

Madiskon

Tavoitearkkitehtuuri ja vaatimusmäärittely

Digital Heritage Studio

Versio 2.1

Luonnos / asiantuntijakommentoitavana

Dokumentin tiedot

Dokumentin nimi	Madiskon - tavoitearkkitehtuuri ja vaatimusmäärittely
Dokumenttityyppi	Arkkitehtuuri- ja vaatimusmäärittely
Versio	2.1
Päivämäärä	17.9.2026
Tila	Luonnos / asiantuntijakommentoitavana
Laatija	Pellervo Kokkonen
Kohderyhmä	Kulttuuriperintötoimijat, IT- ja kehittämiskumppanit, kulttuuriperintöalan asiantuntijat, mahdolliset palveluntarjoajat ja rahoittajat
Tarkoitus	Määritellä Madiskonin tavoitearkkitehtuuri, toimintamalli, tietomallin lähtökohdat, toiminnalliset ja ei-toiminnalliset vaatimukset, ohjelmisto- ja teknologiavaihtoehtojen arvioinnin perusteet sekä vaiheittaisen toteuttamisen periaatteet.
Suhde muihin dokumentteihin	Konkretisoi Madiskon - White Paperissa kuvattua visiota ja kokonaisarkkitehtuuria. Pilotti- ja toteutussuunnitelmat soveltavat tätä määrittelyä rajattuihin kehitysvaiheisiin.
Voimassaolo	Käytetään Madiskonin arkkitehtuurin, pilottien, toteutusten ja IT-kumppanien kanssa käytävien keskustelujen yhteisenä lähtökohtana. Päivitetään kehitystyön ja käyttökokemusten perusteella.
Lisenssi	CC BY-ND 4.0

Yhteystiedot	Pellervo Kokkonen Sähköposti: perintostudio@pm.me www.laxman.academy/kulttuuriperintostudio

0. Sisällysluettelo

Dokumentin tiedot

1. Johdanto

Madiskon on aiemmin työnimellä Kulttuuriperintöstudio kehitetty digitaalisen kulttuuriperinnön toimintamalli ja tekninen ympäristö. Syyskuusta 2026 alkaen kokonaisuudesta käytetään nimeä Madiskon. Kansainvälisessä yhteydessä nimeä voidaan täydentää kuvaavalla määreellä Digital Heritage Studio.

2. Tavoitteet, soveltamisala ja rajaus

3. Käyttötapaukset ja toimintamalli

4. Arkkitehtuuriperiaatteet

5. Toiminnalliset vaatimukset

6. Metatiedot, käyttöoikeudet ja tiedon hallinta

7. Instanssien välinen tiedon jakaminen

8. Ei-toiminnalliset vaatimukset

9. Ohjelmisto- ja teknologiavaihtoehdot

10. Vaiheittainen toteuttaminen, pilotointi ja referenssitoteutukset

1. Johdanto

Madiskon on digitaalisten kulttuuriperintöaineistojen vastaanottoon, hallintaan, rikastamiseen, toimitukselliseen käsittelyyn ja monikanavaiseen julkaisemiseen tarkoitettu toimintamalli ja tekninen ympäristö. Sen lähtökohtia, tavoitteita ja kokonaisarkkitehtuuria on kuvattu tarkemmin Madiskonin White Paperissa.

White Paperissa Madiskon määritellään modulaariseksi kokonaisuudeksi, jossa aineiston vastaanotto, säilyttäminen, kontekstualisointi, toimituksellinen työ ja julkaiseminen muodostavat toisistaan erotettavat mutta yhteiseen tietomalliin ja avoimiin rajapintoihin perustuvat palvelukokonaisuudet. Ratkaisua kehitetään vaiheittain siten, että yksittäiset komponentit voivat tuottaa hyötyä myös itsenäisesti ja että kokonaisuutta voidaan laajentaa tarpeiden ja resurssien kasvaessa.

Tämä dokumentti kuvaa Madiskonin tavoitearkkitehtuurin. Sen tarkoituksena on määritellä toimintamalli, keskeiset toiminnalliset ja ei-toiminnalliset vaatimukset, tietomallin ja

tiedonhallinnan periaatteet, yhteentoimivuus sekä perusteet, joilla ohjelmisto- ja teknologiavaihtoehtoja arvioidaan.

Tavoitearkkitehtuuri ei ole yhden pilotin toteutussuunnitelma. Kaikkia tässä dokumentissa kuvattuja kyvykkyyskäsitteitä ei tarvitse toteuttaa samanaikaisesti eikä yksittäisessä pilotissa. Rajatut pilotit ja referenssitoteutukset toimivat tavoitearkkitehtuurin osien todentamisen ja tarkentamisen välineinä.

Dokumentissa ei valita ennalta yksittäistä teknologiaa, ohjelmistotuotetta tai toteuttajaa. Teknisiä vaihtoehtoja tarkastellaan suhteessa Madiskonin toimintamalliin ja arkkitehtuuriperiaatteisiin. Näitä ovat erityisesti avoimuus, modulaarisuus, yhteentoimivuus, tiedon siirrettävyys, jäljitettävyyden, digitaalinen suvereniteetti ja riippumattomuus yksittäisestä ohjelmistosta tai toimittajasta. Avoimen lähdekoodin ratkaisuja, avoimia tietovarantoja, avoimia rajapintoja ja yhteisiä standardeja hyödynnetään ensisijaisesti silloin, kun se on toiminnallisesti, juridisesti ja tietoturvan kannalta perusteltua.

2. Tavoitteet, soveltamisala ja rajaus

2.1 Tavoite

Madiskonin tavoitteena on mahdollistaa digitaalisten kulttuuriperintöaineistojen hallittu kulku aineiston vastaanotosta rakenteistamiseen, rikastamiseen, toimitukselliseen työskentelyyn, julkaisemiseen ja uudelleenkäyttöön. Arkkitehtuurin tulee tukea sekä paikallista työskentelyä että myöhemmin itsenäisten Studio-instanssien verkostomaista yhteistyötä.

Keskeinen pääoma on tieto ja sen väliset suhteet, ei yksittäinen ohjelmisto. Aineistojen, metatietojen, käyttöoikeuksien, lähteiden ja rikastusten tulee säilyä ymmärrettävinä ja siirrettävinä ohjelmistojen, palveluntarjoajien ja organisaatioiden vaihtuessa.

2.2 Tavoitearkkitehtuurin keskeiset kyvykkyudet

- digitaalinen kulttuuriperintöaineisto voidaan vastaanottaa, säilyttää ja käsitellä hallitusti
- aineistoon voidaan liittää sen käyttöä ohjaavat metatiedot, käyttöoikeudet ja käyttörajoitukset
- toimituksellinen työ ja rikastukset voidaan erottaa alkuperäisaineiston säilyttämisestä
- aineisto voidaan liittää rakenteisesti henkilöihin, paikkoihin, tapahtumiin, ajankohtiin ja muihin kohteisiin
- samaa tietoa voidaan hyödyntää useissa julkaisukanavissa
- Studion omasta tietovarannosta ja tarkoituksenmukaisista ulkoisista tietolähteistä voidaan hakea aineistoa, tunnisteita ja rikastavaa metadataa
- itsenäisten Studio-instanssien välinen tiedon jakaminen voidaan toteuttaa hallitusti siten, että alkuperä, käyttöoikeudet, versiohistoria ja muut keskeiset metatiedot säilyvät
- arkkitehtuurin keskeiset ohjelmistokomponentit voidaan toteuttaa avoimiin rajapintoihin ja mahdollisuuksien mukaan avoimen lähdekoodin ratkaisuihin tukeutuen.

2.3 Soveltamisala

Madiskon soveltuu digitaalisten tai digitoitujen valokuvien, asiakirjojen, karttojen, julkaisujen, videoiden, äänitteiden ja muiden kulttuuriperintöaineistojen käsittelyyn. Aineiston ei tarvitse olla valmiiksi täydellisesti kuvailtua, vaan Studio tukee puutteellisesti rakenteistetun aineiston asteittaista täydentämistä.

2.4 Rajaukset

Madiskon ei lähtökohtaisesti ole pitkäaikaissäilytysjärjestelmä, kaiken kattava museo-, arkisto- tai kokoelmanhallintajärjestelmä eikä raskas tutkimus- tai metadatainfrastruktuuri. Se voi hyödyntää tällaisia järjestelmiä ja niiden tuottamaa tietoa rajapintojen, tunnisteiden ja linkitysten kautta.

Studio ei pyri kopioimaan kaikkia ulkoisia tietovarantoja eikä ylläpitämään niiden rikkaita tietomalleja paikallisesti. Integraatioiden tulee olla tarkoituksenmukaisia ja kevyitä: tieto voidaan hakea, linkittää tai tarvittavilta osin tuoda Studioon siten, että lähde ja alkuperäinen tunniste säilyvät.

2.5 Tavoitetilä ja vaiheittainen toteutus

Tavoitearkkitehtuuri kuvaa kypsää kokonaisuutta. Yksittäinen käyttöönotto tai pilotti voi toteuttaa vain osan sen kyvykkyyksistä. Erityisesti verkostomainen tiedonvaihto, kehittynyt semanttinen validointi, automaattinen rikastaminen ja erikoistuneet analyysipalvelut voidaan toteuttaa myöhemmissä vaiheissa.

3. Käyttötapaudet ja toimintamalli

3.1 Toimintaympäristö

Madiskon tukee kulttuuriperintötoimijoiden omaa aineistohallintaa ja toimituksellista työskentelyä. Tavoitetilassa Studio voi muodostua useista toisistaan riippumattomista instansseista, jotka sisältävät samat keskeiset loogiset palvelukokonaisuudet: aineiston vastaanoton, säilyttämisen, metatietojen hallinnan, rikastamisen, toimituksellisen työskentelyn ja julkaisemisen.

Instanssit hallitsevat omia aineistojaan, käyttäjiään ja käyttöoikeuksiaan itsenäisesti. Ne voivat vaihtaa aineistoa, metatietoja ja muita tietosisältöjä yhteisesti sovittujen periaatteiden mukaisesti. Tavoitteena on keskitetyn järjestelmän sijasta itsenäisten mutta yhteentoimivien Madiskon-instanssien toimintamalli.

3.2 Toimijat

Madiskonia voivat käyttää esimerkiksi paikallisyhdistykset, museot, arkistot, tutkimushankkeet, kulttuuriperintöverkostot ja muut kulttuuriperintöaineistoja hallinnoivat tai käyttävät organisaatiot. Eri toimijoiden aineistot, toimintatavat ja resurssit voivat poiketa huomattavasti toisistaan, minkä vuoksi arkkitehtuurin tulee tukea erikokoisia käyttöönottoja.

3.3 Aineistot

Aineisto voi sisältää valokuvia, asiakirjoja, karttoja, julkaisuja, videoita, äänitteitä ja niihin liittyviä metatietoja. Käyttötilanteissa tulee voida käsitellä sekä avoimesti julkaistavaa että käyttöoikeuksiltaan rajoitettua aineistoa. Näin tekijänoikeudet, lisenssit, henkilötiedot, julkaisurajoitukset ja muut käyttöehdot voidaan hallita osana normaalia työnkulkua.

3.4 Keskeinen työnkulku

1. digitaalinen aineisto vastaanotetaan järjestelmään
2. alkuperäinen aineisto tallennetaan muuttumattomana tai sen eheys muuten varmistetaan
3. aineistolle kirjataan tarvittavat kuvailu-, alkuperä-, käyttöoikeus- ja rajoitustiedot
4. aineistoa rakenteistetaan ja rikastetaan
5. aineisto liitetään henkilöihin, paikkoihin, tapahtumiin, ajankohtiin ja muihin sisältöihin
6. asiantuntija arvioi ja hyväksyy tehdyt täydennykset
7. aineistosta muodostetaan toimituksellisia sisältökokonaisuuksia
8. sisältö julkaistaan valittuun julkaisukanavaan.

Työnkulku voi sisältää myös ulkoisen aineiston ja metatiedon hakua. Käyttäjä voi hakea samaan henkilöön, paikkaan, tapahtumaan, teokseen tai aiheeseen liittyvää tietoa Studion omista aineistoista sekä tarkoituksenmukaisista ulkoisista palveluista ja liittää löydetyn tiedon osaksi kuvailua tai toimituksellista työtä.

Tavoitetilassa aineistoa tai siihen liittyviä metatietoja voidaan välittää instanssista toiseen. Vastaanottava instanssi voi hyödyntää aineistoa omassa toimituksellisessa työssään, yhdistää sen omiin aineistoihinsa, täydentää sitä uusilla tiedoilla ja tuottaa siitä omia julkaisujaan. Vastaanottavassa instanssissa tehdyt täydennykset eivät automaattisesti muuta alkuperäisen instanssin tietoja.

3.5 Monikielisyys

Madiskon suunnitellaan alusta alkaen monikieliseksi. Monikielisyys koskee käyttöliittymän lisäksi aineistoa, metatietoa, hakua ja julkaisemista. Samalla kohteella voi olla useita rinnakkaisia nimiä, kuvauksia ja muita sisältöjä eri kielillä. Järjestelmän tulee pystyä tallentamaan alkuperäinen teksti, ihmisen tekemä käännös ja koneellisesti tuotettu käännös erillisinä versioina.

Tietomallin rakenteet ja tekniset tunnisteet ovat kieliriippumattomia. Käyttäjälle näkyvät kenttien nimet, luokitukset ja sisällöt voidaan esittää valitulla käyttöliittymän kielellä. Hakutoimintojen tulee pystyä hyödyntämään samaan kohteeseen liittyviä eri kielisiä nimimuotoja ja vaihtoehtoisia kirjoitusasuja.

3.6 Käyttäjät ja roolit

- aineiston luovuttaja
- aineiston vastaanottaja ja kuvailija
- asiantuntija tai tutkija

- toimittaja tai kuraattori
- hyväksyjä
- julkaisija
- järjestelmän ylläpitäjä

Sama henkilö voi toimia useassa roolissa. Olennaista on, että järjestelmässä voidaan erottaa aineiston tarkastelu, muokkaaminen, hyväksyminen, julkaiseminen ja välittäminen toiselle instanssille.

3.7 Samanarvoisten instanssien tavoitemalli

Madiskonin verkostomaisessa tavoitetilassa instanssit ovat arkkitehtuuriltaan ja toiminnallisilta mahdollisuuksiltaan samanarvoisia. Kukin voi vastaanottaa, hallita, rikastaa, julkaista, luovuttaa ja vastaanottaa aineistoa. Tavoitteena ei ole muodostaa yhteistä keskitettyä tietovarastoa eikä määritellä yhtä instanssia pysyvästi aineiston tuottajaksi ja toista vastaanottajaksi.

- mitä aineistoa ja tietoa instanssien välillä voidaan välittää
- mitä metatietoja aineiston mukana on siirrettävä
- miten aineiston alkuperä ja käsittelyhistoria säilyvät
- miten käyttöoikeudet, lisenssit ja rajoitukset ilmaistaan vastaanottavassa instanssissa
- miten vastaanottava instanssi voi täydentää ja rikastaa aineistoa
- miten täydennykset voidaan välittää takaisin tai asettaa toisen instanssin hyödynnettäviksi
- miten eri instansseissa syntyvät versiot ja mahdolliset ristiriitaiset tiedot erotetaan toisistaan
- miten tiedonvaihdon tapahtumat dokumentoidaan ja tehdään jäljitettäviksi.

4. Arkkitehtuuriperiaatteet

Madiskonin ohjelmisto- ja teknologiavaihtoehtojen tulee tukea tässä luvussa esitettyjä arkkitehtuuriperiaatteita. Periaatteiden tarkoituksena ei ole rajata käytettäviä teknologioita, vaan varmistaa, että ratkaisut tukevat pitkän aikavälin tavoitteita.

Yhteinen ydin tarkoittaa ensisijaisesti yhteistä tietomallia, rajapintoja, vähimmäistoiminnallisuuksia ja yhteentoimivuuden sääntöjä. Se ei välttämättä tarkoita yhtä itse rakennettua ohjelmistotuotetta tai kaikissa instansseissa täysin identtistä ohjelmistokokoonpanoa.

4.1 Toimintamalli määrittää teknologian

Madiskon on ennen kaikkea toimintamalli digitaalisten kulttuuriperintöaineistojen hallintaan, rikastamiseen ja julkaisemiseen. Ohjelmistot ovat tämän toimintamallin toteuttamisen välineitä, eivät sen lähtökohta. Teknologisia ratkaisuja arvioidaan sen perusteella, kuinka hyvin ne tukevat määriteltyjä työnkuluja, tietomallia ja käyttäjien tarpeita.

4.2 Tietomalli ennen ohjelmistoa

Madiskonin keskeinen pääoma on tieto ja sen väliset suhteet, ei yksittäinen ohjelmisto. Arkkitehtuurin perustana on yhteinen tietomalli, jonka avulla aineistot, metatiedot, käyttöoikeudet ja niiden väliset yhteydet voidaan säilyttää riippumatta siitä, millä ohjelmistolla niitä käsitellään. Yhteinen tietomalli sisältää metatietojen lisäksi keskeiset rakenteiset tietoelementit, kuten ajan, paikan ja toimijat.

4.3 Kevyt käyttäjälle, rakenteinen koneelle

Käyttöliittymien ja työnkulkujen tulee olla mahdollisimman helposti omaksuttavia myös käyttäjille, jotka eivät tunne ontologioita tai linkitetyn tiedon teknisiä yksityiskohtia. Taustalla tieto tallennetaan riittävän rakenteisesti, jotta sitä voidaan hakea, yhdistää, siirtää, validoida ja hyödyntää uusissa palveluissa.

4.4 Verkostomainen arkkitehtuuri

Madiskon voi muodostua itsenäisten mutta yhteentoimivien instanssien verkostosta. Jokainen instanssi hallitsee omia aineistojaan, käyttäjiään ja käyttöoikeuksiaan itsenäisesti sekä vastaa omasta toimituksellisesta työskentelystään. Instanssit voivat vaihtaa aineistoa, metatietoja ja muita tietosisältöjä yhteisesti sovittujen rajapintojen ja toimintamallien mukaisesti.

4.5 Yhteinen ydin ja laajennettava ekosysteemi

Madiskonin yhteinen ydin määrittelee tietomallin, rajapinnat ja keskeiset toiminnalliset periaatteet. Uusia toimintoja toteutetaan ensisijaisesti laajennuksina tai erillisinä palveluina. Tällaisia voivat olla julkaisuratkaisut, ulkoiset integraatiot, haku- ja metahakupalvelut, analysointi- ja tunnistustyökalut, kartta- ja paikkatietopalvelut sekä tekoälyä hyödyntävät avustajat.

4.6 Kansainvälisyys ja kieliriippumaton tietomalli

Ohjelmiston sisäinen tietomalli on kieliriippumaton. Tietokannan rakenteet, tekniset kentät ja rajapinnat noudattavat mahdollisuuksien mukaan kansainvälisiä standardeja ja englanninkielisiä tunnisteita. Käyttäjälle nämä esitetään valitulla käyttöliittymän kielellä.

4.7 Avoimet rajapinnat ja yhteiset standardit

Yhteentoimivuus perustuu avoimiin rajapintoihin ja yleisesti käytettyihin standardeihin. Madiskon hyödyntää ensisijaisesti olemassa olevia tietovarantoja, tunnisteita ja tietomalleja eikä pyri rakentamaan rinnakkaisia ratkaisuja silloin, kun avoimia standardeja voidaan hyödyntää.

4.8 Ulkoisen tiedon linkittäminen ja hakeminen

Ulkoista tietoa ei lähtökohtaisesti kopioida Studioon kokonaisuudessaan. Tietovarantoja hyödynnetään tunnisteiden, linkitysten, rajapintahakujen ja metahaun avulla. Tarvittaessa Studioon tuodaan vain käyttötapauksen kannalta tarpeellinen tieto siten, että lähde, alkuperäinen tunniste ja käyttöehdot säilyvät.

4.9 Modulaarisuus ja vaihdettavat komponentit

Aineiston vastaanotto, säilyttäminen, rikastaminen, haku, toimituksellinen työ ja julkaiseminen muodostavat loogisesti erotettavia palveluita. Yksittäisiä ohjelmistokomponentteja voidaan kehittää tai korvata ilman, että koko järjestelmän toimintamallia tai tietomallia tarvitsee muuttaa.

4.10 Toimittajariippumattomuus, avoin lähdekoodi ja digitaalinen suvereniteetti

Avoimen lähdekoodin ohjelmistoja suositetaan aina, kun ne täyttävät toiminnalliset ja laadulliset vaatimukset. Tavoitteena ei ole avoin lähdekoodi itseisarvona, vaan läpinäkyvä, laajennettava ja hallittava ratkaisu, jossa aineistot, metatiedot ja käyttäjätiedot säilyvät niitä ylläpitävien organisaatioiden määräysvallassa.

4.11 Omistajuus, IPR ja jatkuvuus

Konseptin, yhteisen tietomallin, dokumentaation, rajapintamääritysten, lähdekoodin ja laajennusten omistajuus- ja lisenssiperiaatteet tulee ratkaista riittävän varhain. Kehittämisen jatkuvuus ei saa olla riippuvainen yksittäisestä hankkeesta, tutkimus- tai kehittämisorganisaatiosta tai kaupallisesta toimittajasta.

4.12 Tekoälypalvelut osana laajennettavaa ekosysteemiä

Tekoälypalvelut toteutetaan erillisinä palveluina tai laajennuksina, jotka hyödyntävät yhteistä tietomallia ja avoimia rajapintoja. Ne voivat avustaa aineiston analysoinnissa, metadatan muodostamisessa, entiteettien tunnistamisessa, ulkoisten tietovarantojen linkityksessä, monikielisessä käsittelyssä ja toimituksellisessa työssä. Tekoälyn tuottamat ehdotukset eivät muuta hyväksyttyä tietoa automaattisesti.

5. Toiminnalliset vaatimukset

Toiminnalliset vaatimukset on jäsennetty Madiskonin White Paperissa esitetyn kokonaisarkkitehtuurin mukaisesti. Yhteiseen ytimeen (Core) kuuluvat vaatimukset määrittelevät tiedonhallinnan ja yhteentoimivuuden vähimmäisperustan. Tavoitearkkitehtuurin kyvykkyydet ja laajennukset voidaan toteuttaa vaiheittain käyttötarpeen mukaan. Core ei tarkoita yhtä yhteistä ohjelmistotuotetta.

5.1 Heritage Gateway - aineiston vastaanotto

Heritage Gateway muodostaa Madiskonin sisääntulokerroksen. Sen tehtävänä on vastaanottaa digitaalisia kulttuuriperintöaineistoja erilaisista lähteistä ja valmistella ne myöhempää käsittelyä varten. Vastaanoton tulee säilyttää aineiston alkuperä, tunnistetiedot ja käyttöoikeuksiin liittyvät tiedot muuttumattomina.

Core-vaatimukset

- digitaalisen aineiston vastaanotto
- vastaanotetun aineiston muuttumattomuuden ja eheyden varmistaminen ennen sen tallentamista Repositoryyn
- pysyvän tunnisteiden muodostaminen
- alkuperä- ja provenienssitietojen kirjaaminen
- käyttöoikeuksien ja käyttörajoitusten tallentaminen

Tavoitekyvykkyydet ja laajennukset

- tiedostojen erätuonti
- CSV- ja JSON-tuonti
- IIIF-aineistojen vastaanotto
- ulkoisten aineistopalvelujen hyödyntäminen rajapintojen kautta

5.2 Heritage Repository - perintötietovarasto

Heritage Repository toimii kunkin Madiskonin instanssin perintötietovarastona. Eri instanssien repositoryt perustuvat yhteiseen tietomalliin ja rajapintaperiaatteisiin, mutta kukin organisaatio hallitsee omaa tietovarastoaan itsenäisesti. Se säilyttää alkuperäisen aineiston, siihen liittyvät metatiedot, käyttöoikeudet, versiot sekä aineistojen väliset yhteydet. Kaikki myöhemmät rikastukset ja toimitukselliset muutokset tallennetaan alkuperäisaineistosta erillisinä tietokerroksina.

Core-vaatimukset

- alkuperäisen aineiston säilyttäminen tai hallittu viittaus alkuperäiseen
- yhteinen metatietomalli
- käyttöoikeuksien ja lisenssien hallinta
- provenienssin hallinta
- lähdeviitteiden tallentaminen
- versiohistoria
- käsittelyhistorian säilyttäminen
- aineistojen välisten suhteiden tallentaminen
- paikkatiedon hallinta
- maantieteellisten koordinaattien tallentaminen
- kohteiden geometrian hallinta
- aikatasojen ja ajallisten suhteiden tallentaminen

Tavoitekyvykkyydet ja laajennukset

- automaattiset metatietoehdotukset
- ulkoisten tunnisteiden hyödyntäminen
- vaihtoehtoiset objektivarastot
- pitkäaikaissäilytystä tukevat integraatiot

5.3 Heritage Intelligence - rakenteistaminen ja kontekstualisointi

Heritage Intelligence muodostaa järjestelmän älykkään käsittelykerroksen. Sen tehtävänä on auttaa tunnistamaan aineistojen sisältöä, muodostamaan yhteyksiä aineistojen välille sekä ehdottamaan yhteyksiä ulkoisiin tietovarantoihin. Kaikki automaattisesti muodostetut ehdotukset arvioidaan toimituksellisessa työssä ennen niiden hyväksymistä. Rikastukset voivat sisältää myös paikkatietoa, kuten kohteiden tunnistamista, georeferointia, karttakohteiden digitointia tai sijaintitietojen tarkentamista. Tulokset tallennetaan osaksi yhteistä tietomallia riippumatta siitä, millä työkalulla ne on tuotettu.

Core-vaatimukset

- henkilöiden, paikkojen, tapahtumien ja ajankohtien tunnisteiden hallinta
- aineistojen välisten yhteyksien tallentaminen
- linkitetyn tiedon hyödyntäminen
- ehdotusten hyväksymisprosessi
- ihmisen tekemän arvioinnin erottaminen automaattisista ehdotuksista
- rikastuksen lähteen, tekijän, ajankohdan ja hyväksyntätilan tallentaminen

Tavoitekyvykkyydet ja laajennukset

- OCR ja tekstintunnistus
- puheentunnistus
- kuva-, rakennus- ja kohdetunnistus
- automaattiset metadataehdotukset
- Wikidata-, Finto- ja muiden tietovarantojen linkitysehdotukset
- semanttisten suhteiden ehdottaminen
- paikkatiedon automaattinen tai puoliautomaattinen rikastaminen

5.4 Heritage Workspace - toimitusympäristö

Heritage Workspace muodostaa käyttäjän varsinaisen työympäristön. Sen kautta aineistoa tarkastellaan, kuvaillaan, rikastetaan, yhdistetään muihin sisältöihin, arvioidaan ja valmistellaan julkaistavaksi. Työympäristön tulee tukea eri käyttäjärooleja ja erottaa luonnokset, ehdotukset, hyväksytty tieto ja julkaistava sisältö toisistaan.

Core-vaatimukset

- aineiston tarkastelu
- metatietojen täydentäminen ja muokkaaminen
- rikastusten ja muutosehdotusten tarkastelu ja hyväksyminen
- aineistojen linkittäminen henkilöihin, paikkoihin, tapahtumiin, ajankohtiin ja muihin kohteisiin
- toimituksellisten sisältökokonaisuuksien muodostaminen
- käyttäjäroolien ja käyttöoikeuksien huomioiminen työskentelyssä

Tavoitekyvykkyydet ja laajennukset

- karttapohjaiset työkalut
- visuaaliset linkitystyökalut
- kommentointi ja yhteistyöominaisuudet
- tekoälyä hyödyntävät toimitukselliset avustajat
- muutosten vertailu ja rikastusten yhdistämisen tuki

5.5 Publishing - julkaiseminen

Publishing-palvelukokonaisuuden tehtävänä on erottaa aineiston hallinta ja toimituksellinen työ yksittäisestä julkaisukanavasta. Sama rakenteinen tieto voidaan julkaista eri muodoissa ja eri kanaviin ilman, että sitä ylläpidetään erillisinä kopioina.

Core-vaatimukset

- hallittu julkaisuprosessi
- julkaistavan sisällön ja sen lähdeaineiston jäljitettävyys
- julkaisutilan ja hyväksynnän hallinta
- julkaisukanavan erottaminen perintötietovarastosta

Tavoitekyvykkyydet ja laajennukset

- rajapintapohjainen julkaiseminen
- julkaisuprofiilit
- WordPress ja muut sisällönhallintajärjestelmät
- staattiset verkkosivustot
- PDF-julkaisut
- kartta- ja reittipalvelut
- IIIF-julkaiseminen
- muut sovelluskohtaiset julkaisukanavat

5.6 Käyttäjähallinta, lokitus ja versiointi

Käyttäjähallinta, käyttöoikeudet, lokitus ja versiointi ovat läpileikkaavia toimintoja. Niiden tulee tukea aineiston koko elinkaarta vastaanotosta julkaisemiseen ja mahdolliseen instanssien väliseen jakamiseen.

- käyttäjien ja roolien hallinta
- aineisto- ja toimintokohtaiset käyttöoikeudet
- keskeisten tapahtumien lokitus
- versiohistorian säilyttäminen
- hyväksyntöjen ja julkaisujen jäljitettävyys
- tarvittaessa kertakirjautuminen ja ulkoiset tunnistautumispalvelut

5.7 Haku, metatietohaku ja aineistohaku

Haku on Madiskonin läpileikkaava toiminnallinen kyvykkyys. Sen tulee palvella sekä aineiston löytämistä että rakenteisen tiedon täydentämistä. Haku ei rajoitu Studion omaan tietovarastoon, vaan tavoitearkkitehtuuri mahdollistaa tarkoituksenmukaisten ulkoisten tietolähteiden käyttämisen ilman, että niiden koko sisältö tai tietomalli kopioidaan Studioon.

5.7.1 Studiohaku

Studiohauella käyttäjä etsii aineistoja, kohteita ja toimituksellisia sisältöjä omasta instanssista sekä käyttöoikeuksien salliessa muista käytettävissä olevista Studio-aineistoista. Haun tulee hyödyntää rakenteisia metatietoja, henkilöitä, paikkoja, tapahtumia, ajankohtia, aiheita, ulkoisia tunnisteita ja aineistojen välisiä suhteita.

5.7.2 Metatietohaku

Metatietohauella etsitään ulkoisista auktoriteetti-, sanasto-, ontologia- ja erikoistietopalveluista tunnisteita, nimimuotoja, luokituksia ja muuta aineiston kuvailua tai linkittämistä tukevaa tietoa. Käyttäjän tulee voida arvioida hakutulokset ennen sen liittämistä Studioon. Liitetyn tiedon lähde ja ulkoinen tunniste säilytetään.

5.7.3 Aineistohaku ja metahaku

Aineistohaku etsii samaan henkilöön, paikkaan, tapahtumaan, teokseen, kohteeseen tai aiheeseen liittyviä digitaalisia aineistoja ulkoisista aineistopalveluista. Metahaku voi välittää yhden käyttäjän haun useaan lähteeseen ja esittää tulokset yhtenäisesti siten, että alkuperäinen lähde, tunniste, aineistotyyppi ja käyttöoikeustiedot säilyvät.

Core-vaatimukset

- haku Studion omista aineistoista ja rakenteisesta metadatatista
- haku henkilöiden, paikkojen, tapahtumien, ajankohtien ja aiheiden perusteella
- ulkoisten tunnisteiden käyttäminen hakuehtoina ja linkityksissä
- hakutuloksen lähteen ja tunnisteen säilyttäminen, kun tietoa liitetään Studioon
- käyttöoikeuksien huomioiminen hakutuloksissa

Tavoitekyvykkyudet ja laajennukset

- rajapintahaku ulkoisiin metatietopalveluihin
- rajapintahaku ulkoisiin aineistopalveluihin
- useita lähteitä yhdistävä metahaku
- monikielinen haku ja vaihtoehtoisten nimimuotojen hyödyntäminen
- paikka- ja aikarajaukset
- hakutulosten ehdottaminen rikastuksiksi
- hakuprofiilit eri aineisto- ja metatietolähteille

6. Metatiedot, käyttöoikeudet ja tiedon hallinta

6.1 Metatietomallin lähtökohdat

Madiskonin yhteinen metatietomalli muodostaa instanssien välisen yhteentoimivuuden perustan. Mallin tulee olla riittävän kevyt käytännön kuvailutyöhön mutta riittävän rakenteinen tiedon yhdistämiseen, hakuun, siirtämiseen ja myöhempään rikastamiseen. Käyttäjän ei tarvitse hallita taustalla olevia ontologioita tai linkitetyn tiedon teknisiä yksityiskohtia.

Tietomallia voidaan vaiheittain mapata ulkoisiin käsitteisiin ja suhteisiin. Kohteille hyödynnetään ensisijaisesti pysyviä ulkoisia tunnisteita silloin, kun tarkoituksenmukainen tunniste on olemassa.

6.2 Kuvailutiedot

Aineistosta tallennetaan sen tunnistamisen, löytämisen, ymmärtämisen ja käytön kannalta tarpeelliset tiedot. Näitä voivat olla esimerkiksi nimeke, kuvaus, aineistotyyppi, aihe, avainsanat, ajankohta, paikka, tekijä, aineistoon liittyvät henkilöt, organisaatiot ja tapahtumat sekä lähde- ja viitetiedot.

6.3 Aika ja aikatasojen hallinta

Kulttuuriperintöaineistossa samaan historialliseen ilmiöön voi liittyä eri aikoina syntyneitä aineistoja. Tietomallin tulee erottaa historiallisen tapahtuman tai ilmiön aika, sitä myöhemmin esittävän tai tulkitsevan aineiston aika sekä dokumentoinnin, digitoinnin tai muun käsittelyn aika.

Aineiston suhde historialliseen tapahtumaan voidaan ilmaista esimerkiksi aikalaislähteenä, myöhempänä esityksenä, tulkintana, elävöityksenä tai dokumentaationa. Näin eri aikakausien aineisto voidaan yhdistää samaan historialliseen ilmiöön sekoittamatta aineistojen provenienssia tai merkitystä.

Ensimmäinen toteutus voi olla kevyt tietomalli- ja käyttöliittymäratkaisu. Myöhemmin sen päälle voidaan rakentaa semanttista ajallista validointia, aikajanoja ja kehittyneempiä hakutoimintoja.

6.4 Paikkatieto

Paikkatieto on yhteisen tietomallin peruselementti. Se voi kuvata aineiston kuvauspaikkaa, kuvassa esiintyvän kohteen sijaintia, tapahtuman paikkaa, kartalla esitettävää kohdetta tai muuta maantieteellistä yhteyttä.

Paikkatiedossa tulee voida erottaa reaali maailman sijainti ja sijainti digitaalisessa aineistossa. Tarvittaessa tallennetaan koordinaattijärjestelmä, geometrian tyyppi, tarkkuus, epävarmuus ja tiedon lähde. Paikkatietoa voidaan hyödyntää hallinnassa, haussa, linkityksissä, analytiikassa ja julkaisemisessa.

6.5 Proveniensi, omistajuus ja lähteet

Jokaisen aineiston ja olennaisen rikastuksen alkuperä tulee voida osoittaa. Provenienssitiedot kuvaavat, mistä aineisto tai tieto on saatu, kuka sen on luonut, luovuttanut tai tuottanut, kuka vastaa sen hallinnasta ja miten se on päätynyt Studioon. Myös ulkoisesta tietolähteestä haetun tai tuodun tiedon lähde säilytetään.

6.6 Tekijänoikeudet, lisenssit ja käyttörajoitukset

Aineistolle tulee voida määritellä sen käyttöä koskevat oikeudet ja rajoitukset. Näihin voivat kuulua tekijänoikeudet, lisenssit, julkaisuoikeudet, käyttörajoitukset ja määräaikaisten käyttöehtojen. Oikeustiedon tulee seurata aineistoa viennissä, julkaisemisessa ja instanssien välisessä jakamisessa.

6.7 Henkilötiedot ja muu suojattava tieto

Tietomallin ja työnkulkujen tulee mahdollistaa henkilötietoja tai muuta suojattavaa tietoa sisältävien aineistojen tunnistaminen, käyttörajoitusten hallinta ja julkaisupäätösten dokumentointi.

6.8 Rikastukset ja toimituksellinen tieto

Alkuperäinen aineisto erotetaan sitä täydentävästä tiedosta. Rikastus voi olla asiantuntijan tuottama tieto, ulkoiseen lähteeseen perustuva linkitys tai automaattisesti tuotettu ehdotus. Rikastuksesta tallennetaan sen lähde, tekijä tai tuottanut palvelu, ajankohta, mahdollinen luottamustaso ja hyväksyntätila.

6.9 Hyväksyntä, versiointi ja ristiriitainen tieto

Järjestelmän tulee säilyttää tieto keskeisistä muutoksista, niiden tekijöistä ja hyväksynnöistä. Historiallinen tieto ei aina ole yksiselitteistä, joten tietomallin tulee mahdollistaa tiedon lähteen ja tarvittaessa epävarmuuden ilmaiseminen. Ristiriitaisia tulkintoja ei tarvitse pakottaa yhdeksi arvoksi, jos niiden erottaminen on sisällöllisesti perusteltua.

6.10 Ulkoiset metatieto- ja aineistolähteet

Madiskon hyödyntää ulkoisia tietolähteitä ensisijaisesti haun, tunnisteiden, linkitysten ja rikastusten kautta. Lähteiden valinta riippuu aineistosta ja käyttötapauksesta. Alla oleva kartoitus toimii tavoitearkkitehtuurin viitteenä, ei tyhjentävänä integraatioluettelona.

Lähde	Sisältö	Tunniste / rajapinta	Käyttö Studioissa	Tapa	Huomioitava
Wikidata	Yleiset entiteetit ja rakenteinen linkitetty tieto	Wikidata QID; rajapinnat	Tunnisteet, nimimuodot, suhteet, koordinaatit,	Haku / linkitys / valikoitu tuonti	Avoimen datan ehdot huomioitava

			rikastus		
Finto	Suomalaiset sanastot, ontologiat ja auktoriteettipalvelut	URI-tunnisteet; rajapinnat	Asiasanat, käsitteet, nimimuodot ja semanttinen linkitys	Haku / linkitys	Lähdekohtaiset käyttöehdot
Finna	Museoiden, arkistojen ja kirjastojen digitaaliset aineistot ja metadata	Finna-tunnisteet; API	Aineistohaku, kuvailutieto ja linkitys alkuperäiseen palveluun	Aineistohaku / linkitys / tapauskohtainen tuonti	Objektikohtaiset oikeudet ja lisenssit tarkistettava
Wikimedia Commons	Avoimesti käytettäviä media-aineistoja ja kategorioita	Commons-tiedosto- ja kategoriaturunnisteet; rajapinnat	Kuvat ja muu media, kategoriapohjainen aineiston löytäminen	Aineistohaku / linkitys / tuonti käyttöehtojen mukaan	Tiedostokohtainen lisenssi säilytettävä
Europeana	Eurooppalaisia kulttuuriperintöaineistoja ja aggregoitua metadataa	Europeana-tunnisteet; API	Kansainvälinen aineistohaku ja linkitys lähdekokoelmiin	Metahaku / linkitys	Oikeustiedot voivat vaihdella aineistoittain
Iconclass	Kuvataiteen ja ikonografian luokitus	Iconclass-tunnisteet	Teosten aiheiden ja ikonografian kuvailu	Metatietohaku / linkitys	Käyttöehdot tarkistetaan toteutuskohtaisesti
Organ Index	Urkuja koskeva erikoistunut tietolähde	Palvelun kohde- ja sivutunnisteen	Erikoisalan tunnisteet ja rikastava metadata	Metatietohaku / linkitys	Niche-lähde; sovelletaan vain relevanttiin aineistoon
RKY / Museovirasto	Valtakunnallisesti merkittävät rakennetut kulttuuriympäristöt ja kohdekuvaukset	Kohdekohtaiset tunnisteet ja verkkoviitteet	Rakennetun ympäristön kohteiden tunnistaminen ja lähdetieto	Haku / linkitys	Rajapintojen ja uudelleenkäytön ehdot arvioitava
Kansallisgalleria / Wikidata-	Taidekokoelmien metadata	Teos- ja Wikidata-	Taidekohteiden	Haku /	Julkaisuoikeudet

aineistot	ja Wikidata-linkitykset	tunnisteet	tunnistaminen, linkitys ja metadatan rikastaminen	linkitys	erotettava metadatan käyttöoikeuksista
-----------	-------------------------	------------	---	----------	--

Lähdekohtaiset rajapinnat, lisenssit ja tekniset käyttöehdot tarkistetaan aina toteutusvaiheessa. Arkkitehtuuri ei saa olettaa, että kaikki yllä mainitut lähteet ovat käytettävissä samalla tavalla tai että ne soveltuvat kaikkiin aineistoihin.

7. Instanssien välinen tiedon jakaminen

7.1 Perusperiaate

Madiskonin tavoitearkkitehtuurissa kukin instanssi säilyy itsenäisenä. Aineiston tai metatiedon jakaminen ei siirrä automaattisesti hallintavastuuta eikä muuta alkuperäisen instanssin tietoja. Tiedonvaihdossa säilytetään alkuperä, käyttöehdot, tunnisteet ja käsittelyhistoria.

7.2 Jakamisen laajuus

Jakaminen voidaan määritellä eri laajuiseksi. Aineisto tai sisältö voidaan pitää vain oman instanssin käytössä, jakaa nimetyille käyttäjille tai käyttäjärühmille, toiselle organisaatiolle, valituille verkoston instansseille tai julkaista avoimesti. Jakamispäätös ja vastaanottajan käyttöoikeudet tallennetaan osaksi aineiston hallintatietoja.

7.3 Mitä tietoa ei lähtökohtaisesti jaeta?

- keskeneräiset luonnokset
- organisaation sisäiset kommentit
- paikalliset työnkulut
- käyttäjätiedot
- muut sisäiseen työskentelyyn liittyvät tiedot, ellei niiden jakamisesta sovita erikseen

7.4 Metatiedot tiedon mukana

- aineiston tunnistetiedot
- kuvailutiedot
- aika- ja aikatasotiedot
- maantieteelliset koordinaatit
- geometriat
- paikkatunnisteet
- proveniensi
- omistajuus
- käyttöoikeudet ja lisenssit
- versio- ja hyväksyntätiedot
- rikastusten alkuperä

- ulkoiset tunnisteet ja lähdeviitteet

Näiden tietojen avulla vastaanottava instanssi voi arvioida, miten aineistoa voidaan käyttää ja millä ehdoilla sitä voidaan edelleen hyödyntää.

7.5 Rikastusten palautuminen

Vastaanottava instanssi voi täydentää aineistoa uusilla metatiedoilla tai asiantuntijoiden tuottamilla rikastuksilla. Rikastukset voivat sisältää myös paikkatiedon tai aikatiedon tarkennuksia. Tavoitearkkitehtuurissa tällaiset rikastukset voidaan välittää takaisin alkuperäisen aineiston haltijan arvioitaviksi. Alkuperäinen aineisto ei muutu automaattisesti, vaan aineiston haltija päättää, hyväksytäänkö rikastukset osaksi omaa tietovarantoaan.

7.6 Rajapintaperiaatteet

Instanssien välinen tiedon jakaminen perustuu avoimiin rajapintoihin ja yhteisesti sovittuihin tietorakenteisiin. Rajapinnan tulee välittää itse aineiston lisäksi sen käyttöä ja tulkintaa ohjaavat metatiedot. Toteutuksessa hyödynnetään ensisijaisesti olemassa olevia standardeja.

7.7 Versiot ja ristiriitaiset tiedot

Sama aineisto tai kohde voi saada eri instansseissa erilaisia rikastuksia ja tulkintoja. Järjestelmän tulee säilyttää niiden alkuperä ja versiohistoria siten, ettei ristiriitaisia tietoja yhdistetä automaattisesti yhdeksi totuudeksi.

7.8 Tiedonvaihdon jäljitettävyys

Tiedon jakamiseen liittyvät keskeiset tapahtumat tallennetaan. Vastaanottavan instanssin tulee voida osoittaa, mistä aineisto on saatu, millä ehdoilla se on saatu ja mitä muutoksia siihen on paikallisesti tehty.

7.9 Vaiheittainen toteutus

Instanssien välinen tiedonvaihto on tavoitearkkitehtuurin keskeinen kyvykkyys, mutta sitä ei tarvitse toteuttaa jokaisessa pilotissa tai paikallisessa käyttöönotossa. Sitä voidaan testata erillisessä kehitysvaiheessa sen jälkeen, kun yhteinen tietomalli ja paikallinen aineistosta julkaisuun -työnkulku ovat riittävän vakaita.

8. Ei-toiminnalliset vaatimukset

8.1 Tietoturva

Ratkaisun tulee mahdollistaa turvallinen tunnistaminen, käyttöoikeuksien hallinta, rajapintojen suojaaminen ja aineistojen suojaaminen luvattomalta käytöltä. Tietoturvan tulee kattaa sekä yksittäinen instanssi että mahdollinen instanssien välinen tiedonvaihto.

8.2 Tietosuoja

Henkilötietoja käsitellään vain järjestelmän toiminnan, käyttöoikeuksien, tietoturvan ja jäljitettävyyden kannalta tarpeellisessa laajuudessa. Tietoja ei käytetä profilointiin tai mainontaan. Lokien käyttötarkoitus, säilytys ja käyttöoikeudet määritellään.

8.3 Digitaalinen suvereniteetti

Aineistojen, metatietojen ja käyttäjätietojen tulee säilyä niitä ylläpitävien organisaatioiden hallinnassa. Ratkaisun tulee tukea tietojen siirrettävyyttä ja välttää tarpeetonta riippuvuutta palveluista, joissa käyttäjä- tai sisältötietoja hyödynnetään muihin tarkoituksiin.

8.4 Suorituskyky

Järjestelmän tulee vastata käyttötapauksen aineistomääriin ja käyttäjämääriin kohtuullisilla vasteajoilla. Arkkitehtuurin tulee mahdollistaa suorituskyvyn kasvattaminen ilman yhteisen tietomallin tai rajapintojen olennaista muuttamista.

8.5 Käytettävyys

Käyttöliittymien tulee tukea sujuvaa työskentelyä myös käyttäjille, jotka eivät tunne metatietomalleja tai semanttisen webin teknisiä käsitteitä. Rakenteisen tiedon tuottamisen tulee olla käyttäjälle mahdollisimman kevyttä.

8.6 Saavutettavuus

Käyttäjärajausten ja julkaisupalveluiden suunnittelussa huomioidaan soveltuvat saavutettavuusvaatimukset.

8.7 Ylläpidettävyys

Modulaarinen rakenne, dokumentoidut rajapinnat ja yhteinen tietomalli mahdollistavat komponenttien päivittämisen ja vaihtamisen ilman koko järjestelmän uudelleenrakentamista.

8.8 Dokumentointi

Arkkitehtuurikuvaukset, tietomallit, rajapinnat, asennusohjeet, ohjelmistovalinnat, IPR- ja lisenssiperaatteet sekä olennaiset käyttökokemukset dokumentoidaan jatkokehitystä ja uusia käyttöönottoja varten.

8.9 Asennettavuus ja toistettavuus

Uusi Studio-instanssi tulee voida ottaa käyttöön dokumentoidulla ja mahdollisimman toistettavalla prosessilla. Paikalliset mukautukset eivät saa tarpeettomasti rikkoa yhteentoimivuutta.

8.10 Skaalautuvuus

Arkkitehtuurin tulee tukea erikokoisia organisaatioita, aineistomääriä ja myöhempää verkoston kasvua. Sen tulee tukea myös paikkatietoa sisältävien aineistojen käsittelyä ja julkaisemista ilman, että paikkatieto edellyttää erillistä rinnakkaista perusarkkitehtuuria.

8.11 Lokitus ja jäljitettävyys

Järjestelmän tulee tallentaa aineistojen käsittelyyn liittyvät keskeiset tapahtumat siten, että muutokset, hyväksynnit, julkaisut ja instanssien välinen tiedon jakaminen voidaan jälkikäteen todentaa. Lokituksen tulee tukea sekä tietoturvaa että aineistojen provenienssin säilymistä.

8.12 Varmuuskopiointi ja palautuminen

Aineistojen, metatietojen ja järjestelmän palautettavuus suunnitellaan käyttöönoton riskitason mukaisesti. Yhden instanssin ongelmat eivät saa vaarantaa muiden instanssien toimintaa tai yhteentoimivuutta.

8.13 Arkkitehtuurin laatuominaisuuksien arviointi

Toteutuksissa ja piloteissa arvioidaan, tukeeko valittu ratkaisu Madiskonin keskeisiä periaatteita: verkostomaista toimintamallia, yhteentoimivuutta, toimittajariippumattomuutta, avointa lähdekoodia, digitaalista suvereniteettia sekä vaiheittaista kehittämistä. Arvioinnin tulokset tarkentavat myöhempiä vaatimuksia ja teknologiavaihtoehtoja.

9. Ohjelmisto- ja teknologiavaihtoehdot

Tässä luvussa tarkastellaan ohjelmisto- ja teknologiavaihtoehtoja, joilla Madiskonin toiminnallisia vaatimuksia voidaan toteuttaa. Tarkoituksena ei ole valita yksittäisiä ohjelmistoja, vaan tunnistaa arkkitehtuurin kannalta relevantteja viiteratkaisuja ja teknologialuokkia asiantuntija-arvioinnin tueksi.

Luettelo kuvaa vuoden 2026 tilannetta eikä ole tyhjentävä. Ohjelmistojen ajantasaisuus, lisenssit, ylläpito, rajapinnat ja soveltuvuus arvioidaan aina toteutus- tai pilottikohtaisesti. Ohjelmistoja ei arvioida tunnettuuden perusteella, vaan sen mukaan, kuinka hyvin ne tukevat Madiskonin toimintamallia, arkkitehtuuriperiaatteita ja verkostomaista yhteistyötä.

9.1 Kokonaisalusta vai komponenttiarkkitehtuuri?

Ensimmäinen toteutustapa perustuu valmiiseen kokoelma- tai repositoryjärjestelmään, jota täydennetään tarvittavilla laajennuksilla. Tällöin merkittävä osa toiminnallisuuksista saadaan käyttöön valmiina.

Toinen vaihtoehto perustuu modulaariseen arkkitehtuuriin, jossa eri tehtäviin valitaan erillisiä ohjelmistokomponentteja, jotka yhdistetään avoimien rajapintojen avulla.

On myös mahdollista, että tarkoituksenmukaisin ratkaisu muodostuu näiden yhdistelmästä: jokin valmis järjestelmä toimii perustana, mutta sitä täydennetään Madiskonin omilla komponenteilla ja laajennuksilla.

9.2 Repository- ja kokoelmanhallintaratkaisut

CollectiveAccess

Avoimen lähdekoodin kokoelmanhallinta- ja julkaisu ympäristö, jota voidaan arvioida erityisesti joustavan metatietomallin, suhteiden ja kulttuuriperintöaineistojen hallinnan näkökulmasta.

Omeka S

Avoimen lähdekoodin verkkopohjainen kulttuuriperintö- ja julkaisualue, joka tukee resurssien linkittämistä, sanastoja ja laajennuksia.

Fedora

Repository-alue, jota voidaan arvioida erityisesti digitaalisten objektien, pysyvien rakenteiden ja rajapintapohjaisen arkkitehtuurin näkökulmasta.

InvenioRDM

Avoimen lähdekoodin tutkimus- ja digitaalisten aineistojen repositoryratkaisu, jonka soveltuvuutta voidaan arvioida aineistojen hallinnan, metadatan ja rajapintojen näkökulmasta.

9.3 Säilytys ja objektivarastot

Alkuperäisten digitaalisten objektien säilytys voidaan toteuttaa paikallisella tiedostojärjestelmällä, verkko- tai objektivarastolla tai S3-yhteensopivalla tallennusratkaisulla. Arkkitehtuurissa digitaalisten objektien säilytys voidaan erottaa metatietojen hallinnasta, kun tunnisteet, eheys ja viittaukset säilyvät hallittuina.

9.4 Työympäristö ja sisällönhallinta

CollectiveAccess / Omeka S

Voivat toimia myös käyttäjän työympäristön osana, jos niiden työnkulut ja käyttöliittymät soveltuvat valittuun käyttötapaukseen.

Directus

Yleiskäyttöinen avoimen lähdekoodin data- ja hallintakerros, jota voidaan arvioida kevyen rakenteisen työympäristön ja API-pohjaisen toteutuksen vaihtoehtona.

Räätälöity työympäristö

Kevyt oma käyttöliittymä voi olla perusteltu, jos valmis järjestelmä tekee käyttäjäkokemuksesta tarpeettoman raskaan. Räätälöinnin tulee kuitenkin perustua yhteiseen tietomalliin ja avoimiin rajapintoihin.

9.5 Paikkatieto ja semanttinen tieto

- PostgreSQL/PostGIS paikkatiedon tallentamiseen ja kyselyihin
- GeoJSON tiedonsiirtoon ja rajapintoihin
- GeoServer paikkatietopalveluihin silloin, kun erillinen palvelukerros on tarpeen
- QGIS asiantuntijatyöhön, georeferointiin ja aineistojen valmisteluun
- Wikidata ja Finto tunnisteesiin, linkityksiin ja rikastamiseen
- CIDOC CRM viitekehikseksi silloin, kun tietomallin semanttisia suhteita halutaan verrata kulttuuriperintöalan kansainvälisiin malleihin

9.6 IIIF

IIIF-standardia voidaan hyödyntää kuvien ja muiden digitaalisten objektien jakamisessa, esittämisessä ja eri palveluiden välisessä yhteentoimivuudessa. Sen käyttö arvioidaan aineistotyyppin ja julkaisutarpeen perusteella.

9.7 Haku ja indeksointi

Pienissä toteutuksissa haku voidaan rakentaa suoraan relaatiotietokannan ja sen hakutoimintojen varaan. Aineistomäärien, hakutarpeiden tai metahaun kasvaessa erillinen hakupalvelu, kuten OpenSearch, voi olla tarkoituksenmukainen. Ulkoiset haut toteutetaan lähdepalvelujen rajapintoja käyttäen eikä keskitettyä kopioindeksiä rakenneta ilman erillistä tarvetta.

9.8 Identiteetinhallinta

Käyttäjähallinta voidaan toteuttaa sovelluskohtaisesti tai erillisellä identiteetinhallintaratkaisulla. Keycloak on yksi arvioitava avoimen lähdekoodin vaihtoehto silloin, kun tarvitaan keskitettyä tunnistautumista, rooleja tai kertakirjautumista.

9.9 Julkaiseminen

Julkaisukanavia voivat olla WordPress, Omeka- tai Pawtucket-tyyppiset kulttuuriperintöjulkaisimet, staattiset verkkosivustot, kartta- ja reittipalvelut, PDF-julkaisut sekä rajapintojen kautta toimivat muut palvelut. Julkaisukanavan tulee olla vaihdettavissa ilman, että alkuperäinen aineisto tai yhteinen tietomalli sidotaan siihen.

9.10 Ohjelmistojen arviointikriteerit

Arviointikriteeri	Arvioinnin kohde
-------------------	------------------

Toiminnallinen soveltuvuus	Kuinka hyvin ratkaisu tukee määriteltyjä käyttötapauksia ja toiminnallisia vaatimuksia.
Arkkitehtuurin yhteensopivuus	Soveltuuko ratkaisu modulaariseen ja verkostomaiseen tavoitearkkitehtuuriin.
Avoimet rajapinnat	Tukeeko ratkaisu integraatioita, avointa tiedonsiirtoa ja ulkoisia hakuja.
Avoin lähdekoodi	Onko ratkaisu läpinäkyvä, jatkokehitettävä ja hallittavissa ilman kohtuutonta toimittajalukkoa.
Laajennettavuus	Voidaanko uusia toimintoja toteuttaa laajennuksina tai erillisinä palveluina.
Yhteentoimivuus	Tukeeko ratkaisu yhteisiä tietorakenteita, tunnisteita ja tiedon siirrettävyyttä.
Metatietojen hallinta	Mahdollistaako ratkaisu laajennettavan metatietomallin, provenienssin, versiot ja käyttöoikeudet.
Paikkatiedon tuki	Tukeeko ratkaisu koordinaatteja, geometrioita, paikkatunnisteita ja paikkatietohakuja.
Semanttisen tiedon tuki	Mahdollistaako ratkaisu linkitetyn tiedon, sanastojen, tunnisteiden ja ontologioiden hyödyntämisen.
Haku ja ulkoiset tietolähteet	Tukeeko ratkaisu rakenteista hakua, rajapintahakuja ja tarvittaessa metahakua.
Työnkulut ja käyttöoikeudet	Tukeeko ratkaisu rooleja, hyväksyntää ja toimituksellista työskentelyä.
Tietoturva ja tietosuoja	Voidaanko ratkaisu suojata asianmukaisesti ja käyttää tietosuojavaatimusten mukaisesti.
Digitaalinen suvereniteetti	Voidaanko aineistot, metadata ja käyttäjätiedot pitää organisaation määräysvallassa.
Ylläpidettävyys	Onko ratkaisu dokumentoitu, aktiivisesti ylläpidetty ja teknisiltä riippuvuuksiltaan hallittava.
Kehittäjä- ja käyttäjäyhteisö	Onko ratkaisulla riittävä yhteisö, osaamispohja tai kaupallisen tuen mahdollisuus.
Kulttuuriperintöalan kokemus	Onko ratkaisusta näyttöä kulttuuriperintöaineistojen tai vastaavien tietorakenteiden käsittelyssä.
Kokonaiskustannus	Mitä käyttöönotto, mukauttaminen, ylläpito,

	infrastruktuuuri, päivitykset ja osaaminen maksavat elinkaaren aikana.
Lisenssi, IPR ja tiedon siirrettävyys	Voidaanko tiedot, määrittelyt ja omat laajennukset säilyttää ja siirtää myöhemmin.
Käytettävyys ja saavutettavuus	Tukeeko ratkaisu sujuvaa ja saavutettavaa työskentelyä.
Toteutettavuus	Voidaanko ratkaisu ottaa käyttöön käytettävissä olevilla resursseilla ilman kohtuutonta räätälöintiä tai ylläpitotakkaa.

9.11 Teknologiavalintojen dokumentointi

Toteutuskohtaiset teknologiavalinnat dokumentoidaan erikseen. Dokumentoinnissa kuvataan valinnan perusteet, vaihtoehdot, riippuvuudet, lisenssit, IPR-vaikutukset, tietojen siirrettävyys ja mahdollinen poistumisstrategia. Näin tavoitearkkitehtuuri säilyy ohjelmistotuotteista riippumattomana, vaikka se tarjoaa asiantuntijoille konkreettisia viiteratkaisuja.

10. Vaiheittainen toteuttaminen, pilotointi ja referenssitoteutukset

10.1 Periaate

Madiskonia kehitetään vaiheittain. Tavoitteena ei ole rakentaa koko tavoitearkkitehtuuria yhdellä kertaa, vaan toteuttaa rajattuja kokonaisuuksia, jotka tuottavat käyttäjille itsenäistä hyötyä ja samalla tarkentavat arkkitehtuuria.

10.2 Pilotin tehtävä

Pilotti todentaa yhden tai useamman tavoitearkkitehtuurin osan käytännössä. Se voi kohdistua esimerkiksi aineiston vastaanottoon, rakenteiseen kuvailuun, ulkoiseen aineisto- tai metatietohakuun, paikkatietoon, aikatasojen hallintaan, rikastamiseen, toimitukselliseen työnkulkuun, WordPress-julkaisuun tai instanssien väliseen tiedonvaihtoon.

Pilotin ei tarvitse toteuttaa koko tässä dokumentissa kuvattua arkkitehtuuria eikä rakentaa kahta täysimittaista Studio-instanssia, ellei juuri instanssien välinen yhteentoimivuus ole pilotin tutkimuskohde.

10.3 Pilotin rajausta ja onnistumiskriteerit

Jokaiselle pilotille määritellään erikseen tavoite, aineisto, käyttäjät, toteutettavat tavoitearkkitehtuurin kyvykkyydet, oikeudelliset reunaehdot, tekniset ratkaisut, onnistumiskriteerit ja jatkokäyttö. Onnistumista arvioidaan suhteessa pilotin rajattuun tavoitteeseen, ei koko tavoitearkkitehtuurin toteutumiseen.

10.4 Kevyt referenssityönkulku

Kevyessä ensimmäisen vaiheen toteutuksessa alkuperäisten tiedostojen säilytys- ja työtila voi olla erillinen palvelu, Madiskon voi toimia rakenteisen kuvailun, suhteiden, haun ja rikastamisen kerroksena ja erillinen sisällönhallintajärjestelmä julkaisukanavana. Olennaista on, että metadata ja keskeiset suhteet säilyvät siirrettävinä esimerkiksi CSV- tai JSON-muodossa.

10.5 Tekoälyavusteinen kehittäminen

Tekoälyä voidaan käyttää ohjelmistokehityksen, dokumentoinnin, testauksen ja aineiston käsittelyn apuna. Tekoälyn tuottama koodi, metadata ja muut ehdotukset tarkistetaan ja testataan ennen hyväksymistä. Luottamuksellista, henkilötietoja sisältävää tai muuten suojattavaa aineistoa ei välitetä ulkoiseen tekoälypalveluun ilman erillistä arviointia ja hyväksyntää.

10.6 Omistajuus ja jatkuvuus kehittämishankkeissa

Kehittämishankkeissa sovitaan ennen olennaisen toteutustyön käynnistämistä vähintään lähdekoodin, dokumentaation, tietomallin, rajapintamääritysten ja hankkeessa syntyvien laajennusten omistajuudesta, käyttöoikeuksista ja lisensseistä. Toteutuksen tulee olla jatkettavissa myös hankkeen tai yksittäisen kehittämiskumppanin osallistumisen päätyttyä.

10.7 Arviointi ja jatkokehitys

Piloteista ja käyttöönotoista kerätään kokemuksia tietomallin, työnkulkujen, hakujen, ohjelmistojen, rajapintojen, käytettävyyden, ylläpidettävyyden ja kustannusten toimivuudesta. Kokemusten perusteella tätä tavoitearkkitehtuuria ja erillisiä toteutusmäärittelyjä tarkennetaan.

Madiskonin kehityksen onnistumista ei arvioida uuden ohjelmistokoodin määrällä vaan sillä, kuinka hyvin toimintamalli helpottaa kulttuuriperintöaineistojen hallittua rakenteistamista, rikastamista, löytämistä, yhdistämistä, julkaisemista ja uudelleenkäyttöä sekä mahdollistaa ratkaisujen pitkäjänteisen kehittämisen.