KANSALLISARKISTO

16.12.2019

Massadigitoinnin pilotoinnin loppuraportti
Julkinen versio

Sisällys

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

1 Lähtökohdat

Massadigitoinnin tavoitteena on digitoida valtion viranomaisten pysyvästi ja pitkään säilytettävät asiakirjat niin, että alkuperäiset analogiset asiakirjat voidaan sähköisen säilytyksen ja käytettävyyden varmistamisen jälkeen hävittää. Massadigitoinnin toteuttamista on suunniteltu Kansallisarkistossa vuodesta 2017 lähtien. Hankkeelle asetettu ohjausryhmä linjasi, että massadigitoinnin varsinaisen tuotannon käynnistäminen edellyttää pilotoinnin toteuttamista. Päätös pilotoinnin toteuttamisesta tehtiin hankkeen ohjausryhmässä helmikuussa 2018.

Lähtökohtana pilotoinnille olivat vuoden 2017 massadigitoinnin suunnitteluprojektin tuottamat arviot digitoinnin toteutustavasta sekä nopeus- ja kustannusarvioista. Pilotoinnin tavoitteita täsmennettiin vähitellen niin, että lopullinen versio pilotin tavoitteista laadittiin elokuussa 2018. Pilotoinnin valmistelu oli tässä vaiheessa jo käynnistetty. Pilotoinnin arvioinnin kriteerit määriteltiin ohjausryhmässä huhtikuussa 2019.

2 Tavoitteet ja loppuraportin rakenne

Pilotoinnin tavoitteena oli saada konkreettista näyttöä digitointiprosessin toimivuudesta ja etenemisnopeutta koskevien arvioiden luotettavuudesta. Massadigitoinnin prosessi, siinä käytettävä teknologia, läpimenoajat ja kustannukset oli arvioitu vuonna 2017 toteutetussa massadigitoinnin suunnitteluprojektissa. Pilotoinnin tavoitteena oli selvittää näiden arvioiden realistisuutta todellisessa digitoinnin tuotannossa. Keskeiset tavoitteet pilotille oli selvittää:

- Digitointiprosessin luotettavuus ja sen eri työvaiheisiin kuluva aika sekä kokonaiskesto
- Todellisen tuotannon perusteella digitoinnin laatu (hävittämiseen tähtäävän digitoinnin kriteerien mukaiset tiedosto-objektit)
- Miten sujuvasti ja eheästi digitointiprosessi todellisessa tuotannossa toimii ja etenee
- Miten massadigitoinnin tuotannon työn organisoiminen ja johtaminen on tehokkainta toteuttaa
- Digitoinnin edellyttämän ohjeistuksen selkeys ja kattavuus

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Pilotoinnin tavoitteet, rajaukset ja aikataulu määriteltiin tarkemmin erikseen laaditussa ja hankkeen ohjausryhmän hyväksymässä suunnitelmassa.

Pilotoinnin toteuttaminen edellytti kertaluonteista toimitiloista, laitteistosta ja tuotantotiloista muodostuneen tuotantoympäristön rakentamista, josta saatiin arvokasta tietoa massadigitoinnin tuotannon pystyttämisen suunnitteluun. Samoin pilotissa rakennettiin eri tietokantojen, ohjelmistojen, sovellusten ja palvelinratkaisujen muodostama kokonaisuus, joka muodostaa ensimmäisen version massadigitointiin tarvittavasta tuotannonohjauksen kokonaisuudesta.

Pilotissa myös käytännössä testattiin todellisessa tuotannossa skannerilaitteiston kestävyyttä, toimivuutta ja nopeuksia sekä eri toteutusmalleja. Vastaavia asioita oli jo aikaisemmin ilman tuotantoa testattu massadigitoinnin *Proof of Concept* -selvityksessä vuonna 2018.

Pilotin onnistumiselle asetettiin myöhemmin hankkeen ohjausryhmässä erilliset onnistumisen kriteerit 1–8. Kriteerit käyvät ilmi taulukosta 3. ja niiden mukainen pilotin onnistumisen arviointi esitetään luvussa 12.

Loppuraportti on pyritty pitämään mahdollisimman tiiviinä ja keskittymään siinä tulosten, havaintojen ja johtopäätösten esittämiseen. Loppuraportti painottuu pilotoinnin digitointivaiheeseen.

3 Rajaukset

Pilotoinnin lyhyen suunnittelu- ja digitoinnin toteutusjakson keston vuoksi osa massadigitoinnin varsinaiseen tuotantoon määritellyistä komponenteista ja tavoitteista rajattiin pilotin ulkopuolelle. Lähtökohtana oli, että pilotointi pystytään tiiviissä aikataulussa toteuttamaan ilman järjestelmämuutoksia, joiden toteuttaminen vaatisi merkittävästi enemmän aikaa kuin pilotoinnin valmisteluun oli. Tällaisia olivat arkistoinnin palvelukokonaisuus SAPA, integroitu tuotannonohjausjärjestelmää, sekä metatietojen tuottaminen sisällönanalysoinnin keinoin. Pilotoinnissa haluttiin tietoa mahdollisimman monen digitointiprosessin vaiheen toimivuudesta. Tämän vuoksi toteutus ulotettiin viranomaisvalmistelusta aina digitaalisen aineiston viranomaiskäyttöön. Tarkemmat rajaukset käyvät ilmi alla olevasta taulukosta 1. Yhteenvetona voidaan todeta, että pilotointi toteutettiin alkuperäisten rajausten mukaisesti.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Taulukko 1. Pilotoinnin rajaukset

Tuotanto	Pilotti (tavoite)	Pilotti (toteutuma)
Prosessivaihe/tietojärjestelmä	Sisältyy (kyllä/ei)	
Viranomaisvalmistelu ja metatietojen syöttö AHAA-palveluun	**Kyllä**	**Kyllä**
Erillinen logistiikkaprosessi asiakirjojen digitointiin kokoamiseksi	**Ei** (KA kuitenkin vastaa asiakirjojen noutamisesta)	**Ei**
Digitoinnin vaiheiden ohjaaminen tuotannonohjauksella	**Ei** (hoidetaan manuaalisesti)	**Kyllä** (KA:n toteuttama logistiikkaohjelma)
Aineiston valmistelu digitointia varten digitoinnin toteutuksen yhteydessä	**Kyllä**	**Kyllä**
Pysyvästi säilytettävien asiakirjojen hävittämiseen tähtäävän digitoinnin mukainen suurteho- ja muu skannaus	**Kyllä**	**Kyllä**
Määräajan säilytettävien asiakirjojen hävittämiseen tähtäävän digitoinnin mukainen suurteho- ja muu skannaus	**Ei**	**Ei**
OCR –tekstintunnistus	**Osittain** (Digitointiin sisältyy tekstintunnistus, mutta ei aineistojen sisältöhakua)	**Kyllä**
Digitaalisten ilmentymien validointi, paketointi ja siirto KP-PAS -palveluun (säilyttäminen)	**Kyllä**	**Osittain** (odottaa nauhoilla SAPA-palvelua ja sen myötä tehtävää KP-PAS -siirtoa)
Käyttökappaleiden säilyttäminen SAPA-palvelussa	**Ei**	**Osittain** (digitoinnissa tuotettiin käyttökappaleet, jotka tullaan siirtämään SAPA-palveluun)
Ilmentymien asiakaskäyttö	**Kyllä**	**Ei** (käyttöliittymä tehty, mutta ei vielä asiakaskäytössä)
Automatisoitu digitoinnin validointiprosessi ja integroitu tuotannonohjausjärjestelmä	**Ei**	**Ei**
Automaattinen metatietojen tuottaminen sisällönanalysoinnin keinoin asiakirjoista	**Ei**	**Ei**
Analogisten asiakirjojen hävittäminen	**Ei** (Analogisten asiakirjojen hävittäminen ei kuulu pilotointiin.)	**Ei**

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

4 Aineisto

Pilotoinnissa digitoitavan aineiston valinnan lähtökohtana oli aineistojen soveltuvuus pilotoinnille. Pilottiin valittiin aineistot Tiekartan 1. ryhmästä, eli ensimmäisenä massadigitoinnin tuotantoon aikataulutetuista aineistoista. Aineiston kokoamisessa pyrittiin myös huomioimaan erilaiset aineistotyypit, jotta läpimenonopeuksista ja erilaisten aineistojen soveltuvuudesta massadigitointiin saataisiin konkreettista lisätietoa. Määrälliseen tavoitteeseen vaikuttivat aineistotyyppien ohella arviot digitointinopeudesta sekä käytettävän laitteiston ja henkilöstöresurssien laajuudesta. Lopputuloksena oli, että pilotoinnissa sitouduttiin digitoimaan viiden kuukauden tuotantojaksolla noin 450 hyllymetriä tiekartan 1. ryhmän viiden eri viranomaisen aineistoja, jotka edustivat erilaisia aineistotyyppejä.

Aineistot koostuivat erilaisista paperiasiakirjoista, sidoksista ja kortistoista. Suurin osa aineistosta oli A4-kokoisia asiakirjoja. Sidoksia oli noin 2 % kokonaismäärästä. Aineistot on esitetty alla olevassa taulukossa 2. Aineiston laajuus on ilmoitettu hyllymetreissä.

Taulukko 1. Pilotoinnin aineistot, määrät ja laajuus

Organisaatio	Arkisto ja aineistotyyppi	Rajavuodet	Säilytys-yksiköitä (kpl)	Säilytystila (hyllymetriä =hm)
Verohallinto	**Espoon verovirasto** A4-arkki, paperiliittimet	1994-2011	1119	93,65
Verohallinto	**Pääkaupunkiseudun verotoimisto** A4, paperiliittimet	2014-2015	144	13,75
Fimea	**Fimea** A4, nauhakiinnitys	1995-2016	1238	116,7
Terveyden ja hyvinvoinnin laitos	**Kansanterveyslaitos** A4-vihko, eri arkkikokoja, niittauskiinnitys	1966-2001	993	99,4
Työ- ja elinkeino-ministeriö	**Työministeriö** A4, paperiliittimet	1999-2009	877	71,05
Länsi-Uudenmaan käräjäoikeus	**Espoon tuomiokunta** Kortistot, sidokset, A4, velobind-kiinnitys	1971-1994	286	30,4
Länsi-Uudenmaan käräjäoikeus	**Espoon käräjäoikeus** A4	1993-1994	12	1,3
SUMMA		**1966-2015**	**4669**	**426,25**[1]

[1] Aineiston määrään tehtiin tarkistuslaskenta digitoinnin yhteydessä, jonka mukaan sen määrä on 386,17 hyllymetriä.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

5 Aikataulu ja etenemisvaiheet

Pilotointi ei tarkoita vain digitointia. Pilotointi käynnistyi sen tavoitteiden määrittelystä ja aineistovalinnoista, minkä jälkeen se käsitti kaikki digitointiprosessin vaiheet viranomaisvalmistelusta aineiston asiakaskäyttöön. Pilotoinnin vaiheet olivat osittain limittäisiä. Yleisemmällä tasolla pilotointi voidaan jakaa suunnittelu-, valmistelu-, toteutus- ja analysointivaiheisiin. Näistä kaksi ensiksi mainittua toteutettiin aikavälillä 02/2018 – 03/2019, ja kaksi viimeistä aikavälillä 04 – 10/2019. Suunnittelu- ja valmisteluvaiheessa mm. määritettiin pilotoinnin tavoitteet, toteutettiin viranomaisvalmistelu, sen ohjeistus ja ohjaus, sekä kilpailutettiin pilotoinnin hankinnat. Toteutus- ja analysointivaiheessa mm. pystytettiin tuotantoympäristö, digitoitiin aineisto ja analysoitiin tulokset.

Pilotoinnin toteutus- ja analysointivaiheen ajankohdaksi asetettiin 04–10/2019. Pilotoinnin digitoinnin ja tulosten tuli siten olla valmiit 31.10.2019 mennessä.

Pilotoinnin suunniteltu ja toteutuma-aikataulut käyvät ilmi seuraavalla sivulla olevasta kuvasta 1.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Kuva 1. Pilotoinnin etenemisaikataulu osa-alueittain

Pilotoinnin osa-alueet ja kesto	Vuosi	2018					2019				
Osa-alue	Kuukausi	03-04	05-06	07-08	09-10	11-12	01-02	03-04	05-06	07-08	09-10
Toteutusmallin suunnittelu ja alkubudjetointi	Suunnitelma										
	Toteutuma										
Viranomaisyhteistyö (aineistovalinnat, sopimukset)	Suunnitelma										
	Toteutuma										
Viranomaisohjaus (perehdytys ja aineiston viranomaisvalmistelu)	Suunnitelma										
	Toteutuma										
Teknisten kyvykkyyksien toteutus (sovellukset ja järjestelmät)	Suunnitelma										
	Toteutuma										
Laitteiston ja ohjelmistojen kilpailutukset ja hankinnat	Suunnitelma										
	Toteutuma										
Henkilöstörekrytointi	Suunnitelma										
	Toteutuma										
Infran ja työtilojen valmistelu ja muutostyöt	Suunnitelma										
	Toteutuma										
Aineistologistiikka (kuljetussopimukset ja aineiston kuljettaminen)	Suunnitelma										
	Toteutuma										
Henkilöstöperehdytys, työn organisoiminen, digitoinnin käynnistyminen	Suunnitelma										
	Toteutuma										
Pilotin aktiivinen toteutusvaihe (valmistelu ja digitointi) ja loppuraportointi	Suunnitelma										
	Toteutuma										

6 Viranomaisyhteistyö ja -ohjaus

Jokaisen viranomaisen kanssa laadittiin pilotoinnin sopimukset, joissa määriteltiin vastuut sopijapuolten välillä mm. aineiston digitaalisten ja analogisten versioiden hallinta- ja käyttöoikeudesta, digitoinnista, säilyttämisestä ja hävittämisestä, henkilötietojen käsittelystä ja rekisterinpitäjyydestä, tietopalvelusta sekä tietoturvaan ja -suojaan liittyvistä kysymyksistä.

Aineistovalinnat tehtiin aikataulussa, mutta sopimusten valmistelu venyi suunniteltua pitemmälle erityisesti tietosuojaa koskevien kirjausten tarkentamisen vuoksi.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Pilotointia varten laadittiin uudet versiot Kansallisarkiston siirto-oppaasta ja siirtomääräyksestä, joissa esitettiin massadigitoinnin kriteerit ja vaatimukset. Pilotoinnin viranomaisvalmistelu toteutettiin näiden ohjeiden mukaisesti. Massadigitoinnin myötä viranomaisvalmisteluun tuli jonkin verran kevennystä aikaisempiin siirtokriteereihin verrattuna.

Viranomaisten tuli ennen aineistojen siirtämistä pilotointiin luetteloida sitä koskevat luettelointitiedot Kansallisarkiston AHAA-hakemistopalveluun. AHAA:n tuotantoversio otettiin käyttöön samassa yhteydessä, joten osana pilotointia tehtiin sekä AHAA:n kehitystyötä, että ohjattiin viranomaisia AHAA:n käyttöön. Viranomaisille laadittiin erillinen luettelointiohje AHAA-hakemistopalvelun käytöstä. Viranomaisten luettelointityötä ohjattiin ja työn etenemistä seurattiin ohjauskäyntien ja etäohjauksen avulla.

Viranomaisille annetusta ohjauksesta ja laadituista ohjeista toteutettiin pilotin viranomaisille erillinen palautekysely, jonka pohjalta ohjeita ja ohjausta kehitetään edelleen pilotoinnin päättymisen jälkeen.

Aineistojen kuljettamisesta viranomaisilta Kansallisarkistoon ja kuljetuskustannuksista vastasi Kansallisarkisto. Kuljetukset toteutti Hanselin puitesopimuksen pohjalta kevennetyllä kilpailutuksella valittu kuljetusliike. Kuljetuksissa huomioitiin normaalin arkistoaineiston kuljetuksen ja tietoturvan takaamisen lisäksi viranomaisilta tulleet ehdot aineiston käsittelyyn.

7 Laitteistohankinnat ja infrastruktuuriratkaisut

Pilotoinnin teknisten kyvykkyyksien osalta vaihtoehtoina olivat toteuttaminen Kansallisarkiston omilla resursseilla, kokonaan kilpailutettuina ulkoisina hankintoina ja ostopalveluina tai näiden välimuotona. Parhaaksi toteutusmalliksi valittiin yhdistelmä, jossa skannerit ja niiden ohjelmistokokonaisuus hankittiin ostopalveluna, infrastruktuuriratkaisut hankittiin CSC Oy:ltä ja tuotannonohjaukseen sekä aineiston siirtoon liittyvät sovellukset kehitettiin Kansallisarkiston omana työnä.

Skanneri- ja ohjelmistokokonaisuutta varten Kansallisarkisto laati teknisen vaatimusmäärittelyn, jonka pohjalta toteutettiin avoimen kilpailutuksen menettelyn mukainen hankinta. Kilpailutuksessa käytettiin Hanselin konsulttipalveluita. Skannerikokonaisuuden toimitusprojekti päättyi hankinnan hyväksyntätestaukseen, jolla varmistettiin, että kokonaisuus täytti vaatimusmäärittelyn.

Skannereiden kyky tuottaa laatuvaatimusten mukaisia kuvia vaatimusmäärittelyssä asetetuilla enimmäisnopeuksilla todennettiin käyttöönottotestauksessa. Tuotannossa

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

maksiminopeudet riippuivat ja asetettiin aineiston laadun mukaan. Suurteholaitteella yleinen keskinopeus oli 210 arkkia minuutissa. Dokumenttiskannerilla enimmäisnopeus oli noin 170 arkkia minuutissa. Mastoskannerin keskimääräinen nopeus oli noin 9 arkkia minuutissa.

CSC:n toteuttamat infrastruktuuriratkaisut perustuivat virtuaalipalvelinjärjestelmään, jolla pystyttiin parhaiten ratkaisemaan palvelinten sijoituspaikkaan, määrään ja skaalaamiseen liittyvät tarpeet. CSC:n toteutus perustui Kansallisarkiston ja CSC:n väliseen puitesopimukseen.

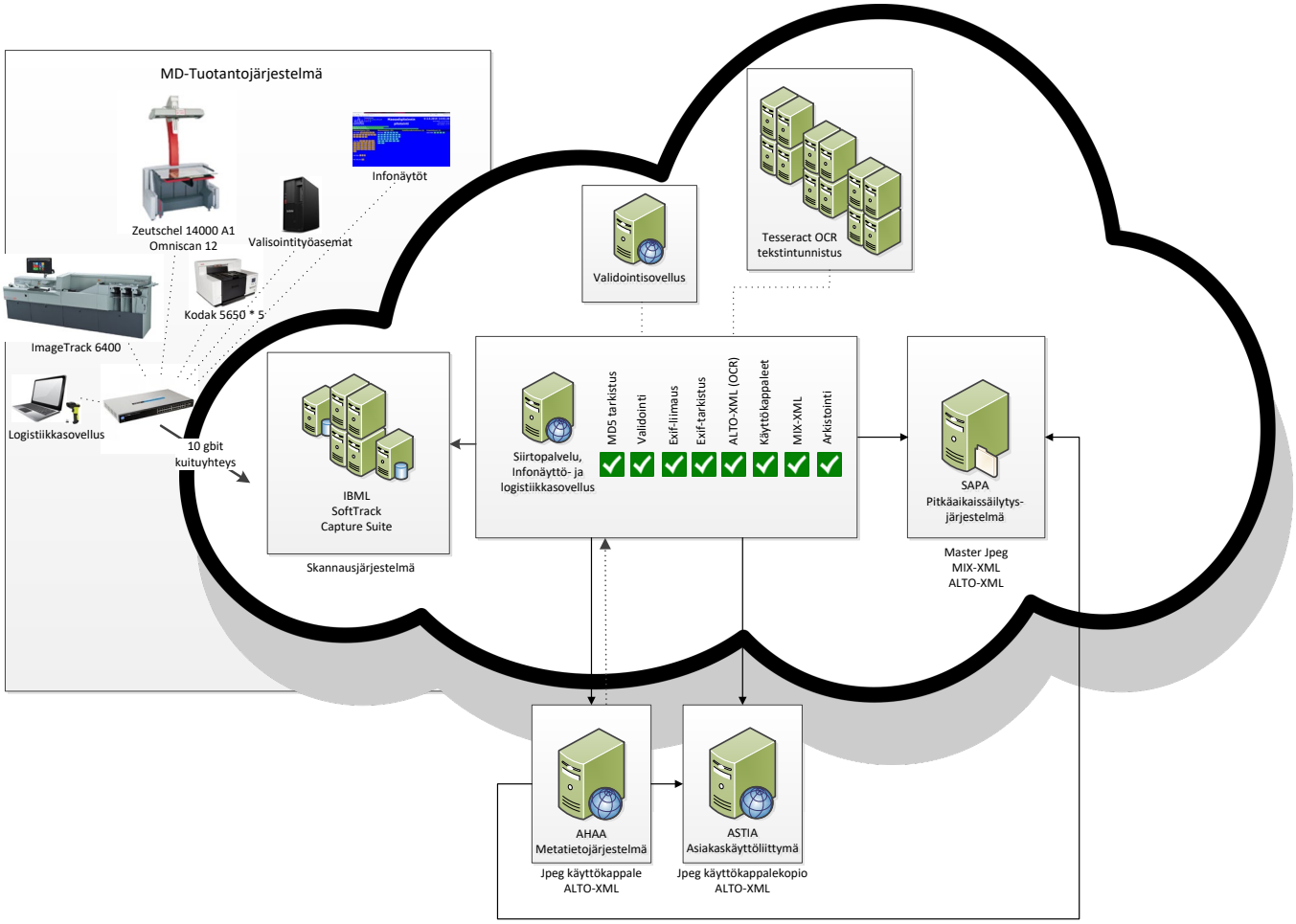
Omana työnään Kansallisarkisto kehitti pilottia varten oman logistiikkasovelluksen, validointisovelluksen, siirtopalvelusovelluksen sekä infonäyttösovelluksen.

Logistiikkasovelluksen tehtävänä oli varmistaa analogisten ja digitaalisten ilmentymien hallinta prosessin eri vaiheissa. Logistiikkasovelluksen avulla aineiston siirtymistä eri vaiheisiin voitiin seurata reaaliaikaisesti. Logistiikkasovelluksella pyrittiin luomaan ketterä menetelmä, jonka avulla voitiin varmistaa laadunvalvontaan, tuotantonopeuteen ja tuotannon tunnuslukuja kuvaaviin raportointiominaisuuksiin liittyvien toiminnallisuuksien saaminen osaksi pilotointia.

Helposti konfiguroitavan *validointisovelluksen* avulla toteutettiin digitoidun aineiston kuvanlaadun tarkistus tietosisällön osalta. *Siirtopalvelusovellus* kehitettiin digitaalisen aineiston jälkikäsittelyyn ja prosessointiin (mm. OCR-tunnistus). *Infonäyttösovelluksen* tehtävänä oli jakaa tuotantohenkilöstölle erilaista ajankohtaista tietoa mm. aineistomääristä eri prosessivaiheissa, tuotantolupien käsittelystä, aikatauluista ja levykapasiteetista. Pilotoinnin järjestelmäkokonaisuus käy yleisellä tasolla ilmi alla olevasta kuvasta 2.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Kuva 2. Pilotoinnin digitoinnin järjestelmäkokonaisuuden sovellukset, ohjelmistot ja rajapintajärjestelmät



8 Tietoturva ja tietosuojaratkaisut

Digitoitavan aineiston joukossa oli runsaasti salassa pidettäviä asiakirjoja. Tietosuojan varmistamiseksi toteutettiin useita erilaisia tietoturvamenettelyjä. Keskeiset ratkaisut kirjattiin pilotointia koskeviin virastojen ja Kansallisarkiston välisiin sopimuksiin. Digitoinnin säilytys- ja toimitiloihin rajoitettiin kulkuoikeudet vain massadigitoinnin henkilöstölle. Digitoinnin ohjelmisto- ja palvelinympäristö sovelluksineen toteutettiin muista verkoista eriytettynä ratkaisuna. Henkilöstölle annettiin tietoa ja perehdytystä tietosuoja- ja tietoturvavaatimuksiin sekä menettelytapoihin. Pilotoinnille laadittiin oma tietoturvasuunnitelma, jonka perusteella tarvittavat tietoturvaratkaisut toteutettiin. Pilotin tietoturvallisuuden tilaa seurattiin säännöllisesti ja poikkeamatilanteiden hallintaan laadittiin oma ohjeistus.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

9 Pilotoinnin resurssit

Henkilöstörekrytoinnit toteutettiin kahdessa päävaiheessa. Kun resurssitarpeet oli määritelty, rekrytoitiin ensimmäisessä vaiheessa asiantuntija- ja ohjaushenkilöstö. Toisessa vaiheessa rekrytoitiin tuotantohenkilöstö. Henkilöille laadittiin perehdytysohjelma, jossa keskityttiin pilotoinnin tavoitteisiin, toteutustapaan, työkaluihin sekä tietosuojavaatimuksiin.

Uusien rekrytointien lisäksi määriteltiin ja ohjattiin myös massadigitoinnin ja muita Kansallisarkiston henkilöresursseja pilotoinnin asiantuntijatehtäviin. Näiden asiantuntijatehtävien osuus pilotoinnissa on ollut huomattava. Yhteensä pilotointiin osoitettiin asiantuntijapanosta vuosien 2018 – 2019 aikana yli 14 henkilötyövuoden verran.

Pilotoinnin tuotantohenkilöstö koostui kokonaan uusista ja tätä tehtävää varten rekrytoiduista henkilöistä. Pilotointiin oli suunniteltu 7+2 henkilön laajuinen tuotannon henkilöstöresurssi. Koska pilotoinnissa käytettiin seitsemää skanneria, oli henkilöstömäärä suppein mahdollinen, jolla tuotannosta arvioitiin voitavan selvitä aikataulussa. Henkilöstössä tapahtui jonkin verran vaihtuvuutta. Keskimääräinen tuotantopäivän aikana paikalla olleen henkilöstön määrä oli 7,8. Skannauksen ohella henkilöstön tuli vastata myös mm. aineiston valmistelusta, validoinnista, tietopalvelusta ja tilojen ylläpidosta. Logistiikasta vastanneet kaksi henkilöä osallistuivat tarpeen mukaan myös muihin tehtäviin. Koko henkilöstö perehdytettiin kaikkiin tuotannon osa-alueisiin. Tuotantohenkilöstön kokonaisresurssi vuonna 2019 oli yli 5 henkilötyövuotta.

Pilotointiin osallistui yhteensä 35 Kansallisarkiston vakituista tai määräaikaista työntekijää.

10 Rahoitus

Pilotin rahoitus oli osa massadigitoinnin kehittämisen rahoitusta, joka koostui osista ja eri tavoin vuosina 2018 ja 2019.

Pilotointia varten jouduttiin tekemään melko suuri määrä kertaluonteisia hankintoja, koska sekä tuotantotilat, -laitteisto sekä suurelta osin myös ohjelmistot, sovellukset ja muu infrastruktuuri jouduttiin suunnittelemaan ja rakentamaan alusta asti pilotointia varten. Pilotointiin kohdistuneen rahoituksen kustannushyödyt tulevat ulottumaan siksi pitemmälle ajalle pilotoinnin jälkeen, kun pilotoinnin tuotantoympäristöä ja luotuja prosesseja ja toimintamalleja voidaan hyödyntää sen jälkeiseen massadigitoinnin tuotantoon.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Lisäksi pilotoinnin yhteydessä kehitettiin laajemmin massadigitoinnin edellytyksiä toteuttamalla mm. massadigitoinnin edellyttämiä muutoksia liittyviin tietojärjestelmiin (AHAA ja AHJ) sekä kehittämällä mm. viranomaisohjeistusta, tarkentamalla tiekarttasuunnitelmaa ja kehittämällä automaattista metatietojen tuottamista sisällönanalysoinnin keinoin. Kaikkia kehitettyjä kokonaisuuksia ei päästy vielä hyödyntämään pilotissa ja osin kehittämistyö jatkuu seuraavina vuosina.

Pilotoinnin keskeiset kulut muodostuivat henkilöstömenoista (noin 31 %) ja laite-, ohjelmisto- ja palveluhankinnoista (noin 52 %).

11 Digitoinnin toteutus

Pilottituotannon keskeiset prosessit olivat aineiston viranomaisvalmistelu, aineiston valmistelu digitoitavaksi, digitointi (sisältäen mm. skannauksen ja validoinnin aliprosessit), digitoidun aineiston prosessointi (sisältäen digitaalisen aineiston OCR- ja muun jälkiprosessoinnin), sekä analogisen aineiston käsittely digitoinnin jälkeen.

Digitoinnin tuotannon toteuttaminen edellytti erillistä tuotantosuunnittelua, jossa huolehditaan tuotannon aikatauluttamisesta ja aineiston jakamisesta eri laitteille, resurssien kohdentamisesta sekä henkilöstön ohjauksesta ja ohjeistamisesta tehtäviin. Hankitun laitteiston pohjalta tuotantosuunnittelussa päädyttiin kolmen tuotantolinjan malliin, joihin aineistoa valmisteltiin ja syötettiin neljällä eri toteutusmallilla. Tuotantolinjoilla tarkoitetaan pilotoinnissa käytettyjä skannerityyppejä ja toteutusmallilla aineiston valmistelun ja skannauksen työvaiheiden erilaisia toteuttamistapoja.

Tuotantosuunnitteluun kuuluivat myös tilasuunnittelu ja tarvikehankinnat. Pilotoinnin tuotantotiloihin tehtiin merkittäviä investointeja hankkimalla siihen tarvittava kalustus, muu laitteisto ja välineistö sekä päivittämällä tilan tietoliikenneyhteyksiä.

Kokonaisuudessaan pilotoinnissa pyrittiin teolliseen linjastotuotantoon siinä määrin kuin käytettävissä olevissa tiloissa ja ilman täysin automatisoituja tuotannonohjausjärjestelmiä oli mahdollista.

Keskeisenä välineenä tuotantosuunnittelussa toimi pilotoinnin *logistiikkasovellus*, jonka avulla seurattiin ja ohjattiin aineiston siirtymistä prosessien välillä. Aineistojen yhteyteen oli viranomaisvalmisteluvaiheessa tulostettu AHAA-järjestelmään syötettyjen metatietojen pohjalta viivakoodi, jonka avulla aineisto luettiin prosessivaiheesta toiseen. Aineistoa säilytettiin tuotantotiloissa ainoastaan valmistelun ja digitoinnin edellyttämä aika. Heti digitoinnin päätyttyä aineisto palautettiin fyysisesti

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

säilytysmakasiiniin odottamaan hävittämistä ja luettiin tähän käsittelyvaiheeseen logistiikkasovelluksessa.

Digitoinnissa tuotettujen digitaalisten ilmentymien muoto, laadunvarmistus ja prosessointi kuvataan luvussa 12.

12 Tulokset

Tässä luvussa tuloksia arvioidaan pilotoinnin kriteerien ja muiden hankkeen ohjausryhmän asettamien mittareiden kautta. Kriteerit, mittarit ja toteutuminen on koottu alle taulukkoon 3. Kunkin kriteerien toteutumista arvioidaan tarkemmin sanallisesti taulukon alla.

Taulukko 2. Pilotoinnin kriteerit ja niiden toteutuminen

Pääkriteeri	**Kriteerin selitys**	**Arviointiperuste / skaala (lihavoitu toteutunut vaihtoehto)**
K1. Digitoinnin laatu	Hävittämiseen tähtäävän digitoinnin mukaisen kuvanlaadun sekä digitoinnin laatutason (skannauksen virheettömyys) saavuttaminen pilotissa vaaditulla nopeudella	***Toteutuu*** *Lieviä poikkeamia* *Merkittäviä poikkeamia* *Ei toteudu*
K2. Digitoidun aineiston siirto säilytykseen	Digitoinnissa pystytään tuottamaan määritelty siirtopaketti vaadituilla pitkäaikaissäilytyksen metatiedoilla	***Toteutuu*** *Lieviä poikkeamia* *Merkittäviä poikkeamia* *Ei toteudu*
K3. Digitoidun aineiston käyttö ja käytettävyys	Viranomaiset ja Kansallisarkisto voivat käyttää digitoitua aineistoa tarvittaviin käyttötarkoituksiin. Aineiston käytettävyys parantuu merkittävästi.	*Toteutuu* *Lieviä poikkeamia* ***Merkittäviä poikkeamia*** *Ei toteudu*
K4. Aineiston hävittäminen on mahdollista	Kriteerien K1 – K3 vaatimukset täyttyvät niin kattavasti, että aineisto on mahdollista hävittää varoajan jälkeen	***Kyllä****/Ei*
K5. Pilotoinnin tuotantotavoitteiden saavuttaminen	Pilotissa saavutetaan tavoitteen mukainen keskimääräinen tuotantonopeus	*Toteutuu* ***Lieviä poikkeamia*** *Merkittäviä poikkeamia* *Ei toteudu*
K6. Pilottiaineiston digitointi onnistuu	Pilotin koko aineisto saadaan digitoitua pilottijakson aikana, aineistoa ei jää digitoimatta	***Toteutuu*** *Lieviä poikkeamia* *Merkittäviä poikkeamia* *Ei toteudu*
K7. Viranomaisvalmistelun toimivuus	Aineisto on saatu viranomaisissa valmisteltua ajallaan ja Kansallisarkiston ohjeiden mukaan. Aineiston valmistelussa ei tule esiin puutteita, jotka haittaavat digitointia tai digitoidun aineiston käyttöä.	***Toteutuu*** *Lieviä poikkeamia* *Merkittäviä poikkeamia* *Ei toteudu*
K8. Kokonaisprosessin toimivuus	Digitoinnin kokonaisprosessi viranomaisvalmistelusta aineiston käyttöön toimiin: prosessin yksittäisiä vaiheita ei tarvitse merkittävästi toistaa tai korjailla. Pilotin alaprosessit ovat toteutettavissa. Pilotin kokonaisaikataulu toteutuu.	*Toteutuu* ***Lieviä poikkeamia*** *Merkittäviä poikkeamia* *Ei toteudu*

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Kriteeri 1: Digitoinnin laatu

Asetettu kriteeri oli hävittämiseen tähtäävän digitoinnin vaatimusmäärittelyn mukaisen kuvanlaadun sekä digitoinnin laatutason (virheettömyys) saavuttaminen pilotissa vaaditulla nopeudella.[2]

- ***Kriteeri toteutuu***

Kriteerin 1. mittareiksi asetetut vaatimukset oli määritelty pakolliseksi vaatimukseksi skanneri- ja ohjelmistokokonaisuuden hankinnan tekniseen vaatimusmäärittelyyn. Niiden toteutuminen hankitussa laite- ja ohjelmistokokoonpanossa varmistettiin osana järjestelmän käyttöönottotestausta. Järjestelmän pystytysvaiheessa laitetoimittaja teki laitteille vaatimusten mukaisen kalibroinnin. Käyttöönottohyväksynnässä vaatimusten voitiin todeta täyttyneen.

Kalibroitujen asennusmääritysten pohjalta pilotoinnin tuotannossa määriteltiin kuvanlaadun referenssiarvot, jotka tarkistettiin päivittäin. Jokaiselle laitteelle annettiin siten päiväkohtaisesti tuotantoluvat. Poikkeuksen muodosti mastoskanneri, jonka laadun varmentaminen toteutettiin vähäisten käyttömäärien vuoksi kerran tuotantoviikossa. Automaattisen kuvanlaadun mittauksissa käytettiin *Universal Test Target (UTT)* -kuvanlaatutargetteja ja *iQ-Analyzer*-ohjelmaa. Targettien kuluminen, ajautuminen vinoon tai pölyisyys olivat tyypillisimpiä tuotantolupamittausten hylkäyssyitä.

Skannereiden kuvanlaatua tarkkailtiin jatkuvasti myös silmämääräisesti skannauksen yhteydessä. Yleisin kuvanlaatua heikentävä häiriö, jota automaattinen analysointi ei tunnistanut, olivat naarmut tai raidat kuvassa. Nämä johtuivat yleensä pölystä tai roskista.

Kolmas kuvanlaadun tarkistamisvaihe toteutettiin validoinnissa, jota varten toteutettiin Kansallisarkiston omana työnä erillinen *validointisovellus*. Pilotoinnissa päädyttiin validoimaan 100 % digitoidusta aineistosta, eli jokainen kuva tarkastettiin visuaalisesti. Validoinnin avulla varmistettiin kuvatiedoston sisällöllisen informaation eheys. Pilotoinnissa määriteltiin erilliset kriteerit kuvien hylkäämiselle.

Jos validoinnissa havaittiin virheellisiä kuvia, pääsääntöisesti arkistoyksikkö skannattiin kokonaan uudestaan. Yksittäisten virheiden kohdalla oli harkinnanvaraisesti skannata

[2] https://arkisto.fi/fi/viranomaisille/Julkishallinnon-asiakirjahallinnon-ja-arkistotoimen-ohjaus/maeaeraeykset/kansallisarkiston-vaatimukset-h%C3%A4vitt%C3%A4miseen-t%C3%A4ht%C3%A4%C3%A4v%C3%A4%C3%A4n-digitointiin

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

uudestaan ainoastaan virheelliset kuvat. Noin 9 % pilotoinnin aineistosta päätyi validoinnin kautta hylätyiksi ja siten uudelleen skannattavaksi.

Validoinnin hylkäyskriteerit ja hylkäysten jakautuminen niiden kesken käyvät ilmi alla olevasta taulukosta 4.

Taulukko 4. Validoinnin hylkäyskriteerit

Kuvien hylkäyskriteerit	**Virheiden jakautuminen (%)**
Tuplasyöttö	5,28
Tietoa piilossa	61,80
Kuva väärinpäin, yli 1 % arkistoyksikköön sisältyvistä kuvista (koskee vain konekirjoitettua tekstiä)	18,94
Huono kuvanlaatu	5,90
Rikkoutunut asiakirja	0,31
Väärä sisältö	0,00
Arkki / sivu puuttuu	5,59
Muu syy	2,17
Virheet, jotka eivät johtaneet kuvien hylkäämiseen	
Tyhjä lomake, joka on väärinpäin.	
Lomakepohjasta puuttuu pieni alue esim. kulma taittunut, mutta sama tietosisältö löytyy muista kuvista.	
Kokonaan käsin kirjoitettu asiakirja, joka on väärinpäin (ei lasketa väärinpäin laskettavien otokseen).	
Skannauksessa virheellisesti skannattu, joka on huomattu skannausvaiheessa ja skannattu uudelleen, minkä seurauksena validoinnissa nähtävillä molemmat kuvat (eli huono kuva ja eheä kuva peräkkäin).	
Väärinpäin oleva lomake, joka toistuu yksikössä koko ajan, ja jonka tieto on täytetty käsin.	
Asiakirja, jossa ei ole tekstisisältöä ja joka on väärinpäin (esim. valokuva).	

Kriteeri 2: Digitoidun aineiston siirto säilytykseen

Asetettu kriteeri oli, että digitoinnissa pystytään tuottamaan hävittämiseen tähtäävän digitoinnin vaatimusmäärittelyjen mukainen siirtopaketti, joka sisältää vaaditut pitkäaikaissäilytyksen metatiedot.

- ***Kriteeri toteutuu***

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Skanneri- ja ohjelmistokokonaisuuden kilpailutuksessa edellytettiin pitkäaikaissäilytyksen vaatimien metatietojen tuottamista skannauksessa. Hankinnan hyväksyntätestauksessa varmistuttiin, että metatiedot muodostuvat digitoinnin yhteydessä.

Skannaussovellusten (Softrack Capture Suite, Omniscan 12) tuottaman vaatimusmäärittelyn mukaisen tiedostorakenteen ja metatietojen jatkuva varmistaminen toteutettiin pilotoinnissa automaattisesti Kansallisarkiston toteuttaman *siirtopalvelusovelluksen* avulla. Sovellus tarkisti vaatimusmäärittelyn mukaisen tiedostorakenteen määrällisen vastaavuuden sekä laski MD5-tarkistesummat. Tämän jälkeen siirtopalvelu ohjasi aineiston *validointisovellukseen*, joka on kuvattu kriteerin 1. yhteydessä. Tämän jälkeen aineisto siirtyi optiseen tekstintunnistukseen (OCR), jossa avoimen lähdekoodin välineillä (Tesseract) tuotettiin jokaisesta tietosisältöä sisältäneestä asiakirjasivusta ALTO 3.0 -standardin mukaiset tekstintunnistustiedostot (XML). Digitaalisten ilmentymien tallekappaleiksi siirtopalvelusovelluksessa tuotettiin kuvatiedostojen osalta JPEG (90 %, 300 ppi) ja digitointiprosessia kuvaavien metatietojen osalta MIX-skeeman mukainen XML-tiedosto kriteerin 1. yhteydessä mainittujen Kansallisarkiston hävittämiseen tähtäävän digitoinnin vaatimusmäärittelyn mukaisesti.

Menettely takasi kriteerin 2. mukaisen pitkäaikaissäilytyksen vaatimusten mukaisen metatietojen muodostumisen ja säilymisen.

Kriteeri 3: Digitoidun aineiston käyttö ja käytettävyys

Asetettu kriteeri oli, että viranomaiset ja Kansallisarkisto voivat käyttää digitoitua aineistoa tarvittaviin käyttötarkoituksiin ja aineiston käytettävyys parantuu merkittävästi.

- ***Kriteeri sisältää toistaiseksi merkittäviä poikkeamia***

Pilotoinnin tuottamien digitaalisten kuvien katselu, lataus, luettelointitietojen haku ja selaus sekä sisältötunnistetusta tekstistä haku oli suunniteltu toteutettavaksi asiakkaille Astia-käyttöliittymän kautta. Palvelun kehittämistä tehtiin Kansallisarkistossa virkatyönä erillisessä projektissa yhtäaikaisesti pilotoinnin suunnittelun ja toteutuksen kanssa.

Astia-käyttöliittymästä on tuotantoversio, jossa pääsee tarkastelemaan massadigitoitua aineistoa sen käyttörajoitukset huomioiden. Astia-käyttöliittymää ei ole kuitenkaan voitu ottaa suunnitellusti viranomaiskäyttöön johtuen kahdesta syystä. Ensinnäkin pilotin aikana selvisi, että kaikilla pilottiviranomaisilla ei ole käytössä VIRTU-tunnistautumisessa roolitiedon välitystä. VIRTU-roolitieto edellytetään Astiaan

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

kirjautumisessa. Kun ongelma tuli laajasti Kansallisarkiston tietoon, päätettiin Astian tunnistautumisratkaisua muuttaa siten, että viranomaiselta ei vaadita roolitietoa. Tämä muutostyö on edelleen kesken. Toinen ongelma havaittiin järjestelmän testauksessa, jossa suurten kuvamäärien lataaminen kerralla käyttöliittymässä aiheutti virheitä kuvien esittämisessä. Tämän ongelman korjaus on työn alla.

Huolimatta palvelun toiminnallisista puutteista viranomaiskäyttäjät on perehdytetty Astian käyttöön erillisen koulutustilaisuuden yhteydessä. Valmiudet palvelun käynnistämiseen ovat olemassa, kun keskeneräiset ratkaisut on saatu riittävän testauksen jälkeen viimeisteltyä.

Kriteeri 4: Aineiston hävittäminen on mahdollista

Asetettu kriteeri oli, että kriteerit 1–3 ovat toteutuneet niin kattavasti, että aineisto on mahdollista hävittää varoajan jälkeen.

- ***Kriteeri toteutuu – aineisto voidaan hävittää varoajan jälkeen***

Käytännössä pilotoinnin tarkoituksena oli toteuttaa hävittämiseen tähtäävän digitoinnin vaatimusten mukainen suurteho- ja muu skannaus. Tämän taustalla hävittämiseen tähtäävän digitoinnin kriteerejä määrittää arkistolaki, jonka 14 a §:n mukaan hävittäminen on mahdollista, jos se tapahtuu vaarantamatta asiakirjan tai siihen sisältyvän tiedon säilymistä, eheyttä ja autenttisuuden toteamista sekä heikentämättä asiakirjan kulttuurihistoriallista arvoa tai oikeudellista todistusvoimaa.

Tarkoituksena ei kuitenkaan ollut toteuttaa analogisten kappaleiden hävittämistä osana pilotoinnin toteutusjaksoa, vaan digitoida analoginen aineisto niin, että se on myöhemmin varoajan jälkeen hävitettävissä. Tämä edellyttää, että aineiston pitkäaikaissäilytys on varmistettu siirtämällä digitoitu aineisto SAPA-palveluun ja sitä kautta KP-PAS säilytykseen.

Pilotoinnin onnistumisen kriteereissä 1–2 edellytetyt kuvanlaadulle ja metatiedoille asetetut vaatimukset toteutuivat pilotissa digitoidun aineiston osalta. Osa hylätyistä aineistoista saatiin uudelleenskannattua vasta varsinaisen pilottijakson jälkeen. (Katso kriteeri 6. ja sitä koskevat havainnot.) Digitaalisen aineiston asiakaskäyttöä ei myöskään saatu toteutettua aikataulussa.

Kriteerin 4 vaatimukset täyttyivät validoinnissa jo hyväksytyn aineiston osalta. Yhteenvetona voidaan todeta, että aineiston hävittäminen on mahdollista sitten, kun digitaalisen aineiston asiakaskäytöstä on saatu kokemusta ja aineistot on siirretty onnistuneesti pitkäaikaissäilytysjärjestelmään.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Kriteeri 5: Pilotoinnin tuotantotavoitteiden saavuttaminen

Asetettu kriteeri oli, että pilotissa saavutetaan tavoitteen mukainen keskimääräinen tuotantonopeus 4,5 hm/päivä.

- ***Kriteeri sisältää lieviä poikkeamia***

Kriteerin mittariksi asetettu päiväkohtainen tuotantonopeus oli laskettu pilotoinnin suunnittelussa käytetyn laskennallisen aineistomäärän, 450 hyllymetriä, ja viiden kuukauden tuotanto-olettaman mukaan. Aineistomääriin tehtiin pilotoinnin yhteydessä tarkistuslaskenta, jonka seurauksena varsinaisen skannattavan aineiston määräksi tarkentui 386,17 hm (katso taulukko 2). Tämän määrän mukainen vastaava päiväkohtainen tuotantonopeus olisi 3,86 hm. Pilotoinnissa saavutettu päiväkohtainen eteneminen oli 3,91 hm. Alkuperäinen ja toteutunut tavoiteluku eivät ole täysin vertailukelpoisia, johtuen edellä mainitusta erosta aineistomäärässä sekä oletetuissa ja toteutuneissa tuotantopäivien määrässä.[3]

Kriteeri sisältää lieviä poikkeamia, koska tuotannon etenemisnopeus ei vakiintunut ja noussut kumulatiivisesti loppua kohti suunnitellulla tavalla. Keskeisesti tähän vaikuttivat aineiston vaihteleva luonne ja käytettävä laitteisto. Heinäkuussa digitoitiin nopeimmin skannattavaa aineistoa tehokkaimmalla suurtehoskannerilla, mikä näkyy luvuissa. Lisäksi tuotantovauhtiin vaikutti henkilöresurssien määrä. Arvioitujen vähimmäisresurssien määrässä ei ollut riittävästi huomioitu poissaoloja, joista johtuen laitteita ei voitu ajaa täydellä kapasiteetilla. Henkilöstöä kuormitti myös uudelleen skannattavien aineistojen suhteellisen korkea määrä (9 %), jota ei ollut huomioitu pilotoinnin aikataulusuunnittelussa. Tämä johtui siitä, että suunnitteluvaiheessa ajatuksena oli toteuttaa validointia otantaperusteisesti. Skannausvaiheen alussa päädyttiin kuitenkin 100 % validointiin, koska kuvanlaadun toteutuminen haluttiin varmistaa. Validointiin kului siten skannaukseen suunniteltua työaikaa. Validoinnista seurannut arvioitua suurempi uudelleenskannaustarve vaikutti lähinnä työn kokonaiskestoon.

Pilotoinnin etenemisnopeudella ja käytössä olleilla resursseilla pystyttäisiin digitoimaan vuodessa noin 978 hyllymetriä. On kuitenkin syytä korostaa, että nopeuteen vaikuttaa ennen kaikkea digitoitavana olevan aineiston muoto ja kunto.

Eri toteutusmallien absoluuttiset ja keskiarvotuotantonopeudet sekä datamäärät käyvät ilmi alla olevasta taulukosta 5.

[3] Etenemisnopeudessa on huomioitu todelliset tuotantopäivät.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Taulukko 5. Pilotoinnin tuotantomäärät ja -nopeudet sekä datamäärät

Toteutusmalli	**Linjasto**	**Huhti-toukokuu**		**Kesäkuu**		**Heinäkuu**		**Elokuu**		**Syyskuu**		**Lokakuu**		**Yhteensä**	
		Päivä (ka)	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**	**Päivä (ka)**	**Yhteensä (hm)**
Toteutusmalli 1 (suurtehoskanneri)	**Valmistelu** (etukäteen)	2,27	61,28	1,13	18,00	1,83	42,20	1,34	28,04	1,92	34,54	1,27	24,08	1,68	208,14
	Skannaus (erillinen)	2,61	18,30	3,81	39,96	3,90	52,70	3,54	33,62	2,98	25,33	2,95	26,51	3,39	196,42
Toteutusmalli 2 (5 dokumenttiskanneria)	**Valmistelu** (etukäteen)	1,19	32,20	0,38	6,14	0,02	0,39	0,05	0,99	0,04	0,70	0,34	6,46	0,38	46,88
	Skannaus (erillinen)	1,12	30,18	0,55	8,81	0,02	0,46	0,01	1,14	0,25	4,57	0,56	10,64	0,45	55,80
Toteutusmalli 3 (5 suurtehoskanneria)	**Valmistelu ja skannaus** (yhtäaikainen)	0,18	4,96	1,05	16,87	1,80	41,31	1,80	37,88	2,11	38,06	1,52	28,83	1,35	167,91
Toteutusmalli 4 (mastoskanneri)	**Valmistelu**	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Skannaus (yhtäaikainen)		0,00	0,02	0,11	0,04	0,80	0,04	0,68	0,01	0,10	0,00	0,00	0,02	1,69
Yhteensä	**Valmistelu**	3,46	93,48	1,51	24,14	1,85	42,59	1,38	29,03	8,81	35,24	1,61	30,54	2,06	255,02
	Skannaus (skannaus, valmistelu+skannaus)	4,45	53,44	4,11	65,75	4,14	95,27	3,49	73,32	3,78	68,06	3,47	65,98	3,91	421,82
Määrät (kpl, Gt)		Pv	Kk	Pv	Kk	Pv	Kk	Pv	Kk	Pv	Kk	Pv	Kk	Ka pv	Touko-lokakuu
Kuvamäärä (kpl)		23 808	642 809	61 294	980 702	64 607	1 485 959	46 425	974 935	46 545	837 804	62 472	1 436 851	51 283	6 359 060
Datamäärä (Gt)		50,29	1 357,80	148,29	2 372,61	152,88	3 516,23	113,25	2 378,26	95,97	1 727,46	161,93	3 076,68	116,36	14 429,04

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Kriteeri 6: Pilottiaineiston digitointi onnistuu

Asetettu kriteeri oli, että pilotin koko aineisto saadaan digitoitua pilottijakson aikana ja aineistoa ei jää digitoimatta.

- ***Kriteeri toteutuu***

Kaikki pilotoinnin aineisto saatiin digitoitua pilotoinnin taka-aikarajaan mennessä vähintään kerran. Tämä työ valmistui 28.10.2019.

Kaikki erilaiset aineistotyypit onnistuttiin valmistelemaan ja digitoimaan eli mitään aineistoa ei tarvinnut jättää digitoimatta sen ominaisuuksien vuoksi. Sidosaineistoa lukuun ottamatta digitointinopeudet vastasivat pääpiirteittäin suunniteltua.

Lisäksi osassa digitointia ilmeni tekninen ongelma, jonka seurauksena osasta digitoituja yksiköitä puuttui yksittäisiä kuvia. Virhe saatiin paikannettua yhden skannerityypin ajurin toimintaan, ja virheen tunnistamiseen saatiin työkalu. Yksittäisten virheiden esiintyminen, erityisesti ilman digitaalisen aineiston käyttöliittymää, olisi kuitenkin edellyttänyt osan aineistosta käymistä läpi manuaalisesti, mitä ei pidetty työekonomisesti kannattavana. Parhaaksi vaihtoehdoksi katsottiin osan aineistosta skannaaminen uudelleen. Tämä uudelleenskannaus saatiin valmiiksi 18.11.2019.

Kriteeri 7. Viranomaisvalmistelun toimivuus

Kriteerin mittarina oli, että aineisto on saatu viranomaisissa valmisteltua ajallaan ja Kansallisarkiston ohjeiden mukaan. Aineiston valmistelussa ei tule esiin puutteita, jotka haittaavat digitointia tai digitoidun aineiston käyttöä.

- ***Kriteeri toteutuu***

Kansallisarkisto laati pilotoitavan aineiston siirtokuntoon saattamista koskevan ohjeen (Siirto-opas) sekä sen keskeisintä tehtävää, aineiston luettelointia AHAA-palveluun, ohjaavan AHAA-luetteloinnin viranomaisohjeen. AHAA:n käytöstä järjestettiin lisäksi koulutustilaisuus. Lisäksi viranomaisten luettelointityötä ohjattiin viranomaisten järjestelmään syöttämiä tietoja seuraamalla sekä antamalla lisäohjeita virastokäynneillä, puhelimitse ja tätä varten perustetun palautesähköpostin kautta.

Ennen viranomaisvalmistelun käynnistymistä aineiston vastaavuus lähtötietoihin todennettiin paikan päällä virastokäynneillä. Kansallisarkisto vastasi myös aineiston kuljettamisesta digitointiin sekä siihen liittyneistä kustannuksista. Viranomaiset antoivat myönteistä palautetta pilotoinnin ohjeistuksesta.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Viranomaisvalmistelu onnistui hyvin, eikä aineistosta tai sen AHAA-palveluun viedyistä luettelotiedoista löytynyt virheitä. Viranomaisten sitoutuminen ja asiantuntemus olivat keskeinen tekijä siinä, että aineisto saatiin ajoissa ja ilman puutteita digitointiin.

Kriteeri 8. Kokonaisprosessin toimivuus

Kriteerin mittarina oli, että digitoinnin kokonaisprosessia tai sen yksittäisiä vaiheita viranomaisvalmistelusta aineiston käyttöön toimii, sitä ei tarvitse merkittävästi toistaa tai korjailla, pilotin alaprosessit ovat toteutettavissa ja pilotin kokonaisaikataulu toteutuu.

- ***Kriteeri sisältää lieviä poikkeamia***

Pilotti edellytti laaja-alaista kokonaisuuden hallintaa, jossa viranomaisten valmistelutyö, aineistojen kuljetukset sekä digitoinnin tuotanto piti koordinoida, jotta aineistot saatiin digitointiin oikea-aikaisesti ja oikein valmisteltuna. Pilotissa tähän kokonaisuuteen kuului lisäksi samanaikaisesti koko tuotantoympäristön rakentaminen alusta lähtien sekä henkilöstön rekrytointi ja perehdyttäminen tehtäviinsä.

Kokonaisuutena katsoen pilotoinnin aikataulu ja varsinkin sen tuotantovaihe pitivät jopa odotettua paremmin. Viranomaisvalmistelu syksyllä 2018 sekä skanneri- ja ohjelmistohankinnan ja palvelinympäristön toteuttaminen keväällä 2019 tehtiin erittäin kireässä aikataulussa. Varsinkin jälkimmäinen asetti suuria riskejä järjestelmän pystyttämiselle ja käyttöönottohyväksynnälle suunnitellussa aikataulussa. Digitointivaihe käynnistyi tästä huolimatta alle kaksi viikkoa suunniteltua ajankohtaa myöhemmin. Aineiston manuaalinen valmistelu skannaukseen saatiin käyntiin jo huhtikuun aikana.

Laitteiston toimivuus oli hyvä, aineiston valmistelu ja skannaus etenivät pääpiirteittäin suunnitellulla tavalla ja aikataulussa. Tuotantosuunnittelu toimi tehokkaasti ja joustavasti, laitteistoa ja henkilöstöresursseja käytettiin tehokkaasti, eikä merkittäviä pullonkauloja tuotantoon muodostunut. Merkittävää on osana suunnittelua tehty huolellinen perehtyminen aineistoon, jonka ansiosta viranomaisvalmistelu sujui odotettua nopeammin, aineistoa koskeva tuotantosuunnittelu eli aineistojen ohjaaminen eri valmistelutapojen mukaisille linjastoille sujui pääpiirteittäin suunnitellulla tavalla. Pilotoinnin aliprosessit toimivat pääsääntöisesti suunnitellulla tavalla eikä niihin tarvinnut tehdä merkittäviä muutoksia. Pienimuotoista kehittämistä ja hienosäätöä tehtiin tarvittaessa prosesseihin.

Palvelinympäristössä olleet tekniset häiriöt ja niiden aiheuttamat tuotantokatkokset hidastivat kuitenkin jonkin verran digitointia. Huomioiden myös validoinnin suhteellisen

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

korkea hylkäysprosentti, yksittäisestä ohjelmistovirheestä johtuneesta puuttuvista kuvista seurannut uudelleenskannaustarve sekä digitaalisten kuvien käyttöliittymän puutteen aiheuttamat haasteet, on kokonaisuutena tarkastellen aliprosesseissa ja varsinkin poikkeustilanteista palautumisen käytännöissä hiottavaa sujuvasti etenevän kokonaisprosessin saavuttamiseksi.

12.2 Muut tulokset ja havainnot

Aineiston luetteloinnin onnistuminen

Aineiston luetteloiminen osana viranomaisvalmistelua onnistui hyvin. Tilanne oli haastava, koska AHAA-palvelu otettiin tuotantokäyttöön samaan aikaan viranomaisvalmistelun käynnistymisen kanssa. Osa aineistoista syötettiin ohjelmaan manuaalisesti, osa vietiin rajapintasiirtoina. Jälkimmäisessä tapauksessa siirto jouduttiin suunnittelemaan tapauskohtaisesti, koska valmiita vaatimusmäärittelyjä ei ollut käytettävissä. Huomioiden palvelun tuoreus, luettelointi onnistui erinomaisesti.

AHAA-palvelussa ilmeni joitakin virheitä aineiston käyttörajoitusmerkintöjen esittämisessä, vaikka ne oli luetteloinnin yhteydessä syötetty oikein, joten palvelua tulee vielä kehittää. Pilotoinnin jälkeiseen digitointiin tarkoituksena on ottaa käyttöön valmisteilla oleva AHAA-palvelun viranomaisversio, jonka käyttö on nykyistä yksinkertaisempaa.

Kansallisarkiston viranomaisohjeistuksen toimivuus ja selkeys

Viranomaisvalmisteluun laadittu ohjeistus toimi hyvin. Siitä pilotoinnin digitointivaiheessa eli valmistelun päätyttyä laadittiin palautekysely, jonka pohjalta ohjeistusta on edelleen tarkoitus kehittää. Kyselyn mukaan erityisesti virastokäynteinä ja sähköpostitse toteutettua ohjausta pidettiin hyvänä. Keskeisenä kehittämiskohteena nähtiin AHAA-palvelun käytettävyys. Kokonaisuutena ohjeistusta ja ohjausta koskevien arvioiden keskiarvo oli 4 (Asteikolla 1 – 5, Huono – Hyvä).

Käytettyjen ohjelmistojen laajuus ja toimivuus

Kokonaisuutena pilotoinnissa käytetyt skannausohjelmistot toimivat hyvin. Ohjelmistoilla pystyttiin tuottamaan luotettavasti vaatimusmäärittelyn mukaisia (kuvanlaadun ja pitkäaikaissäilytyksen kriteerit) kuvia. Ohjelmistokoodien

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

pienimuotoinen muokkaus tuotannon aikana onnistui joustavasti ja hyvässä yhteistyössä alihankkijan kanssa pääosin etäyhteydellä.

Suurimmat haasteet liittyivät enemmän virheisiin, jotka eivät sinällään tuottaneet alle kriteerien jäävää kuvanlaatua, vaan joiden johdosta osa kuvista jäi kokonaan muodostumatta tai niiden lukusuunta oli virheellinen sisällöntunnistuksen kannalta. Virheet liittyvät pelkästään dokumenttiskannereihin. Automaattisen käännön luotettavuus vaikutti ennen kaikkea niin, että aineistoja jouduttiin valmistelemaan manuaalisesti oikeaan lukusuuntaan skannereille, mikä hidasti kokonaistyöaikaa. Yksittäisten puuttuvien kuvien tunnistaminen jälkikäteen vaatii taas joko aikaa vieviä tarkistuksia tai tietyn aikavälin aineiston skannaamista kokonaan uudelleen.

Kuvat käännettiin päälukusuuntaan kaikilla laitemalleilla, osalla automaattisesti, osalla manuaalisen valmistelun yhteydessä. Yhdellä laitetyypillä automaattinen kääntö tapahtui jälkiprosessina ja johti kuvan kertaluonteiseen uudelleen pakkaamiseen.

Haasteita oli jonkin verran ollut myös palvelinympäristön toiminnassa. Nämä johtivat pahimmillaan jopa kokonaisten päivien mittaisiin tuotannon keskeytyksiin digitoinnissa, koska kuvia ei voitu viedä siirtopalveluun. Levyjärjestelmissä havaittiin myös virhe, jonka johdosta kaikki tiedostotyypit eivät siirtyneet siirtopalvelussa eteenpäin, mikä voi pahimmillaan johtaa uudelleen skannauksiin.

Suuri osa edellä kuvatuista virheistä olisi luultavasti pystytty estämään, jos ennen digitointia olisi tehty pidempi testausjakso ja jos digitoinnin tuotantohenkilöstöllä olisi suora pääsy tarkastelemaan digitoitua aineistoa validointivaiheen jälkeen.

Tiedonsiirto- ja tietojenkäsittelykapasiteetti

Pilotoinnissa käytettiin 10 gbit tiedonsiirtokapasiteettia, joka ei osoittanut minkäänlaista hidastetta datansiirrossa. Täsmällisiä siirtonopeuksia olisi saatavissa ainoastaan verkontarjoajalta.

Infrastruktuuri toteutettiin virtuaalipalvelinratkaisuna, joka yleisesti ottaen oli onnistunut ratkaisu järjestelmän toimivuuden kannalta. Data siirtyi käytännössä skannereilta suoraan virtuaalipalvelinympäristöön. Palvelinten määrää ja levytilaa on kasvatettu vähitellen tarpeiden mukaan, kapasiteetin riittävyyttä tarkastellen. Arvio levytilatarpeesta (23–25 TT) on riittänyt hyvin toteutumaan (14,4 TT). OCR:n osalta päädyttiin 12 palvelimen kokonaisuuteen. Nykyinen ratkaisu skaalautuu myös suuremmille määrille. Suurimmat haasteet ovat liittyneet seurantaan ja hallintaan, joka olisi edellyttänyt nyt toteutettua tarkempaa palvelukuvausta ja roolien määrittämistä.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Odotusten ja suunnitelmien mukaisesti tasalaatuisen A4-aineiston digitointi oli toteutettavissa kaikkein nopeimmin. Aineiston digitointiajassa ei tule seurata vain skannausnopeuksia, vaan keskittyä enemmän koko digitointiprosessin kattavaan työaikaan. Suurin osa ajasta kuluu aineiston valmisteluun, varsinkin liittimien, tarra-arkkien ja teippien irrottamiseen, kun taas keskeisten skannerityyppien väliset nopeuserot eivät poikkea toisistaan suuresti. Näin on erityisesti siksi, että mitään skannerilaitetta ei voi ajaa täydellä nopeudella, jolloin nopeimpien laitteiden nopeushyöty muihin nähden tasoittuu. Suurteho- ja dokumenttiskannereiden välillä tehtävässä arvioinnissa tulevissa hankintapäätöksissä ratkaisevaa on, miten hyvin aineistot pystytään etukäteen käymään läpi, tunnistamaan ja luokittelemaan oikeille laitetyypeille.

Pilotoinnin aikana skannattiin onnistuneesti aineiston joukossa olevia niin sanottuja ”erotettuja arkkeja”, jotka ominaisuuksiensa perusteella jouduttiin väliaikaisesti erottamaan muun aineiston joukosta ja skannaamaan erikseen erikoisskannerilla (mastoskannerilla). Pilotointiin pyrittiin saamaan aineistoa, jossa näitä erikoisaineistoskannerille ohjautuvia arkkeja olisi runsaasti, mutta viranomaisvalmistelun edetessä suurin osa tästä aineistosta seuloutui määräajan säilytettävinä pois. Määrän vähyyden vuoksi pilotoinnin perusteella ei voida tehdä kattavia johtopäätöksiä siitä, miten paljon erikoiskäsiteltävää aineistoa digitoitava kokonaisuus voi enimmillään sisältää, ennen kuin työhön tarvittavan ajan vuoksi se rajautuu massadigitoinnin ulkopuolelle.

Massadigitoinnin ja siten myös pilotoinnin ulkopuolelle oli rajattu aineistotyypit, joiden käsittely edellyttää runsaasti poikkeavia, erityisiä tai varovaisuutta vaativia toimenpiteitä. Pilottiin sisältyi 60 sidosta, jotka digitoitiin joko mastoskannerilla avaten digitoinnin yhteydessä tarvittaessa sidosrakennetta, tai sidokset purettiin ja digitoitiin irtoarkkeina dokumenttiskannerilla. Pilotin tulosten mukaan sidokset, jotka on purettava kokonaan käsin sekä sidokset, joiden purkamisessa ei juurikaan voida hyödyntää sähköleikkuria, eivät sovellu massadigitointiin työn hitauden vuoksi. Jatkossa tulee arvioida, pitäisikö sidosaineisto rajata massadigitoinnin ulkopuolelle. Jos sidos on pääosin purettavissa sähköleikkurilla, tai jos purkamisessa hyväksytään marginaalin leikkautumisessa tapahtuvaa mahdollista tiedonmenetystä, voidaan sidokset sisällyttää massadigitointiin.

Sidosten skannaamisessa ratkaisevaa on myös, käytetäänkö perinteistä mastoskanneria, jonka osalta laitteiden nopeuksissa on eroja, vai räätälöidympiä laitteita, jotka soveltuvat paremmin tällaiseen erikoisaineistoon. Näiden skannereiden käsittelyyn ja

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

aineiston liikutteluun niillä kuluu huomattavasti aikaa, minkä merkitys korostuu suuriin määriin skaalatessa.

13 Johtopäätökset ja jatkotoimenpiteet

Massadigitoinnin pilotointi toteutettiin vuosina 2018 – 2019 noin puolentoista vuoden mittaisena suunnittelun ja tuotannon kokonaisuutena. Digitointi- eli tuotantovaiheen kesto oli noin kuusi kuukautta.

Pilotointi täytti ensisijaisen tavoitteensa eli osoitti, että suunniteltu massadigitoinnin prosessi ja toteutusmalli toimivat ja digitointia voidaan tehdä suunnitellulla tehokkuudella. Pilotoinnissa toteutettu digitointi oli hävittämiseen tähtäävien kriteerien mukainen, mikä mahdollistaa analogisten asiakirjojen hävittämisen varoajan jälkeen.

Pilotti osoitti, että massadigitoinnin kokonaisuuden koordinointi toimi käytännössä. Viranomaisohjauksella varmistettiin, että aineistot siirtyivät oikea-aikaisesti ja oikein valmisteluna digitointiin. Koko prosessia ja aineiston siirtymistä vaiheesta toiseen on voitu seurata luotettavasti, vaikka käytössä ei ole vielä kaikkia vaiheita kattavaa tuotannonohjauksen järjestelmää.

Massadigitoinnin pilotti täytti sille asetetut keskeiset kriteerit ja osoitti, että massadigitoinnille suunniteltu toteutustapa toimii käytännössä. Pilotissa onnistuttiin kokonaisuuden hallinnassa, digitointituotannon tehokkuudessa sekä vaaditun laatutason saavuttamisessa. Pilotin perusteella massadigitoinnissa voidaan edetä tuotantoon nykyisen toteutustavan pohjalta.

Pilotointi pystyttiin toteuttamaan pääpiirteittäin suunnitellulla tavalla, vaikka suunnittelu ja toteutus vedettiin tiettyjen osa-alueiden osalta läpi todella kiireellisellä aikataululla. Aliprosessien prosessikuvaukset ja arkkitehtuurikuvaus tulee pilotoinnin päätyttyä päivittää ja arvioida, mitä osia näistä tulee kehittää.

Pilotoinnin päiväkohtainen eteneminen vastasi melko hyvin sille asetettuja tavoitteita. Tulevassa digitoinnin suunnittelussa ja resursoinnissa on huomioitava tarkemmin poissaolojen vaikutus ja todellisen digitointiin käytettävän työajan laskenta resurssien asettamisessa. Tähän tarvittavat vertailuluvut saadaan pilotoinnin tuloksista.

Pilotointi osoitti, että aineiston tuntemus eli tiekarttasuunnittelu maksaa itsensä takaisin digitointivaiheessa. Viranomaisohjaus ja viranomaisten näkemysten ja palautteen kuuleminen ovat kivijalka, jolle toimiva massadigitoinnin toteutus rakennetaan. Mitä paremmin aineistot tunnetaan, sitä mutkattomampaa on viranomaisvalmistelun ohjaaminen ja digitoinnin suunnittelu ja toteuttaminen.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi

Pilotoinnin jälkeen keskeinen kehittämiskohde on aineistoluokittelun tarkentaminen pilotoinnin havaintojen perusteella. Tällä on myös suoria vaikutuksia kustannuslaskennan luotettavuuteen. Aineistoluokkia tulee kehittää, ja viime kädessä yksi tavoite on linjata, mitkä aineistot eivät fyysisten ominaisuuksien ja niistä seuraavan valmistelun ja skannauksen hitauden vuoksi sovellu digitoitavaksi massadigitointiin.

Pilotoinnissa havaittiin seuraavia kehittämiskohteita:

1. Pilotoinnissa osa hankittiin ulkoisina hankintoina, osa CSC Oy:ltä ja osa Kansallisarkiston omana työnä. Kokonaisuus oli toimiva, mutta edellytti mittavaa suunnittelua. Jos digitointi toteutetaan eri toimittajien yhteisprojektina, tulee sen pohjautua huolellisiin palvelukuvauksiin, roolien ja vastuiden määrittämiseen sekä alustoihin, joilla dokumentaatiota ja tietoa jaetaan, ja päätöksiin siitä, ketkä koordinoivat kokonaisuutta.
2. Digitointituotannon tulee perustua mahdollisimman pitkälle automatisoituihin toimintoihin. Kun pilotoinnissa ja sitä ennen *Proof of Concept* -testauksessa on keskitytty skanneri- ja ohjelmistoratkaisuihin sekä aineiston luokitteluun, tulee kehittämisessä niiden rinnalle jatkossa tuoda tuotannonohjausjärjestelmän tiiviimpi integrointi eri järjestelmiin sekä tuotannonohjauksen ulottaminen massadigitoinnin prosessin kaikkiin vaiheisiin.
3. Pilotointi oli alun alkaen tarkoituksena toteuttaa niin, että ensin luodaan tuotannon kehittämisympäristö tai laboratorio, jonka päälle varsinainen tuotanto pystytetään. Aikataulusyistä pilotoinnissa toteutettiin tuotantoympäristö. Tulevaisuudessa tuotannon tulee aina nojata sen rinnalla käytössä olevaan täysimittaiseen testausympäristöön, joka ulottuu ensimmäisestä viimeiseen digitoinnin aliprosessiin. Keskeinen kehittämiskohde on täysimittaisen pilotoinnin testausympäristön rakentaminen.
4. Pilotoinnin digitointityön ohjeistus koottiin yhdeksi kattavaksi ohjepaketiksi. Kokonaisvaltaista dokumentaatiota tulee jatkossa edistää edelleen niin, että se kattaa paremmin myös päivittäisen tuotannonohjauksen, ja tulos- ja henkilöohjauksen toimivuus sekä ajanmukainen tilannekuva saadaan varmistettua kaikissa vaiheissa ja helposti esiin. Kaikessa ohjeistuksessa tulee tuoda huolellisesti esiin roolitus, tehtäväkuvat ja vastuut. Tuotannon keskeisiä tavoitteita koskevat mittarit tulee määritellä etukäteen ja ohjelmistoihin tai niitä tukeviin sovelluksiin tulee rakentaa raportointiominaisuudet, joista keskeiset tunnusluvut saadaan kerättyä automaattisina eräajoina sovituin määräajoin.

Kansallisarkisto Rauhankatu 17 PL 258, 00171 Helsinki Puh. Tel. 029 533 7000 kirjaamo@arkisto.fi
Riksarkivet Fredsgatan 17 PB 258, 00171 Helsingfors Fax (09) 176 302 http://www.arkisto.fi