Taulukko 6 kuvaa toiminnalliset vaatimukset kategoriatasolla.

Yleiset toiminnalliset vaatimukset	
Tietoaalustan perustoiminnallisuudet	Arkkitehtuurin kokonaiskattavuus. (Tiedon tuonti, jalostus, tarjoaminen, haku)

Tietoalustan vastaavuus JOD-hankkeen arkkitehtuurilinjauksiin	Vaatimukset kuvattu tavoitearkkitehtuurin dokumentissa
Digitaalisten palveluiden hyödyntäminen osana JOD-tietoalustaa	Kustannusten muodostuminen, hankinta vs. itse tekeminen, kustannus suhteessa kapasiteettiin ym.
Datan hallinnan toiminnallinen prosessi	
Tiedon jalostaminen	Laaja tiedon jalostus ja yhdistelemistarve eri tietolähteistä. Kehittyvät vs. kypsä teknologia. Koneoppimisen tai tekoälyratkaisujen hyödyntäminen.
Tiedon anonymisointi	Tietoalustassa tulee olla mahdollista anonymisoida tietoa. Tietoalustan tulee hyödyntää tähän liittyviä teknologioita. Arvioitu teknologian elinikä, anonymisoinnin tyyppi (anonyymi/pseudonyymi), GDPR, prosessi. Tietoalustan tietokartta yksilöllisen tiedon käsittelyyn.
Tiedon hyödyntäminen alustalta dynaamisesti	Tietoalustalta tulee olla mahdollista hyödyntää tietoa dynaamisesti ja reaaliaikaisesti yhdistämällä eri tietolähteitä.
Tiedon tallentaminen palveluista syntyvän datan pohjalta	Käyttäjien tuottaman datan tallentaminen ja hyödyntäminen toiminnallisuuden kehittämisessä tai uusien toiminnallisuuksien luomisessa.
Mallintaminen ja analyysi	Helppokäyttöiset mahdollisuudet datajoukon mallintamiseen ja analyysien tekemiseen.
Tiedon jakaminen palveluna	Mahdollisuudet tarjota tietoja ulkopuolisille käyttäjille rajapintojen avulla tai palveluina niin, että organisaatio isännöi tietojaan pilvessä ja tarjoaa niihin pääsyn käyttäjille.

<p>Tekstianalytiikka ja muu tekoälyn hyödyntäminen</p>	<p>Mikä on tietoalustan valmius tällä hetkellä sekä yleinen tekoälyn hyödyntämisen tiekartta?</p> <p>Nopeasti kehittyvä tekstianalytiikan hyödyntäminen ja sen ohjaaminen datalla sekä digitaalisella ohjeistuksella.</p> <p>Huomiona myös nopeasti etenevä ohjelmistorobotiikan hyödyntäminen.</p>
--	---

Taulukko 6. Toiminnalliset vaatimukset

4.3. Toiminnallisen vertailun tulokset

Seuraavissa taulukoissa on kuvattu vertailuun valittavien tietoalustojen vastaavuus toiminnallisiin vaatimuksiin.

Mylly	Vastaavuus vaatimukseen
<p>Tietoalustan perustoiminnallisuudet</p>	<p>Moderni MS Azure -pohjainen tietoalusta, jolla on kattavat tietoalustan perustoiminnallisuudet.</p>
<p>Tietoalustan vastaavuus JOD-hankkeen arkkitehtuurilinjauksiin</p>	<p>Vastaa valtiohallinnon arkkitehtuurilinjauksia.</p>
<p>Digitaalisten palveluiden hyödyntäminen osana JOD-tietoalustaa</p>	<p>Myllyn hyödyntäminen tuo synergiaa suhteessa vaihtoehtoihin, joissa tietoalusta toteutettaisiin itse tai hyödynnetään kumppaneita, joilla on tietoalusta mutta ei JOD-ekosysteemin vaatimia tietosisältöjä.</p> <p>Mylly vastaa jo osin suoraan JOD-palvelujen tarpeisiin mm. työelämätiedon tuottamisessa ja hyödyntämisessä, mutta tämän lisäksi tarvitaan myös kokonaan uusia palveluja.</p> <p>JOD-palvelujen vaatima työmarkkinatieto ja kohtaanto voivat hyödyntää tietoalustan tietoa. Alustalla ei ole kaikkea tietoa, mutta JOD-palvelun keräämä ja jalostama tieto todennäköisesti hyödyntäisi myös muita Myllyn olemassa olevia tuotteita.</p>

Tiedon jalostaminen	Monipuoliset mahdollisuudet hyödyntää ETL-yhdistelyjä. Tekoäly ja koneoppiminen puuttuvat.
Tiedon anonymisointi	Tietoa voidaan anonymisoida valmiiden algoritmien avulla.
Tiedon hyödyntäminen alustalta dynaamisesti	Tietoalusta mahdollistaa itsepalveluna tehtävän analytiikan ja valmisanalyysien hyödyntämisen. Lisäksi tietoalusta pitää sisällään valmiita raportteja.
Tiedon tallentaminen palveluista syntyvän datan pohjalta	Käytetään työmarkkinatorin ratkaisussa.
Mallintaminen ja analyysi	Data Vault 2.0 (Ei helppo, mutta laadukas)
Tiedon jakaminen palveluna	On valmis rajapintapalvelukokonaisuus
Tekstianalytiikka ja muu tekoälyn hyödyntäminen	Ei valmiina, mutta kehityskartalla.

Tietojohtaja	Vastaavuus vaatimukseen
Tietoalustan perustoiminnallisuudet	Moderni parhaita ratkaisuja hyödyntävä tietoalusta, saatavilla pilvi- ja OnPremise-ratkaisuna. Sisältää kattavat tietoalustan perustoiminnallisuudet.
Tietoalustan vastaavuus JOD-hankkeen arkkitehtuurilinjauksiin	Vastaa valtiohallinnon arkkitehtuurilinjauksia.

Digitaalisten palveluiden hyödyntäminen osana JOD-tietoalustaa	Synergiaa suhteessa itse tehtyyn. Vältetään ohjelmisto- ja toteutuskilpailutukselta. Mahdollista saada lisärahoitusta 200 000 euroa per käyttötapaus vuoden 2023 aikana käynnistettyyn kehitykseen. Hallinnointipalvelut resursoidaan käyttötapauskohtaisesti. Toteutuksen kustannuksia ei voi arvioida.
Tiedon jalostaminen	Tietovarastoautomaatio (WhereScape). Tietojohtajassa ei ole tekoäly- tai koneoppimista, mutta toimittajalla on osaamista näiden toteutukseen.
Tiedon anonymisointi	Ratkaisu on rakennettu henkilötietojen käsittelyä varten (ml. turvaluokittelu), yleiskäyttöistä avoimen datan tietoa on valmiina ja palvelussa on vahva tietoturva sekä anonymisointi- ja pseudonymisointiosaaminen. Henkilötietoja ei tallenneta pilvipalveluihin henkilötietoja vaan tiedot tallennetaan pseudonymisoituina, joka voidaan tehdä kaksisuuntaisena.
Tiedon hyödyntäminen alustalta dynaamisesti	On mahdollista laatia itsepalveluanalytiikkaa, valmisanalyyseja ja raportteja sekä erilaisia ratkaisuja.
Tiedon tallentaminen palveluista syntyvän datan pohjalta	Ei valmiina, mutta DigiFinland on toteuttanut Omaolo-ratkaisussa. DigiFinland on myös Aurora AI:ssa tehnyt suosittelumoottorin, joka suosittelee valmiita palvelupolkuja.
Mallintaminen ja analyysi	Kyllä
Tiedon jakaminen palveluna	On palveluita, jotka pystyvät jakamaan valmisanalyysejä, jalostettua dataa ja raakadataa.
Tekstianalytiikka ja muu tekoälyn hyödyntäminen	Aurora AI:ssa tehty rippikoululaisille suosittelumoottori, joka suosittelee valmiita palvelupolkuja. Tekstianalytiikka tulossa myös ASPA:aan.

Antero	Vastaavuus vaatimukseen
Tietoalustan perustoiminnallisuudet	Vipunen on ollut tuotannossa jo pitkään, ja se on toteutettu sekä avoimen lähdekoodin että Microsoftin OnPremise-ratkaisuilla, jotka sisältävät kaikki perustoiminnallisuudet.
Tietoalustan vastaavuus JOD-hankkeen arkkitehtuurilinjauksiin	Vastaa valtiohallinnon arkkitehtuurilinjauksia.
Digitaalisten palveluiden hyödyntäminen osana JOD-tietoalustaa	<p>Synergiaa suhteessa itse tehtyyn. Vältetään ohjelmisto- ja mahdollisesti toteutuskiipailutukselta. Toteutuksen kustannuksia ei voida arvioida.</p> <p>JOD-palvelujen tarkoituksiin tarvitaan laajempaa analytiikkaa kuin mitä tällä hetkellä on valmiina. JOD-palvelut koulutustarjonnasta tulevat nojaamaan osittain tietoalustan tietoon. Anteroon on läheisesti yhteydessä Lampi-tietoallas, johon Anteron tiedot siirretään. Lampi on toteutettu AWS:n data lakeen, josta pystytään jakamaan Anteron tietoja tehokkaasti. Mahdollisesti tätä alustaa voitaisiin hyödyntää myös muiden tietojen varastoinnissa ja jakelussa.</p>
Tiedon jalostaminen	Toteutettu skripteillä (Ei ETL-työkalua): Datankäsittely MS Powershellillä ja DAX:lla. Työ on manuaalista ja aikaa vievää. Tekoäly tai koneoppiminen puuttuu.
Tiedon anonymisointi	Pitkälle mietityt anonymisointi- ja pseudonymisointityökalut, joita on mahdollista hyödyntää.
Tiedon hyödyntäminen alustalta dynaamisesti	Ei dynaamista hyödyntämismahdollisuutta. Tieto tarjoillaan Tabular Modelissa Exceliin.
Tiedon tallentaminen palveluista syntyvän datan pohjalta	Omien Excel-tiedostojen tallentaminen kansioihin on mahdollista, mutta ei suoraan JOD-palvelujen vaatimalla tavalla, jossa palveluista syntyvä tieto tallennettaisiin suoraan tietoalustaan.

Mallintaminen ja analyysi	Jenkins open source, Python ja R-skriptit, joilla koodataan skriptejä. MS Analytics Services rakentaa kuutiot manuaalisesti.
Tiedon jakaminen palveluna	Dataa pystytään jakamaan rajapintojen kautta.
Tekstianalytiikka ja muu tekoälyn hyödyntäminen	Ei käytössä.

Opintoallas	Vastaavuus vaatimukseen
Tietoalustan perustoiminnallisuudet	Tietoalusta tarjoaa laajan joukon datan hallinnan peruspalveluita, ja palvelut eivät sinänsä ota kantaa siihen, mitä tietoa käsitellään. Palvelut ovat enemmän ”välittäjä” tiedon käsittelyssä kuin tiedon omistajia. Palvelut, kuten ETL, ovat joustavia erilaisiin tarpeisiin.
Tietoalustan vastaavuus JOD-hankkeen arkkitehtuurilinjauksiin	Vastaa JOD-hankkeen linjauksia.
Digitaalisten palveluiden hyödyntäminen osana JOD-tietoalustaa	Opintoaltaan tietoalusta on teknologian näkökulmasta skaalautuva eikä ota kantaa siihen, mitä dataa tai palveluja tietoalusta hyödyntää. Opintoallas on kuitenkin aikaisessa kehitysvaiheessa muun kuin korkeakoulujen hyödyntämän datan tiedon osalta, eli kehitystyötä ja dataintegrointeja tarvitaan paljon. JOD-palvelut korkeakoulutustarjoomasta tulevat nojaamaan osittain Opintoaltaan tietoon.
Tiedon jalostaminen	Korkeakoulujen tarjoamadataa kerätään ja yhdistetään.
Tiedon anonymisointi	Määritelty ratkaisu kohtauttamiseen. Kaikki data on pseudonymisoitua.

Tiedon hyödyntäminen alustalta dynaamisesti	Dataa voi hyödyntää integraatorajapinnan kautta.
Tiedon tallentaminen palveluista syntyvän datan pohjalta	Tällaisia palveluita on toteutuksessa. Tietomallien määrittäminen on kriittistä.
Mallintaminen ja analyysi	Kysytty tarkennusta, analytiikka on syksyllä menossa toteutukseen.
Tiedon jakaminen palveluna	Onnistuu rajapinnan kautta jaettavina mikropalveluina.
Tekstianalytiikka ja muu tekoälyn hyödyntäminen	Suunnitteilla.

4.4. Ei-toiminnalliset vaatimukset

Ei-toiminnalliset vaatimukset arvioivat JOD-tietoalustan käytännön toteutettavuutta suhteessa tunnistettuihin toteutusvaihtoehtoihin. Toteutettavuudessa arvioidaan isoa kuvaa ja sitä, kuinka helppo ratkaisu on toteuttaa suunnitellussa aikataulussa ja budjetissa.

Taulukossa 7 on kuvattu JOD-hankkeen ei-toiminnalliset vaatimukset ja näiden selite.

Ei-toiminnallinen vaatimus	Kuvaus
Toteutuksen aikataulu	Arvio toteutettavuudesta JOD-hankkeen vaatimassa aikataulussa siten, että vaadittavat 24 JOD-palvelua olisivat toteutettuina vuoden 2025 loppuun mennessä.
Resurssit ja tarvittavat osaaminen	Huomioidaan toteutustavan ja tietoalustan vaatimat resurssit ja osaaminen sekä toteuttajilta että JOD-hankkeelta yleisesti.

Pitkän tähtäimen kustannukset	Huomioidaan syntyvät pitkän tähtäimen kustannukset tällä etenemistavalla.
Teknologian jatkuvuus	Arvioidaan valittavien tietoalustojen teknologian jatkuvuutta ja teknologian elinkaarta.
Operatiivinen ohjausmalli ja tämän ketteryys	Huomioidaan toimittajaympäristö ja kehityksen ohjausmalli JOD-hankkeen näkökulmasta. Kuinka hyvin JOD-hanke pääsee vaikuttamaan kehitykseen?
Innovatiivisuus	Mitä tunnistettuja mahdollisuuksia tai uutta toiminnallisuutta tietojärjestelmä tarjoaa?
Tietoturva	Huomioidaan erityisesti yksilön tietosuojan huomiointi, sekä tietoalustan yleisten tietoturva-vaatimusten huomiointi.

Taulukko 7. Ei-toiminnalliset vaatimukset

4.5. Ei-toiminnallisten vaatimusten vertailun tulokset

Taulukossa 8 on kuvattuna vertailuun otettujen tietoalustojen arvio sopivuudesta ei-toiminnallisten vaatimusten näkökulmasta.

Ei-toiminnallinen vaatimus	Tietojohtaja	Mylly	Antero	Opintoallas
Toteutuksen aikataulu	Nopea käynnistysmahdollisuus. Valmiit mallit ja puitesopimukset. Välttää kilpailutuksen.	Nopea käynnistysmahdollisuus. Valmiit mallit ja puitesopimukset. Välttää kilpailutuksen	Vaatii mallista sopimisen. Teknologian uusiminen vaatii kilpailutuksen.	Vaatii mallista sopimisen, mutta teknologia ei ole este nopealle käynnistymiselle.
Resurssit ja tarvittavat osaaminen	On valmis teknologia-alusta, hallintomalli ja puitesopimukset tekijöille.	On valmis teknologia-alusta, hallintomalli ja puitesopimukset tekijöille.		

Pitkän tähtäimen kustannukset	Skaalautuva pilvialusta, alusta-ajattelu ja modernit työkalut mahdollistavat kustannustehokkaan kehityksen.	Skaalautuva pilvialusta, alusta-ajattelu ja modernit työkalut mahdollistavat kustannustehokkaan kehityksen. On kuitenkin uusi tietoalusta ostajan kannalta. Toteuttajat alihankintana tuovat yhden hallinnollisen kerroksen lisää.	Kustannushyödyt saadaan, jos Anteroa voidaan samalla modernisoida OPH:n uudeksi data-alustaksi, joka korvaa olemassa olevia tietoalustoja. Lyhyellä aikavälillä kustannukset korkeammat, koska myös korvattavien alustojen toiminnallisuus kehitettävä.	Kustannusten hallinta vaatii mallin uusimista. Nykymallilla kustannukset jaettu Korkeakoulujen kesken. Saattaa olla muita alustoja työvaltaisempi malli.
Teknologian jatkuvuus	Moderni Microsoft-pilvialusta.	Moderni parhaiden ratkaisujen pilvialusta	Edellisen sukupolven Microsoft OnPremise-alusta. Teknisesti elinkaarensa loppupuolella.	Moderni virtuaalinen mikropalvelualusta. Avoimen lähdekoodin laaja hyödyntäminen.
Operatiivinen ohjausmalli ja tämän ketteryys	Erinomainen saatavuus osajista markkinoilla.	Hyvä saatavuus osaamisesta markkinoilla.	Hyvä saatavuus osaamisesta markkinoilla.	Hyvä saatavuus markkinoilla.
Innovatiivisuus	Alusta-ajattelu ja valmiit prosessit mahdollistavat nopean innovoinnin.	Alusta-ajattelu ja valmiit prosessit mahdollistavat nopean innovoinnin.	Vaatii mallista sopimisen. Teknologian uusiminen vaatii kilpailutuksen.	Mikropalvelurakenne, avoin lähdekoodi, Owasp Zap -työkalu.
Tietoturva	Toteutettu tietoturvaohjeiden mukaisesti.	Toteutettu tietoturvaohjeiden mukaisesti. Saatavilla pilvi- ja OnPremise-palveluna.	Toteutettu tietoturvaohjeiden mukaisesti.	Toteutettu tietoturvaohjeiden mukaisesti.

Taulukko 8. Ei-toiminnallisten vaatimusten vertailu

5. SWOT-analyysi

5.1. SWOT-analyysin kuvaus

SWOT-analyysi on strateginen työkalu, joka auttaa organisaatioita arvioimaan vahvuuksia, heikkouksia, mahdollisuuksia ja uhkia. SWOT-analyysi koostuu neljästä osasta:

1. Vahvuudet (*Strengths*): Alustan sisäiset tekijät, jotka ovat positiivisia ja joilla on vaikutus hankkeen menestykseen.
2. Heikkoudet (*Weaknesses*): Alustan sisäiset tekijät, jotka ovat negatiivisia ja joilla on vaikutus hankkeen menestykseen.
3. Mahdollisuudet (*Opportunities*): Alustan ulkoiset tekijät, jotka voivat tuoda positiivisia vaikutuksia hankkeen menestykseen.
4. Uhat (*Threats*): Alustan ulkoiset tekijät, jotka voivat aiheuttaa negatiivisia vaikutuksia hankkeen menestykseen.

SWOT-analyysi auttaa organisaatioita ymmärtämään nykytilannettaan ja tunnistamaan tulevia haasteita ja mahdollisuuksia. Sen avulla organisaatiot voivat kehittää strategioita, jotka hyödyntävät vahvuuksiaan, korjaavat heikkouksiaan, hyödyntävät mahdollisuuksiaan ja torjuvat uhkiaan.

5.2. SWOT-analyysin tulokset

TIETOJOHTAJA (Toteutustapa B: ostopalveluna ostettava tietoalusta)	
VAHVUUDET	HEIKKOUEDET
<ul style="list-style-type: none"> • Nopea tapa lähteä liikkeelle, koska palvelun tilaamisen käytännöt olemassa • Yleiskäyttöistä avoimen datan tietoa valmiina • Moderni data-alusta • Puolueeton toimija • Vahva tietoturva ja anonymisointi- tai pseudonymisointiosaaminen • Rakennettu henkilötietojen käsittelyä varten (ml. turvaluokittelu) 	<ul style="list-style-type: none"> • On tilaajalle kuitenkin uusi järjestelmä • Ei valmiina opetus- eikä työelämädataa • Ei valmiina tekstianalytiikkaa tai tekoälyä
MAHDOLLISUUDET	UHAT
<ul style="list-style-type: none"> • Laajentaisi datapohjaa "valtiokonsernille" • JOD-hanke tuottaisi lisädataa sotelle (koulutuksen ja työn merkitys terveydelle) • Valmista sote-dataa • Ollut osa Aurora AI-ohjelmaa. Osaamista tekoälyn hyödyntämisestä tekstianalytiikassa • VM:ltä tulisi projektille tänä vuonna mahdollisuus saada lisärahoitusta 200 000 euroa per käyttötapaus • Voisi olla laajennettavissa, kun mietitään opetusdatojen jatkoa • Sote-alan asiakkaat maksavat myös 	<ul style="list-style-type: none"> • JOD-hanke joutuisi maksamaan datankäsittelystä, ja synergiat syntyisivät vasta myöhemmin • Toteuttajat alihankintana, tuo yhden hallinnollisen kerroksen lisää • Luo toimittajariippuvuuden (kuitenkin DigiFinland voi vaihtaa toimittajia)

<p>teknologiakehityksestä, joka voi hyödyntää myös koulutusalan toimijoita</p>	
--	--

MYLLY (Toteutustapa A: hyödyntää muita tietoalustoja)	
VAHVUUDET	HEIKKOUEDET
<ul style="list-style-type: none"> • Moderni tietoalusta • Valmis organisaatio ja prosessit • Pohjalla alusta-ajattelu • Työmarkkinatietoa valmiina ja valmiiksi jalostettuna • Hyvin organisoitunut valmis monivirastomalli, jossa tuotteita • On muita vastaavia ulkoisella rahoituksella toimivia ratkaisuja 	<ul style="list-style-type: none"> • Ei valmiina opetus- tai koulutustietoja • Ei valmiina tekstianalytiikkaa • Onko henkilötietojen käsittely riittävällä tasolla?
MAHDOLLISUUDET	UHAT
<ul style="list-style-type: none"> • Tarkoitus panostaa tekoälyratkaisuihin, synergiaetuja saavutettavissa. • Laajentaisi Myllyn tietopohjaa muihin analyyseihin • Voisi olla laajennettavissa, esim. Opetustietoaltaiden korvaajaksi 	<ul style="list-style-type: none"> • Paljon muita kehitysprojekteja. Riittävätkö resurssit? • Tietokantateknologian vaihto

ANTERO (Toteutustapa A: hyödyntää muita tietoalustoja)	
VAHVUUDET	HEIKKOUEDET

<ul style="list-style-type: none"> • Valmis organisaatio ja prosessit • Paljon opetus- ja koulutustietoa valmiina ja valmiiksi jalostettuna • OPM:n ja OPH:n yhteishanke • Henkilödatan käsittely- ja anonymisointiosaamista 	<ul style="list-style-type: none"> • Ei valmiina työmarkkinatietoa • Ei-standardoitu ratkaisu • Vaatii osaajien kouluttamista • Paljon manuaalista skriptausta • Ei valmiina tekstianalytiikkaa • Toimittajariippuvuus
MAHDOLLISUUDET	UHAT
<ul style="list-style-type: none"> • Projektin aikana olisi mahdollista kehittää myös Anteron teknologialustaa • Laajentaisi datapohjaa Anteron muihin analyysiin • Hyödyntää koottua koulutusdataa valmiiksi käsiteltynä 	<ul style="list-style-type: none"> • Paljon muita kehitysprojekteja. Riittävätkö resurssit? • Osaamista ei löydy • Saattaa osoittautua kalliiksi ja hitaaksi kehittää ja ylläpitää • Saattaa vaatia teknologioiden vaihtoa

OPINTOALLAS (Toteutustapa A: hyödyntää muita tietoalustoja) - arvio varauksin	
VAHVUUDET	HEIKKOUEDET
<ul style="list-style-type: none"> • Moderni mikropalveluarkkitehtuuri • Palvelurakenteet, helppo datan jakaminen • Vahva tietoturva ja anonymisointi- sekä pseudonimisointiosaaminen 	<ul style="list-style-type: none"> • Ei tietoallas • Ei valmiita tietovarasto-ohjelmistoja • Organisoitu korkeakoulujen ympärille. • On tilaajalle uusi järjestelmä • Ei valmiina opetus- eikä työelämädataa • Ei valmiina tekstianalytiikkaa tai tekoälyä
MAHDOLLISUUDET	UHAT

<ul style="list-style-type: none">• Hyödyntää heidän kokoamaansa korkeakouludataa suoraan• Voisi olla laajennettavissa esim. Opetustietoaltaiden korvaajaksi.	<ul style="list-style-type: none">• Fokus ja prioriteetti Digivision toteutuksessa• Raskas tietojen yhdistely ongelmallista• Vaatii toimintamallin muutoksia, jos halutaan laajentaa mittakaavaa
--	--

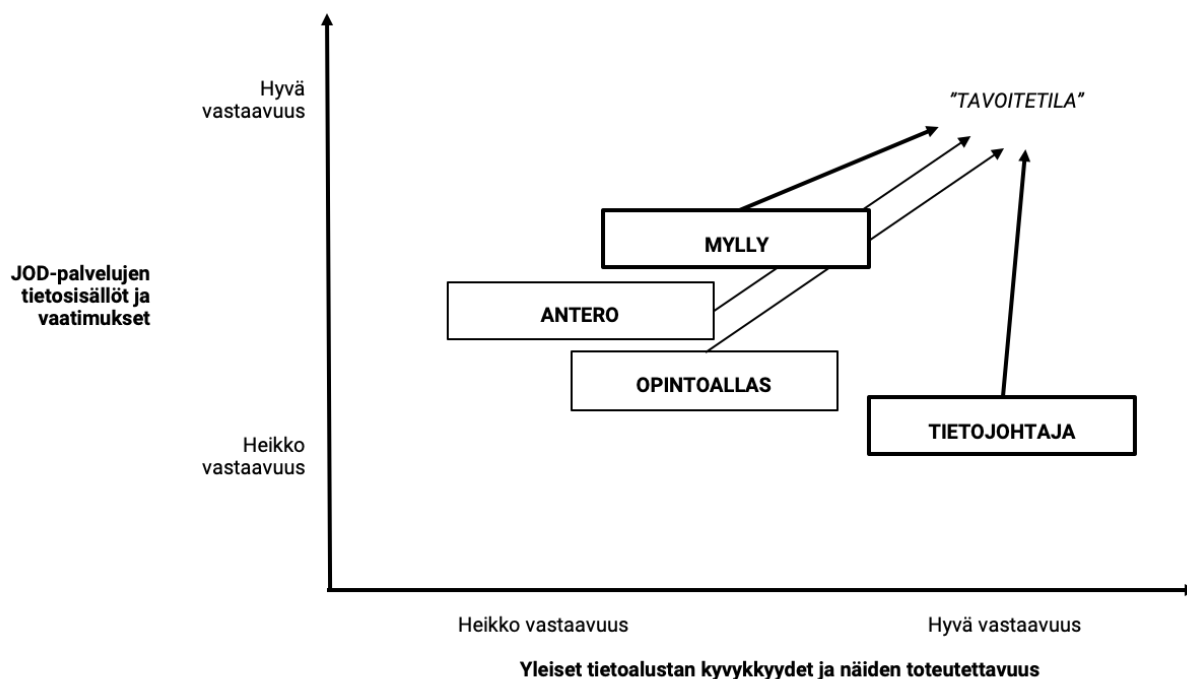
6. Yhteenveto

6.1. Tietoalustojen vertailujen yhteenveto

Esiselvitys ei voinut tällä hetkellä käytävissä olevilla tiedoilla antamaan yksiselitteistä vastausta JOD-hankkeen kannalta parhaaseen tietoalustaratkaisuun. JOD-hankkeen kannalta tarpeisiin soveltuvia vaihtoehtoja kyettiin kuitenkin rajaamaan ja lisäksi nostettiin esiin tarpeisiin soveltuvien vaihtoehtojen hyviä ja huonoja puolia. Tarkemman tason toteutettavuus on mahdollista selvittää tarjousmenettelyn avulla.

Esiselvitys kykeni myös rajaamaan pois toteutusvaihtoehdon, jossa JOD-hanke toteuttaisi kokonaan oman tietoalustansa. Tämä vaihtoehto ei ole JOD-hankkeelle annetulla aikataululla realistinen, ja olisi lisäksi päällekkäinen ratkaisu jo olemassa olevien JOD-tarpeisiin soveltuvien tietoalustojen kanssa.

Alla olevassa kuvassa 6 on havainnollistettu toteutusvaihtoehtojen vertailuun otettujen tietoalustojen suhdetta JOD-tavoitettiin.



Kuva 6. Tietoalustojen vertailun yhteenveto

Toteuttamiskelpoisimmat JOD-tietoalustan vaihtoehdot ovat **Myllly** ja **Tietojohtaja**. Mylllyn osalta päästään parhaiten liikkeelle, ja alustassa on jo nykyisellään JOD-palveluiden tarvitsemää tietoa. Tietojohtaja ei tällä hetkellä pidä sisällään JOD-palvelujen tarvitsemää

tietoa. Toisaalta Tietojohtaja voi pitkällä tähtäimellä antaa synergiaa sote-palvelujen puolelta tulevasta kehityksestä, erityisesti yksilön tietojen hallinnan osalta. Tietojohtaja on vienyt yksilön tiedonhallinnan kyvykkyksiä eteenpäin sote-palveluissa ja näitä voidaan osittain hyödyntää myös JOD-palveluissa. Lisäksi Tietojohtaja tarjoaa vakautta eli rahoitusta tietoalustaan tulee muualtakin kuin koulutuksen ja työllisyyden ekosysteemistä. Tietojohtajan osalta osaamisen siirrossa DigiFinlandin puolelle täytyy varautua jonkinlaiseen viiveeseen suhteessa Kehaan, joka on osa JOD-ekosysteemiä ja myös tuleva palveluiden käyttäjä.

Aikataulupaine nostaa Myllyn tai Tietojohtajan päälle rakennettavat ratkaisut parhaiksi vaihtoehdoiksi, koska näissä on valmis alusta-ajattelu ja kilpailutetut puitesopimukset, jotka mahdollistavat nopean toteutuksen käynnistämisen. Molemmissa tapauksissa Mylly tai Tietojohtaja olisivat JOD-hankkeen näkökulmasta vaadittava ”oma” tietoalusta, johon muut sidosryhmien tietoalustat tarjoavat dataintegraation.

Antero on tietoalusta, jossa on paljon JOD-hankkeen näkökulmasta valmista tietoa. Lisäksi Anterossa on henkilötietojen käsittelyn osalta valmiita prosesseja. Anteron osalta suurimmaksi esteeksi muodostuu tietoalustan toteutustapa skriptauksella, joka on JOD-vaatimusten toteutettavuuden ja aikataulun osalta riski. Toteutustapa vaatii myös resursseja, joilla tähän soveltuvaa osaamista. Tästä syystä Antero olisi vaihtoehto, mikäli samalla uudistettaisiin teknologinen arkkitehtuuri, ja JOD-hanke voisi pitkällä tähtäimellä tarjota myös Anteron tarjoamat palvelut. Anteron osalta työ alkaisi yhteisen tiekartan määrittämisellä, joka tulee huomioida varauksena toteutuksen aikataulussa.

Opintoaltaassa on JOD-palvelujen näkökulmasta relevanttia tietoa, mutta toteutustapa ei ole varsinainen tietoalusta, vaan enemmän kokoelma palveluja. Opintoaltaan muuttaminen tietoalustaksi, jossa olisi sekä koulutus- että työelämä tietoja vaatisi ison muutoksen toimintamalliin ja ohjauksen. Myös tietojen omistajuuteen liittyvät kysymykset tulisi selvittää.

6.2. Suositus

JOD-hanke on laaja hanke ja sen toteuttamisen kannalta suosittelemme painottamaan nopeaa tietoalustatyön käynnistymistä toimijan kanssa, jolla on valmiudet aloittaa nopeasti. Useat tietoalustat voivat toteuttaa tarvittavat kyvykkyudet mutta annettu aikataulu vuoden 2025 loppuun muodostuu keskeisimmäksi reunaehdoksi. Tämän johdosta selvityksen suositus on valita joko Mylly tai Tietojohtaja JOD-hankkeen tietoalustaksi. Antero ja Opintoallas varauksin.

Toteutettaessa palveluja kannattaa kuitenkin huomioida, mitä kaikkia palveluita voidaan saada muilta toimijoilta. Anterossa on valmiiksi jalostettuna ja yhdisteltynä laajasti koulutustarjontaa, ja Opintoallas tulee toteuttamaan korkeakoulutarjonnan valmiiksi jalostettuna ja yhdisteltynä. Anteron ja Opintoaltaan valmiiksi hankkimia ja jalostamia tietoja kannattaa käyttää sellaisenaan JOD-palveluiden toteuttamiseksi ja ajan säästämiseksi. Eli Anteroa ja Opintoallasta ei kannata nykyisessä tilassaan ”kuormittaa” vaatimalla heiltä koko JOD-tietoalustan toteutusta, vaan sen sijaan hyödyntää keskeisenä sidosryhmänä JOD-tietoalustalle.

6.3. Selvitettävät asiat

Esiselvitys suosittaa tarkentamaan tunnistettujen tietoalustojen (Mylly, Tietojohdaja ja Antero) osalta seuraavia asioita:

- Käytön ja ylläpidon kustannukset
- Resurssit
- Toteutuksen aloitusaikataulu käytännössä
- Operatiivinen kehityksen hallintamalli ja se, miten tässä hallintamallissa varmistetaan JOD-hankkeen tarpeiden priorisointi

Lisäksi Anteron osalta tietokartan tarkentaminen ja keskustelu mahdollisuudesta, jossa JOD-tietoalusta ja Antero ”yhdistyvät”, ja jossa myös JOD-hankkeen toteutusta käytetään mahdollisuutena uudistaa Anteron datan hallinnan prosesseja.

7. Liitteet

Esiselvitykseen osallistuneet henkilöt

Seuraavat henkilöt ovat osallistuneet ja/tai olleet kutsuttuna esiselvityksen työpalaveriin ja ohjausryhmään.

Henkilö	Rooli
Jussi Järvinen, JJS Konsulttiverkko	Selvityksen laatija
Atte Roine, JJS Konsulttiverkko	Selvityksen laatija
















Jari Kauppi, OKM	Hanketoimisto, projektipäällikkö
Henriikka Kokkola, TEM	Hanketoimisto, erityisasiantuntija
Tomi Kytölä, OKM	JOD-hankevastaava
Tiina Salminen, TEM	JOD hankevastaava
Sami Mäkinen, OPH	JOD KA-ryhmä, It-arkkitehti
Lea Pitkänen, KEHA	Mylly sovelluksen edustaja
Heikki Heikkilä, KEHA	Mylly sovelluksen edustaja
Juho Mikkonen, KEHA	Mylly-sovelluksen edustaja s
Jukka Haapamäki, OKM	Antero-alustan edustaja
Kari Korhonen, OKM	Antero-alustan edustaja
Juho Mikkonen, OPH	Lampi-sovelluksen edustaja
Toni Iltanen, CSC,	Digivisio-arkkitehti
Matti Riihimäki, CSC	JOD KA-ryhmä, Digivisio-arkkitehti
Antti Kanen, Jotpa	JOD KA-ryhmä, it-arkkitehti
Jaakko Pentti, DigiFinland	Tietojohtaja sovelluksen edustus
Sanna-Maria Kumpula, DigiFinland	Tietojohtaja sovelluksen edustus
Meriluoto Satu, Jotpa	JOD KA-ryhmä, projektipäällikkö

Joonas Mäkinen, OPH	Hankeryhmän edustaja
Merja Väistö, OPH	Hankeryhmän edustaja
Ilmari Hyvönen, OKM	Hankeryhmän edustaja
Nokkanen Erja, OPH,	Hankeryhmän edustaja
Orenius Riikka, OPH	Hankeryhmän edustaja
Jonna Korhonen, OPM	Ohjausryhmän edustaja
Pauliina Smolander, Ely-keskus	Ohjausryhmän edustaja
Petri Lempinen, OKM	Ohjausryhmä, puheenjohtaja
Markku Virtanen TEM	Ohjausryhmän edustaja
Hanna Nordlund, CSC	Ohjausryhmän edustaja
Ida Mielityinen	Ohjausryhmän edustaja
Bäckman Marika (OKM);	Hanketoimisto, asiantuntija
Lehtovaara Heidi (OKM)	Hanketoimisto, koordinaattori
Päivi Pietarila (ely-keskus.fi)	JOD KA-ryhmä, it-arkkitehti
Ari Rouvari, CSC	JOD KA-ryhmä, it-arkkitehti

Taulukko 9. Esiselvitykseen osallistuneet henkilöt

Esiselvityksessä käytetyt lähdemateriaalit

Esiselvitys on hyödyntänyt muun muassa seuraavia dokumentteja lähdemateriaaleina.

	Arkkitehtuuri livetyöpaja 19.1.2023 muistio.pdf
	Arkkitehtuuri livetyöpaja 19.4.2023 muistio.pdf
	JOD kerroskomponentit ja integraatiot.xlsx
	JOD tietoalusta esiselvitys_työpajaesitysluonnos 9.5.2023.pptx
	JOD-hankesuunnitelma_28.4.2023.pdf
	JOD-tietoalusta pohjamateriaalia (002).pptx
	JOD-tietoalusta, pohjamateriaalia.pptx
	OPH tietoalustaselvitys.docx
	RRF-tavoitearkkitehtuuriraportti_20.1.2023.pdf
	Tietoalusta MYLLY esittelysetti 29032023.pptx
	Tietoalustan esiselvitykseen liittyvää tarkastelua.pptx
	Tietojohtaja.fi esittely 26.5.2023.pptx
	Työmarkkinatieto_lähteet_työversio.xlsx
	Vipunen-Antero-nykytilan arkkitehtuuri.pdf
	Vipusen ja Lampi-tietoaltaan arviointi - Vipunen yhteistyöryhmä 2023-03-20 (2).pptx

Saatujen dokumenttien lisäksi on hyödynnetty seuraavia tietoportaaaleja:

JOD-tietoalustan kuvaus:

<https://miro.com/app/board/uXjVMRh0vNM=/>

JOD-käyttöliittymien luonnos:

<https://app.conceptboard.com/board/dmfz-c5yr-m6ee-61fg-7b26>

CSC-yksilöpolun rautalankamalli:

https://miro.com/app/board/uXjVPk0p2Hs=/?share_link_id=616460924278

Työmarkkinatieto, työpajamateriaali:

<https://miro.com/app/board/uXjVPCJxvTI=/>

QPR-tietovirtojen mallinnukset:

https://arkkitehtuuripankki.onqpr.com/QPR/Portal/QPR.Isapi.dll?QPRPORTAL&*pudev&FLS=FLS&FORCELOGIN=1

JOD-arkkitehtuurikuvauksia:

<https://wiki.eduuni.fi/pages/viewpage.action?spaceKey=OKMjodigikokonaisarkkitehtuuri&title=JOD-digi+kokonaisarkkitehtuuri+Home>