

SAPA-KOKONAISARKKITEHTUURI

Päivitetty: 7.6.2019
Versio 1

TABLE OF CONTENTS

1	Johdanto	2
1.1	Dokumentin tarkoitus	2
1.2	Dokumentin kohderyhmä	2
1.3	Tausta ja lähtökohdat	2
2	Periaatteelliset linjaukset	3
2.1	Arkkitehtuuria ohjaavat periaatteet	3
2.2	Sidosarkkitehtuurit	5
2.3	Erityisvaatimukset	7
2.4	Arkkitehtuurin yleiskuvaus	7
3	Toiminta-arkkitehtuuri	9
3.1	Palvelut	9
3.1.1	Vastaanoton palvelut.....	10
3.1.2	Hallinnan palvelut.....	10
3.1.3	Käsittelyn palvelut	12
3.1.4	Säilyttämisen palvelut	13
3.1.5	Käyttöliittymä- ja rajapintapalvelut.....	14
3.1.6	Käytön palvelut.....	15
3.2	Käyttäjät ja käyttäjäryhmät.....	15
4	Tietoarkkitehtuuri	16
4.1	Loogiset tietovarannot	16
4.2	Tietoturva.....	17

VERSIOHISTORIA

Versio	Päivämäärä	Muuttaja	Muutosloki
0.1	29.3.2018	Mikko Laukkanen	Ensimmäinen versio
0.2	17.4.2018	Mikko Laukkanen	Lisätty sisältöä arkkitehtuurikappaleisiin ja rajapintakuva, muutettu rakennetta (toiminta-arkkitehtuurin alle prosessit ja työkulut).
0.3	2.5.2018	Kimmo Koivunen	Rakennetta uudistettu vastamaan arkkitehtuurityön tarpeita
0.4	15.5.2018	Mikko Laukkanen, Kimmo Koivunen, Vili Haukkoara, Mikko Eräkaski	Rakennetta uudistettu, sisältöä päivitetty. Ensimmäinen versio ohjausryhmälle.
0.5	22.5.2018	Mikko Laukkanen, Kimmo Koivunen	Ohjausryhmän kommentit otettu huomioon.
1.0	11.6.2018	SAPA-ohjausryhmä	SAPA-ohjausryhmän hyväksymä versio 1.0.

1 JOHDANTO

1.1 DOKUMENTIN TARKOITUS

Tämä dokumentti kuvaa SAPA-palvelukokonaisuuden ratkaisuarkkitehtuurin, joka määrittää palvelukokonaisuuden toiminnalliset tarpeet, hahmottaa toiminnan suhteen ohjaaviin arkkitehtuureihin ja tunnistaa tarvittavat (osa)palvelut. Tätä dokumenttia tarkentaa SAPA-tekninen kuvaus, jossa esitettyyn arkkitehtuuriin perustuu tietojärjestelmätoteutuksen projektisuunnitelma, joka edelleen jalostaa arkkitehtuurin toteutettaviksi toiminnoiksi.

1.2 DOKUMENTIN KOHDERYHMÄ

Dokumentti on kohdennettu ensisijaisesti SAPA-projektin ohjausryhmälle, jotta voidaan muodostaa yhteinen näkemys SAPA-palvelukokonaisuudesta ja ohjata SAPA-tietojärjestelmätoteutusta varsinaisessa toteutustyössä.

1.3 TAUSTA JA LÄHTÖKOHDAT

Tätä ratkaisuarkkitehtuuria ovat edeltäneet toisaalta Kansallisarkiston kehittämishankkeet ja -suunnitelmat (mm. massadigitoinnin suunnittelu, käytössä olevat digitaalisen aineiston hallinnointi- ja käyttöön tarjoamisen ratkaisut sekä Kansallisarkiston arkkitehtuurityö) ja toisaalta mm. KDK:n kokonaisarkkitehtuurin laadinta ja palveluiden sekä Kulttuuriperintö-PAS-palvelun (aiempi KDK-PAS -palvelu) laajennuksen (PAS-EKA -projektin tulos) ja AHAA-palvelun toteuttaminen sekä näyttö- ja käyttörajoitteisten (NKR-) aineistojen hallinnan arkkitehtuurityö.

Kansallisarkiston tarpeita on avattu Kansallisarkiston itsensä työstämässä kuvauksessa SAPA-palvelukokonaisuuden tavoitetilan arkkitehtuurista, joka kuvaa tavoitetilan Kansallisarkiston tavoitteiden ja tarpeiden kautta. Kuvauksessa tunnistettuja, Kansallisarkiston keskeisimpiä tavoitteita ovat:

- Mahdollistaa eri tavoin muodostuneiden digitaalisten aineistojen vastaanotto yhtenäisen digitaalisen tiedonhallinnan piiriin Kansallisarkistossa.
 - Digitaalisen aineiston vastaanotolla tarkoitetaan Kansallisarkiston takautuvassa digitoinnissa tuotetun aineiston, massadigitoinnissa tuotetun aineiston ja alkujaan digitaalisen aineiston vastaanottoa.
- Yhtenäisellä tiedonhallinnalla tarkoitetaan yleisellä tasolla:
 - yhdenmukaista kuvailevien, hallinnollisten ja teknisten metatietojen hallintaa,
 - yhdenmukaista digitaalisten arkisto-, käyttö- ja natiivikappaleiden hallinnointia sekä
 - yhdenmukaisia toimenpiteitä, joita aineistolle tehdään sen vastaanottotavasta riippuen (mm. aineiston sisällön tunnistus, sisällön analysointi metatietojen rikastamisen näkökulmasta, tunnisteiden hallinta).

- Mahdollistaa SAPA:an siirrettyjen aineistojen käyttö mahdollisimman jouhevasti olemassa olevan lainsäädäntö ja aineistoon kohdistuvat rajoitukset huomioiden:
 - Tarjota aineisto luovuttaneen tahon käyttöön vakiomuotoisten rajapintojen ja yksinkertaisen graafisen käyttöliittymän kautta.
 - Tarjota yhdenmukainen viranomaisten toimintaa varten tarvittavien käyttöoikeuksien hallintamalli.
 - Tarjota yhdenmukainen kansalaiskäyttäjien (esim. tutkijat, media, satunnaiskäyttäjät) tarvitsemien käyttöoikeuksien hallintamalli.
 - Tarjota aineisto käyttöön viranomaisille ja kansalaisille yksinkertaisen graafisen käyttöliittymän kautta.

Näiden Kansallisarkiston tavoitteiden myötä SAPA-palvelun piirissä eivät ole vain tulevaisuuden arkistoaineistot (tulevaisuudessa Kansallisarkistoon vastaanotettavat alkujaan digitaaliset arkistoaineistot tai tulevaisuudessa digitoitavat aineistot), vaan palveluun siirrettään arkistoaineistoa myös Kansallisarkiston arkkitehtuurissa kuvatuista, tämän arkkitehtuuridokumentin yleiskuvassa mainituista nykyisistä palveluista niiden elinkaari ja toteutussuunnitelmassa ehdotettu aikataulu huomioon ottaen. Kansallisarkiston tavoitteena on, että kaikkien, myös nykyisen Digitaaliarkiston digitaalisten käyttökappaleiden säilytys keskitetään SAPA-ratkaisuun.

Säilyttämisen vaatimukset kasvavat, kun digitoitujen aineistojen analogiset ilmentymät hävitetään. Kansallisarkiston tavoitteena on arkistokappaleiden hallinnon keskittäminen PAS-palveluun, jolloin kaikki arkistokappaleet hallinnoidaan samalla tavalla aineistotyyppistä riippumatta. Kaikki digitaaliset aineistot, myös Digitaaliarkistosta, siirretään tavoitetilassa SAPA-palvelun kautta Kulttuuriperintö-PAS-palveluun, eikä tarvetta erillisten siirtokanavien tai Digitaaliarkiston tallekappaleiden erilliseen ylläpitoon enää ole. Tästä voidaan kuitenkin päättää vasta myöhemmin SAPA-projektin edistyessä.

2 PERIAATTEELLISET LINJAUKSET

2.1 ARKKITEHTUURIA OHJAAVAT PERIAATTEET

Periaate	Vaikutus
Noudatetaan JHKA:n periaatteita	Julkisen hallinnon kokonaisarkkitehtuuri (JHKA) on rakenne, jonka avulla koordinoidaan ja kehitetään julkisen hallinnon organisaatioiden ja palveluiden välistä yhteentoimivuutta ja joka määrittää julkisen hallinnon kokonaisarkkitehtuurityössä käytettävän arkkitehtuurikokonaisuuden rakenteen (http://vm.fi/julkisen-hallinnon-yhteinen-kokonaisarkkitehtuuri).

Noudatetaan OKM:n ja KDK:n periaatteita	OKM:n arkkitehtuuriperiaatteet määrittelevät yleiset puitteet hallinnonalan arkkitehtuurityölle, ja KDK:n arkkitehtuuriperiaatteet pohjautuvat niihin. KDK:n periaatteet puolestaan ohjaavat SAPA:n arkkitehtuurityötä (Kansallisen digitaalisen kirjaston kokonaisarkkitehtuuri, http://www.digime.fi/wp-content/uploads/2018/05/KDK_kokonaisarkkitehtuuri_3_1.pdf).
Noudatetaan yhteisesti sovittuja standardeja	SAPA:n arkkitehtuuri rakennetaan aineistojen sisällön, muodon, metatietojen ja siirtomenettelyiden osalta OKM:n, KDK:n ja YTI:n tieto- ja käsitemallien perustalle sekä Kansallisarkiston siirtorakennemääritykset huomioon ottaen.
Noudatetaan KDK-standardisalkun määrityksiä	SAPA:n arkkitehtuurissa tukeudutaan KDK:n standardisalkun määrityksiin ja tarpeen mukaan täydennetään niitä. (http://www.digime.fi/wp-content/uploads/2018/05/KDK-standardisalkku-1.2.0.pdf).
Asiakas- ja toimintälähtöisyys	SAPA:n arkkitehtuurissa otetaan huomioon arkistonmuodostajien ja arkistoitujen aineistojen käyttäjien tarpeet ja mahdollisuudet siten, että palveluun liittyminen ja aineiston siirtäminen voidaan toteuttaa.
Hyödynnetään ja täydennetään yhteisiä ja olemassa olevia palveluita	SAPA:n arkkitehtuuri rakennetaan mahdollisuuksien mukaan hyödyntäen KDK:n ja julkiselle hallinnolle tarjottuja ja käyttövelvoitettuja yhteisiä palveluita sekä kansallista palveluarkkitehtuuria välttäen päällekkäisiä ratkaisuja.
Toteutetaan SAPA-palvelun tietosuoja ja tuetaan tietoturvallisia toimintatapoja	SAPA:n arkkitehtuuri toteuttaa rekisterinpitäjän antamat ohjeet tietojen käsittelystä. Kansallisarkisto tuottaa SAPA-palvelulle tietosuoja-asetuksen 35 artiklan mukaisen vaikutustenarvioinnin SAPA-palvelukokonaisuudesta.
Noudatetaan toimittajariippumattomia periaatteita	Vältetään toimittajalukkoon joutuminen. Palveluun ei tehdä räätälöityjä järjestelmäkohtaisia liittymiä.

Huomioidaan aineistojen ja toiminnan pitkä elinkaari	<p>Suunnittelussa noudatetaan hyvää tiedonhallintatapaa.</p> <p>SAPA-arkkitehtuurin tulee olla modulaarinen ja palveluorientoitunut.</p> <p>Toiminnan elinkaari ylittää kaikkien teknologisten ratkaisujen elinkaaren, jolloin teknologia tulee uudistumaan toiminnan aikana. Täten jatkuvuudenhallinta tulee huomioida suunnittelussa ja toteutustyössä.</p>
--	---

2.2 SIDOSARKKITEHTUURIT

Sidosarkkitehtuurit ovat SAPA:an liittyviä arkkitehtuureja, joilla on oma vaikutuksensa SAPA:an. Yleisten julkisen hallinnon ja OKM:n kokonaisarkkitehtuurien, kansallisen palveluarkkitehtuurin ja JHS-suositusten lisäksi erityisen merkittäviä ovat Kulttuuriperintö-PAS-arkkitehtuuri, Kansallisarkiston massadigitoinnin kehittämishanke sekä NKR-ratkaisu (metatietojen käyttörajoitusten ja aineistojen näyttörajoitusten toteutus).

Arkkitehtuuri	Vaikutus
Julkisen hallinnon ja OKM:n kokonaisarkkitehtuurit	Julkisen hallinnon (JHKA) ja OKM:n kokonaisarkkitehtuurit määrittelevät yleisen viitekehyksen, yleiset periaatteet ja yhteiset komponentit digitoinnin ja pysyvissäilytyksen arkkitehtuurille.
Kulttuuriperintö-PAS -arkkitehtuuri	Kulttuuriperintö-PAS -arkkitehtuuri määrittelee tiedon pysyvissäilytyksen ratkaisun sekä sitä täydentävät tukipalvelut (http://www.digime.fi/wp-content/uploads/2018/05/KDK_kokonaisarkkitehtuuri_3_1.pdf).
NKR-ratkaisuarkkitehtuuri: metatietojen käyttörajoitukset ja aineistojen näyttörajoitukset	Metatietojen käyttörajoitukset ja aineistojen näyttörajoitukset toteuttava käyttäjä- ja käyttövaltuushallinnan arkkitehtuuri sekä rajoitustietojen metatiedoissa kuvaamisen määritykset.
Kansallisarkiston massadigitoinnin kehittämishanke	Massadigitoinnin kehittämishankkeessa suunnitellaan asiakirjallisten aineistojen suurimittaista digitointia. Hanke tuottaisi keskeisen aineistomassan, joka siirretään SAPA-palveluun. Oheisissa dokumenteissa on kuvattu

	<p>massadigitoinnin prosessi sekä hahmoteltu alustavasti käytettävää tietojärjestelmäkokonaisuutta:</p> <p>https://www.arkisto.fi/uploads/Kansallisarkisto/Hankkeet/Massadigitointi/Liite_3_1_Digitointiprosessin_toteuttamissuunnitelma.pdf</p> <p>https://www.arkisto.fi/uploads/Kansallisarkisto/Hankkeet/Massadigitointi/Liite_3_2_Massadigitoinnin_prosessikuvaukset.pdf</p>
YTI- tieto- ja käsittemallit	<p>Yhteisen tiedon hallinta (YTI) -hankkeen tavoitteena on parantaa tietojärjestelmien ja niiden tietojen yhteentoimivuutta, luoda edellytykset hallinto- ja sektorirajoista riippumattomalle toimintojen kehittämiselle sekä tehostaa olemassa olevan tiedon hyödyntämistä (https://yhteentoimiva.suomi.fi).</p>
Kansallisarkiston arkkitehtuuri	<p>SAPA-arkkitehtuuri toteuttaa Kansallisarkiston kokonaisarkkitehtuurissa määriteltyä tavoitetilaa. Kansallisarkiston arkkitehtuurista ei ole saatavilla julkista kuvausta (valmistuu vuoden 2018 aikana).</p>
Arkistokuvailun kansallisen käsittemallin luonnos 0.2	<p>Kansallisen käsittemallin rooli kuvailujärjestelmässä on toimia perustana elinkaarensa eri vaiheissa olevan niin analogisen kuin siitä digitoidun, samoin kuin erityyppisen alkujaan digitaalisen tietovarannon arkistokuvailun metatietomalleille.</p> <p>Kansallisen käsittemallin tarkoituksena on yhdenmukaistaa arkistokuvailua kansallisella tasolla. Esimerkiksi AHAA:n tietomalli perustuu kansalliseen käsittemalliin ja noudattaa sitä.</p> <p>Kansallinen käsittemalli julkaistaan kansainvälisen arkistokuvailutyön aikataulu huomioiden.</p> <p>Käsittemallin omistaa Kansallisarkisto.</p> <p>Kansallinen käsittemalli on saatavissa: https://www.arkisto.fi/kasitemalli</p>

2.3 ERITYISVAATIMUKSET

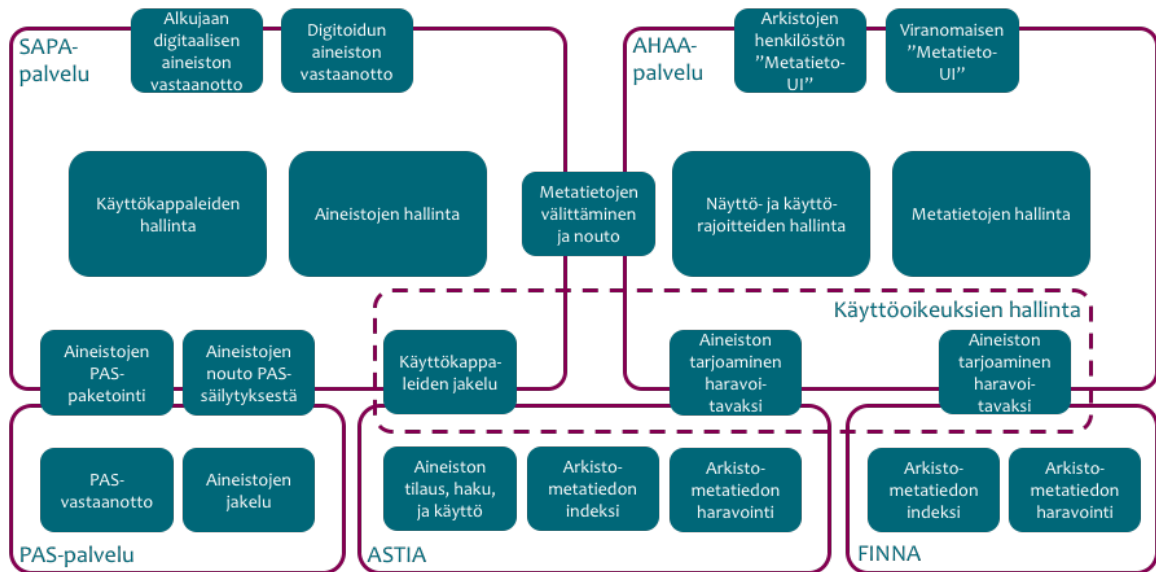
Vaatus	Tarkennus
Skaalautuvuus	<p>SAPA-palvelun arkkitehtuurin tulee mahdollistaa palvelun tallennuskapasiteetin skaalautuvuus useisiin petatavuihin; massadigitoinnin arvioitu maksimi digitointinopeus on:</p> <ul style="list-style-type: none">- arviolta n. 2 376 000 tiedostoa päivässä,- arviolta n. 13,6 Tt päivässä. <p>Kansallisarkiston takautuvassa digitoinnissa tai alkujaan digitaalisten aineiston vastaanotossa ei käsitellä lähellekään tämän suuruisia aineistomääriä. Näistä aiheutuva lisäys kapasiteettiin on yhteensä n. 5–10% massadigitoinnin määrään nähden. Tämän lisäksi huomioida, että alkujaan digitaalisten aineistojen vastaanotto ei ole samalla tavalla ns. tasainen virta, vaan tarkoittaa ajoittaisia siirtoja.</p> <p>SAPA-palvelun käsittelyn kapasiteetti suunnitellaan näiden vaatimusten pohjalta.</p>
Saatavuus	<p>SAPA-palvelun saatavuustavoitteeksi asetetaan 99% palveluaikana klo 8-16.</p>
Tiedon omistajuus	<p>SAPA-palveluun siirretyn aineiston omistajuus siirtyy Kansallisarkistolle.</p>

2.4 ARKKITEHTUURIN YLEISKUVAUS

Ratkaisuarkkitehtuuri kuvaa SAPA-palvelun toiminnan ja liittymät (myös ulkoiset palvelut), joilla mahdollistetaan Kansallisarkistolle siirrettävien valtionhallinnon asiakirjallisten aineistojen hallinta ja käsittely. Arkkitehtuurin yleiskuva on esitetty Kuva 1. Kuvassa esitellään arkkitehtuuriin kuuluvat keskeiset palvelut ja hahmotellaan miten nämä palvelut liittyvät jo olemassa oleviin tietojärjestelmiin. SAPA-palvelun toteutustyön edistyessä siihen jotkin tai osa siihen sisältyvistä palveluista tai toteuttavista järjestelmästä voi vaihtua, mutta arkkitehtuurissa esitellyt tarpeet säilyvät niin kauan kuin määritellyt toiminnallisuudet pysyvät ennallaan.

SAPA-palvelukokonaisuudella tarkoitetaan kaikkia niitä valtionhallinnon asiakirjallisten aineistojen arkistoinnin palveluita (tai osapalveluita), joilla tuotetaan asiakirjallisten aineistojen luotettava vastaanotto, säilyttäminen ja aineistojen käyttöön saattaminen. SAPA-palvelukokonaisuus sisältää myös ulkoisia palveluita, joista keskeisimpiä ovat: Kansallisarkiston aineistokuvailun AHAA-palvelu sekä pitkäaikaissäilytyksen PAS-palvelu.

SAPA-palvelulla tarkoitetaan niitä palveluita, jotka kehitetään erikseen mahdollistamaan valtionhallinnon asiakirjallisten aineistojen vastaanotto, säilyttäminen ja käyttöön saattaminen osana SAPA-palvelukokonaisuutta.



Kuva 1: Arkkitehtuurin yleiskuva.

SAPA-palvelukokonaisuus sisältää seuraavat toiminnot:

- Alkujaan digitaalisten asiakirjallisten aineistojen vastaanottaminen sähköisesti valtion viranomaisilta.
- Digitoitujen asiakirjallisten aineistojen vastaanottaminen valtion viranomaisilta.
- Vastaanotettujen aineistojen kuvailevien metatietojen yhdistäminen ja lisääminen Kansallisarkiston aineistokuvailun hallintaan.
- Vastaanotettujen aineistojen käyttökappaleiden tuottaminen, hallinta, säilyttäminen ja jakelu Kansallisarkiston määrittelemiін tarpeisiin näyttö- ja käyttörajoitteet huomioiden.
- Vastaanotettujen aineistojen paketointi pitkäaikaissäilytykseen ja siirtojen hallinta.
- Palvelun käyttöön ja hallintaan soveltuvat käyttöliittymät sekä Kansallisarkistolle että aineistoja siirtäville viranomaisille.
- Aineistojen säilyttäminen sekä määrääjän että pysyvästi.
- Aineiston eheyden varmentaminen.
- Palvelun käytön seuranta ja raportointi erillisen kuvauksen mukaisesti.
- Aineiston arkistollisten kuvailutietojen (kontekstimetatiedot) hallinta.
- Aineistojen näyttö- ja käyttörajoitteiden hallinta.
- Metatietojen välittämien AHAA-palveluun ja metatietojen nouto AHAA-palvelusta.
- Aineistojen pitkäaikaissäilytys PAS-palvelussa PAS-määritysten mukaan.

- Aineistojen rikastaminen; metatietojen rikastaminen ja käyttökappaleiden rikastaminen.

Lisäksi kuvaan on piirretty käyttöoikeuksien hallinta, joka liittyy SAPA-, AHAA-, ASTIA- ja FINNA-palveluihin, ja jota toteutetaan erillisessä projektissa. Projektin tavoitteena on tuottaa toiminnot, joiden avulla voidaan tehdä päätös, saadaanko metatieto tai käyttökappale välittää ulkopuolisille tahoille (esim. tutkijat) käyttö- ja näyttörajoitteet huomioiden.

3 TOIMINTA-ARKKITEHTUURI

3.1 PALVELUT

Ratkaisuarkkitehtuuri koostuu Kuva 2:n palvelukartassa esitellyistä palveluista, jotka on jaettu palveluryhmiin:

- Vastaanoton palvelut
- Hallinnan palvelut
- Käsittelyn palvelut
- Säilyttämisen palvelut
- Käyttöliittymä- ja rajapintapalvelut
- Käytön palvelut



Kuva 2: SAPA-palvelukartta.

Palvelukartta esittää palvelut, jotka on tunnistettu SAPA-palvelukokonaisuuden tähänastisessa määrittelytyössä. Kartta voi täydentyä ajan mittaan uusilla palveluilla tai siinä nyt kuvatut palvelut voivat korvautua uusilla palveluilla.

Seuraavissa aliluvuissa esitellään kunkin palveluryhmän sisältämät palvelut.

Ratkaisuarkkitehtuuriin sisältyvät palvelut kuvataan loogisella tasolla, eikä kuvaus ota kantaa siihen, miten tai mikä taho ne toteuttavat. Kuvatut palvelut voidaan toteuttaa osin

tai kokonaan myös muiden kuin SAPA-palvelun, esim. AHAA- tai Kulttuuriperintö-PAS-palvelun avulla.

3.1.1 Vastaanoton palvelut

SAPA tarjoaa valtionhallinnon organisaatioille mahdollisuuden siirtää palveluun digitaalisessa muodossa olevia aineistoja, jotka luovutetaan Kansallisarkistolle. SAPA-palvelukokonaisuuteen voidaan vastaanottaa alkujaan digitaalisessa muodossa olevia aineistoja tai analogisesta aineistosta digitaaliseen muotoon muutettuja aineistoja (digitoitujen aineistojen vastaanotto).

SAPA-palvelu tukee vain määritellyissä muodoissa ja rakenteissa olevan digitaalisen aineiston vastaanottoa. Vastaanotettavan aineiston tulee täyttää aineiston tiedostomuotoja ja metatietoja koskevat minimivaatimukset. Näillä vaatimuksilla varmistetaan, että palvelussa voidaan luotettavasti säilyttää aineistoa ja hallita sen käyttöön saattamista. Palveluun voidaan siirtää niin määrääjän kuin pysyvästi säilytettävää aineistoa.

3.1.1.1 Alkujaan digitaalisten aineistojen vastaanotto

Palvelu toteuttaa suoraan valtion viranomaisista vastaanotettavien alkujaan digitaalisten asiakirjallisten aineistojen vastaanottamiseen. Se mahdollistaa sekä SÄHKE2- että tietokantamuotoisen aineiston vastaanoton omissa siirtorakenteissaan. Palvelu tarkastaa aineistolle ja siirrolle asetettujen minimivaatimusten täyttymisen. Palvelu raportoi vastaanotetuista aineistoista aineistojen hallinnalle.

3.1.1.2 Digitoidun aineistojen vastaanotto

Palvelu toteuttaa digitoidun aineiston vastaanottamisen: 1) suoraan valtion viranomaisilta, 2) Kansallisarkiston massadigitoinnin toiminnosta sekä 3) Kansallisarkiston takautuvasta digitoitotoiminnasta. Palvelu tarkastaa aineistolle ja siirrolle asetettujen minimivaatimusten täyttymisen. Palvelu raportoi vastaanotetuista aineistoista aineistojen hallinnalle.

3.1.1.3 Vastaanoton tukipalvelut

Vastaanoton tukipalveluilla ratkaistaan aineiston luovuttajien aineistonsiirtoprosessiin liittyviä erilaisia tarpeita ja tuetaan siirron eri vaiheita. Tukipalveluiden tavoitteena on tukea määräyksenmukaisen siirtorakenteen muodostamista. Palvelut sisältävät mm. siirtopaketin muodostamiseen liittyvän käyttöliittymän ja siirtopaketin määritystenmukaisuuden validoinnin.

3.1.2 Hallinnan palvelut

Hallinnan palvelut mahdollistavat SAPA-kokonaisuuden toimivuuden ja erityisesti huolehtivat aineistoihin liittyvien metatietojen linkittymisestä toisiinsa. Tyypillisesti hallinnan palvelut huolehtivat riittävien tunnisteiden liittämistä aineistojen metatietoihin.

Osa hallinnan palveluista voidaan toteuttaa esimerkiksi AHAA-palvelun avulla, jos se osoittautuu kokonaisuuden kannalta tarkoituksenmukaiseksi.

3.1.2.1 Käyttökappaleiden hallinta

Palvelu tuottaa vastaanotetuista aineistoista käyttökappaleet, jos ne puuttuvat vastaanotetusta aineistosta. Käyttökappaleet tuotetaan Kansallisarkiston kanssa määritellyllä tavalla. Palvelu tuottaa käyttökappaleiden hyödyntämisen kannalta tarvittavat tunnisteet ja tallentaa ne osaksi aineiston metatietoja sekä siirtää käyttökappaleet käyttökappaleiden säilytykseen. Lisäksi palvelun avulla käyttökappaleita tulee voida poistaa esimerkiksi määräajan säilytettävän aineiston hävittämisen myötä tai tilanteissa kun käyttökappaleen formaattia muutetaan. Palvelu siirtää käyttökappaleita käyttökappaleiden säilytyksestä käyttökappaleiden jakeluun aineistojen hallinnan ja näyttö- ja käyttörajoitteiden hallinnan ohjauksessa.

3.1.2.2 Aineistojen hallinta

Palvelu varmistaa, että SAPA-kokonaisuus onnistuu hallitsemaan kaikkia aineistojaan ja että aineistot ja metatiedot siirtyvät tarvittaviin palveluihin. Palvelu huolehtii mm.:

- aineistojen vastaanoton seurannasta ja aineistojen siirroista vastaanoton palveluista käsittelyn ja hallinnan muihin palveluihin,
- pysyvässä säilytyksessä tai määräaikaissäilytyksessä siirrettyjen aineistojen määrän, laadun ja tyyppien seurannasta ja raportoinnista,
- määräaikaaisesti säilytettävien aineistojen poistopyynnöistä määräaikaissäilytykselle ja käyttökappaleiden hallinnalle,
- hakupalvelusta tulevien käyttöpyyntöjen välittämisestä näyttö- ja käyttörajoitteiden hallinnalle ja tarvittaessa käyttökappaleiden hallinnalle,
- aineistojen PAS-paketointiin liittyvien tunnisteiden tallentamisesta osaksi aineistojen metatietoja,
- aineistojen PAS-siirtojen seurannasta ja PAS-vastaanottoraporttien liittämistä osaksi aineistojen metatietoja sekä
- raporttien koostamisesta ja tarjoamisesta käyttöliittymä- ja rajapintapalveluille.

3.1.2.3 Metatietojen hallinta

Palvelu vastaa kontekstimetatietojen hallinnasta ja käsittelystä. Se sisältää myös siirron yhteydessä vastaanotettujen metatietojen oikeellisuuden tarkistamista. Palvelu siirtää aineistojen metatiedot *metatietojen välittäminen*-palvelulle.

3.1.2.4 Näyttö- ja käyttörajoitteiden hallinta

SAPA-kokonaisuus sisältää aineistoja, joiden hyödyntämiseen liittyy näyttö- ja käyttörajoitteita. Tällaiset rajoitustiedot sisältyvät AHAA-palvelussa hallinnoitaviin metatietoihin. Tältä osin AHAA täyttää SAPA-kokonaisuuden tarpeet.

3.1.2.5 Käyttöoikeuksien hallinta

Palvelun tarjoaja tavan hallinnoida käyttäjien oikeuksia käyttää aineistoja. Lisäksi palvelu tarjoaa tavan hallinnoida käytösääntöjä, joilla voidaan määrittää esimerkiksi tietyn organisaation tai käyttäjäryhmän edustajien käyttöoikeudet tiettyihin aineistoihin. Käytösääntö kuvaa ehdot, joiden täytyessä käyttöoikeuksia voidaan myös myöntää automaattisesti.

Aineistojen käyttö lupien hakemiseen, myöntämiseen ja hallintointiin rakennetaan yhtenäistä ratkaisua erillisessä NKR-projektissa. SAPA-kokonaisuus hyödyntää tulevaisuudessa NKR-projektin tuotoksia. Käyttöoikeuksien hallinta on tunnistettu tulevaisuuden tarpeisiin. Se vaatii laajamittaisen ja organisaatioiden rajojen ylittävän kehitystyöhön.

3.1.2.6 Asiakkuuksien hallinta

Palvelu sisältää Kansallisarkistolle aineistoja luovuttavien viranomaisten SAPA-palveluun luotavien asiakastilien hallinnan. Palvelun avulla Kansallisarkiston henkilökunta voi luoda uusia organisaatioasiakkuuksia ja hallinnoida niihin liitetyjä käyttäjätunnuksia.

3.1.3 Käsittelyn palvelut

Käsittelyn palveluihin sisältyvät toiminnallisuudet, joissa säilytettäviä aineistoja työstetään tai niistä tuotetaan laajempaa tai koneellisesti työstettyä metatietoa.

Näiden palveluiden avulla Kansallisarkiston henkilökunta voi määritellä ja suorittaa vastaanotetulle tai palvelussa säilytettävälle aineistolle tehtäviä toimenpiteitä.

Kansallisarkisto vastaa aineistolle tehtävistä metatietojen muutoksista (esim. virheellisten metatietojen korjaaminen tai kuvailun rikastaminen) ja aineiston noudosta PAS-palvelusta, kun on tarve saada käyttöön aineiston arkistokappaleet.

3.1.3.1 Aineiston PAS-paketointi

Palvelu sisältää PAS-säilytykseen viennin edellyttämät toiminnot. Palvelu muodostaa vastaanotetusta aineistosta PAS-määritysten mukaisen PAS-vastaanottokelpoisen paketin. Aineistojen PAS-paketoinnissa hyödynnetään PAS-palvelun tarjoamaa paketointikomponenttia, joka pakatoi aineistot määritysten mukaisesti METS- ja PREMIS-formaatteihin. Palvelu palauttaa paketin tunnisteiden aineistojen hallinnalle liitettäväksi aineistojen metatietoihin.

3.1.3.2 Aineistojen nouto PAS-säilytyksestä

Palvelu mahdollistaa aineiston palauttamisen PAS-säilytyksestä. Aineistot haetaan metatietoihin tallennettujen PAS-paketoinnin tuottamien tunnisteiden avulla. Palvelu saa tunnisteiden ja noutopyynnön aineistojen hallinnalta. Palvelu ilmoittaa aineistojen hallinnalle noudon suorituksesta.

3.1.3.3 *Metatietojen tarjoaminen haravoitavaksi*

Palvelu tarjoaa aineistoihin liittyvää metatietoa haravoitavaksi SAPA:n ulkopuolisille tahoille. Haravoinnissa pitää ottaa huomioon metatietojen näyttörajoitukset: julkisten metatietojen tulee olla haravoitavissa sellaisenaan, mutta rajoitetut metatiedot ovat haravoitavissa vain tahoille, joilla on aineiston saamiseksi riittävä valtuutus. SAPA-kokonaisuus hyödyntää AHAA-palveluun sisältyvää toiminnallisuutta metatietojen tarjoamisessa haravoitavaksi.

3.1.3.4 *Aineiston rikastaminen*

Aineiston rikastamisella tarkoitetaan automatisoituja menetelmiä, joilla aineistoa käsittelemällä voidaan tuottaa esimerkiksi uutta metatietoa tai muuta aineistojen hyödynnettävyyttä parantavaa tietoa tai parantavia versioita. Tällaisen rikastamisen tulokset liitetään aina osaksi aineistoja.

SAPA-kokonaisuuteen sisältyviä aineistoja voidaan rikastaa kahdella tavalla:

- Aineistojen arkistollisten kuvailutietojen (kontekstimetatiedot) täydentäminen
- Aineistosta uudenlaisten versioiden tuottaminen (esim. OCR-analyysi kuville)

Palvelu tarvitsee rikastamista varten pääsyn aineistoon, suorituskykyä aineistojen analysointiin sekä riittävän tallennustilan tulosten tallentamiseen. Palvelun tulokset toimitetaan aineistojen hallinnalle, joka vastaa tuotosten liittämistä aineistojen metatietoihin tai muutoin osaksi aineistoa.

Palvelu toteuttaa rajapinnan, jonka avulla aineistojen rikastamisessa voidaan hyödyntää myös ulkoisia palveluita tai komponentteja. Suunnittelun tässä vaiheessa vaikuttaa kuitenkin siltä, että jokainen ulkoinen palvelu tai komponentti vaatii räätälöityä rajapintaa.

Aineistojen rikastaminen on tunnistettu liittyvän SAPA-kokonaisuuden tulevaisuuden tarpeisiin eikä sitä toteuteta SAPA-palvelun alkuvaiheessa.

3.1.3.5 *Metatietojen välittäminen*

Palvelun avulla välitetään siirtopaketissa vastaanotetut metatiedot AHAA-palveluun ja yhdistetään siinä osaksi olemassa olevan arkistollisen kokonaisuuden metatietoja. Palvelun avulla voidaan myös noutaa AHAA-palvelusta aineiston metatiedot ja liittää ne osaksi PAS-säilytykseen siirrettävää pakettia.

3.1.4 **Säilyttämisen palvelut**

SAPA-palveluun siirretään sekä pysyvästi että määräajan säilytettävää aineistoa, joten SAPA-arkkitehtuuriin sisältyvät näitä molempia vastaavat toiminnalliset tarpeet. Lisäksi säilyttämisen palveluihin kuuluu käyttökappaleiden säilytys.

3.1.4.1 Pysyvässäilytys

SAPA-kokonaisuus sisältää pysyvästi säilytettäviä aineistoja. Tällainen säilytys on toteutettava aineistojen autenttisuudesta ja eheydestä varmistuen. Kulttuuriperintö-PAS-palvelu täyttää SAPA-kokonaisuuden pysyvässäilytyksen tarpeet.

3.1.4.2 Määräaikaissäilytys

Palvelu huolehtii määräajan säilytettävän aineiston säilytyksestä ja varmistaa sen muuttumattomuuden. Suunnittelussa on täsmennettävä, miten määräaikainen säilytys toteutetaan ja miten säilytysaikoja hallinnoidaan (AHAA-palvelussa vai SAPA:n omissa tietokannoissa). Palvelu saa aineistojen poistopyynnöt aineistojen hallinnalta.

3.1.4.3 Käyttökappaleiden säilytys

Palvelu toteuttaa käyttökappaleiden säilytyksen. Palveluun siirretään aineistot käyttökappaleiden hallinnasta, mistä tulee myös mahdolliset käyttökappaleiden poistamispyynnöt.

3.1.5 Käyttöliittymä- ja rajapintapalvelut

3.1.5.1 Kansallisarkiston käyttäjien käyttöliittymä

Palvelu tarjoaa käyttöliittymän, joka kokoaa kaikki SAPA-palveluun liittyvät Kansallisarkiston henkilökunnalle tarkoitetut hallinnointitoiminnot. Mikäli Kansallisarkisto kokee tarkoituksenmukaiseksi määrittää henkilökunnalle tarkempia rooleja, toteutetaan tämä käyttöliittymä näihin rooleihin pohjautuen.

Hallintatoiminnot ovat erityisesti hallinnan palveluiden kautta tapahtuvaa aineistojen hallintaa. Palvelu vastaa käyttäjäystävällisen käyttöliittymän tarjoamisesta näihin muiden palveluiden toteuttamiin toiminnallisuuksiin.

3.1.5.2 Viranomaisten käyttöliittymä

Palvelu tarjoaa käyttöliittymän, jolla viranomainen pystyy seuramaan organisaationsa asiakkuuteen liittyvää raportointia ja organisaationsa aineistojen tilaa. Palvelu saa näytettävät tiedot ja raportit hallinnan palveluilta. Viranomainen voi myös siirtyä palvelusta vastaanoton tukipalveluihin sisältyvään siirtopaketin muodostamiseen.

3.1.5.3 Käyttökappaleiden jakelu

Palvelu tarjoaa rajapinnan, jolla ulkoiset tahot voivat pyytää käyttökappaleita SAPA-kokonaisuuteen sisältyvistä aineistoista, jos ovat niiden hyödyntämiseen oikeutettuja. Palvelu saa tarjottavat aineistot käyttökappaleiden hallinnalta. Palvelu raportoi käyttäjien suorittamista aineistonoudoista aineistojen hallinnalle.

Käyttökappaleiden jakelussa huomioidaan aineistojen suojaustasojen asettamat rajoitteet, ja käyttökappaleita tarjotaan vain riittävälle tietoturvasuustasolle auditoituihin järjestelmiin.

3.1.6 Käytön palvelut

3.1.6.1 Hakupalvelu

Palvelu tarjoaa tavat hakea, löytää ja identifioida aineistoja näiden metatietojen avulla. Hakupalvelu pohjautuvat AHAA-palvelusta haravoitujen tietojen päälle toteutettuun ratkaisuun, jossa käyttäjille tarjotaan erilaisia haku- ja navigointivaihtoehtoja aineistojen metatietoihin. Palvelutarve toteutetaan FINNA:n ja NKR-projektin avulla. Palvelu välittää aineistojen käyttöpyynnöt aineistojen hallinnalle, joka palauttaa vastaukset käyttöpyyntöön palvelulle.

3.1.6.2 Tilauspalvelut

Palvelun avulla asiakkaat voivat tehdä aineistotilauksia. Tilaukset voivat olla erilaisia tietopyyntötilauksia, joissa asiakkaalle toimitetaan aineisto tai tämän jäljenteet eri muodoissa. Nykyisin tilauspalvelut keskittyvät analogisen aineiston ja tästä tuotettujen jäljenteiden tilaamiseen, mutta tulevaisuudessa tilaukset voivat kohdistua esim. aineiston saattamiseksi digitaaliseksi tai erikoisaineistojen (esim. rekisterit) tilaamiseen käyttöön.

3.1.6.3 Käyttökappaleiden käyttö

Palvelu tarjoaa virtuaalisen työtilan, jossa voi katsella ja jalostaa aineistoja, joihin on riittävät oikeudet. Palvelu on esimerkiksi virtuaalinen työpöytä, jolla voi hyödyntää aineistojen käyttökappaleita esimerkiksi kuvina, teksteinä ja tarvittaessa yhdistellä näitä.

Palvelu hakee aineistot käyttökappaleiden jakelun tarjoamasta rajapinnasta ja käyttää tunnistamisratkaisuja, sekä käyttökappaleiden hallintaa.

Käyttökappaleiden käyttö on tunnistettu tulevaisuuden tarpeisiin. Se vaatii laajamittaisempaa ja organisaatioiden rajat ylittävää kehitystyötä.

3.2 KÄYTTÄJÄT JA KÄYTTÄJÄRYHMÄT

SAPA-palvelun käyttäjät jakaantuvat seuraaviin käyttäjäryhmiin:

- a. Järjestelmäkohtaiset tunnukset
 - I. Valtionhallinnon järjestelmät (tietojärjestelmien välinen integraatio)
- b. Henkilökohtaiset tunnukset
 - I. Valtionhallinnon viranomaiset (nimetyt edustajat)
 - II. Kansallisarkiston henkilökunta
 - III. Tekniset ylläpitäjät

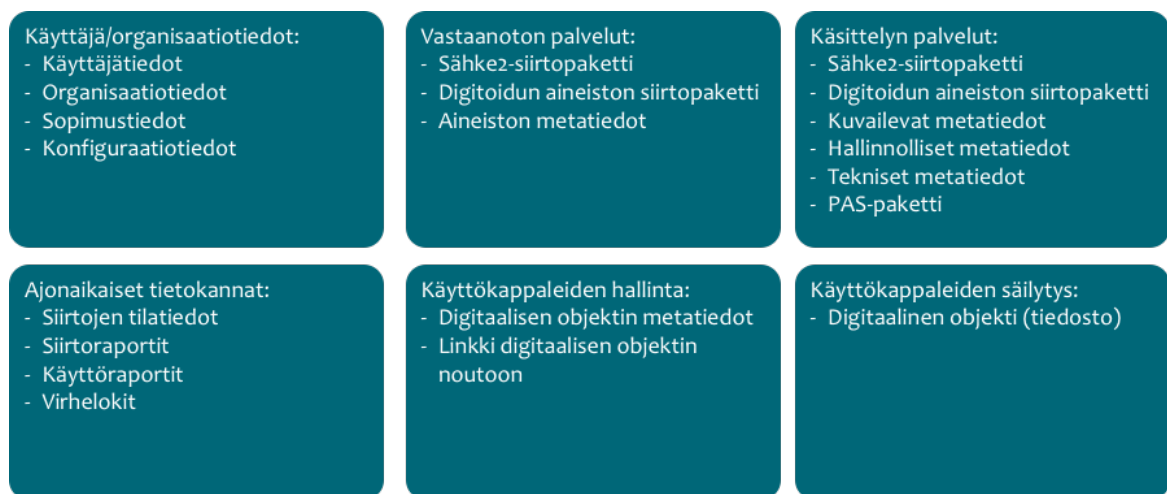
Järjestelmäkohtaiset tunnukset mahdollistavat valtionhallinnon tietojärjestelmien integraation SAPA-palveluun ja esimerkiksi aineistojen säännölliset luovutukset Kansallisarkistolle.

Henkilökohtaisia ja yksilöiviä tunnuksia käyttävät kaikki muut SAPA-palvelun käyttäjät (edellä olevassa listassa b.I., b.II. ja b.III.). Käyttäjistä kerättävästä henkilötiedosta vastaa tietosuojalainsäädännön tarkoittamana rekisterinpitäjänä Kansallisarkisto. Käyttäjistä kuvataan järjestelmään vähintään käyttäjän nimi, organisaatio, käyttäjärooli sekä käyttäjätunnus ja salasana. Tätä tarvittavien tietojen listaa ja käsittelyn periaatteita voidaan tarkentaa suunnittelu- ja toteutusvaiheessa.

4 TIETOARKKITEHTUURI

4.1 LOOGISET TIETOVARANNOT

Looginen tietovaranto kattaa toiminnan tarpeista kootun ja yhteisesti hallinnoidun joukon tietoja tai tietoaineistoja, jotka ovat olennaisia toiminnassa ja palveluissa. SAPA:n loogiset tietovarannot on esitetty Kuva 3.



Kuva 3: SAPA:n loogiset tietovarannot.

Käyttäjä- ja organisaatitiedot sisältävät kaikki tarvittavat tiedot SAPA:a käyttävistä organisaatioista ja niiden käyttäjistä ja näiden rooleista. Organisaatitietoihin liittyvät myös tehdyn sopimuksen tiedot ja sopimustunnus, jota käytetään mm. aineiston vastaanotossa. Tähän tietovarantoon sisältyvät myös organisaatio- ja käyttäjäkohtaiset konfiguraatitiedot.

Vastaanoton palvelussa talletetaan SAPA:an tulevat, alkujaan digitaalisten SÄHKE2-muotoisten ja digitoitujen aineistojen siirtopaketit, jotka sisältävät myös aineiston metatietoa.

Käsittelyn palvelussa käsitellään vastaanoton päätteeksi hyväksytysti vastaanotetut siirtopaketit (alkujaan digitaaliset SÄHKE2-muotoiset ja digitoidun aineiston siirtopaketit), sekä muodostetaan PAS-paketti pitkäaikaissäilytystä varten. Käsittelyn palveluun liittyy myös metatietoja, joista aineistoa kuvailevat ja hallinnolliset metatiedot päätyvät AHAA-palveluun ja tekniset metatiedot PAS-paketin mukana PAS-palveluun.

Ajonaikaiset tietokannat sisältävät kaikki dynaamiset tiedot liittyen SAPA-palvelukokonaisuuden työkulkuihin ja toimintoihin. Ajonaikaisiin tietoihin kuuluvat mm. siirtojen tilatiedot, siirto- ja käyttöraportit, virhestatukset ja asynkronisten toimintojen etenemisen ja valmistumisen statustiedot.

Käyttökappaleiden hallinta sisältää käyttökappaleiden, eli digitaalisten objektien metatiedot sekä linkin käyttökappaleeseen, siis tiedon siitä, mistä käyttökappaleen voi noutaa.

Käyttökappaleiden säilytys sisältää käyttökappaleet digitaalisina tiedostoina.

4.2 TIETOTURVA

Toteutus tullaan auditoimaan voimassa olevia tietoturvamääräyksiä ja -ohjeita vasten. Käyttöliittymiä ja rajapintoja toteutettaessa otetaan huomioon myös OWASP¹-tietoturvaohjeet.

¹ Open Web Application Security Project (OWASP). https://www.owasp.org/index.php/Main_Page