

Specificatiedocument

Aanlevering gedigitaliseerd materiaal in het kader van Metamorfoze Archieven en Collecties aan de Koninklijke Bibliotheek

Opgesteld door de Koninklijke Bibliotheek december 2014
Martina Hoffmann
Martijn van der Kaaij
Versie 1.0
Definitief

Inleiding

De afdeling digitalisering van de Koninklijke Bibliotheek is binnen de projecten van Metamorfoze Archieven en collecties verantwoordelijk voor de technische controle van gedigitaliseerd materiaal door externe leveranciers en het doorleveren van dit materiaal naar de duurzame eindopslag in het e-depot van het Nationaal Archief.

Ten behoeve van een soepele en efficiënte doorloop van deze controles heeft de Koninklijke Bibliotheek de volgende aanleveringsrichtlijn opgesteld.

Kader

Dit specificatiedocument beschrijft de eisen aan de opdrachtgever inzake het aanleveren van een toegang tot het te digitaliseren materiaal, en de eisen aan de aanlevering van gedigitaliseerd materiaal door externe leveranciers.

Aanlevering door de opdrachtgever:

Toegang

Om het gedigitaliseerde materiaal op te kunnen nemen in het depot van het Nationaal Archief, is het noodzakelijk dat de eigenaar van het materiaal informatie verschaft in de vorm van een toegang tot dit materiaal. De gevraagde structuur van de toegang is als volgt.

M = mandatory (verplicht), **O** = optional; tenzij anders aangegeven komen velden 1x voor

Archief:

Instellingscode (ISIL)	M	
Instellingsnaam	M	
Archiefkenmerk	M	
Archiefnaam	M	
Archiefvormer	M	
Type archief	M	'Persoon', 'Organisatie' etc.
Archief_begindatum	M	
Archief_einddatum	M	
Beschrijving archief	M	
Openbaarheid/tekst archief	M	'Openbaar', 'Beperkt openbaar' etc.
Openbaarheid/geldigheid archief	M	
Openbaarheid/grondslag	M	

Serie:

	O, herhaalbaar;	
SerieID	M	
Serienamen	M	
Seriebeschrijving	M	
Openbaarheid/tekst serie	O	Gegevens m.b.t. openbaarheid alleen
Openbaarheid/geldigheid serie	O	invullen, als ze afwijken van de openbaar-
Openbaarheid/grondslag serie	O	heidsgegevens van het archief.

Per Inventarisnummer:

Inventarisnummer	M	
Beschrijving inventarisnummer	O	
Openbaarheid/tekst inventarisnummer	O	Gegevens m.b.t. openbaarheid alleen
Openbaarheid/geldigheid inventarisnummer	O	invullen, als ze afwijken van de openbaarheidsgegevens van het archief.
Openbaarheid/grondslag inventarisnummer	O	Link naar eventuele serie (= waarde SerieID)
refID Serie	O	
<i>Subdossier:</i>	O, herhaalbaar;	
SubdossierID	M	
Subdossiernaam	M	
Subdossierbeschrijving	M	
Openbaarheid/tekst subdossier	O	Gegevens m.b.t. openbaarheid alleen
Openbaarheid/geldigheid subdossier	O	invullen, als ze afwijken van de openbaarheidsgegevens van het
Openbaarheid/grondslag subdossier	O	inventarisnummer.

De KB levert een Excelbestand waarin de opdrachtgever de gevraagde gegevens aanlevert aan de KB. Wordt de toegang niet aangeleverd, dan kan niet tot digitalisering worden overgegaan. De toegang wordt na controle door de KB doorgeleverd aan de gekozen digitaliseerder.

Acceptatiecriteria toegang

De door de opdrachtgever aangeleverde toegang wordt door de KB gecontroleerd. Als de toegang niet aan de hierboven gegeven specificaties voldoet, kan het materiaal niet gedigitaliseerd en duurzaam opgeslagen worden.

Aanlevering van materiaal door de digitaliseerder

De onderstaande specificaties zijn vereist en leveringen kunnen worden geweigerd bij niet naleving van deze richtlijnen.

De aanlevering geschiedt op externe harde schijven ('HD') van minimaal 2 TB die geformatteerd zijn op NTFS onder toekenning van volledige rechten voor alle gebruikers van de schijf.

Een batch is een levering van alle gevraagde producten per HD. In het vervolg kunnen de benoemingen 'HD' en 'batch' als synoniemen gelezen worden.

De harde schijf zelf, de interne naam van de harde schijf en de verpakking van de harde schijf bevatten als naam de batchcode die bij de start van een project door de Koninklijke Bibliotheek uitgegeven wordt en volgt onderstaand stramien binnen een project.

Naamgeving batches

De naam van de batch bestaat uit een projectcode gevolgd door het nummer van de batch (9 cijfers), het deelnummer van de batch (1 cijfer) en het versienummer van de batch (2 cijfers). Bijvoorbeeld MMKB03_000000009_1_04, de vierde versie van het eerste deel van batch 9 voor het project MMKB03. Een gedigitaliseerd object/inventarisnummer mag niet over meerdere batches verdeeld zijn.

Wanneer een batch om wat voor reden dan ook opnieuw wordt aangeleverd moet het versienummer voor de gecorrigeerde batch met 1 worden opgehoogd. De nieuwe versie van de

batch moet precies dezelfde objecten bevatten als de voorgaande versie. Wanneer het ontbreken van bestanden een reden was om de batch af te keuren, moet de batch de oorspronkelijke bestanden plus de ontbrekende bestanden bevatten.

Producten

De aan te leveren producten per batch zijn:

- Dagtargets
- Meetresultaten dagtargets
- Masterbestanden in TIFF met objectleveltargets
- Masterbestanden in TIFF zonder objectleveltargets
- Afgeleiden (afwisselend per project)
- Rapportages onregelmatigheden
- Checksum [SHA512, hexadecimaal] per batch op het hoogste niveau van de HD (root). Deze checksum moet alle bestanden omvatten die op de HD staan.

De masterbestanden worden geproduceerd in TIFF met de volgende specificaties: Baseline, Tagged Image File Format, versie 6.0, zonder compressie, zonder ingesloten thumbnail, met de extensie "tif".

De afgeleiden kunnen in overeenstemming met de instelling verschillende bestandsformaten hebben.

De dagtargets en masterbestanden met objectleveltargets worden opgenomen volgens de Richtlijn Preservation Imaging versie Metamorfoze 1.0 (2012).

Mappenstructuur



Metadata - headerinformatie

De leverancier dient in elk TIFF-bestand headerinformatie volgens de onderstaande tabel op te nemen. Headerinformatie moet (wanneer mogelijk) in XMP-tags te worden opgeslagen: in de tabel zijn de namen van de XMP-tags opgenomen, gevolgd door de XMP-namespace waar de tag deel van uit maakt.

WAT	WAAR
Directory	= Directory (volledig pad)
Filename	= FileName
Resolution horizontal	= XResolution (XMP tiff)
Resolution vertical	= YResolution (XMP tiff)
Resolution unit	= ResolutionUnit (XMP tiff)
Pixels horizontal	= ImageWidth (XMP tiff)
Pixels vertical	= ImageHeight (XMP tiff)
Filesize	= FileSize (in bytes)
Bitdiepte	= BitsPerSample (XMP tiff)
Software camera	= CreatorTool (XMP xmp)
Software bewerking	= Software (XMP tiff)
Kleurruimte	= ColorSpace (XMP exif)
Kleurprofiel	= ICCProfileName (XMP photoshop)
Datum en tijd opname	= DateTime (XMP tiff)
Model scanner of camera	= Model (XMP tiff)
Type scanner of camera	= Make (XMP tiff)
Compressie	= Compression (XMP tiff)
<i>De velden hieronder zijn alleen van toeval in het geval van One-Shot-digitalisering.</i>	
Sluittijd (indien van toepassing)	= ShutterSpeedValue (XMP exif)
	= ExposureTime (XMP exif)
Diafragma (indien van toepassing)	= ApertureValue (XMP exif)
ISO waarde (indien van toepassing)	= ISO (XMP exif)
Brandpuntsafstand (indien van toepassing)	= FocalLength (XMP exif)
Belichtingsprogramma (indien van toepassing)	= ExposureProgram (XMP exif)
Speciale camera of RAW conversie instellingen (indien van toepassing)	ook in XMP-tags
<i>De velden hieronder zijn niet verplicht: de opdrachtgever mag beslissen, of deze informatie opgenomen wordt.</i>	
Naam van het archief (indien van toepassing)	= Headline (XMP photoshop)
Copyright (indien van toepassing)	= Copyright (XMP tiff)
Naam van instelling (indien van toepassing)	= Credit (XMP photoshop)

Versleutelen

De opdrachtgever kan per project of deel van een project er voor kiezen om de informatiedrager(s) te laten versleutelen. De aangeleverde informatiedrager dient in dat geval met BitLocker versleuteld te zijn (zie afzonderlijke gebruikershandleiding).

Bestandsnaamgeving

De masterbestanden (TIFF) worden duurzaam opgeslagen in het e-depot van het Nationaal Archief te Den Haag. Het Nationaal Archief heeft voor het opnemen van gedigitaliseerd materiaal een bestandsnaamgevingsconventie die te vinden is in 'Richtlijn 1 – Bestandsnamen 1.1'. De richtlijn kan bij de afdeling digitalisering van de Koninklijke Bibliotheek aangevraagd worden.

Kort abstract:

Een bestandsnaam bestaat uit meerdere onderdelen die van elkaar gescheiden worden door middel van een underscore ('_').

De volgende onderdelen moeten minimaal deel uitmaken van de bestandsnaam:

ISIL-code (of vergelijkbaar uniek identificatiekenmerk voor Bibliotheken OCLC)

Toegang (naam of nummer van archief/collectie)

Inventarisnummer (object- of dossiernummer)

Volgnummer

Extensie

De opbouw is: `ISILcode_toegang_inventarisnummer_volgnummer.extensie`

Voorbeeld: `NL-HaNA_1.01.01.01_123_0001.tif`

Er mogen in bestandsnamen geen spaties en geen diakritische tekens voorkomen. Alleen de karakters A-Z, a-z, 0-9, '_', '' en '-' mogen gebruikt worden.

Ten behoeve van de bestandsnaamgeving heeft de Koninklijke Bibliotheek een formulier ontwikkeld dat door de instellingen ingevuld wordt en na eerste check door de Koninklijke Bibliotheek wordt doorgestuurd aan de digitaliserende partij. Middels deze informatie zijn de bouwstenen voor de bestandsnaamgeving voor de digitaliseerder beschikbaar.

Instructie meetresultaten

De meetresultaten van de dagtargets (testopnamen) afkomstig uit bijvoorbeeld IE Analyzer en Imcheck worden aangeleverd als screenshots én als data files (jpg/xls/txt). Voor een vlot en goed verloop van de beeldkwaliteitscontrole is het van belang de gegeneerde screenshots en data steeds op een eenduidige manier aan te leveren. Het is belangrijk dat de meetresultaten eenduidig gelinkt kunnen worden aan de betreffende targetbestanden.

1. Screenshots van bijv. IE Analyzer en Imcheck

De screenshots van de grafieken uit IE analyzer en Imcheck kunnen worden opgenomen in de tabbladen van het door de KB beschikbaar gestelde excelverzamelbestand.

De naamgeving van dit excel-bestand is als volgt:

- `MMABC_AfkortingDigitaliseerder_jjjjmmdduumm_werkstation_resolutie_meetresultaten.xls`

Dit excel-bestand krijgt 5 tabbladen, genaamd: OECF, Uitlichting, SFR, Ruis en Delta E. Hierin komen de desbetreffende screenshots uit bijvoorbeeld:

- IE Analyzer - tabblad: Delta E (met kleurwaarden en kleurvergelijking);
- Imcheck - tabbladen: OECF (met de luminantiewaarden uit de ruismeting), Uitlichting, SFR en Ruis.

2. Data van metingen uit bijvoorbeeld IE Analyzer en Imcheck

De data van de metingen van IE Analyzer en Imcheck bestaande uit 5x SFR, 1x Ruis en 1x Uitlichting worden in 7 aparte excel-bestanden aangeleverd. De data van de Delta E-meting wordt in een txt-bestand aangeleverd.

De naamgeving van deze 7 excel-bestanden is gelijk aan de naamgeving van de individuele target met als toevoeging:

- _SFR_M.xls (SFR midden)
- _SFR_LO.xls (SFR linksonder)
- _SFR_LB.xls (SFR linksboven)
- _SFR_RO.xls (SFR rechtsonder)
- _SFR_RB.xls (SFR rechtsboven)
- _Ruis.xls
- _Uitlichting.xls

De naamgeving van het txt-bestand met de Delta E meting wordt automatisch gegenereerd en is gelijk aan de naamgeving van de individuele target met als toevoeging:

- _color.txt

Neem bij onduidelijkheden contact op met de Senior productiemanager digitalisering van de KB.

N.B. Het aanleveren van de meetresultaten van de dagtargets op de hierboven omschreven wijze is essentieel voor het controleproces van de beeldkwaliteit binnen Metamorfoze ABC.

Acceptatiecriteria Beeldkwaliteit technisch en output

De dagtargets en objectleveltargets worden gecontroleerd volgens Richtlijn Metamorfoze Preservation Imaging 1.0 versie 2012.

De afbeeldingen zelf worden op basis van een steekproef gecontroleerd. Zie onderstaande lijst:

Voor het vaststellen van de grootte van de steekproef en beoordelen van de batch wordt gebruik gemaakt van ISO 2859 deel 1: 1999 EN "Sampling procedures for inspection by attributes - Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection". Uitgegaan wordt van de volgende parameters:

- grootte van deze steekproef is conform het Algemene Testniveau II.
- acceptabel kwaliteitsniveau is 1

Zie voor de steekproefgrootte en de acceptatie van fouten op het Algemene Testniveau II de onderstaande tabel conform ISO 2859 deel 1.

Algemeen testniveau II		Kwaliteitsniveau 1	
batchgrootte	steekproef grootte	acceptatie	afkeuren
		aantal fout	aantal fout
21-150	20	0	≥ 1
151-280	32	≤ 1	≥ 2
281-500	50	≤ 1	≥ 2
500-1200	80	≤ 2	≥ 3
1201-3200	125	≤ 3	≥ 4
3201-10000	200	≤ 5	≥ 6
10001-35000	315	≤ 7	≥ 8
35001-150000	500	≤ 10	≥ 11

De KB behoudt zich het recht voor om bij bewezen hoog kwaliteitsniveau binnen een project de steekproef volgens Algemeen testniveau I uit te voeren (zie onderstaande tabel).

Zie voor de steekproefgrootte en de acceptatie van fouten op het Algemene Testniveau I de onderstaande tabel conform ISO 2859 deel 1.

Algemeen testniveau I		Kwaliteitsniveau 1	
batchgrootte	steekproef grootte	acceptatie	afkeuren
		aantal fout	aantal fout
21-500	20	0	≥ 1
500-1200	32	≤ 1	≥ 2
1201-3200	50	≤ 1	≥ 2
3201-10000	80	≤ 2	≥ 3
10001-35000	125	≤ 3	≥ 4
35001-150000	200	≤ 5	≥ 6

Tijdens de steekproef worden zowel technische als ook output gerelateerde afspraken en eisen gecontroleerd. Artefacten, scheef liggen, clippen en dergelijke vallen daaronder maar ook oriëntatie zoals afgesproken bij start van het project als ook croppen en het gebruik van de juiste achtergrond.

Acceptatiecriteria metadata

De Headerinformatie wordt door de KB op volledigheid en juistheid gecontroleerd. Is de informatie incorrect, dan moet het bestand (en daarmee de gehele batch) opnieuw aangeleverd worden. Er geldt een 0% foutenmarge op metadata.

Inhoudelijke controle door opdrachtgever

Na technische goedkeuring van zowel metadata, aanleveringscriteria en beeldkwaliteit wordt door de opdrachtgever (erfgoedinstelling) een inhoudelijke controle uitgevoerd. De KB verzorgt de levering van de benodigde bestanden aan de opdrachtgever. De acceptatiecriteria voor de inhoudelijke controle worden per project afgesproken bij het startgesprek.