

Digitoidun aineiston validointityökalun käyttöohje

Sisällys

Mikä on digitoidun aineiston validointityökalu?	2
Validointityökalun perustoiminta	3
Validointityökalun käyttöohje	5
Tarkistettavan kohteen lataaminen työkaluun	5
Tarkistusraportin tutkiminen ja tulostaminen	7
Esimerkkiraportteja	10

Mikä on digitoidun aineiston validointityökalu?

Digitoidun aineiston validointityökalu on tarkoitettu viranomaisille ja palveluntuottajille, jotka digitoivat arkistolain (831/1994) nojalla arkistoitavaksi määrättyjä asiakirjoja.

Työkalun avulla voit tarkistaa, vastaavatko digitoitujen tiedostojen ja muiden dataobjektien metatiedot *Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin* (KA/3357/07.01.02.04.02/2019) -määrityksen vaatimuksia.

Voit tarkistaa työkalulla seuraavia tietoja:

- tallekappale (bittikarttakuva TIFF tai JPEG)
 - KA/3357/07.01.02.04.02/2019-määrityksessä esitetyt bittikarttakuvan metatiedot
- tallekappaleen (bittikarttakuva) syntyhistoriaan liittyvät tekniset metatiedot
 - MIX-metatietoskeeman mukaisesti, versio 2.0
- bittikarttakuvaan liittyvä tekstintunnistustieto (OCR)
 - formaatissa Analyzed Layout and Text Object (ALTO XML), versio 3.0 tai uudempi
- tiedoista muodostetun siirtopaketin rakenteeseen liittyvät vaatimukset

Mikäli tarkistetun kohteen metatiedot eivät vastaa Kansallisarkiston vaatimuksia, on asiakirja digitoitava uudestaan määritysten edellyttämällä tavalla.

Validointityökalun perustoiminta

Digitointiprosessissa tuotettavien kohteiden pitää sisältää koneluettavassa muodossa kaikki Kansallisarkiston vaatimuksissa pakollisiksi määritellyt metatiedot. Metatiedon arvo voi olla:

- dataobjektikohtainen (esimerkiksi kuvatiedoston koko)
- tai sen tulee vastata Kansallisarkiston vaatimuksissa ilmoitettua arvoa (esim. väritila = RGB)

Työkalu tarkistaa, että tarkistettavaksi ladatun kohteen kaikissa pakollisissa metatiedoissa on arvo. Mikäli arvon tulee vastata Kansallisarkiston vaatimuksissa ilmoitettua arvoa, työkalu tarkistaa, että metatietokentän arvo on oikea.

Työkalu tuottaa käsitellyn kohteen metatiedoista erilliselle tulossivulle raportin taulukkomuodossa. Työkalu ei tallenna tietoja, vaan kaikki tiedot poistuvat palvelusta 12 tuntia viimeisen tiedoston lataamisesta. Voit halutessasi tulostaa ja tallentaa istunnon aikana muodostuneet tarkistusraportit itsellesi PDF-muodossa.

Löydät validointityökalun käyttöohjeen, Kansallisarkiston digitointivaatimukset ja Esimerkkiarvot kaikkien sivujen alalaidassa näkyvästä *Ohjeet ja muut dokumentit* -osiosta (*Kuva 1*).

Työkaluun ei tarvitse kirjautua.

Älä vie työkaluun turvallisuusluokiteltavia asiakirjoja tai arkaluonteisia asiakirjoja.

Ongelmatilanteissa saat apua osoitteesta viranomaisohjaus@arkisto.fi

Validointityökalu

Tervetuloa Kansallisarkiston validointityökaluun. Työkalulla voit tarkistaa, vastaavatko digitoitujen tiedostojen ja muiden dataobjektien metatiedot Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin (KA/3357/07.01.02.04.02/2019) -määrityksen vaatimuksia.

Voit validoida tallekappalekuvan, tuotetun tekstitiedoston (OCR) tai siirtopaketin rakenteen. Validointityökalu ei tallenna tietoja, joten muista ladata itsellesi tarkistusraportit validoinnin päätteeksi.

Tämä sivusto toimii uusimmilla Edge-, Chrome- ja Firefox-selainversioilla.

Kuvatiedosto ja kuvan metatiedot

[Bittikarttakuva](#)
[MIX-metatiedot](#)

Tunnistettu teksti

[AltoXML](#)

Siirtopaketti

[Siirtopaketti](#)

Omat raporttini

[Seuraa vietyjen tiedostojen raportteja](#)

Kansallisarkisto © 2020-2021

Ohjeet ja muut dokumentit

Validointityökalun käyttöohje
Kansallisarkiston vaatimukset hävittämiseen
tähtäävään digitointiin
Esimerkkiarvot
Saavutettavuusseloste

Yhteystiedot

viranomaisohjaus@arkisto.fi

Kuva 1. Digitoidun aineiston validointityökalun etusivu. Pääset käyttöohjeeseen, Kansallisarkiston digitointivaatimuksiin ja Esimerkkiarvoihin kaikkien sivujen alalaidassa näkyvästä Ohjeet ja muut dokumentit -osiosta.

Validointityökalun käyttöohje

Tarkistettavan kohteen lataaminen työkaluun

Toimi näin:

1. Valitse työkalun etusivulla, millaisen dataobjektin haluat ladata tarkistettavaksi.
2. Lataussivu avautuu.
3. Voit nyt ladata tarkistettavat tiedostot/kohteet yksitellen tai useamman tiedoston kerralla. Huomaathan, että bittikarttakuvien kohdalla on ennen lataamista lisäksi valittava tarkistettavan tiedoston tiedostomuoto (TIFF- tai JPEG-tiedosto).
4. Voit raahata ja pudottaa tarkistettavat tiedostot valkoiselle latausalueelle.
5. Vaihtoehtoisesti voit klikata latausaluetta hiirellä ja valita tarkistettavat tiedostot selaamalla tiedostoitasi.




Kuva 2. Bittikarttakuvan lataus lataussivulla.

6. Työkalu ilmoittaa tallennuksen onnistumisesta tiedoston päällä näkyvällä vihreällä merkillä (Kuva 3).
7. Onnistuneen latauksen jälkeen pääset työkalun **tarkistusraporttisivulle/omat raporttini - sivulle** latausruudun alla näkyvästä **Tänne**-linkistä (Kuva 3)

Onko tallekappaleesi TIFF vai JPEG? Voit tuoda useamman tiedoston yhtä aikaa.

TIFF

JPEG



Tiedosto "kuva.jpg" on tallennettu palvelimelle onnistuneesti. Validoinnin tulokset tulevat [Tänne](#) kun tarkistukset ovat valmiita.

Kuva 3. Vihreä merkki tiedoston päällä kertoo, että tallennus on onnistunut. Latausalueen alla olevasta Tänne-linkistä pääset tarkistusraporttisivulle.

Tarkistusraportin tutkiminen ja tulostaminen

Tarkistusraporttisivulla voit tutkia ja tallentaa raportteja 12 tuntia viimeisen tiedoston lataamisesta.

Toimi näin:

1. Valitse tarkistusraporttisivulla tiedosto, jonka tarkistusraporttia haluat katsella (*Kuva 4*).



Kuva 4. Omat raporttini -näkyvä, jossa näkyy kaikki ladattujen tiedoston valmistuneet tarkistusraportit.

2. Raportti avautuu omalle sivulleen.
3. Huomaathan, että tiedoston onnistuneen lataamisen jälkeen raportti ei aina näy raporttisivulla heti (*Kuva 5*). Voit raportin muodostumista odottaessasi palata lataamaan muita tiedostoja (*Validointityökalu*-linkki vasemman yläaidan murupolussa) tai tutkia muita raportteja *Omat raporttini* -sivulta (*Omat raporttini* -linkki vasemman yläaidassa tai *Takaisin omiin latauksiin tästä* -linkki tekstin alla).



Kuva 5. Omat raporttini -näkyvä, kun ladatun tiedoston tarkistusraportti ei ole vielä valmistunut.

4. Kun raportti on valmis, se kertoo, onko tiedostosta luettu arvo vaadittujen arvojen mukainen (**OK**) vai virheellinen (**VIRHE**) (Kuva 6).
5. Voit tulostaa tai tallentaa raportit yksittäin PDF-muodossa vasemman yläkulman *Tulosta raportti PDF* -linkistä (Kuva 6).

Tehtävä tarkistus: metatiedot

[Tulosta raportti PDF](#)

Elementti	Vaadittu arvo	Metatieto skeema	Metatietokenttä	Luettu arvo	#
Tiedosto				kuva.jpg	OK
Formaatti	image/jpeg tai image/tiff		MIME Type	image/jpeg	OK
Versio	1.00 tai 1.01 tai 1.02		JFIF Version	1.01	OK
Päiväys		Exif:FD	DateTimeOriginal (36867)	Arvo puuttuu	Virhe
Väritila	RGB	Exif:Image	PhotometricInterpretation (262)	Väärä arvo. Pitää olla RGB	Virhe
ICC -profiili	sRGB, eciRGB v2, ProPhoto RGB, Adobe RGB (1998)	ICC	profileDescription	Väärä arvo: 0	Virhe

Kuva 6. JPEG-kuvatiedoston tarkistusraportti. Voit tulostaa raportin vasemmasta yläkulmasta "Tulosta raportti PDF" -kohdasta.

6. Voit aina palata kaikkiin istunnon aikana muodostuneisiin omiin tarkistusraportteihisi validointityökalun etusivulta (Kuva 7)

Validointityökalu

Tervetuloa Kansallisarkiston validointityökaluun. Työkalulla voit tarkistaa, vastaavatko digitoitujen tiedostojen ja muiden dataobjektien metatiedot Kansallisarkiston vaatimukset hävittämiseen tähtäävään digitointiin (KA/3357/07.01.02.04.02/2019) -määrityksen vaatimuksia.

Voit validoida tallekappalekuvan, tuotetun tekstitiedoston (OCR) tai siirtopaketin rakenteen. Validointityökalu ei tallenna tietoja, joten muista ladata itsellesi tarkistusraportit validoinnin päätteeksi.

Tämä sivusto toimii uusimmilla Edge-, Chrome- ja Firefox-selainversioilla.

Kuvatiedosto ja kuvan metatiedot

Tunnistettu teksti

Siirtopaketti

Omat raporttini

Kuva 7. Pääset omiin tarkistus raportteihisi aina myös validointityökalun etusivulta.

Esimerkkiraportteja

Kuva 8. MIX-metatietojen raportti

Validointityökalu

Tehtävä tarkistus: mixTiedosto

Tulosta raportti [PDF](#)

Tarkistus	Tietue	Vastaus	#
Skeematarkistus	mix-2-0.xsd	On skeeman mukainen	OK
Tiedosto		0004.xml	OK
Asiakirjan yksilöivä ID	objectIdentifierType	Arvo puuttuu	Virhe
Identifioiva merkkisarja	objectIdentifierValue	Arvo puuttuu	Virhe
Tiedoston koko	fileSize	653002	OK
Formaatti	formatName	image/jpeg	OK
Formaattiversio	formatVersion	1.02	OK

Kuva 9. Alto-tunnistettu tekstin raportti

Validointityökalu

Tehtävä tarkistus: tunnistettuTeksti

Tulosta raportti [PDF](#)

Tiedosto 0005alto.xml täyttää alto-4-0.xsd skeeman.

Kansallisarkisto © 2020-2021

Ohjeet ja muut dokumentit

Validointityökalun käyttöohje
Kansallisarkiston vaatimukset hävittämiseen
tähtäävään digitointiin
Esimerkkiarvot
Saavutettavuusseloste

Yhteystiedot

viranomaisohjaus@arkisto.fi

Kuva 10. Siirtopaketin raportti

Tehtävä tarkistus: siirtopaketti

Tulosta raportti [PDF](#)

Tarkistus	Tietue	Vastaus	Pakollisuus	#
Tiedostonnimi		1399431832.tar	Pakollinen	<input type="button" value="OK"/>
Tunnus (Identifier)	Numeerinen	1399431832		<input type="button" value="OK"/>
Aineisto	Käsittelyssä	Aineisto on generoitu käsittelyyn		<input type="button" value="OK"/>
Hakemistorakenne	master	Hakemisto löytyy	Pakollinen	<input type="button" value="OK"/>
Hakemistorakenne	ocr	Hakemisto löytyy	Ei pakollinen	<input type="button" value="OK"/>
Hakemistorakenne	mix	Hakemisto löytyy	Pakollinen	<input type="button" value="OK"/>
Tiedostomäärät	Hakemisto: master	7		<input type="button" value="OK"/>
Tiedostomäärät	Hakemisto: ocr	4		<input type="button" value="OK"/>
Tiedostomäärät	Hakemisto: mix	7		<input type="button" value="OK"/>
Tiedostomäärät	Alto ja Master -tiedostot eivät täsmää	4 / 7(M)	Ei pakollinen	<input type="button" value="OK"/>