

PAKKET VAN EISEN

van de KB betreffende

Europese openbare aanbesteding

Digitaal gedrukt tekstmateriaal

Inhoud

1	Inleiding	5
2	Bronmateriaal en aanlevering	6
2.1	Bronmateriaal	6
2.2	Aanlevering materiaal en metadata door de KB.....	6
2.3	Aanleveren digitale bestanden door de KB.....	7
2.3.1	Herkennen afleveringen bij aanlevering digitale moederbestanden	7
2.3.2	Directorystructuur harddisk.....	7
2.3.3	Teruglevering harddisk met digitale moederbestanden	7
2.4	Zorgplicht met betrekking tot het Bronmateriaal.....	7
2.4.1	Verantwoordelijkheid	7
2.4.2	Aanvullend op het vorige artikellid.....	7
2.4.3	Schade.....	7
2.4.4	Doeleinden Bronmateriaal	7
2.4.5	Instructielijst.....	7
2.5	Eisen met betrekking tot transport van het Bronmateriaal	7
2.5.1	Transport	8
2.5.2	Verpakking Bronmateriaal	8
2.5.3	Vervoer Bronmateriaal.....	8
2.5.4	Meegestuurd documenten.....	8
2.5.5	Transport Bronmateriaal en risico's.....	8
2.5.6	Vervoer Zending	8
2.5.7	Transportplanning.....	8
2.5.8	Locatie Bronmateriaal.....	8
2.5.9	Extra transporteisen	8
2.5.10	Aflevering Bronmateriaal	9
2.6	Eisen met betrekking tot opslag van Bronmateriaal en gebruik van Bronmateriaal.....	9
2.6.1	Inspectie Bronmateriaal.....	9
2.6.2	In- en uitladen Bronmateriaal	9
2.6.3	Opslagruimte Bronmateriaal.....	9
2.6.4	Indeling opslagruimte	9
2.6.5	Plaats Bronmateriaal	9
2.6.6	Brand- en rookmeldinstallaties	9
2.6.7	Bescherming tegen water en rookgassen	9
2.6.8	Gebruik Bronmateriaal zonder schade.....	9
2.6.9	Rampenplannen	10
2.7	Teruggave Bronmateriaal en oplevering Batch	10
2.7.1	Teruggave Bronmateriaal	10
2.7.2	Opleverschema Batches	10
2.7.3	Niet tijdig leveren Batch.....	10
2.7.4	Leverwijze Batch.....	10
2.7.5	Documentatie Batch	10

2.8	Acceptatie van een Batch.....	10
2.8.1	Acceptatie van een Batch tijdens Pilotfase	10
2.8.2	Afkeuring eerste Pilotbatch.....	11
2.8.3	Afkeuring tweede Pilotbatch.....	11
2.8.4	Reguliere acceptatie van een Batch.....	11
2.8.5	Gebrek in Batch	11
2.8.6	Afkeuring Batch	11
2.8.7	Rapportage.....	11
2.8.8	Acceptatietermijn voor één Batch.....	11
2.8.9	Herlevering	11
2.8.10	Correctie door een derde.....	11
2.8.11	Bewaarplicht Batch.....	12
3	Algemene instructies en instructies per materiaalsoort	13
3.1	Masterbestanden	13
3.2	Volledigheid	13
3.3	Blanco pagina's	13
3.4	Segmentatie.....	13
3.5	Auteurherkenning	13
3.6	Supplementen.....	13
3.7	Aparte afleveringen bij tijdschriften.....	13
3.8	Per pagina	13
4	Teruglevering van materiaal en metadata door opdrachtnemer	14
4.1	Directorystructuur en naamgeving.....	14
4.1.1	Naamgeving en omvang batches	14
4.1.2	Object identifiers	14
4.1.3	Directories.....	14
4.1.4	Bestandsnamen.....	16
4.1.5	Checksumbestanden.....	16
4.1.6	Concordantietabel.....	16
4.1.7	Technische targets	16
4.1.8	Pakbon.....	17
4.2	Checksums	17
4.3	Technische specificaties en beeldkwaliteit.....	17
4.3.1	Algemene specificaties	17
4.3.2	Scaninstructies	17
4.3.3	Technische specificaties master- en accessbestanden	18
4.3.4	Belichtingstolerantie Kodak Gray Scale	23
4.3.5	Dagtargets, naamgeving, volgorde en beeldopbouw	24
4.3.6	Intern testrapport	25
4.4	PDF-bestanden.....	25
4.5	Metadatabestanden	27
4.5.1	Algemene eisen aan XML-bestanden	27

4.5.2	Metadatadump en schema	28
4.5.3	METS	33
4.5.4	Schema's	33
4.5.5	Paginanummering	33
4.5.6	Auteursinformatie per tijdschriftartikel	33
4.5.7	Toelichting bij de tabel	34
4.5.8	MODS	41
4.5.9	PREMIS	41
4.5.10	MIX	43
4.5.11	OCR	48
4.5.12	ALTO	49
4.5.13	Pakbon	55
4.5.14	Concordantietabel	60
5	Acceptatiecriteria	62
	Bijlage A: Omgang met het materiaal ten behoeve van digitalisering	63
	Bijlage B: Schade-afhandeling na digitaliseren	64
	Bijlage C: JPEG 2000 Formaat parameters	66

1 Inleiding

Dit Pakket van Eisen specificeert de producten die binnen deze opdracht dienen te worden opgeleverd. Het beschrijft allereerst wat de KB aanlevert. Vervolgens specificeert het de eisen met betrekking tot de teruglevering van materiaal door de opdrachtnemer, onderverdeeld in:

- aanlevering van materiaal en metadata door de KB;
- algemene instructies en instructies per materiaalsoort;
- teruglevering van materiaal en metadata door de opdrachtnemer:
- teruglevering van de digitale bestanden
 - o directorystructuur en naamgeving
 - o checksums
 - o master- en accessbestanden (technische specificaties)
 - o PDF-bestanden
 - o metadata (XML-bestanden en concordantietabellen)

Tot slot worden de acceptatiecriteria toegelicht.

De Minimumeisen die de KB aan de uitvoering van de opdracht stelt van een Inschrijving onvoorwaardelijk in met deze Minimumeisen.

2 Bronmateriaal en aanlevering

2.1 Bronmateriaal

Het Bronmateriaal is niet altijd homogeen. De te digitaliseren objecten zijn voornamelijk gedrukt in Nederland, hoofdzakelijk in het Nederlands. Historische varianten van het Nederlands kunnen in de tekst voorkomen. Incidenteel kunnen (delen van) teksten een andere taal bevatten (bijvoorbeeld Engels, Frans, Duits, Hebreeuws, en Indonesisch). Lay-out, formaten, gehanteerde fonts en lettertypen vertonen differentiatie. Zo kunnen artikelen van tijdschriften en kranten zonder onderbreking doorlopen over meerdere pagina's en kan er sprake zijn van uitklapbladen of (kleur)afbeeldingen in de tekst. Het formaat van het materiaal is divers. Het kan gaan om gebonden en losbladig materiaal. De Koninklijke Bibliotheek kiest afhankelijk van het materiaal om de batches in kleur of gray scale te laten uitvoeren.

Oprachtnemer dient er rekening mee te houden dat de originelen in matige, slechte of zeer slechte staat kunnen verkeren. Vaak voorkomende schades bij papieren originelen zijn: scheuren, vlekken, spikkels, slecht drukwerk, doordrukken, bros (brittle) papier en schade aan de rug of band van het boek. In de meeste gevallen zijn de gedrukte tekstmaterialen gebonden. Het is opdrachtnemers niet toegestaan om de banden zonder expliciete toestemming van de Koninklijke Bibliotheek te verwijderen of open te drukken ("kraken") ten behoeve van digitalisering. Om beschadiging van het object te voorkomen zal bij het scannen op verzoek van de Koninklijke Bibliotheek soms een boekenwieg toegepast moeten worden. Voor materiaal ouder dan 100 jaar, of materiaal dat uniek van aard is kunnen bijzondere maatregelen worden afgesproken. Het bronmateriaal dient te allen tijde zorgvuldig gehanteerd te worden, conform de door de Koninklijke Bibliotheek opgestelde voorschriften. De eisen aan de omgang met het originele materiaal vindt u in bijlage A.

Voorafgaand aan het transport naar de opdrachtnemer zal het te digitaliseren materiaal door de Koninklijke Bibliotheek worden voorbereid. De meest acute schades die digitalisering in de weg staan, worden hierbij gerepareerd. Objectspecifieke bijzonderheden zoals bijvoorbeeld de aanwezigheid van uitklapbladen worden gemarkeerd. De staat van het materiaal wordt in een database beschreven. Transport, verzekering van transport, en verzekering tijdens verblijf is voor rekening van de opdrachtnemer. Voor elk transport wordt een Metadata-export uit de database aangeleverd. Indien het materiaal om een aparte behandeling vraagt, worden instructies meegegeven.

2.2 Aanlevering materiaal en metadata door de KB

De KB zal het te verwerken materiaal in zendingen sturen.

Een zending kan bestaan uit meerdere 'Transporten' van verschillende deelprojecten.

Bij iedere zending die de KB stuurt ter verwerking, wordt voor elk Transport een XML-bestand (metadatadump) meegeleverd waarin staat beschreven wat er in het Transport wordt gestuurd. Hierin staan zowel metadata die in eindproducten moeten worden teruggeleverd als gegevens die alleen van belang zijn voor het digitaliseren en/of genereren van metadata. Bij de beschrijving van de eisen aan de te leveren metadata wordt aangegeven welke gegevens uit de metadatadump moeten worden teruggeleverd, en hoe dat moet gebeuren.

Daarnaast levert de KB een deels ingevuld METS-bestand mee voor elk digitaal object dat de KB verwacht te ontvangen. De specificaties van de metadatadump en van de in te vullen onderdelen van het METS-bestand komen later ter sprake. De eisen in dit document betreffen twee verschillende soorten materiaal: kranten en tijdschriften. Dit materiaal kan een aparte behandeling vereisen. Materiaal kan aangeleverd worden op papier of in de vorm van digitale bestanden (scans). Tevens levert de KB een verzendlijst mee.

Wijzigingen in de aangeleverde metadata moeten altijd verwerkt worden in de database van de KB. Teveel of te weinig issues en andere correcties. De KB levert dan altijd een nieuwe metadatadump met METS met een opgehoogd versienummer. De foutmarge die de KB voor zichzelf hanteert is max. 8 fouten per zending.

2.3 Aanleveren digitale bestanden door de KB

2.3.1 *Herkennen afleveringen bij aanlevering digitale moederbestanden*

Worden er digitale bestanden aangeleverd, dan levert de KB “moederscans” aan op harddisk, geordend per barcode in directories. Deze bestanden moeten door de opdrachtnemer gecropt, gesplitst (indien nodig) en geconverteerd worden naar access- en master-JP2's.

De opdrachtnemer moet in het geval van kranten of tijdschriften daarbij zelf de afleveringen herkennen. Ook bij digitale bestanden wordt een metadatadump en METS geleverd.

2.3.2 *Directorystructuur harddisk*

Een zending zal geleverd worden op een harddisk. Er kunnen meerdere Transporten van verschillende deelprojecten op een harddisk staan. De directorystructuur op de harddisk is als volgt: Hoofddirectory met de batchnaam. Als de digitale bestanden afkomstig zijn van een gescande microfilm, is er een directory voor elke film met als directorynaam het filmsignatuur. Op het laagste niveau staan de terug te leveren scans met als directorynaam een barcode.

2.3.3 *Teruglevering harddisk met digitale moederbestanden*

De harddisk dient ongewijzigd met het retourtransport terug te worden geleverd aan de KB.

De KB zal per batch een checksumbestand meeleveren. Hierin staan de checksums (SHA-512) van alle bestanden uit de batch. Dit bestand kan gebruikt worden om te controleren of alle bestanden in het transport ongewijzigd zijn ontvangen door de opdrachtnemer en ook weer ongewijzigd retour zijn geleverd aan de KB.

2.4 Zorgplicht met betrekking tot het Bronmateriaal

2.4.1 *Verantwoordelijkheid*

Opdrachtnemer is verantwoordelijk voor het Bronmateriaal dat aan hem ter beschikking wordt gesteld. Opdrachtnemer is gehouden tot behoud van het Bronmateriaal volgens wat blijkt uit het door de KB ingevoerde gegevens in de database over dat Bronmateriaal.

2.4.2 *Aanvullend op het vorige artikellid*

Opdrachtnemer draagt er zorg voor dat:

- uitvoering van de Opdracht tot digitalisering plaatsvindt in een veilige en schone omgeving;
- het Bronmateriaal bij retourzending vrij is van vuil, schimmels en ongedierte, indien dat Bronmateriaal bestaat uit fysieke documenten.

2.4.3 *Schade*

Indien Opdrachtnemer gedurende de uitvoering van een Opdracht constateert dat schade ontstaat of is ontstaan aan het door de KB ter beschikking gestelde Bronmateriaal, dan is Opdrachtnemer verplicht om de KB hiervan onverwijld in kennis te stellen. Zie bijlage B voor schadeprotocol.

2.4.4 *Doeleinden Bronmateriaal*

Opdrachtnemer mag het Bronmateriaal niet voor andere dan genoemde doeleinden gebruiken.

2.4.5 *Instructielijst*

De KB kan aan Opdrachtnemer bij iedere Opdracht een instructielijst verstrekken waarbij aandacht wordt besteed aan de correcte handling van het desbetreffende Bronmateriaal. Opdrachtnemer is gehouden deze lijst met het personeel door te nemen dat uitvoering geeft aan de Opdracht. De lijst is altijd op de plaats aanwezig waar uitvoering wordt gegeven aan de Opdracht.

2.5 Eisen met betrekking tot transport van het Bronmateriaal

Dit artikel heeft enkel betrekking op Bronmateriaal dat bestaat uit fysieke documenten, tenzij uitdrukkelijk in dit artikel is vastgelegd dat onderdelen van dit artikel ook betrekking hebben op Bronmateriaal dat bestaat uit digitale bestanden.

2.5.1 Transport

Het transport van het Bronmateriaal geschiedt in een goed onderhouden, gesloten, vervoermiddel. Opdrachtnemer is verplicht ten behoeve van het transport van Bronmateriaal voldoende, door de KB goedgekeurd, deugdelijk en afsluitbaar en gesloten (dus zonder open handvaten) transportmateriaal aan de KB ter beschikking te stellen. Fysiek materiaal mag niet kunnen schuiven tijdens transport. De verpakkingsmaterialen worden tijdens het transport niet geopend of verwijderd. Het Bronmateriaal dient binnen Nederland te blijven.

2.5.2 Verpakking Bronmateriaal

De KB verpakt het Bronmateriaal ten behoeve van het transport naar Opdrachtnemer. Opdrachtnemer verpakt op dezelfde wijze het Bronmateriaal ten behoeve van het retourtransport (zie verzendlijst met indeling kratnummers) vanaf Opdrachtnemer.

2.5.3 Vervoer Bronmateriaal

Het vervoeren van het Bronmateriaal dient op dusdanige wijze te geschieden dat deze niet beschadigd kan raken en/of het risico loopt van diefstal. Bij een tussenstop wordt het Bronmateriaal niet onbeheerd gelaten. Dat betekent, dat indien tussenstops nodig zijn er twee begeleiders moeten zijn waarvan één steeds bij het transportmiddel blijft.

2.5.4 Meegestuurde documenten

Ten behoeve van het transport van Bronmateriaal stelt de KB aan Opdrachtnemer een bestand met metadata van de titels van het Bronmateriaal beschikbaar. In vervolg daarop bij overhandiging van het Bronmateriaal (zowel fysiek als digitaal) van de KB aan Opdrachtnemer, tekenen Partijen een verzendlijst in tweevoud. Vanaf dat moment is Opdrachtnemer verantwoordelijk voor het Bronmateriaal. Indien Opdrachtnemer daarna constateert dat sprake is van onvolkomenheden tussen het verkregen Bronmateriaal enerzijds en de verzendlijst anderzijds, is Opdrachtnemer gehouden dat binnen twee (2) Werkdagen te berichten aan de KB.

2.5.5 Transport Bronmateriaal en risico's

Opdrachtnemer verzorgt het transport van Bronmateriaal, zoals beschreven in het Projectplan en draagt vanaf dat moment het risico met betrekking tot het Bronmateriaal, behoudens voor zover schade aan dat Bronmateriaal voortvloeit uit onvolkomenheden aan de verpakking daarvan door de KB. Ter dekking van deze aansprakelijkheid is Opdrachtnemer gehouden, zoals ook blijkt uit de afspraken vastgelegd in artikel 12, het Bronmateriaal voor de periode van het transporteren van de locatie van de KB naar de locatie waar de Opdrachtnemer de digitalisering van het Bronmaterieel verricht en vice versa, te verzekeren tegen verlies en beschadiging.

2.5.6 Vervoer Zending

Het transport van en naar de KB door Opdrachtnemer omvat minimaal Bronmateriaal behorende tot één Transport. Het is Opdrachtnemer niet toegestaan om een Zending in delen te transporteren. Opdrachtnemer is voorts gehouden het Bronmateriaal behorende tot één Zending bijeen te houden bij de opslag van dat Bronmateriaal.

2.5.7 Transportplanning

Opdrachtnemer voert het transport uit op een vaste dag in de week zoals vastgelegd in het Projectplan. Opdrachtnemer is voor iedere Opdracht gehouden op basis van wat in het Projectplan is opgenomen een transportplanning aan de KB ter beschikking te stellen. Opdrachtnemer is gehouden bij de uitvoering van Opdrachten in overeenstemming met die transportplanning het Bronmateriaal bij de KB op te halen en weer te retourneren. Indien Opdrachtnemer voorziet dat afwijkingen plaatsvinden van die transportplanning, informeert Opdrachtnemer onverwijld de KB daarover.

2.5.8 Locatie Bronmateriaal

Het Bronmateriaal mag in de periode gelegen tussen heen- en retourtransport van de locatie van de KB, het gebouw niet verlaten waar de digitalisering door Opdrachtnemer plaatsvindt.

2.5.9 Extra transporteisen

De KB is bevoegd om in bijzondere gevallen, zoals het transport van zeldzaam, kostbaar en/of kwetsbaar Bronmateriaal, extra eisen te stellen aan het transport van het Bronmateriaal, die al dan niet

beschreven zijn in het Projectplan. Opdrachtnemer is gehouden aan die eisen te voldoen. Eventuele kosten die met deze extra eisen gemoeid zijn, zijn voor rekening van Opdrachtgever.

2.5.10 Aflevering Bronmateriaal

Opdrachtnemer geeft bij aflevering bij de KB van zowel fysiek als digitaal Bronmateriaal duidelijk aan voor welke natuurlijke persoon van de KB het Bronmateriaal bestemd is. Aflevering bij de KB kan enkel tijdens Werkdagen, uiterlijk tot 16:30 uur. Indien niet is vermeld voor welke natuurlijke persoon van de KB het Bronmateriaal bestemd is, mag de KB het in ontvangst nemen van het Bronmateriaal weigeren.

2.6 Eisen met betrekking tot opslag van Bronmateriaal en gebruik van Bronmateriaal

2.6.1 Inspectie Bronmateriaal

De KB is gerechtigd te allen tijde het Bronmateriaal te inspecteren, ter plaatse waar het Bronmateriaal is opgeslagen en wordt gebruikt, zodat de KB kan controleren in hoeverre aan de vereisten van dit artikel is voldaan.

Indien volgens de KB niet aan die vereisten is voldaan, is de KB bevoegd:

- onmiddellijk het Bronmateriaal terug te nemen en de Opdracht te beëindigen; of
- bindende aanwijzingen te geven aan Opdrachtnemer over de wijze van opslag en behandeling van het Bronmateriaal, waaraan Opdrachtnemer onverkort en onvoorwaardelijk medewerking zal verlenen. De kosten daarvan komen voor rekening van Opdrachtnemer.

2.6.2 In- en uitladen Bronmateriaal

Na aankomst op de bestemming wordt het Bronmateriaal direct naar de daartoe bestemde opslagruimte gebracht. Het in- en uitladen van het materiaal dient onder een overkapping plaats te vinden zodat het materiaal niet aan weersomstandigheden wordt blootgesteld. Na het uitpakken dient Opdrachtnemer het Bronmateriaal conditioneel te controleren. Deze controle dient schriftelijk te worden vastgelegd.

2.6.3 Opslagruimte Bronmateriaal

Opdrachtnemer dient te beschikken over een beveiligde, geconditioneerde opslagruimte – brandwerend en vergrendeld – die voldoet aan de normen voor archief ruimten (RLV 50-55%, temp. 16°-20°C) en zal enkel die ruimte gebruiken voor opslag van het Bronmateriaal. Deze vereiste geldt ook voor ruimten die worden gebruikt voor het scannen van het Bronmateriaal.

2.6.4 Indeling opslagruimte

Opdrachtnemer is verplicht om (i) in de desbetreffende opslagruimte een voldoende aantal kisten te reserveren voor het Bronmateriaal van de KB, en (ii) ervoor te zorgen dat de afmetingen van de planken in die kisten niet kleiner zijn dan de afmetingen van de kisten waarin het Bronmateriaal wordt aangeleverd.

2.6.5 Plaats Bronmateriaal

Opdrachtnemer mag niet (i) de door de KB aangeleverde kisten met Bronmateriaal op de grond van de opslagruimte plaatsen en (ii) het Bronmateriaal van de KB uit de opslagruimte verwijderen, behoudens voor zover noodzakelijk voor digitalisering van het Bronmateriaal of voor retourtransport daarvan naar de KB.

2.6.6 Brand- en rookmeldinstallaties

Opdrachtnemer is verplicht te waarborgen dat huisvesting en opslagruimte die worden gebruikt voor de uitvoering van de Opdracht voorzien zijn van gecertificeerde brand- en rookmeldinstallaties met rechtstreekse doormelding naar een meldkamer van de brandweer.

2.6.7 Bescherming tegen water en rookgassen

Opdrachtnemer is verplicht om de huisvesting en opslagruimte die worden gebruikt voor de uitvoering van de Opdracht, veilig in te richten en afdoende te beschermen tegen instroom van water en rookgassen.

2.6.8 Gebruik Bronmateriaal zonder schade

Opdrachtnemer moet het Bronmateriaal zodanig gebruiken dat geen schade kan optreden aan het Bronmateriaal, anders dan onvermijdelijke schade voortvloeiend uit het scanproces.

2.6.9 Rampenplannen

Oprachtnemer is verplicht te beschikken over rampenplannen en procedures voor incidenten als inbraak, brand, terroristische aanslagen en andere calamiteiten met betrekking tot de ruimten die worden gebruikt voor de uitvoering van de Opdracht. Deze rampenplannen en procedures dienen up-to-date te zijn.

2.7 Teruggave Bronmateriaal en oplevering Batch

2.7.1 Teruggave Bronmateriaal

Oprachtnemer is verplicht tot teruggave van het Bronmateriaal op de plaats, datum en overeenkomstig het schema zoals opgenomen in het Projectplan, na voltooiing van de Opdracht, bestaande uit werkpakket A of werkpakket B, zoals beschreven in de Opdracht, uiterlijk binnen één week na acceptatie van de Batch met betrekking tot dat Bronmateriaal. In ieder geval dient het Bronmateriaal (per Batch) in dezelfde volgorde te worden teruggegeven, als in welke volgorde de KB het Bronmateriaal aan Oprachtnemer heeft aangeleverd. Voor transport van Bronmaterieel van Oprachtnemer naar de KB, ter nakoming van de verplichting tot teruggave van Bronmateriaal.

2.7.2 Opleverschema Batches

Het Resultaat van een Opdracht kan bestaan uit de oplevering van meerdere Batches. Oprachtnemer is verplicht iedere Batch op te leveren overeenkomstig het opleverschema zoals opgenomen in het Projectplan en volgens de vereisten voortvloeiend uit de Overeenkomst. De in het opleverschema opgenomen data voor oplevering van iedere Batch zijn fatale termijnen in de zin van 6:83, onderdeel a, Burgerlijk Wetboek. Indien Oprachtnemer verwacht de datum van oplevering niet te kunnen halen, vermeld Oprachtnemer dat onverwijld aan de KB.

2.7.3 Niet tijdig leveren Batch

Indien Oprachtnemer niet tijdig een Batch heeft opgeleverd, zal Oprachtnemer in overleg met de KB een nieuwe datum van oplevering vaststellen, zonder dat die nieuwe datum de einddatum zoals opgenomen in de planning volgens het Projectplan mag overschrijden. Die vaststelling voor een nieuwe datum laat onverlet dat sprake is van een tekortkoming in de nakoming van de verplichting van Oprachtnemer tot tijdige oplevering, waaruit de gevolgen voortvloeien zoals vastgelegd in de Overeenkomst.

2.7.4 Leverwijze Batch

Oprachtnemer is verplicht iedere Batch afzonderlijk op een externe harddisk en in vaste directorystructuur met een vaste set metadata op te leveren. In de toekomst gaat dit mogelijk via een andere manier in overleg met de leverancier, zie ook hoofdstuk 4.

Daarbij neemt Oprachtnemer de volgende cumulatieve vereisten in acht dat:

- o één Batch niet over meerdere harde schijven mag worden verdeeld;
- o één Batch niet meer opslagruimte dan 495 GB mag innemen;
- o één of meer Batches per harde schijf zijn toegestaan;
- o de harde schijven een USB 3.0 of eSATA interface moeten bevatten.

2.7.5 Documentatie Batch

Indien op enig tijdstip tijdens het gebruik van een goedgekeurde Batch door de KB (zoals beschreven in het volgende artikel) blijkt dat de daarbij behorende documentatie onjuist, onvolledig of niet meer actueel is, dan zal Oprachtnemer op eerste verzoek van de KB ervoor zorgen dat de door hem geleverde documentatie zo spoedig mogelijk, en op zijn kosten, zal worden aangepast of vervangen.

2.8 Acceptatie van een Batch

2.8.1 Acceptatie van een Batch tijdens Pilotfase

De eerste Batch van een Project geldt als een Pilotbatch. Doel van deze pilotfase is te beoordelen in hoeverre Oprachtnemer in staat is om de Opdracht overeenkomstig de eisen van deze Overeenkomst

uit te voeren. De beoordeling van deze Pilotbatch vindt plaats binnen de termijn zoals opgenomen in de Projectaanvraag.

2.8.2 Afkeuring eerste Pilotbatch

Bij afkeuring door de KB van de eerste Pilotbatch krijgt Opdrachtnemer de mogelijkheid om die afgekeurde Pilotbatch te corrigeren, binnen de termijn zoals opgenomen in de Projectaanvraag. Indien die gecorrigeerde Pilotbatch opnieuw door de KB wordt afgekeurd, dan kan de KB gebruik maken van haar bevoegdheid als bedoeld in het tiende lid van dit artikel. Voor een afgekeurde Pilotbatch kan de KB geen Boete berekenen.

2.8.3 Afkeuring tweede Pilotbatch

Nadat de eerste Pilotbatch door de KB wordt goedgekeurd, dient Opdrachtnemer een tweede Pilotbatch op te leveren. Bij afkeuring door de KB van de tweede Pilotbatch, kan de KB besluiten het contract per direct en zonder consequenties voor de KB te beëindigen.

2.8.4 Reguliere acceptatie van een Batch

Na oplevering door Opdrachtnemer van een Batch, beoordeelt de KB binnen de termijn zoals opgenomen in het Projectplan of die Batch voldoet aan de eisen voortvloeiend uit de Overeenkomst. Bij die beoordeling mag de KB betrekken dat de kwaliteit van de Batch ten minste gelijk is aan de kwaliteit van de testset, behorende bij de Inschrijving. Die beoordeling vindt steekproefsgewijs plaats, waarbij de omvang en inhoud van die steekproef voor iedere Batch afzonderlijk door de KB wordt bepaald. Uitkomst van die beoordeling is de acceptatie van die Batch of de afkeuring van die Batch.

2.8.5 Gebrek in Batch

Mocht uit de in het vorige artikellid bedoelde steekproef blijken dat sprake is van een Gebrek, dan is de KB bevoegd de gehele Batch af te keuren, aan welke afkeuring Opdrachtnemer is gebonden.

2.8.6 Afkeuring Batch

De acceptatie van de KB van een andere batch dan de ter beoordeling voorliggende Batch is niet van invloed op de bevoegdheid van de KB tot het al dan niet afkeuren van die Batch. De KB is dus bevoegd tot afkeuring vanwege een Gebrek geconstateerd tijdens een steekproef, ongeacht of de omstandigheid die tot afkeuring leidt zich eerder heeft voorgedaan en ongeacht of dat door de KB is opgemerkt. Een eenmaal geaccepteerde Batch blijft onaangetast.

2.8.7 Rapportage

De KB stuurt een rapport met bevindingen over de uitkomst van de beoordeling van een Batch aan Opdrachtnemer. Indien de KB een Batch afkeurt, vermeldt de KB de reden(en) voor afkeuring, bestaande uit een beschrijving van de door de KB geconstateerde Gebreken.

2.8.8 Acceptatietermijn voor één Batch

De acceptatietermijn voor één Batch is ten minste vijftien (15) werkdagen, ook indien een kortere termijn uit het Projectplan zou kunnen volgen. De acceptatietermijn voor een Batch vangt aan nadat de acceptatietermijn van de Batch die daarvoor is opgeleverd, is geëindigd, ook indien dat niet uit het Projectplan blijkt. De data van oplevering van Batches zoals opgenomen in het Projectplan wijzigen niet, door toepassing van de regels over de acceptatietermijn volgens de vorige volzinnen.

2.8.9 Herlevering

Indien de KB een Batch afkeurt, dient Opdrachtnemer de gehele Batch opnieuw op te leveren, met herstel van de gebreken. De datum van oplevering van een gecorrigeerde Batch wordt in onderling overleg tussen de KB en Opdrachtnemer vastgesteld, welke datum uiterlijk is gelegen twee (2) maanden na verzending van het bericht van de KB inhoudende afkeuring van die Batch. Voor de beoordeling van een gecorrigeerde Batch geldt een acceptatietermijn van vijftien (15) werkdagen. Bij oplevering van een gecorrigeerde Batch, gebruikt Opdrachtnemer hetzelfde batchnummer, maar met ander versienummer.

2.8.10 Correctie door een derde

Indien de KB een Batch voor de tweede keer of daarop volgende keer afkeurt, dan is de KB bevoegd, onverminderd zijn verdere rechten, om de in de Batch geconstateerde Gebreken door een derde te laten corrigeren. Opdrachtnemer is verplicht hieraan zijn medewerking te verlenen en zo nodig daarvoor

benodigde informatie op eerste verzoek te verstrekken aan die derde. De kosten die hiermee gemoeid zijn komen voor rekening van Opdrachtnemer.

2.8.11 Bewaarplicht Batch

Opdrachtnemer is verplicht om van ieder geaccepteerde Batch ten minste 1 kopie van het bestand van die Batches te bewaren gedurende ten minste zes (6) maanden, ingaande vanaf de datum van verzending van het bericht van de KB inhoudende acceptatie van die Batch. Gedurende deze periode is Opdrachtnemer gehouden op eerste verzoek van de KB een kopie van die Batch kosteloos aan de KB beschikbaar te stellen.

Partijen stellen vast dat de afspraken over intellectuele eigendomsrechten zoals vastgelegd in de Overeenkomst, onverkort van toepassing zijn op het bestand met de Batches die Opdrachtnemer bewaart.

3 Algemene instructies en instructies per materiaalsoort

De instructies en opmerkingen in dit document m.b.t. het scannen zijn uiteraard niet van toepassing op digitale bestanden.

3.1 Masterbestanden

Masterbestanden worden in grayscale of kleur opgeleverd door opdrachtnemer. Dit wordt in de Projecten aangegeven. Accessbestanden 'volgen' de masterbestanden: wordt het masterbestand in kleur opgeleverd, dan geldt dat ook voor het accessbestand.

3.2 Volledigheid

Alle pagina's uit een krant of tijdschrift moeten volledig en in juiste volgorde (conform origineel) worden gescand, tenzij anders aangegeven door de KB.

3.3 Blanco pagina's

Bij tijdschriften moeten blanco pagina's en omslagen van banden worden meegescand, tenzij anders aangegeven door de KB. Bij kranten worden alleen de issues zelf gescand, blanco's en de omslagen van de banden worden niet gescand.

3.4 Segmentatie

Bij kranten wordt altijd op artikelniveau ontsloten (gesegmenteerd). Bij tijdschriften kan er op artikel- of paginaniveau ontsloten worden. Wat gewenst is, zal per Project aangegeven worden.

3.5 Auteurherkenning

Bij tijdschriften kan er, wanneer er op artikelniveau ontsloten wordt, gevraagd worden om per artikel de auteur(s) te herkennen en deze op te nemen in de metadata. Wat gewenst is, zal per Project aangegeven worden.

3.6 Supplementen

Bij krantennummers kan het voorkomen dat er één of meer supplementen bij horen. De supplementen moeten tijdens het digitaliseren aan het juiste krantennummer worden toegevoegd. Dat wil zeggen dat er één digitaal krantennummer wordt geleverd, bestaande uit het hoofdnummer en alle eventuele supplementen. In het origineel zijn de supplementen vaak direct achter het krantennummer in de band ingebonden, maar het kan ook voorkomen dat ze aan het einde van de band zijn opgenomen of in een andere band. In het door de KB aangeleverde XML-bestand staat opgenomen of er sprake is van een supplement en zo ja, bij welk krantennummer het supplement hoort. Dit kan tijdens het digitaliseren gebruikt worden om de supplementen samen te voegen met de juiste krant.

3.7 Aparte afleveringen bij tijdschriften

Bij tijdschriften worden losse inhoudsopgaven, indexen, omslagen en bijlagen als aparte afleveringen behandeld.

3.8 Per pagina

Kranten worden altijd per pagina gescand: ook als de tekst van krantenartikelen doorloopt over meerdere pagina's wordt er toch een afzonderlijke scan per pagina gemaakt (dus geen scans "per opening"). Bij tijdschriften moet er per opening gescand worden indien een artikel/afbeelding over twee pagina's loopt.

4 Teruglevering van materiaal en metadata door opdrachtnemer

De resultaten worden teruggeleverd in batches. Een batch is een hoeveelheid gedigitaliseerd materiaal die door de opdrachtnemer aan de KB wordt aangeleverd.

Het digitale materiaal in een batch moet identiek zijn aan het materiaal van het Transport aangeleverd door de KB (Transport in = batch out). Een Transport kan bij teruglevering uit meerdere batches (delen) bestaan.

De batches worden in overleg aangeleverd op een harddisk van de opdrachtnemer.

Deze harddisk zal na goed- of afkeuring retour geleverd worden aan de opdrachtnemer.

Alle digitale bestanden dienen door leverancier per Zending in één keer aangeleverd te worden.

Nadat een batch door de KB geaccepteerd is dient leverancier de bestanden nog tenminste zes maanden te bewaren. Leverancier vraagt via het KB sharepoint of en wanneer de bestanden verwijderd kunnen worden.

In de toekomst gaat teruglevering mogelijk verlopen via een andere manier (digitaal in plaats van fysiek), in overleg met de leverancier.

4.1 Directorystructuur en naamgeving

4.1.1 Naamgeving en omvang batches

De naam van de batch bestaat uit een deelprojectcode gevolgd door het nummer van de batch (9 cijfers), het deelnummer van de batch (1 cijfer) en het versienummer van de batch (2 cijfers), gescheiden door underscores. MMKB03_000000009_1_04 is dus de vierde versie van het eerste deel van batch 9 voor het deelproject MMKB03. Een gedigitaliseerd object (barcode band) mag niet over meerdere (deel)batches verdeeld zijn.

De basis voor de batchnaam (=deelprojectcode plus het 9-cijferige batchnummer) wordt door de KB aangeleverd in de metadatadump. Deel- en versienummers moet de opdrachtnemer toevoegen.

Een batch mag niet meer materiaal bevatten dan 495 GB. Wanneer het materiaal van één aangeleverde batch een terug te leveren batch zou opleveren die groter is dan 495 GB, moet deze batch worden gesplitst in twee batches met deelnummers 1 en 2: MMKB03_000000009_1_01 en MMKB03_000000009_2_01.

NB: de KB kan in specifieke projecten de hier genoemde maximumgrootte naar beneden bij te stellen.

Wanneer een batch om wat voor reden dan ook opnieuw wordt aangeleverd moet het versienummer voor de gecorrigeerde batch met 1 worden opgehoogd. Alleen de afgekeurde batch moet opnieuw geleverd te worden. De nieuwe versie van de batch moet precies dezelfde bestanden bevatten als de voorgaande versie, tenzij het ontbreken of teveel aanwezig zijn van bestanden een reden was om de batch af te keuren.

Het niet ophogen van een versienummer bij een herlevering leidt tot directe afkeuring en de nieuwe batch moet binnen 48 uur opnieuw geleverd worden.

4.1.2 Object identifiers

Elk gedigitaliseerd object heeft een unieke identifier, die gebaseerd is op de in de metadatadump aangeleverde barcode. Voor **tijdschriften** en **kranten** eindigt de barcode op drie nullen. Deze drie nullen worden vervangen door het volgnummer van de aflevering binnen de band:

Barcode: MMKB03002256000

Objectidentifiers: MMKB03_002256001, MMKB03_002256002 etc.

Deze object identifiers worden door de KB gegenereerd, en staan in de METS-bestanden die door de KB worden aangeleverd.

4.1.3 Directories

Een batch is volgens een vaste structuur ingedeeld in directories. Hierbij wordt gebruik gemaakt van standaard benamingen voor bestandstypen.

master	Jp2 bestand voor duurzame opslag
targets-jp2_master	Targets volgens de jp2 master specificaties
targets-jp2_access	Targets volgens de jp2 access specificaties
access	Jp2 bestand voor ontsluiting op het web in lagere resolutie
alto	Text/layout/coördinaten van scan in ALTO 2.0 formaat (XML)
text	Fulltext volgens het KB fulltext formaat (XML)
mets	METS-bestand (XML)
concordantietabel	Tekstbestand
pdf	PDF-bestand
checksums	Checksumbestanden
pakbon	Pakbonbestand (XML)

De specificatie van de directorystructuur is als volgt:

- [batchnaam]
 - o [object-id]
 - o [bestandstype]
 - o targets-jp2_master
 - o [targetset]
 - o targets-jp2_access
 - o [targetset]
 - o concordantietabel
 - o checksums
 - o pakbon

Directory Technische targets

De subdirectories voor targetsets kennen de volgende naamgeving:

targetset_[datum(yyymmdd)]_[shiftnummer (2 cijfers)]_[cameranummer(2 cijfers)]

Voorbeeld: targetset_20101005_01_05.

Hanteer de volgende shiftnummering:

- 01= ochtendshift
- 02= middagshift
- 03= avondshift
- 04= nachtshift

Een voorbeeld van de indeling is:

- o MMKB03_00000006_1_01
 - o MMKB03_000006127
 - o master
 - o access
 - o mets
 - o alto
 - o text
 - o pdf
- o MMKB03_000006128
 - o master
 - o access
 - o ... etc.
- o targets-jp2_master
 - o targetset_20101005_01_05
 - o targetset_20101005_02_05
- o targets-jp2_access
 - o targetset_20101005_01_05
 - o targetset_20101005_02_05
- o concordantietabel

- o checksums
- o pakbon

4.1.4 Bestandsnamen

Voor bestandsnamen gelden de volgende eisen:

- Bestandsnamen moeten altijd beginnen met de objectidentificer.
- In de bestandsnamen moeten volgnummers gebruikt worden. Deze volgen na de identificer en een underscore. Een volgnummer is alleen nodig voor bestanden die een pagina of een artikel bevatten, niet voor bestanden die een heel document (aflevering) bevatten. Dus bv. wel voor master- en accessbestanden maar niet voor een PDF- of metadatabestand dat het hele document bevat.
- Na identificer en eventueel volgnummer volgt in de bestandnaam een underscore een aanduiding van het bestandstype (kleine letters).
- Tenslotte volgt de extensie, volgens de gebruikelijke aanduidingen van het bestandstype (jp2, xml, pdf).

Voorbeelden:

- | | | |
|---|--------|---------------------------------------|
| - | master | bv. MMKB03_123456789_00001_master.jp2 |
| - | access | bv. MMKB03_123456789_00001_access.jp2 |
| - | alto | bv. MMKB03_123456789_00001_alto.xml |
| - | pdf | bv. MMKB03_123456789_pdf.pdf |
| - | mets | bv. MMKB03_123456789_mets.xml |
| - | text | bv. MMKB03_123456789_00001_text.xml |

Voor een aantal bestanden gelden andere regels:

4.1.5 Checksumbestanden

- Batchnaam
- De naam van het bestandstype waar het checksumbestand betrekking op heeft, nl. "mets", "concordantietabel", "pakbon" of "targets".
- De extensie .csv

Voorbeelden:

- MMKB03_000000001_1_01_mets_sha512.csv
- MMKB03_000000001_1_01_targets_sha512.csv
- MMKB03_000000001_1_01_pakbon_sha512.csv
- MMKB03_000000001_1_01_concordantietabel_sha512.csv

4.1.6 Concordantietabel

- [batchnaam].csv

Bijvoorbeeld: MMKB03_000000001_1_01.csv

4.1.7 Technische targets

- Voor targets no 1,2 en 4:

[deelprojectcode]_(TON/COL/GEO)_[datum(yyymmdd)]_[shift_nummer(2cijfers)]_[cameranummer(2 cijfers)].extensie

Bijvoorbeeld: MMKB03_TON_20101005_01_05.jp2

- Voor target no 3:

[deelprojectcode]_MTF_(GRAY/RGB)_[datum(yyymmdd)]_[shift_nummer(2cijfers)]_[cameranummer(2 cijfers)].extensie

Bijvoorbeeld: MMKB03_MTF_GRAY_20101005_01_05.jp2
MMKB03_MTF_RGB_20101005_01_05.jp2

4.1.8 Pakbon

- [batchnaam]_pakbon.xml

Bijvoorbeeld: MMKB03_00000001_1_01_pakbon.xml

4.2 Checksums

Om de integriteit van de data te controleren hanteert de KB checksums voor de individuele bestanden: voor alle aangeleverde bestanden in een batch moet een checksum (SHA-512, hexadecimaal) worden aangeleverd.

De meeste checksums worden in de METS-bestanden opgenomen. Alleen de checksums van de technische targets, van de METS-bestanden zelf, van de concordantietabel en van de pakbon worden in losse checksumbestanden geleverd, die in een afzonderlijke directory “checksums” worden gezet.

NB: er wordt per batch één checksumbestand voor alle technische-targetbestanden geleverd, één checksumbestand voor alle METS-bestanden.

Het formaat van het checksumbestand is csv (comma separated values). Het checksumbestand bevat op iedere regel eerst een bestandsnaam en vervolgens de checksum (SHA-512). Als separator wordt een puntkomma gebruikt (semicolon “;”). Bij de bestandsnaam wordt het relatieve directorypad (ten opzichte van de checksumfile) opgenomen, dus beginnend met ../

Er mogen geen andere regels, commentaarregels of overbodige karakters zoals spaties in het bestand staan.

Er mag geen BOM (Byte Order Mark) in de csv bestanden staan.

Van de csv-checksumbestanden zelf hoeven geen afzonderlijke checksums te worden geleverd.

4.3 Technische specificaties en beeldkwaliteit

4.3.1 Algemene specificaties

Masterbestanden dienen in grayscale of in kleur aangeleverd te worden (gespecificeerd per Project).

Masterbestanden dienen in JP2 formaat aangeleverd te worden.

De sampling rate van masterbestanden is afhankelijk van het formaat van de originelen en van de lettergrootte (zie specificaties hieronder).

Accessbestanden dienen in grayscale of in kleur aangeleverd te worden: accessbestanden volgen hierin de masterbestanden.

Accessbestanden dienen in JP2 formaat aangeleverd te worden.

De sampling rate van accessbestanden is gelijk aan die van de masterbestanden.

4.3.2 Scaninstructies

- Pagina's dienen in de volgorde van het origineel te worden gescand;
- Er wordt zó gecropt dat het hele origineel zichtbaar blijft. Rondom het origineel is tussen de 10 en 20 pixels ruimte zichtbaar;
- Indien de tekst (of afbeelding) binnen een pagina meer dan 3 graden scheef staat moet de tekst rechtgezet worden, ongeacht de positionering van het beeld;
- Er moet geroteerd worden als meer dan 50% van de tekst op de pagina verticaal uitgelijnd is;
- Schaduw in de kneep moet voorkomen worden;
- Artefacten mogen niet voorkomen evenals clipping;
- Opnames op egale zwarte achtergrond, tenzij zeer doorzichtig origineel dan andere kleur achtergrond in overleg met opdrachtgever. Bij kleinere pagina's achterliggende gedrukte pagina onzichtbaar maken;
- Wanneer een object niet verder open mag dan 110° wordt dit door de KB aangegeven in de metadatadump.

4.3.3 Technische specificaties master- en accessbestanden

Specificaties Beeldkwaliteit Richtlijnen [Preservation Imaging Metamorfoze versie 1.0 \(2012\)](#)

- De KB gaat uit van de algemeen geldende theoretische L*a*b* waarden van de Colorchecker SG en de Q13 Kodak grijstrap bij de beeldkwaliteit controles;
- Indien er een afwijking boven de tolerantiewaarde voor één of meer van de beeldkwaliteitscriteria wordt gevonden, worden de meetresultaten van de leverancier opgevraagd, waar mogelijk inclusief de referentiewaarden van de eigen technische testkaarten;
- **NB:** Voor de Technische Targets gelden de JP2 Masterspecificaties.

Nr	Criteria / Omschrijving	Bepaling / tolerantie
0	Formaat parameters (toegelicht in bijlage B)	
	File format	JP2 (JPEG 2000 Part 1)
	Compression type	Master: Lossless (reversible 5-3 wavelet filter) Access: Lossy (irreversible 9-7 wavelet filter)
	Colour transform	Yes (only for colour images)
	Number of decomposition levels	5
	Progression order	RPCL
	Tile size	1024 x 1024
	Code block size	64 x 64 (2 ⁶ x 2 ⁶)
	Precinct size	256 x 256 (2 ⁸) for 2 highest resolution levels; 128 x 128 (2 ⁷) for remaining resolution levels
	Number of quality layers	Master: 11 Access: 8
	Target compression ratio layers	Master: 2560:1 [1]; 1280:1 [2]; 640:1 [3]; 320:1 [4]; 160:1 [5]; 80:1 [6]; 40:1 [7]; 20:1 [8]; 10:1 [9]; 5:1 [10]; - [11] Access: 2560:1 [1]; 1280:1 [2]; 640:1 [3]; 320:1 [4]; 160:1 [5]; 80:1 [6]; 40:1 [7]; 20:1 [8]
	Error resilience	Start-of-packet headers; end-of-packet headers; segmentation symbols
	Sampling rate	Stored in JP2's "Capture Resolution" fields
	ICC profiles	Embedded using "Restricted ICC" method; original profile from source TIFF should not be modified by encoder
	Capture metadata	Embedded as XMP metadata in JP2 XML box
Specs reference	<i>(embedded as a codestream comment)</i> Master: KB_MASTER_LOSSLESS_01/01/2015 Access: KB_ACCESS_LOSSY_01/01/2015	
1	Kleurruimte	Adobe RGB (1998), Gray Gamma 2,2. Voor grayscale bestanden is het vereiste icc profiel Gray Gamma 2,2.. Adobe RGB (1998) is de vereiste capture kleurruimte indien er een digitaliseringsworkflow wordt gehanteerd waarbij RGB wordt omgezet in grayscale.
2	Sampling rate	De vereiste sampling rate is afhankelijk van het formaat van het origineel en van de grootte van de letter:

Nr	Criteria / Omschrijving	Bepaling / tolerantie
		<p>Formaat</p> <p>≥ A6 (formaat v/d pagina): 300 ppi</p> <p>< A6 (formaat v/d pagina): 600 ppi</p> <p>Indien formaat origineel ≥ A1, dan wordt ppi in overleg met de KB vastgesteld.</p> <p>Lettergrootte</p> <p>kleine letter “e “ < 1 mm: 600 ppi</p>
3	Bitdiepte per kanaal	8 bits
4	Witbalans & tonale weergave	<p>De witbalans wordt uitgedrukt in ΔC^* met als tolerantie $\Delta C^* \leq 2$</p> $\Delta C^*_{ab(i)} = \sqrt{a^{*2}_{s(i)} + b^{*2}_{s(i)}} - \sqrt{a^{*2}_{ref(i)} + b^{*2}_{ref(i)}}$ <p>Toelichting</p> <p>Indien de witbalans niet goed is afgesteld kan er kleurzweem optreden.</p> <p>De tolerantie voor de witbalans $\Delta C^* \leq 2$ wordt gemeten in de neutrale vakjes van de Digital ColorChecker SG.</p> <p>Niveau Metamorfoze Light: tot en met vakje K8.</p> <p>De toleranties voor de kleurzweem zijn niet van toepassing op grayscale bestanden.</p>
5	Belichtingstolerantie.	<p>De belichtingstolerantie wordt uitgedrukt in ΔL^*, met als tolerantie $\Delta L^* \leq 2$</p> $\Delta L^*_{(i)} = L^*_{s(i)} - L^*_{ref(i)}$ <p>Toelichting</p> <p>Voor een juiste tonale weergave moet de belichting optimaal zijn.</p> <p>De tolerantie voor de belichtingstolerantie $\Delta L^* \leq 2$ wordt gemeten in de neutrale vakjes van de Digital ColorChecker SG.</p> <p>Niveau Metamorfoze Light: tot en met vakje K8.</p>
4 & 5	Witbalans en Belichting aanvullend	$\Delta E^* \leq 2,83$

Nr	Criteria / Omschrijving	Bepaling / tolerantie
		$\Delta E^* = \sqrt{(L_s - L_{ref})^2 + (a_s - a_{ref})^2 + (b_s - b_{ref})^2}$ <p>Toelichting</p> <p>Voor de beoordeling van de belichting en witbalans wordt aanvullend naar de ΔE^* waarden in de neutralen gekeken. De tolerantiewaarde voor witbalans en belichting, uitgedrukt in ΔE^* in de neutralen is $\Delta E^* \leq 2,83$. Voor deze berekening wordt de formule CIE 1976 gebruikt.</p> <p>De tolerantie voor de belichtingstolerantie $\Delta E^* \leq 2,83$ wordt gemeten in de neutrale vakjes van de Digital ColorChecker SG.</p> <p>Niveau Metamorfoze Light: tot en met vakje K8.</p>
6	<p>Gain modulation De gain modulation wordt beoordeeld in de hoge lichten.</p>	<p>De gain modulation wordt uitgedrukt in procenten: 0,8 – 1,08 80% - 108%</p> <p>Toelichting</p> <p>Om te zorgen dat informatie in de hoge lichten (de helderste gedeeltes binnen een opname) niet verloren gaat moet de gain modulation minimaal 80% tot maximaal 108% zijn.</p> <p>De gain modulation wordt berekend over vakje E5-J6 van de Digital ColorChecker SG. Bij de technische controles van de KB wordt deze berekend met behulp van de visueel gewogen OECF (Opto Electronic Conversion Function) Y waarden.</p>
7	<p>Ruis</p>	<p>Ruis wordt omschreven als een standaardafwijking. De hoeveelheid wordt uitgedrukt in pixelwaarden: STD ≤ 4</p> <p>Toelichting</p> <p>Ruis is informatie in een afbeelding, die geen deel uitmaakt van het originele object.</p> <p>Bij de technische controles van de KB wordt de ruis beoordeeld in de neutrale vakjes van de Digital ColorChecker SG</p> <p>Niveau Metamorfoze Light: tot en met vakje K8.</p> <p>Niveau Metamorfoze: tot en met vakje E6</p>

Nr	Criteria / Omschrijving	Bepaling / tolerantie								
8	Uitlichting	<p>De uitlichtingstoleranties zijn afhankelijk van de grootte van het opnamevlak. De toleranties zijn uitgedrukt in 8 bit pixelwaardes.</p> <p>Om clipping te voorkomen dienen de maximale pixelwaardes onder of gelijk aan 248px te zijn over de gehele uitlichting.</p> <table border="1" data-bbox="304 499 603 656"> <tr> <td data-bbox="304 499 603 533">≤ DIN A3</td> <td data-bbox="608 499 1407 533">8 bit pixelwaarde Adobe RGB (1998) ≤ 8,</td> </tr> <tr> <td data-bbox="304 533 603 566">> DIN A3 en ≤ DIN A2</td> <td data-bbox="608 533 1407 566">8 bit pixelwaarde Adobe RGB (1998) ≤ 12,</td> </tr> <tr> <td data-bbox="304 566 603 600">> DIN A2 en ≤ DIN A1</td> <td data-bbox="608 566 1407 600">8 bit pixelwaarde Adobe RGB (1998) ≤ 14,</td> </tr> <tr> <td data-bbox="304 600 603 656">> DIN A1 en ≤ DIN A0</td> <td data-bbox="608 600 1407 656">8 bit pixelwaarde Adobe RGB (1998) ≤ 18,</td> </tr> </table>	≤ DIN A3	8 bit pixelwaarde Adobe RGB (1998) ≤ 8,	> DIN A3 en ≤ DIN A2	8 bit pixelwaarde Adobe RGB (1998) ≤ 12,	> DIN A2 en ≤ DIN A1	8 bit pixelwaarde Adobe RGB (1998) ≤ 14,	> DIN A1 en ≤ DIN A0	8 bit pixelwaarde Adobe RGB (1998) ≤ 18,
≤ DIN A3	8 bit pixelwaarde Adobe RGB (1998) ≤ 8,									
> DIN A3 en ≤ DIN A2	8 bit pixelwaarde Adobe RGB (1998) ≤ 12,									
> DIN A2 en ≤ DIN A1	8 bit pixelwaarde Adobe RGB (1998) ≤ 14,									
> DIN A1 en ≤ DIN A0	8 bit pixelwaarde Adobe RGB (1998) ≤ 18,									
		<p>Toelichting</p> <p>Bij de technische controles van de KB wordt de dagtargetopname van de uitlichting gemeten op een aantal punten binnen het opnamevlak met een selectiekader van 13x13 pixels, het aantal kaders is afhankelijk van de grootte:</p> <ul style="list-style-type: none"> ≤ DIN A3 – 9 punten > DIN A3 en ≤ DIN A2 – 18 punten > DIN A2 en ≤ DIN A1 – 36 punten > DIN A1 en ≤ DIN A0 – 72 punten <p>De waarden uitgedrukt in de visueel gewogen OECF (Opto Electronic Conversion Function) Y worden onderling vergeleken tegen de tolerantiewaardes.</p>								
9	Kleurnauwkeurigheid	<p>Gemiddelde $\Delta E \leq 5$</p> <p>Maximale $\Delta E \leq 18$</p> $\Delta E^* = \sqrt{(L_s - L_{ref})^2 + (a_s - a_{ref})^2 + (b_s - b_{ref})^2}$ <p>Toelichting</p> <p>De volgende tolerantiewaarden gelden voor alle vakken van de Digital ColorChecker SG.</p> <p>Voor deze berekening wordt de formule CIE 1976 gebruikt.</p> <p>Wit van het ingesloten profiel wordt als witpunt gebruikt.</p>								
10	Vershil tussen claimed sampling rate & Obtained sampling rate	<p>Toelichting</p> <p>Met het begrip obtained sampling rate wordt de echte, of beter gezegd, de werkelijke sampling rate van het bestand bedoeld. Obtained betekent verkregen.</p> <p>De obtained sampling rate wordt vastgesteld door het uitvoeren van een MTF meting (Modulatie Transfer Function, Spatial</p>								

Nr	Criteria / Omschrijving	Bepaling / tolerantie
		<p>Frequency Respon) met behulp van de dagtarget opname van de MTF testkaarten.</p> <p>De controle van de KB vergelijkt de in de metadata vermelde sampling rate met de echte, werkelijke sampling rate. Deze mogen niet meer dan 2% van elkaar verschillen.</p>
11	Sampling Efficiency	<p>≥ 85%</p> <p>Toelichting Op MTF10 wordt de maximale scherpte beoordeeld, het scheidend vermogen, beoordeeld in het helderheidssignaal Y.</p> <p>Op MTF10 wordt beoordeeld in het helderheidssignaal Y. De scherpte wordt uitgedrukt in lijnenpaar per mm (lp/mm). Het minimale vereiste aantal lp/mm op MTF10 is afhankelijk van de sampling rate. Bij het digitaliseren met een sampling rate van 300 ppi is een minimaal vereiste van 5 lp/mm vereist. Dit komt overeen met een sampling efficiency van 85%.</p>
12	MTF50	<p>Niveau Metamorfoze Light: ≥ 45% van de vereiste MTF10 performance</p> <p>Toelichting Op MTF50 wordt de contourscherpte beoordeeld.</p> <p>Afhankelijk van de sampling rate is voor MTF50 een minimale prestatie, uitgedrukt in lp/mm, voorgeschreven:</p> <ul style="list-style-type: none"> • Voor een sampling rate van 300 ppi betekent dit: MTF10 ≥ 5 lp/mm. MTF50 ≥ 2,25 lp/mm (niveau Metamorfoze Light) en ≥ 2,50 (niveau Metamorfoze). • Voor een sampling rate van 600 ppi betekent dit: MTF10 ≥ 10 lp/mm. MTF50 ≥ 4,5 lp/mm (niveau Metamorfoze Light)
13	Modulatie (verscherping)	<p>≤ 1,05</p> <p>Verscherping is <u>niet</u> toegestaan in het masterbestand.</p>
14	Kleurmisregistratie. De kleurmisregistratie wordt per kleurkanaal beoordeeld	<p>< 0,50 pixel per kleurkanaal.</p> <p>Toelichting Kleurmisregistratie kan zichtbaar zijn in contrastrijke delen in een afbeelding door een gekleurd randje.</p> <p>De KB controleert op kleurmisregistratie tijdens het uitvoeren van een MTF meting in de opnames van de MTF testkaarten.</p> <p>De tolerantie voor kleurmisregistratie is niet van toepassing op grayscale bestanden.</p>
15	Geometrische vervorming	<p>≤ 2%</p> <p>Toelichting</p>

Nr	Criteria / Omschrijving	Bepaling / tolerantie
		Alle soorten van vervorming veranderen de verhouding tussen de horizontale en verticale lijnen in het beeld en moeten voorkomen worden. De tolerantie voor geometrische vervorming is een verandering van deze verhouding van maximaal 2%. Geometrische vervorming wordt door de KB gemeten in de dagtargetopname van het QA-2 Image Evaluation Test Target.
16	Artefacten	Artefacten zijn niet toegestaan. Artefacten zoals: moire, newtonringen, reflecties, weerspiegelingen, horizontale strepen, verticale strepen, pixelverstoringen, haloing, verspringingen, golfbewegingen, krommingen e.d. kunnen enkel visueel worden vastgesteld bij een beoordeling van het digitale bestand. De digitaliseerder dient dit zoveel mogelijk op te vangen door het gebruik van de juiste opnameapparatuur.
17	Schaduw in de kneep.	Schaduw in de kneep moet voorkomen worden.

4.3.4 Belichtingstolerantie Kodak Gray Scale

Overzicht van de theoretische¹ 8 bit pixelwaarden in Adobe RGB (1998) met een belichtingstolerantie van ΔL^* 2. Technische testkaarten kunnen onderling afwijken. Dit wordt beoordeeld op target 2 ("Tweede opname: tonale weergave, witbalans en kleurbetrouwbaarheid").

Vak	L*	8 bit pixelwaarde Adobe RGB (1998)	Vak	L*	8 bit pixelwaarde Adobe RGB (1998)
A	97,63	248	10	37,82	89
	95,63	242		35,82	85
	93,63	236		33,82	81
1	89,39	224	11	33,99	81
	87,39	218		31,99	77
	85,39	212		29,99	72
2	81,75	202	12	30,45	73
	79,75	196		28,45	69
	77,75	191		26,45	65
3	74,68	182	13	27,16	66
	72,68	177		25,16	62
	70,68	171		23,16	58
4	68,12	165	14	24,12	60
	66,12	159		22,12	56
	64,12	154		20,12	52
5	62,06	149	15 ²	21,31	54
	60,06	143		19,31	50

¹ De 8 bit pixelwaarden in de tabel zijn gebaseerd op de theoretische reflectiewaarden van de grijstrap van de Kodak Gray Scale en de Metamorfoze belichtingstolerantie van ΔL^* 2. De reflectiewaarden en dus ook de $L^*a^*b^*$ waarden van grijstrappen van testkaarten kunnen onderling afwijken.

² Te digitaliseren materiaal bestaat overwegend uit gedrukte tekst. De KB acht het toereikend om witbalans, belichting en ruis te beoordelen t/m vak 15 van de Kodak Gray Scale.

	58,06	138		17,31	47
	56,55	134		18,70	49
6	54,55	129	16	16,70	45
	52,55	124		14,70	42
	51,24	121		16,28	45
7	49,24	116	17	14,28	41
	47,24	111		12,28	37
	46,42	110		14,04	40
8	44,42	105	18	12,04	37
	42,42	100		10,04	33
	41,95	99		11,97	37
9	39,95	94	19	9,97	33
	37,95	90		7,97	30

4.3.5 Dagtargets, naamgeving, volgorde en beeldopbouw

Alle opnamen van de technische targets moeten als jp2 te worden aangeleverd.

Om te kunnen monitoren wat er gebeurt bij het generen van de JP2 master en access bestanden moeten de dagtargets zowel in een set met master en in een set met access JP2 specificaties te worden aangeleverd.

De dagtargets moeten door dezelfde workflow gaan als de master- en accessbestanden.

De sampling rate bij het opnemen van de dagtargets moet identiek zijn aan de sampling rate die gehanteerd wordt bij het opnemen van de originelen.

Alle technische testkaarten moeten op een zwarte ondergrond worden gedigitaliseerd.

In de loop van het project kunnen mogelijk bepaalde targets worden vervangen en/of worden toegevoegd.

Witbalans, belichting en ruis worden beoordeeld t/m vak 15 van de Kodak Gray Scale. Hierbij zijn de waarden gemeten in Tweede opname: tonale weergave, witbalans en kleurbetrouwbaarheid (COL) leidend.

Bij alle hieronder besproken opnamen van technische testkaarten moet onderin beeld, of boven in beeld een Kodak Gray Scale of gelijkwaardig worden opgenomen.

De beeldopbouw en volgorde van de dagtargets voor bestanden in kleur is als volgt:

- **Eerste opname: tonale weergave en uitlichting.**

De Q-13 wordt op een wit vel karton beeldvullend opgenomen. De opzichtsensiteit van het witte karton moet liggen tussen de 0.05 en 0.15.



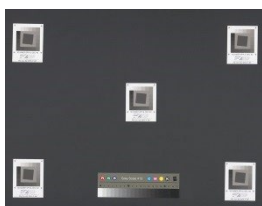
- **Tweede opname: tonale weergave, witbalans en kleurbetrouwbaarheid.**

Aansluitend aan de Q-13 wordt een zwart vel karton beeldvullend opgenomen. In het midden van het beeldvlak, op het zwarte karton, ligt een kleurenkaart: de Digital Color Checker SG (140 velden).



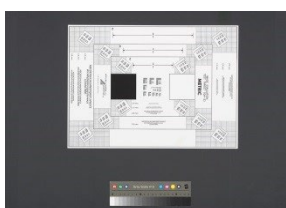
- **Derde opname: tonale weergave en MTF meting.**

Aansluitend aan de Q-13 wordt een zwart vel karton beeldvullend opgenomen. In het midden van het beeldvlak, op het zwarte karton, ligt een QA-62 en in elke hoek van het beeldvlak ligt een QA-62.



- **Vierde opname: tonale weergave en vervorming.**

Aansluitend aan de Q-13 wordt een zwart vel karton beeldvullend opgenomen. In het midden van het beeldvlak, op het zwarte karton, ligt de scherpte testkaart Image Evaluation Test Target (QA-2).



4.3.6 Intern testrapport

De opdrachtnemer is verplicht om documentatie over de eigen interne controle op de gevraagde specificaties zoals hierboven beschreven bij te houden, en moet deze documentatie desgevraagd gelijktijdig met de betreffende batch(es) aan kunnen leveren.

4.4 PDF-bestanden

Er dient een PDF-bestand per tijdschriftnummer of krantenummer te worden gemaakt.

Er dient per object identifier één PDF-bestand te worden gemaakt.

Het PDF bestand bevat alle pagina's van het object en de paginavolgorde is gelijk aan het origineel
Voor de PDF-bestanden gelden de volgende specificaties:

Algemeen	'PDF Searchable Image'. Dit is een PDF met images, gecombineerd met verborgen OCR-tekst. Het PDF-bestand moet alle bij het document behorende opeenvolgende pagina's en/of uitvouwbladen bevatten.
Afbeeldingen	grijswaarden afbeeldingen Resolutie: 150 ppi. Kleurruimte: een apparaatafhankelijke kleurruimte als CalGray, Compressie; Jpeg (kwaliteitsniveau 50).
Compatibiliteit	Minimaal backwards compatible met Adobe Acrobat 6.0.
Optimalisatie weergave	Naast de compressie voor de afbeeldingen mogen eventueel de in de PDF toegestane compressiemethoden worden gebruikt (zoals Flate).

	Moet geoptimaliseerd zijn voor 'Fast Web View'
OCR tekst	Tekst moet gelijk zijn aan die van OCR en ALTO, inclusief paginanummers, kopteksten, voetteksten ed in de marges. Ook als daar eventueel correcties in zijn aangebracht.
Weergave / afdrukken	Fit Page Sizing-instelling. Het PDF-bestand mag niet openen met thumbnails.
Ongewenste onderdelen	Het PDF-bestand mag geen embedded objecten (anders dan de scans), externe links, commentaren, javascript actions, annotaties of data van de dienstverlener bevatten.
Bookmarks	Indien er artikelen onderscheiden worden, worden de artikelkoppen in bookmarks geplaatst.
Beveiliging	Geen beveiliging in de vorm van versleuteling, wachtwoorden, digitale handtekening of watermerk.
Metadata	<p>De volgende velden moeten in het Document Information Dictionary worden gecreëerd:</p> <ul style="list-style-type: none"> • Creator: de naam van de software die de oorspronkelijke scan heeft gecreëerd. • Producer: de software die het document naar PDF heeft geconverteerd. • CreationDate: de datum waarop de PDF is gecreëerd. • ModDate: de datum waarop het document voor het laatst is gewijzigd (komt doorgaans overeen met CreationDate). • Copyright-information: "Gedigitaliseerd door de Koninklijke Bibliotheek, de Nationale Bibliotheek van Nederland / Digitised by the Koninklijke Bibliotheek, the National Library of the Netherlands" <p>Hieronder volgen instructies per materiaalsoort. Attributen uit de metadatadump waaruit data moet worden overgenomen zijn tussen vierkante haken gegeven. Die vierkante haken hoeven niet overgenomen worden. Als er wel vierkante haken opgenomen moeten worden, zoals bij Tijdschriften, wordt dit aangegeven door dubbele vierkante haken te gebruiken.</p> <p>NB: niet alle genoemde attributen zijn voor elk object aanwezig. Als een attribuut ontbreekt, dan vervallen ook de komma die er normaliter achter gezet zou worden, en de eventuele vierkante haken. Achter het laatste overgenomen attribuut moet sowieso geen komma komen.</p>
Metadata	<p>Voor Kranten: Title: [/shipment*/record/@title], jrg. [/shipment*/entity/issue/@volumeNo], nummer [/shipment*/entity/issue/@issueNo], [/shipment*/entity/issue/@publicationDate], editie [/shipment*/entity/issue/@edition], [/shipment*/entity/@sourceProvider], [/shipment*/entity/@shelfmark]</p> <p>NB: de inhoud van [/shipment*/entity/issue/@publicationDate] moet 'omgekeerd' worden, van yyyy-mm-dd naar dd-mm-yyyy.</p> <p>Voorbeelden: Algemeen Handelsblad, jrg. 6, nummer 168, 15-07-1835, editie Dag, KB, 1623 C 4 Nationaale Courant, nummer 84, 03-08-1795, editie Dag, KB, 452 A 21 Leeuwarder Courant, 12-11-1938, editie Dag, KBDK Het Centrum, 04-03-1929, editie Dag, KB, MM03T-101072 (dit is een voorbeeld van een microfilm) Tijdinghe uit verscheyde quartieren, 01-04-1628, Kungl. Biblioteket Stockholm</p> <p>Keywords: Gedigitaliseerd door de Koninklijke Bibliotheek; Nederlandse geschiedenis; kranten, historische kranten, oude kranten, archief kranten, krantenarchieven, krant online, familieberichten, stamboom familie, dagblad,</p>

	<p>overlijdensberichten, nieuwsberichten, Nederlandstalige kranten, namen familie, familie Nederland, oorlogskranten, kranten van toen, Surinaamse kranten, Indische kranten, Antilliaanse kranten, databank kranten</p> <p>Voor Tijdschriften: <i>Title:</i> [/shipment*/record/@title], jrg. [/shipment*/entity/issue/@volumeNo], [/shipment*/entity/issue/@volumeYear], [/shipment*/entity/issue/@part], nummer [/shipment*/entity/issue/@issueNo], [/shipment*/entity/issue/@publicationDay], [/shipment*/entity/issue/@publicationMonth], [/shipment*/entity/issue/@publicationYear], [/shipment*/entity/issue/@publicationType], volgnr. [/shipment*/entity/issue/@SequenceNo]], [/shipment*/entity/@sourceProvider], [/shipment*/entity/@shelfmark]</p>
Metadata	<p>De Kunst Algemeen geïllustreerd en artistiek weekblad, jrg. 9, 1916-1917, nummer 473, 17 02 1917, KB, T 1893 Revue der Sporten, jrg. 6, 1913, nummer 37, 21 01 1913, [Aflevering, volgnr. 2], KB, T 12621 Nieuwe gids, jrg. 1, 1885, deel 1, KB, 12345 Nederland, 1855, deel 1, KB, 12345 Nederland, 1855, deel 2, KB, 12345 De jonge gids, 1900, KB, 12345</p> <p>Bijdragen ter bevordering onderwijs, 1810, 01 01 1810, [Index, volgnr. 1], KB, 675 C 12 Bijdragen ter bevordering onderwijs, 1810, 01 01 1810, [Aflevering, volgnr. 2], KB, 675 C 12</p> <p>Keywords: Gedigitaliseerd door de Koninklijke Bibliotheek; Nederlandse geschiedenis; tijdschriften, historische tijdschriften, oude tijdschriften, archief tijdschriften, tijdschrift online, cultuur, letterkunde, religie, wetenschap, politiek, sport, economie</p>

4.5 Metadatabestanden

De onderstaande specificaties geven aan, wat de KB qua metadata maximaal van een opdrachtnemer kan vragen. In specifieke Projecten kan het zo zijn, dat de KB minder vraagt dan hier gespecificeerd is. Dit zal in de afzonderlijke Projecten worden aangegeven.

Als de inhoud moet worden overgenomen uit het door de KB aangeleverde bestand, dan is dit per te vullen element of attribuut aangegeven in de laatste kolom van de tabellen in de specificaties hieronder. Daar staat dan de naam van het overeenkomstige element of attribuut in uit het XML-bestand. Voor metadata-elementen die niet door de KB zijn aangeleverd, staan in de tabel standaardwaarden, of is aangegeven dat de leverancier de gegevens moet genereren.

Cardinaliteit en datatypes worden in de tabellen niet gegeven. Deze zijn af te leiden uit de xml-schema's.

4.5.1 Algemene eisen aan XML-bestanden

Alle metadata- en OCR-bestanden worden geleverd in xml.

De encoding is utf-8, wat ook in de xml-declaratie wordt gespecificeerd.

De bestanden moeten wellformed en valide zijn.

Er mogen geen BOM-marks (Byte Order Marks) in de xml-bestanden staan.

Alle in de documenten gebruikte namespace-prefixes moeten voorzien zijn van de voor de betreffende standaard geldende namespace-uri.

In de aangeleverde XML-bestanden moeten geen schemalocations worden opgenomen.

De KB zal schema's meeleveren ter validatie.

De karakters die standaard xml-entiteiten kennen moeten als entiteit worden gecodeerd:

Karakter	Entiteit
&	&
<	<
>	>
"	"
'	'

Andere character-entiteiten dan bovengenoemde mogen niet worden gebruikt.

Niet-ASCII-tekens moeten consequent in utf-8 codering worden aangeleverd.

Voorbeeld: 'assemblée' in plaats van 'assemblée', 'assemblée', 'assemblée' of 'assembée'.

4.5.2 Metadatadump en schema

Dit bestand wordt door de KB toegezonden aan de opdrachtnemer.

De metadatadump bevat een shipment (zending) bestaande uit records (bibliografische gegevens), entities en issues. Voor kranten en tijdschriften dienen issues met records gematched te worden (element /shipment*/entity/issue/@referredRecordID).

Kolom KT geeft de scope van een element aan (Kranten of Tijdschriften) en, wanneer van toepassing, D=digitaal brontype, M=microfilm brontype (een deelverzameling van D!).

Schema: De metadatadump valideert tegen het door de KB meegeleverd schema kbdg_metadatadump.xsd.

Element	KT	Beschrijving	Toelichting
/shipment	KT		
/shipment/@ID	KT	ID van het transport van de KB naar de leverancier.	
/shipment/@dateFromKB	KT	Datum van het transport van de KB naar de leverancier.	Format: yyyy-mm-dd
/shipment/@diskNameKB	KT D M	Naam van de schijf waarop het transport door de KB aan de leverancier wordt geleverd.	Alleen voor digitaal bronmateriaal.
/shipment/@batchName	KT	Naam van de batch.	
/shipment/@numberOfEntities	KT	Aantal banden (barcodes) dat in het transport is opgenomen. Voor digitale aanlevering, geldt hier het aantal directories met als naam de barcode.	
/shipment/@deelprojectCode	KT	Code van het deelproject.	
/shipment/@deelprojectOwner	KT	Bibliotheek of instelling die het digitaliseringsproject heeft ingediend. Kan verschillen van de eigenaar van het te digitaliseren exemplaar.	
/shipment/@collection	KT	Beschrijft de collectie waar de publicatie onderdeel van uitmaakt.	
/shipment/@material	KT	Materiaalsoort (kranten tijdschriften).	
/shipment/@project	KT	Aanbestedingsronde = naam van het project	
/shipment/@sourceType	KT	Aard van het bronmateriaal dat gedigitaliseerd wordt (origineel microfilm digitaal).	
/shipment/@contractor	KT	Leverancier die de opdracht heeft aangenomen.	
/shipment/@KBcontractorID	KT	Door de KB uitgegeven code voor de leverancier die de opdracht heeft aangenomen.	
/shipment/@specificationsVersion	KT	Versienummer van deze metadataspecificaties.	
/shipment/(books booksMicrofilm journals journalsMicrofilm newspapers newspapersMicrofilm)	KT		
/shipment/*/record	KT		
/shipment/*/record/@ID	KT	ID uit de KB-materiaalvoorbereidingsdatabase van de titelbeschrijving van de periodiek waar de afzonderlijke banden en afleveringen bijhoren.	

Element	KT	Beschrijving	Toelichting
/shipment/*/record/@PPN	KT	Verwijzing naar de titelbeschrijving in de bibliotheekcatalogus.	
/shipment/*/record/@title	KT	Titel van de periodiek waartoe de afzonderlijke banden en afleveringen behoren.	
/shipment/*/record/@publisher	KT	Naam van de uitgever.	
/shipment/*/record/@issuedIn	KT	Plaats van uitgave.	
/shipment/*/record/@rightsStatus	KT	Status van de auteursrechten (unknown publicdomain copyrighted) .	
/shipment/*/record/@rightsJurisdiction	KT	Land van uitgave van de auteursrechten.	ISO 3166 code (zie http://www.iso.org/iso/country_codes.htm)
/shipment/*/record/@rangeOfFilmShelfmarks	KT M	De reeks aanvraagnummers van alle microfilms, waarop afleveringen van een bepaalde titel staan.	Alleen voor digitaal bronmateriaal.
/shipment/*/record/@shelfmarkFilm	KT M	Aanvraagnummer van de microfilm.	Alleen voor digitaal bronmateriaal.
/shipment/*/record/@barcodeFilm	KT M	Barcode van de microfilm in de materiaalvoorbereidingsdatabase van de KB.	Alleen voor digitaal bronmateriaal.
/shipment/*/record/@ISSN	KT	International Standard Serial Number. Een uniek identificatienummer voor periodieke publicaties.	
/shipment/*/record/@alternativeName	KT	Alternatieve titel.	
/shipment/*/record/@issued	KT	Periode of jaar van verschijning.	
/shipment/*/record/@spatial	K	Verspreidingsgebied van de periodiek (Landelijk Regionaal/lokaal Suriname Nederlandse Antillen Nederlands-Indië / Indonesië Huidig België).	
/shipment/*/record/@yearsDigitized	KT	Periode die wordt gedigitaliseerd.	
/shipment/*/entity	KT		
/shipment/*/entity/@barcode	KT	Barcode van de band in de materiaalvoorbereidingsdatabase van de KB. Voor convoluten kunnen meerdere barcodes in een band voorkomen. Bij digitale aanlevering geldt een map met scans als een band.	
/shipment/*/entity/@barcodeSourceProvider	KT	Code voor de band in het eigen systeem van de instelling.	
/shipment/*/entity/@shelfmark	KT	Planknummer van het exemplaar in het magazijn van de instelling.	

Element	KT	Beschrijving	Toelichting
/shipment/*/entity/@EPN	B	Verwijzing naar het exemplaar in de bibliotheekcatalogus.	
/shipment/*/entity/@sourceProvider	KT	Bibliotheek of instelling die eigenaar is van het exemplaar.	
/shipment/*/entity/@numberOfPages	KT	Aantal pagina's in de band inclusief blanco's (schatting).	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@size	KT	Geeft het formaat van het object aan (in onderling overleg tussen KB en leverancier zullen hierover afspraken worden gemaakt).	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@numberOfTargets	KT	Aantal targets dat is toegevoegd in de band. Als dit attribuut leeg is, zijn er geen targets toegevoegd tijdens de materiaalvoorbereiding.	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@foldouts	KT	Of er uitvouwbladen aanwezig zijn. Als dit attribuut leeg is, zijn er geen uitvouwbladen aangetroffen tijdens de materiaalvoorbereiding.	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@loosePages	KT	Of er losbladige toevoegingen aanwezig zijn die ook gedigitaliseerd moeten worden. Als dit attribuut leeg is, zijn er geen losbladige toevoegingen aangetroffen tijdens de materiaalvoorbereiding.	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@maxOpening110	BT	Of het object niet verder open mag dan 110°	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@damaged	KT	Of er onvermijdbare schade aanwezig is.	Alleen bij fysiek bronmateriaal.
/shipment/*/entity/@numberOfFrames	KT M	Aantal frames op de film per barcode (schatting).	Alleen bij digitaal bronmateriaal.
/shipment/*/entity/@pagesPerFrame	KT M	Of de microfilm per pagina (1) of per opening (2) opgenomen is.	Alleen bij digitaal bronmateriaal.
/shipment/*/entity/OCRLanguage	KT	Taal die in de band voorkomt. Bij periodieken kunnen dit meerdere talen per band zijn.	ISO 639-2B codes (Zie http://www.loc.gov/standards/iso639-2/php/code_list.php)
/shipment/*/entity/@yearsRangeStart	KT	Datum eerste aflevering in de band.	
/shipment/*/entity/@yearsRangeEnd	KT	Datum laatste aflevering in de band.	
/shipment/*/entity/comments	KT	Opmerkingen die van belang kunnen zijn voor het digitaliseren.	
/shipment/*/entity/issue	KT		
/shipment/*/entity/issue/@ID	KT	ID uit de KB-materiaalvoorbereidingsdatabase van de aflevering.	
/shipment/*/entity/issue/@EPN	KT	Verwijzing naar de aflevering in de bibliotheekcatalogus.	

Element	KT	Beschrijving	Toelichting
/shipment/*/entity/issue/@referredRecordID	KT	Verwijst naar het ID van het record waar de aflevering bij hoort.	
/shipment/*/entity/issue/@sequenceNo	KT	Volgnummer in de band. Dit is alleen van toepassing op banden waarin meerdere titels zijn opgenomen zoals convoluten en periodieken. In dat geval wordt hier opgenomen "de hoeveelste" titel of aflevering dit in de band is.	
/shipment/*/entity/issue/@publicationDay	T	Dag van publicatie van de aflevering (bibliografische notatie).	
/shipment/*/entity/issue/@publicationMonth	T	Maand van publicatie van de aflevering (bibliografische notatie).	
/shipment/*/entity/issue/@publicationYear	T	Jaar van publicatie van de aflevering (bibliografische notatie).	
/shipment/*/entity/issue/@publicationDate	K	Volledige publicatiedatum datum van de aflevering.	yyyy-mm-dd
/shipment/*/entity/issue/@sortingDate	T	(Vaak fictieve) volledige publicatiedatum van een aflevering in datumformaat.	yyyy-mm-dd
/shipment/*/entity/issue/@volumeYear	KT	Jaartal van de jaargang van de aflevering.	
/shipment/*/entity/issue/@volumeNo	KT	Nummer van de jaargang van de aflevering.	
/shipment/*/entity/issue/@issueNo	KT	Originele nummer van de aflevering.	
/shipment/*/entity/issue/@edition	K	Dagdeel van verschijning (Ochtend Middag Avond Dag).	
/shipment/*/entity/issue/@issueType	K	Type aflevering van een krant. 'I' (issue) staat voor een krantnummer en 'S' voor een supplement. NB: indien het om een supplement gaat, moet dit worden toegevoegd aan de aflevering van de krant waar het bij hoort (voor behandeling van supplementen, zie hoofdstuk 3).	
/shipment/*/entity/issue/@referredIssueID	K	Dit wordt alleen gebruikt bij supplementen en verwijst naar het ID van het krantnummer waar het supplement moet worden bijgevoegd.	
/shipment/*/entity/issue/@publicationType	T	Type aflevering voor tijdschrift (Aflevering Bijlage Index Inhoudsopgave).	
/shipment/*/entity/issue/@part	T	Deelaanduiding tijdschriften (bijvoorbeeld 'Slachtmaand', 'Extra nr.' etc.).	

4.5.3 METS

Structurele metadata en beschrijvende metadata moeten worden teruggeleverd in een METS-bestand. De KB levert dit bestand deels gevuld aan, de opdrachtnemer moet het voltooien.

De METS-file is een XML-bestand dat moet voldoen aan het bijgeleverde XML-schema en de algemene eisen voor XML-bestanden.

Er is één METS-bestand voor elk tijdschriftnummer of krantenummer (dus bv. voor de Volkskrant van vandaag of voor de avondeditie van Het Vaderland van 2 januari 1920). Als een -band of aflevering niet gescand kan worden, levert de opdrachtnemer het bijbehorende METS-bestand ook niet terug aan de KB.

Als er een METS-bestand ontbreekt, dient de opdrachtnemer contact op te nemen met de KB.

De naam van de door de KB aangeleverde METS-bestanden is als volgt opgebouwd: <object identifier>_mets_template.xml, bijvoorbeeld KBRA02KR_000028001_mets_template.xml.

4.5.4 Schema's

Deze door de KB aangeleverde METS-bestanden valideren tegen een eveneens door de KB meegeleverd schema waarvan de naam als volgt is opgebouwd:

kbdg_mets_<materiaalsoort>_SIP_template.xsd,

kbdg_mets_kranten_SIP_template.xsd.

De door de opdrachtnemer teruggeleverde METS-bestanden moeten valideren tegen een meegeleverd schema waarvan de naam als volgt is opgebouwd:

kbdg_mets_<materiaalsoort>_SIP.xsd,

kbdg_mets_kranten_SIP.xsd.

De METS-schema's gebruiken een aantal van de andere meegeleverde schema's (o.a. voor de MODS-, PREMIS- en MIX-onderdelen).

4.5.5 Paginanummering

Paginanummering en -typering worden in het METS-bestand vastgelegd.

Daarbij moet op de volgende zaken gelet worden:

- Oorspronkelijke paginanummers zoals op de originele pagina. Dit moet in de metadata worden overgenomen uit het oorspronkelijke materiaal (zie //mets:div/@ORDERLABEL in de tabel hieronder). Wanneer de pagina geen paginanummer bevat mag ook geen waarde in @ORDERLABEL zijn opgenomen in de metadata. Voor de titelpagina (alleen bij Tijdschriften) wordt in //mets:div/@ORDERLABEL de term "titlepage" opgenomen. Als de titelpagina ook een oorspronkelijk paginanummer bevat, wordt in //mets:div/@ORDERLABEL dit nummer opgenomen, gevolgd door een spatie en de term "titlepage".
- In //mets:div/@ORDER (zie tabel hieronder) moet een nummer worden opgenomen dat de absolute numerieke volgorde van de scans aangeeft. Dit nummer moet overeenkomen met het in de concordantietabel opgenomen volgnummer voor de betreffende scan. Alle gescande pagina's en dus alle volgnummers moeten in dit bestand voorkomen. NB: eventuele scans met dubbele pagina's krijgen dus nog steeds maar één nummer.
- In het element //mets:div/@LABEL moet worden opgenomen of een pagina een enkele of dubbele pagina betreft of een uitklapper. Voor dubbele pagina's worden alle oorspronkelijke paginanummers (indien aanwezig) gescheiden door een spatie in //mets:div/@ORDERLABEL opgenomen.

4.5.6 Auteursinformatie per tijdschriftartikel

Bij tijdschriften die op artikelniveau ontsloten worden kan er gevraagd worden om de auteur(s) van een artikel op te nemen in het METS-bestand.

Daarbij moet op de volgende zaken gelet worden:

- Genereer per artikel met auteursinformatie een mets:dmdSec-sectie met auteursinformatie in MODS van betreffend artikel. Wanneer er geen auteursinformatie aanwezig is wordt er ook geen mets:dmdSec-sectie voor betreffend artikel gegenereerd.
- Voeg een link toe in betreffende artikelen: /mets:structMap[2]/mets:div/mets:div[@TYPE='article:']/@DMDID naar betreffende mets:dmdSec-sectie met auteursinformatie in MODS. Wanneer er geen dmdSec-sectie is gegenereerd wordt er ook geen DMDID attribuut aangemaakt.

4.5.7 Toelichting bij de tabel

- **Kolom “Element”**: geeft, in Xpath notatie, de elementen en attributen die de opdrachtnemer moet toevoegen, invullen etc.
- **Kolom “Toelichting”**: geeft onder meer aan waar de door de opdrachtnemer gemaakte elementen ingevoegd moeten worden
- **Kolom “KT”**: deze kolom geeft de scope van een element of attribuut aan (Kranten of Tijdschriften)
- **Kolom “Waarde”**:
 - standaardwaardes zijn vermeld tussen dubbele aanhalingstekens. Deze waardes moeten worden overgenomen
 - moet de opdrachtnemer een waarde opbouwen of genereren, dan is de instructie tussen < en > geplaatst. De tekens < en > moeten in het resultaat niet worden overgenomen
 - een – in de kolom “waarde” betekent, dat er in het element niet direct een waarde hoeft te worden ingevuld. Het gaat dan meestal om ‘parent’-elementen die andere elementen bevatten. Een parent-element kan wel attributen hebben
 - moeten er data uit een reeds eerder gemaakt element of attribuut worden overgenomen, dan wordt dit middels een XPath-expressie aangegeven
 - eventuele instructies – die dus niet moeten worden overgenomen in het METS-bestand - zijn tussen ronde haken geplaatst.

Element	KT	Toelichting	Waarde
/mets:mets/mets:metsHdr			
@LASTMODDATE (dit element is reeds aanwezig, alleen de inhoud moet worden ingevuld)	KT	Vervang de inhoud van @LASTMODDATE (in het aangeleverde bestand gelijk aan @CREATEDATE) door de waarde beschreven in de kolom "Waarde" hiernaast.	<moment van opslag van het aangevulde METS-bestand> (Datum/tijdsaanduiding volgens standaard ISO 8601: YYYY-MM-DDThh:mm:ss.SSSZ)
/mets:mets/mets:dmdSec (per artikel met auteursinformatie) NB: ALLEEN VAN TOEPASSING VOOR TIJDSCHRIFTEN MET ARTIKELSEGMENTERING			
//mets:dmdSec	T	Vervang de tekst "<!--invullen #1: dmdSec per article: artikel -->" door een dmdSec voor elk betreffend artikel	•
//mets:dmdSec/@ID	T	Deze identifier moet uniek zijn binnen het METS-document	• "DMD"<volgnummer> Volgnummer telt 5 posities, inclusief voorloopnullen.
//mets:dmdSec/mets:mdWrap	T		-
//mets:dmdSec/mets:mdWrap/@MDTYPE	T		"MODS"
//mets:dmdSec/mets:mdWrap/@MDTYPEVERSION	T		"3.5"
//mets:dmdSec/mets:mdWrap/mets:xmlData	T		-
• //mets:dmdSec/mets:mdWrap/mets:xmlData /mods:mods	T		(Zie aparte specificaties in paragraaf 4.5.84 MODS)
/mets:mets/mets:amdSec			
//mets:techMD	KT	Vervang de tekst "<!--invullen #2: techMD per bestand-->" door een techMD voor elk bestand behorend bij dit object, m.u.v. het mets-bestand zelf.	-
/mets:mets/mets:amdSec			
//mets:techMD/@ID	KT	Deze identifier moet uniek zijn binnen het METS-document	"TMD"<volgnummer> (Volgnummer telt 5 posities, inclusief voorloopnullen. Begin bij volgnummer

Element	KT	Toelichting	Waarde
			TMD00002, er is namelijk altijd al een TMD00001.)
//mets:techMD/mets:mdWrap	KT		-
//mets:techMD/ mets:mdWrap/@MDTYPE	KT		"PREMIS:OBJECT"
//mets:techMD/mets:mdWrap/@MDTYPEVERSION	KT		"2.3"
//mets:techMD/mets:mdWrap/mets:xmlData	KT		-
//mets:techMD/mets:mdWrap/mets:xmlData/premis:object	KT		(Zie aparte specificaties in paragraaf 4.5.9)
//mets:techMD	KT	Vervang de tekst "<!--invullen #3: techMD voor MIX voor masterbestand-->" door een techMD voor elk masterbestand behorend bij dit object.	-
//mets:techMD/@ID	KT	Deze identifier moet uniek zijn binnen het METS-document	"TMD<volgnummer>" (Volgnummer telt 5 posities, inclusief voorloophulpen. Sluit aan op het laatste volgnummer dat is uitgegeven aan een techMD voor een bestand.)
//mets:techMD/mets:mdWrap	KT		-
//mets:techMD/mets:mdWrap/@MDTYPE	KT		"NISOIMG"
//mets:techMD/mets:mdWrap/@MDTYPEVERSION	KT		"2.0".
//mets:techMD/mets:mdWrap/mets:xmlData	KT		-
//mets:techMD/mets:mdWrap/mets:xmlData/mix:mix	KT		(Zie aparte specificaties in paragraaf 4.5.10)
/mets:mets/*/premis:event	KT		
//premis:eventDateTime (dit element is reeds aanwezig, alleen de inhoud moet worden ingevuld)	KT	Vervang de inhoud (in het aangeleverde bestand gelijk aan @CREATEDATE) door de waarde beschreven in de kolom "Waarde" hiernaast.	<moment/datetimestamp waarop productie van het gehele object (alle master- en accessbestanden van de gehele premis representation) voltooid is> (Datum/tijdsaanduiding volgens standaard ISO 8601: YYYY-MM-DDThh:mm:ss.SSSZ)

Element	KT	Toelichting	Waarde
/mets:fileSec			
//mets:file	KT	Het METS-bestand bevat 5 fileGrp-elementen. Vervang in elke fileGrp de tekst “<!--invullen #4*: file-element voor elk *-bestand-->” door een file-element voor elk bestand dat in de betreffende groep hoort.	-
//mets:file/@ID	KT	Deze identifier moet uniek zijn binnen het METS-document	“F<volnummer>” (Volnummer telt 5 posities, inclusief voorloopnullen.)
//mets:file/@ADMID	KT	@ID van de corresponderende techmd-sectie met het premis:object van het type premis:file en (in geval van master images) @ID van de corresponderende techmd-sectie met MIX. In het laatste geval worden beide ID's gescheiden door een spatie. NB: volgnummers in //mets:mets/mets:fileSec/mets:fileGrp/mets:file/@ID hoeven niet overeen te komen met de volgnummers in de @ID's waar naar wordt verwezen. Het zijn puur interne identifiers in het METS-document.	(In te vullen door opdrachtnemer)
//mets:file/@MIMETYPE	KT		“image/jp2” (master en acces bestanden) “text/xml” (ALTO en OCR) “application/pdf” (pdf)
//mets:file/@SIZE	KT		<grootte van bestand in bytes>
//mets:file/@CHECKSUMTYPE	KT		“SHA-512”
//mets:file/@CHECKSUM	KT	Moet overeenkomen met de waarden in PREMIS (zie paragraaf 4.5.5)	<SHA-512-checksum -waarde van het bestand>
//mets:file/mets:FLocat	KT		-
//mets:file/mets:FLocat/@LOCTYPE	KT		“URL”
//mets:file/mets:FLocat/@xlink:href	KT	Relatief pad in de batch, t.o.v. het METS-bestand, gevolgd door de bestandsnaam	“file://<relatief pad>/<bestandsnaam>”
/mets:structMap[1]/mets:div			

Element	KT	Toelichting	Waarde
//mets:fptr	KT	Vervang de tekst <!--invullen #5: fptr PDF-bestand--> door een Filepointer voor het PDF-bestand dat hoort bij het object.	-
//mets:fptr/@FILEID	KT		//mets:file/@ID (voor PDF-bestand)
//mets:div	KT	Vervang de tekst <!-- invullen #6: div per pagina--> door een div-element voor elke pagina.	
//mets:div/@ID	KT	Deze identifier moet uniek zijn binnen het METS-document	"divpage<volgnummer>" (Volgnummer telt 5 posities, inclusief voorloop-nullen.)
//mets:div/@TYPE	KT	Authority: marcgt	"page"
//mets:div/@ORDER	KT	'scannummer' (zie inleiding)	Is gelijk aan 'volgnr' in concordantietabel voor dezelfde scan. Altijd 5 posities, dus eventueel met voorloopnullen.
//mets:div/@ORDERLABEL	KT	Origineel paginanummer (range)	Nummer, twee nummers gescheiden door spatie (voor een range), "titlepage" of nummer gevolgd door spatie en "titlepage" (voor genummerde titlepages)
//mets:div/@LABEL	KT	Soort bladzijde	"single", "double" of "foldout"
//mets:div/mets:fptr	KT	3 of (als er niet in artikelen gesegmenteerd wordt) 4 keer, voor master-, access-, ALTO- en (eventueel) text-OCR-bestanden die horen bij de pagina.	-
//mets:div/mets:fptr/@FILEID	KT		//mets:file/@ID (voor het corresponderende bestand)
/mets:structMap[2]/mets:div		NB: ALLEEN VAN TOEPASSING ALS ER IN ARTIKELEN WORDT GESEGMENTEERD	
//mets:div	KT	Vervang de tekst <!-- invullen #7: div per artikel--> door een div-element voor elk artikel	-
//mets:div/@TYPE	KT	Authority: marcgt, gevolgd door eigen artikel-genre. <u>Mogelijke genres voor kranten zijn:</u> "artikel", "illustratie met onderschrift", "advertentie" en "familiebericht". De waarde "illustratie met onderschrift" wordt gebruikt voor op zichzelf staande illustraties met bijschrift (bv. losse foto's, fotopagina's, cartoons) en niet in het geval dat een illustratie onderdeel is van een	<u>Mogelijke waarden voor kranten:</u> "article: artikel" "article: illustratie met onderschrift" "article: advertentie" "article: familiebericht"

Element	KT	Toelichting	Waarde
		<p>artikel. De waarde "advertentie" wordt gebruikt voor commerciële advertenties en niet voor geboorteaankondigingen, huwelijksaankondigingen en rouwadvertenties. De waarde "familiebericht" wordt gebruikt voor geboorteaankondigingen, huwelijksaankondigingen en rouwadvertenties. De waarde "artikel" wordt gebruikt voor alle overige berichten, als nieuwsartikelen, opiniestukken, redactionelen etc.</p> <p><u>Mogelijke genres voor tijdschriften die op artikel niveau worden ontsloten zijn:</u> "artikel", "illustratie met onderschrift", "advertentie", "inhoudsopgave", "register", "colofon" en "omslag". De waarde "illustratie met onderschrift" wordt gebruikt voor op zichzelf staande illustraties met bijschrift (bv. losse foto's, fotopagina's, cartoons), alle overige afbeeldingen die niet onder een andere categorie vallen zoals advertentie en niet in het geval dat een</p>	<p><u>Mogelijke waardes voor tijdschriften die op artikel niveau worden ontsloten zijn:</u> "article: artikel" "article: illustratie met onderschrift" "article: advertentie" "article: inhoudsopgave" "article: register" "article: colofon" "article: omslag"</p>
		<p>illustratie onderdeel is van een artikel. De waarde "advertentie" wordt gebruikt voor commerciële advertenties. De waarde "inhoudsopgave" wordt gebruikt voor inhoudsopgaven (overzichten met verwijzingen naar artikelen in een tijdschrift). "register" wordt gebruikt voor indices of registers (een lijst met trefwoorden die voorkomen in het tijdschriftnummer met een verwijzing naar de pagina waar dit trefwoord te vinden is). De waarde "colofon" wordt gebruikt voor een colofon (een opsomming van technische gegevens zoals uitgever, jaar van uitgave, drukker, oplage etc.). De waarde "omslag" wordt gebruikt voor een omslag. De waarde "artikel" wordt gebruikt voor alle overige berichten.</p> <p>Tijdens het digitaliseren dient per artikel de juiste categorie te worden bepaald. Indien de leverancier niet in</p>	

Element	KT	Toelichting	Waarde
		staat is de juiste categorie te bepalen, of twijfelt aan de juistheid van de categorie kan de KB desgewenst algemene richtlijnen geven met betrekking tot de herkenning en classificatie van het materiaal. Indien men afwijkt van deze richtlijnen behoudt de KB zich het recht voor materiaal af te keuren.	
//mets:div/@LABEL	KT	Kop van het artikel. Als het artikel geen kop heeft, moet de eerste regel worden gebruikt (NB: de eerste regel moet in dat geval ook nog steeds in de tekst worden opgenomen). Meerdere (sub)koppen worden samen in één @LABEL opgenomen. Als een artikel bestaat uit een illustratie met onderschrift zonder kop, dan wordt de 1 ^e regel van het onderschrift als kop genomen. Bij een advertentie, met of zonder tekst, wordt 'Advertentie' als kop genomen. Bij een familiebericht wordt 'Familiebericht' als kop genomen. Bij een omslag wordt 'Omslag' als kop genomen. Bij een colofon wordt de titel van het colofon als kop genomen (bijvoorbeeld 'Van onze redactie'), tenzij er geen titel aanwezig is dan wordt de waarde 'Colofon' als kop genomen. Bij een register wordt de titel van het register als kop genomen (bijvoorbeeld 'Index'), tenzij er geen titel aanwezig is dan wordt de waarde 'Register' als kop genomen. Bij een inhoudsopgave wordt de titel van de inhoudsopgave als kop genomen (bijvoorbeeld 'Inhoud'), tenzij er geen titel aanwezig is dan wordt de waarde 'Inhoudsopgave' als kop genomen.	(Door opdrachtnemer over te nemen uit origineel)
//mets:div/@DMDID	T	Bevat verwijzing naar de bijbehorende dmdSec met MODS voor betreffend artikel met auteursinformatie.	/mets:mets/mets:dmdSec/@ID
//mets:div/mets:fptr	KT		-
//mets:div/mets:fptr/@FILEID	KT	@ID van het text-OCR-bestand van het artikel.	//mets:mets/mets:fileSec/mets:fileGrp/mets:file/@ID

Element	KT	Toelichting	Waarde
			(voor corresponderende fulltext-OCR-bestand)
//mets:div/mets:fptr	KT	NB: dit is een nieuw (tweede) fptr-element!	-
//mets:div/mets:fptr/mets:seq	KT		-
//mets:div/mets:fptr/mets:seq/mets:area	KT	Eén area voor elk TextBlock of ComposedBlock dat bij het artikel hoort. Onderdelen moeten in leesvolgorde worden opgenomen.	-
//mets:div/mets:fptr/mets:seq/mets:area/@FILEID	KT		//mets:mets/mets:fileSec/mets:fileGrp/mets:file/@ID (voor corresponderend ALTO-bestand)
//mets:div/mets:fptr/mets:seq/mets:area/@BETYPE	KT		"IDREF"
//mets:div/mets:fptr/mets:seq/mets:area/@BEGIN	KT		ID van het betreffende TextBlock of ComposedBlock in de bijbehorende ALTO-file.

4.5.8 MODS

Zie voor opname van de MODS-elementen paragraaf 4.5.3 METS, tabel ([/mets:mets/mets:dmdSec \(per artikel met auteursinformatie\)](#)).

Element	BKT	Toelichting	Waarde
//mods:mods	T		-
//mods:mods/mods:name	T	Per auteursnaam een afzonderlijk mods:name element	-
//mods:mods/mods:name/mods:displayForm	T	Tekststring met auteursnaam zoals deze voorkomt in de tekst.	In te vullen door opdrachtnemer.

4.5.9 PREMIS

Zie voor opname van de PREMIS-elementen in METS paragraaf 4.5.3. De tabel is in opzet gelijk aan de tabel in dit paragraaf, alleen de KT-kolom ontbreekt.

Element	Toelichting	Waarde
//premis:object		-

//premis:object/@xsi:type		"premis:file"
//premis:object/premis:objectIdentifier		-
//premis:object/premis:objectIdentifier/premis:objectIdentifierType		"local"
//premis:object/premis:objectIdentifier/premis:objectIdentifierValue	Moet uniek zijn binnen de totale verzameling van gedigitaliseerde bestanden	Bestandsnaam exclusief de extensie (zie paragraaf 4.1.4).
//premis:object/premis:objectCharacteristics		-
//premis:object/premis:objectCharacteristics/premis:compositionLevel		"0" (nul)
//premis:object/premis:objectCharacteristics/premis:fixity		-
//premis:object/premis:objectCharacteristics/premis:fixity/premis:messageDigestAlgorithm		"SHA-512"
//premis:object/premis:objectCharacteristics/premis:fixity/premis:messageDigest	SHA-512-checksum	(Te genereren en in te vullen door opdrachtnemer)
//premis:object/premis:objectCharacteristics/premis:fixity/premis:messageDigestOriginator	Organisatie die de checksum heeft berekend	//mets:agent[1]/mets:name
//premis:object/premis:objectCharacteristics/premis:size	Bestandsgrootte in bytes	In te vullen door opdrachtnemer
//premis:object/premis:objectCharacteristics/premis:format		-
//premis:object/premis:objectCharacteristics/premis:format/premis:formatDesignation		-
//premis:object/premis:objectCharacteristics/premis:format/premis:formatDesignation/premis:formatName	Bestandsformaat	"image/jp2" (master en acces bestanden) "text/xml" (ALTO en OCR) "application/pdf" (pdf)
//premis:object/premis:objectCharacteristics/premis:format/premis:formatDesignation/premis:formatVersion	De versie van het hierboven genoemde bestandsformaat.	In te vullen door opdrachtnemer tekst/xml heeft de standaardwaarde "1.0". Voor image/jp2 moet dit veld leeg blijven. Voor application/pdf wordt de gebruikte versie ingevuld.

4.5.10 MIX

Voor ieder masterbestand moeten technische metadata in MIX-format in het METS-bestand worden opgenomen. Als uitgangspunt voor de vereiste informatie is uitgegaan van de NISO-standaard, NISO Z39.87. De gevraagde versie van mix is 2.0. Validatie vindt plaats als onderdeel van de METS-validatie.

Zie tabel in paragraaf 4.5.3 voor de inpassing van de hier beschreven MIX-elementen in METS.

NB: hoewel het in de tabel niet staat, moet in de XML bij elk mix-element de namespace “mix” worden opgenomen: dus <mix:ObjectIdentifier> i.p.v. <ObjectIdentifier>. Deze namespace is al gedeclareerd in de METS-bestanden, dat hoeft dus niet binnen de mix te gebeuren.

Hieronder zijn de elementen uit deze standaard aangegeven die opgenomen moeten worden. De kolom ‘Obligation’ heeft twee opties:

M = mandatory / verplicht

MA = mandatory if applicable / verplicht indien van toepassing

NISO Z39.87			Obligation
6	Basic Digital Object Information		
6.1	ObjectIdentifier		M
6.1.1	objectIdentifierType	URN	M
6.1.2	objectIdentifierValue	[deelprojectcode]_[negencijferig nummer]_[vijfcijferig volgnr]_master NB: de waarde van objectIdentifierValue is dus gelijk aan de bestandsnaam van het masterbestand exclusief de extensie (“.jp2”; zie paragraaf 4.1.4), en is gelijk aan premis:object/premis:objectIdentifier/premis:objectIdentifierValue van het master bestand.	M
6.5	byteOrder	big endian	MA
6.6	Compression		M
6.6.1	compressionScheme	JPEG 2000 Lossless (reversible 5-3 wavelet filter)	M
6.6.2	compressionSchemeLocalList	-	MA
6.6.3	compressionSchemeLocalValue	-	MA
7	Basic Image Information		
7.1	BasicImageCharacteristics		M
7.1.1	imageWidth	[Breedte in pixels]	M
7.1.2	imageHeight	[Hoogte in pixels]	M
7.1.3	PhotometricInterpretation	-	M

7.1.3.1	colorSpace	[bij master image gespecificeerde kleuruimte]	M
7.1.3.2	ColorProfile		MA
7.1.3.2.1	IccProfile		MA
7.1.3.2.1.1	iccProfileName	[bij master image gespecificeerde profiel]	MA
7.1.3.2.1.2	iccProfileVersion	[bij master image gespecificeerde versie]	MA
7.1.3.2.1.3	iccProfileURI		MA
7.1.3.2.2	LocalProfile		MA
7.1.3.2.2.1	localProfileName		MA
7.1.3.2.2.2	localProfileURL		MA
7.1.3.3	YcbCr		MA
7.1.3.3.1	yCbCrSubSampling		MA
7.1.3.3.2	yCbCrPositioning		MA
7.1.3.3.3	yCbCrCoefficients		MA
7.1.3.4	ReferenceBlackWhite		MA
7.2	SpecialFormatCharacteristics		MA
7.2.1	JPEG2000		MA
7.2.1.2	EncodingOptions		MA
7.2.1.2.1	tiles		MA
	tileWidth	1024	
	tileHeight	1024	
7.2.1.2.2	qualityLayers	11	MA
7.2.1.2.3	resolutionLevels	6	MA
7.2.2	MrSID		MA
7.2.2.1	zoomLevels		MA
7.2.3	Djvu		MA
7.2.3.1	djvuFormat		MA

8	Image Capture Metadata		
8.1	SourceInformation		M
8.1.1	sourceType	Kranten: 'newspaper'; tijdschriften: 'journal'; microfilm: 'microfilm'	M
8.1.2	SourceID		M
8.1.2.1	sourceIDType	URN	M
8.1.2.2	sourceIDValue	Zie 6.1.2	M
8.2	GeneralCaptureInformation		MA
8.2.1	dateTimeCreated	[datum]	MA
8.2.2	imageProducer	Koninklijke Bibliotheek; [leverancier]	M
8.3	ScannerCapture	NB: in geval van aanlevering van digitaal bronmateriaal door de KB worden de bij 8.3.* in te vullen waarden door de KB aangeleverd.	MA
8.3.1	scannerManufacturer	[producent]	M
8.3.2	ScannerModel	[model]	M
8.3.2.1	scannerModelName	[naam]	M
8.3.2.2	scannerModelNumber	[nummer]	M
8.3.4	scannerSensor	[sensor type]	M
8.3.5	ScanningSystemSoftware		M
8.3.5.1	scanningSoftwareName	[naam]	M
8.3.5.2	scanningSoftwareVersionNo	[versie]	M
8.4	DigitalCameraCapture		MA
8.4.1	digitalCameraManufacturer	[producent]	M
8.4.2	DigitalCameraModel	[model]	M
8.4.2.1	digitalCameraModelName	[naam]	M
8.4.2.2	ditigalCameraModelNumber	[nummer]	M
8.4.2.3	digitalCameraModelSerialNo	[nummer]	M
8.4.3	cameraSensor	[sensor type]	M

8.5	Orientation	normal*	M
9	Image Assessment Metadata		
9.1	SpatialMetrics		MA
9.1.1	samplingFrequencyPlane		MA
9.1.2	samplingFrequencyUnit		MA
9.1.2.1	xSamplingFrequency		MA
9.1.2.2	ySamplingFrequency		MA
9.2	ImageColorEncoding		M
9.2.1	BitsPerSample		M
9.2.1.1	bitsPerSampleValue	8	M
9.2.1.2	bitsPerSampleUnit	integer	M
9.2.2	samplesPerPixel	1	M
9.2.3	extraSamples		MA
9.2.4	Colormap		MA
9.2.4.1	colormapReference		MA
9.3	TargetData		M
9.3.1	targetType		M
9.3.2	TargetID		M
9.3.2.1	targetManufacturer	[producent]	M
9.3.2.2	targetName	[naam]	M
9.3.2.3	targetNo	[nummer]	M
9.3.2.4	targetMedia	[materiaal]	M
9.3.3	externalTarget	[verwijzing naar bestand met externe target, elk in een apart element]	M
10	ChangeHistory/ Image Processing		
10.1	ImageProcessing		MA
10.1.1	dateTimeProcessed		MA

10.1.2	sourceData		MA
10.1.3	processingAgency		M
10.1.4	processingRationale		M
10.1.5	ProcessingSoftware		M
10.1.5.1	processingSoftwareName		M
10.1.5.2	processingSoftwareVersion		M
10.1.5.3	processingOperatingSystemName		M
10.1.5.4	processingOperatingSystemVersion		M
10.1.6	processingActions		M
10.2	Previous Image Metadata (retains previous data types)		MA

4.5.11 OCR

Voor tijdschriften die in artikelen worden gesegmenteerd wordt een tekstbestand per artikel gemaakt; in andere gevallen wordt een tekstbestand per pagina gemaakt.

Voor OCR gelden de volgende eisen:

De juistheid en volledigheid van de Segmentatie en classificatie wordt mede beoordeeld a.d.h.v. het origineel. (b.v. lay-out, fysieke staat etc).

Wij vragen een juistheid en volledigheid per steekproef van deelbatches behorend bij een transport van minimaal 98%.

- Het xml-bestand omvat de volledige tekst van een pagina of artikel, in een eenvoudig xml-formaat (de xml-tags zijn overgenomen uit de TEI-standaard).
- Paragraafaanduiding moet overeenkomen met wat menselijke lezers als paragraaf beschouwen.
- De tekst moet in leesvolgorde worden opgenomen. Let daarbij speciaal op eventueel aanwezige verticale tekst aan de linker- en de rechterzijde van een pagina.
- Dit geldt ook voor tabellen die tekst bevatten (vaak heeft OCR-software speciale instellingen om de tekenherkenning in tabellen optimaal te doen).
- Verticale tekst moet ook worden opgenomen.
- Afbreekstreepjes moeten worden verwijderd. Als een streepje niet voor het afbreken van een woord wordt gebruikt, maar onderdeel van het woord uitmaakt, dan moet het blijven staan.
- Wordt er in artikelen opgeleverd, dan moet het bestand alle tekst van het artikel bevatten, ook als het artikel doorloopt op volgende pagina('s).
- Paginanummers, kopteksten en voetteksten in de marges e.d. moeten worden weggelaten.
- Voor kranten en tijdschriften die op artikelniveau worden ontsloten gelden de paginanummers, de titel van de krant of het tijdschrift (bijvoorbeeld "Het Vaderland") en gegevens over het nummer of de krant/tijdschrift die in de metadata zijn opgenomen, niet als artikel. Deze tekst wordt wel in de ALTO en PDF, maar niet in de text opgenomen.
- Het xml-formaat bevat, als er op artikelen ontsloten wordt, een <title>-element voor de koppen. Dit is verplicht. Als het artikel geen kop heeft, wordt de eerste regel van het artikel als kop genomen (NB: de eerste regel moet in dat geval ook nog steeds in de tekst worden opgenomen). Meerdere (sub)koppen boven een artikel kunnen samen in één title-element worden opgenomen.
- Bij kranten en eventueel tijdschriften: Als een artikel bestaat uit een illustratie met onderschrift zonder kop, dan wordt de 1e regel van het onderschrift als kop genomen. Bij een advertentie, met of zonder tekst, wordt 'Advertentie' als kop genomen.
- Specifieke titels voor kranten: Bij een familiebericht wordt 'Familiebericht' als kop genomen.
- Specifieke titels voor tijdschriften: Bij een omslag wordt 'Omslag' als kop genomen. Bij een colofon wordt de titel van het colofon als kop genomen (bijvoorbeeld 'Van onze redactie'), tenzij er geen titel aanwezig is dan wordt de waarde 'Colofon' als kop genomen. Bij een register wordt de titel van het register als kop genomen (bijvoorbeeld 'Index'), tenzij er geen titel aanwezig is dan wordt de waarde 'Register' als kop genomen. Bij een inhoudsopgave wordt de titel van de inhoudsopgave als kop genomen (bijvoorbeeld 'Inhoud'), tenzij er geen titel aanwezig is dan wordt de waarde 'Inhoudsopgave' als kop genomen.

Groepering advertenties en familieberichten

Advertenties (KT) en familieberichten (K) van eenzelfde breedte die onder elkaar staan mogen gegroepeerd worden en moeten dan als titel "Advertentie" of "Familiebericht" krijgen (hoofdletter A of F en enkelvoud).

Alleen advertentiepagina's en een kolomadvertenties en familieberichten met dezelfde breedte onder elkaar mogen gegroepeerd.

OCR correctie

- De OCR van koppen van de artikelen dient gecorrigeerd te worden tot een niveau van minstens 99,8% per batch. Dat wil zeggen dat minstens 99,8% van alle tekens in de koppen correct dient te zijn.

- Bij tijdschriften die in artikelen worden gesegmenteerd dient de OCR van de auteursnamen van artikelen gecorrigeerd te worden tot een niveau van minstens 99,8% per batch. Dat wil zeggen dat 99,8% van alle tekens in de <author> elementen correct dient te zijn.
- Voor de overige delen van de tekst hoeft alleen correctie te worden toegepast als daarom in een specifiek project gevraagd is.
- De definitieve OCR-tekst (na eventuele correcties of andere nabewerkingen) moet worden opgenomen. Indien er correcties op ge-OCR-de tekst worden gedaan, dan moeten die overal waar die tekst staat, worden opgenomen: in de ALTO, in de PDF en in de tekstbestanden. Correcties op kranten- en tijdschriftkoppen en namen van gemarkeerde auteurs moeten bovendien ook in de metadata worden opgenomen.

Schema's

De text-bestanden die pagina's bevatten moeten valideren tegen het schema waarvan de naam als volgt is opgebouwd: kbdg_text_page.xsd. De text-bestanden die artikelen bevatten moeten valideren tegen het schema waarvan de naam als volgt is opgebouwd: kbdg_text_article.xsd.

Element	BKT	Beschrijving	Card
/text			1
/text/author	T		0-N
/text/title	KT	Voor T alleen als er op artikel wordt ontsloten	1
/text/author	T		0-N
/text/p			1-N
/text/author	T		0-N

4.5.12 ALTO

Door middel van ALTO-bestanden wordt de layout van de originele pagina's vastgelegd. Van alle tekstblokken, tekstregels en woorden worden lettertype, fontsize etc. vastgelegd.

Tevens worden illustraties en grafische elementen vastgelegd.

Van alle elementen op de pagina worden de coördinaten vastgelegd ten opzichte van de afbeelding.

Dit maakt het ondermeer mogelijk om zoekresultaten binnen een tekst te highlighten op de digitale afbeelding van de pagina.

Schema

De ALTO-bestanden moeten valideren tegen een door de KB meegeleverd schema waarvan de naam als volgt is opgebouwd: kbdg_alto.xsd.

Voor ALTO gelden de volgende eisen

- De KB hanteert voor deze specificaties ALTO versie 2.0 (<http://www.loc.gov/standards/alto/>)
- Per masterbestand wordt één ALTO-bestand geleverd.
- Het ALTO-bestand moet worden gecreëerd nadat eventuele nabewerkingen op het bijbehorende masterbestand zijn gedaan, zoals rechtzetten, croppen e.d. Dit om te voorkomen dat de coördinaten niet meer overeenkomen.
- De coördinaten moeten overeenkomen met het master- en accessbestand. Dit betekent dat de accessbestanden qua beeld, op de resolutie na, één op één gelijk dienen te zijn aan de masters; d.w.z. dat de afgeleiden inhoudelijk niet afwijken van de masters.
- De ALTO-bestanden moeten zo volledig mogelijk zijn met informatie over stijl, layout en procesinformatie. In het layout-deel moeten alle elementen van een pagina volledig en juist zijn opgenomen: tekst, illustraties en grafische elementen.
- Alle tekst van de pagina is opgenomen, inclusief paginanummers, kopteksten, voetteksten e.d. in de marges.
- Alle tekst moet in de juiste leesvolgorde worden opgenomen.
- Eventueel aanwezige verticale tekst moet ook worden opgenomen.
- Indien er correcties op ge-OCR-de tekst worden gedaan, dan moeten die overal waar die tekst staat, worden opgenomen: in de ALTO, in de PDF en in de tekstbestanden.

- Correcties op krantenkoppen (en eventuele tijdschriftkoppen) moeten bovendien ook in alle relevante metadatabestanden worden opgenomen

Dubbele pagina's (alleen voor tijdschriften)

Indien de layout over meerdere pagina's loopt (bijvoorbeeld een tabel die over twee pagina's is afgedrukt), dienen master- en accessbestanden aangeleverd te worden per 'opening' (= opname van dubbele pagina's). De text die voor deze dubbele pagina's wordt gemaakt moet in dit geval ook als één pagina worden beschouwd en in één ALTO-bestand worden opgenomen.

De "Double Pages" worden na verwerking als één master image geleverd en moeten als één pagina gefactureerd worden.

Interpretatie leestekens

- Koppeltekens zijn onderdeel van de string als ze niet aan het einde van een regel staan. Koppeltekens aan het einde van de regel moeten geïnterpreteerd worden als afbreekstreepje.
- Leestekens moeten bij de string ervoor worden opgenomen als er geen spatie is tussen de string en het koppelteken. Als er een spatie is, moet het leesteken als afzonderlijke string worden opgenomen.

Hieronder een specificatie van de gevraagde elementen. Waar nodig staat in de laatste kolom uitgelegd hoe deze elementen gevuld moeten worden.

Element	Path	Attributen	Beschrijving
MeasurementUnit	Description/MeasurementUnit		De meeteenheid voor de coördinaten moet 'pixel' zijn. De coördinaten moeten overeenkomen met het bijbehorende beeld
fileName	Description/sourceImageInformation/fileName		Bestandsnaam van de bijbehorende afbeelding.
OCRProcessing	Description/OCRProcessing	ID	
processingStepDescription	Description/OCRProcessing/preProcessingStep/processingStepDescription		'An ordinal listing of the image processing steps performed. For example, "image despeckling." or "align" '
processingDateTime	Description/OCRProcessing/ocrProcessingStep/processingDateTime		" Date or DateTime the image was processed."
processingAgency	Description/OCRProcessing/ocrProcessingStep/processingAgency		"Identifies the organizationlevel producer(s) of the processed image."
processingStepSettings	Description/OCRProcessing/ocrProcessingStep/processingStepSettings		"A description of any setting of the processing application. For example, for a multi-engine OCR application this might include the engines which were used. Ideally, this description should be adequate so that someone else using the same application can produce identical results."
softwareCreator	Description/OCRProcessing/ocrProcessingStep/processingSoftware/softwareCreator		De naam van de organisatie of het bedrijf dat de applicatie heeft gemaakt.
softwareName	Description/OCRProcessing/ocrProcessingStep/processingSoftware/softwareName		De naam van de applicatie.
softwareVersion	Description/OCRProcessing/ocrProcessingStep/processingSoftware/softwareVersion		De versie van de applicatie.
Textstyle	Styles/Textstyle	ID, FONTFAMILY, FONTSIZE, FONTSTYLE	
ParagraphStyle	Styles/ParagraphStyle		
Page	Layout/Page	ID, HEIGHT, WIDTH, PHYSICAL_IMG_NR, PC	De waarde van ID moet zijn Page + volgnummer binnen het object, dus bijvoorbeeld 'Page1'. De waarde van PHYSICAL_IMG_NR is altijd '1'. In het attribuut PC de Page Confidence wordt het betrouwbaarheidspercentage opgenomen als een reëel getal van minimaal 0 en maximaal 1.

Element	Path	Attributen	Beschrijving
TopMargin	Layout/Page/TopMargin	HEIGHT, WIDTH, HPOS, VPOS	Deze attributen identificeren de 'bounding box' van een rechthoek. HPOS en VPOS zijn de x en y-coördinaten van de linkerbovenhoek van de 'bounding box'. HPOS en VPOS zijn relatief aan de linkerbovenhoek van de gehele pagina, die een HPOS en VPOS heeft van (0,0). HEIGHT en WIDTH zijn de hoogte en breedte van de 'bounding box'.
LeftMargin	Layout/Page/LeftMargin	HEIGHT, WIDTH, HPOS, VPOS	
RightMargin	Layout/Page/RightMargin	HEIGHT, WIDTH, HPOS, VPOS	
BottomMargin	Layout/Page/BottomMargin	HEIGHT, WIDTH, HPOS, VPOS	
PrintSpace	Layout/Page/PrintSpace	HEIGHT, WIDTH, HPOS, VPOS	
TextBlock	Layout/Page/PrintSpace/Textblock	ID, HEIGHT, WIDTH, HPOS, VPOS, STYLEREFS	De waarde van ID moet zijn 'TextBlock' + volgnummer van het textblock binnen de pagina, dus bijvoorbeeld 'TextBlock15'. STYLEREFS bevat twee verwijzingen, de eerste naar het ID van een TextStyle element, de tweede naar het ID van een ParagraphStyle element, allebei opgenomen onder element Styles in de xml.
TextLine	Layout/Page/PrintSpace/Textblock/TextLine	ID, WIDTH, HPOS, VPOS, HEIGHT	De waarde van ID moet zijn 'TextLine' + volgnummer van het textline binnen de pagina, dus bijvoorbeeld 'TextLine20'.
String	Layout/Page/PrintSpace/Textblock/Textline/String	ID, HEIGHT, WIDTH, HPOS, VPOS, CONTENT, SUBS_TYPE, SUBS_CONTENT, WC, CC, STYLEREFS	De waarde van ID moet zijn 'String' + volgnummer van het string binnen de pagina, dus bijvoorbeeld 'String111'. Afbreekstreepjes aan het einde van de regel moeten als volgt worden behandeld: -het eerste deel van het woord wordt opgeslagen in het CONTENT attribuut.

Element	Path	Attributen	Beschrijving
			<p>-het SUBS_TYPE attribuut bevat de waarde 'HypPart1'.</p> <p>-het SUBS_CONTENT attribuut bevat het hele woord (eerste en tweede deel zonder afbreekstreepje)</p> <p>Voor het volgende STRING element bevat het SUBS_TYPE attribuut de waarde 'HypPart2' en moet het SUBS_CONTENT attribuut ook worden toegevoegd.</p> <p>Koppeltekens die niet aan het einde van een regel staan, moeten worden beschouwd als onderdeel van een reguliere 'string', en moeten worden opgenomen in de CONTENT zoals in geval van 'Staten-Generaal'. Dit geldt ook voor afbreekstreepjes achter het laatste woord op een pagina.</p> <p>CC bevat "Confidence level of each character in that string. A list of numbers, one number between 0 and 9 for each character.", waarbij 0 het meest zeker aangeeft.</p> <p>WC bevat "Word Confidence: Confidence level of the ocr for this string. A value between 0 and 1.", waarbij 1 aangeeft dat met 100% zekerheid het juiste woord is herkend.</p> <p>STYLEREFS bevat een verwijzing naar het ID van een textstyle element opgenomen onder element Styles in de XML. Dit attribuut wordt alleen opgenomen als de stijl van een string afwijkt van de bij het Textblock gedefinieerde stijl.</p>
SP	Layout/Page/PrintSpace/Textblock/Textline/SP	ID, WIDTH, HPOS, VPOS	De waarde van ID moet zijn 'SP' + volgnummer van de space binnen de pagina, dus bijvoorbeeld 'SP26'. Het SP element wordt gebruikt voor 'white space'.

Element	Path	Attributen	Beschrijving
HYP	Layout/Page/PrintSpace/Textblock/Textline/HYP	WIDTH, HPOS, VPOS, CONTENT	Dit element bevat het afbreekstreepje (zie ook het String element).
ComposedBlock /GraphicalElement	Layout/Page/PrintSpace/ComposedBlock En Layout/Page/Printspace/ComposedBlock/ GraphicalElement	<u>ComposedBlock</u> ID, HEIGHT, WIDTH, HPOS, VPOS, TYPE <u>GraphicalElement</u> ID, HEIGHT, WIDTH, HPOS, VPOS	Plaatjes en grafische elementen worden opgenomen in een ComposedBlock met GraphicalElement. Eventuele bijschriften worden ook binnen het ComposedBlock in een TextBlock opgenomen. <u>ComposedBlock</u> De waarde van ID moet zijn 'Picture_' + volgnummer van het ComposedBlock binnen de pagina, dus bv. 'Picture_2'. De waarde van TYPE moet zijn 'Illustration'. <u>GraphicalElement</u> De waarde van ID moet zijn 'Picture_' + volgnummer van het ComposedBlock binnen de pagina + '_SUB', dus bv. 'Picture_2_SUB'.

4.5.13 Pakbon

De opdrachtnemer levert bij elke versie van een batch met digitaal materiaal een pakbon, die beschrijft wat er in de batch wordt geleverd.

De pakbon is een XML-bestand dat moet voldoen aan het bijgeleverde XML-schema en de algemene eisen voor XML-bestanden.

Indien een batch opnieuw geleverd wordt, dient ook een nieuwe pakbon te worden aangeleverd, met onder meer de nieuwe batchnaam, aangepaste disknaam e.d.

De inhoud van de pakbon moet deels worden overgenomen uit de door de KB aangeleverde metadatadump en deels worden samengesteld uit de tijdens het digitaliseringsproces gecreëerde informatie of de eigen administratie van de opdrachtnemer.

De aantallen in de pakbon dienen correct te zijn en worden gecontroleerd. Omdat deze aantallen de basis vormen voor de facturering, wordt er om een nieuwe pakbon verzocht in geval de aantallen op de pakbon niet correct zijn. Onjuiste informatie op de pakbon is reden tot afkeuring van de batch.

Hieronder staat beschreven welke xml-elementen er in de pakbon moeten worden opgenomen. Met "publication" wordt bedoeld een aflevering van een krant of tijdschrift.

Schema

De pakbon valideert tegen het door de KB meegeleverd schema kbdg_pakbon.xsd.

Toelichting:

Kolom KT beperkt, inden gevuld, de scope van een element of attribuut: als er niets in deze kolom staat, is het element of attribuut op alle materiaalstromen (Kranten, Tijdschriften) van toepassing. Staat er 'D', dan hoeft het element of attribuut alleen ingevuld te worden als er digitale bestanden worden aangeleverd door de KB. D=digitaal brontype. Elementen en attributen gemarkeerd met 'M' worden alleen ingevuld als het digitale bestanden afkomstig van een microfilm betreft.

Element of attribuut	KT	Beschrijving	Waarde / formaat	Metadatadump / genereren door leverancier
/deliveredBatch	KT			
/deliveredBatch/@shipmentID	KT	ID van het transport van de KB naar de leverancier.		/shipment/@ID
/deliveredBatch/@dateFromKB	KT	Datum van het transport van de KB naar de leverancier.	yyyy-mm-dd	/shipment/@dateFromKB
/deliveredBatch/@dateToKB	KT	Datum van het transport van de leverancier aan de KB.	yyyy-mm-dd	Genereren door leverancier
/deliveredBatch/@diskNameKB	KT DM	Naam van de schijf waarop het transport door de KB aan de leverancier wordt geleverd.		/shipment/@diskNameKB
/deliveredBatch/@diskName	KT	Naam van de schijf waarop de batch door de leverancier wordt aangeleverd.		Genereren door leverancier
/deliveredBatch/@addressee	KT	Geadresseerde van het transport.	Standaardwaarde: "Afdeling Digitalisering"	Genereren door leverancier.
/deliveredBatch/@batchName	KT	Naam van de batch.		/shipment/@batchName Deel- en versienummer toe te voegen door leverancier
/deliveredBatch/@deelprojectCode	KT	Code van het deelproject.		/shipment/@deelprojectCode
/deliveredBatch/@numberOfEntities	KT	Aantal banden (barcodes) dat in het transport is opgenomen. Voor digitale aanlevering geldt hier het aantal directories met als naam de barcode.		/shipment/@numberOfEntities
/deliveredBatch/@deelprojectOwner	KT	Bibliotheek of instelling die het digitaliseringsproject heeft ingediend. Kan verschillen van de eigenaar van het te digitaliseren exemplaar.		/shipment/@deelprojectOwner
/deliveredBatch/@collection	KT	Beschrijft de collectie waar de publicatie onderdeel van uitmaakt.		/shipment/@collection
/deliveredBatch/@material	KT	Materiaalsoort.	(kranten tijdschriften)	/shipment/@material
/deliveredBatch/@project	KT	Aanbestedingsronde.		/shipment/@project
/deliveredBatch/@sourceType	KT	Aard van het bronmateriaal dat gedigitaliseerd wordt.	(origineel microfilm digitaal)	/shipment/@sourceType

Element of attribuut	KT	Beschrijving	Waarde / formaat	Metadatadump / genereren door leverancier
/deliveredBatch/@KBcontractorID	KT	Door de KB uitgegeven code voor de leverancier die de opdracht heeft aangenomen.		/shipment/@KBcontractorID
/deliveredBatch/@contractor	KT	Leverancier die de opdracht heeft aangenomen.		/shipment/@contractor
/deliveredBatch/@specificationsVersion	KT	Versienummer van deze metadataspecificaties.		/shipment/@specificationsVersion
/deliveredBatch/files	KT			
/deliveredBatch/files/@totalFileSize	KT	Totale bestandsomvang van alle bestanden in de batch in bytes.		Genereren door leverancier
/deliveredBatch/files/@numberOfFiles	KT	Totaal aantal bestanden in de batch.		Genereren door leverancier
/deliveredBatch/files/@averageFileSize	KT	Gemiddelde bestandsomvang van alle bestanden in de batch in bytes.		Genereren door leverancier
/deliveredBatch/files/fileType/ @fileName	KT	Naam van het bestandstype.	(access alto checksums concordantietabel master mets pdf targets-jp2_access targets-jp2_master text)	Genereren door leverancier
/deliveredBatch/files/fileType/@totalFileSize	KT	Omvang van het totaal aantal bestanden van het bestandstype in bytes.		Genereren door leverancier
/deliveredBatch/files/fileType/ @numberOfFiles	KT	Aantal bestanden van het bestandstype.		Genereren door leverancier
/deliveredBatch/files/fileType/ @averageFileSize	KT	Gemiddelde omvang van het bestandstype per bestand in bytes.		Genereren door leverancier
/deliveredBatch/(books booksMicrofilm journals journalsMicrofilm newspapers newspapersMicrofilm)	KT			/shipment/(books booksMicrofilm journals journalsMicrofilm newspapers newspapersMicrofilm)
/deliveredBatch/*/publication	KT			
/deliveredBatch/*/publication/@ID	KT	Unieke object identifier van de digitale publicatie (dit is de identifier van het digitale object, die kan verschillen van de		Genereren door leverancier

Element of attribuut	KT	Beschrijving	Waarde / formaat	Metadatadump / genereren door leverancier
		identificer(s) in de materiaalvoorbereidingsdatabase van de KB).		
/deliveredBatch/*/publication/@recordID	KT	ID uit de KB-materiaalvoorbereidingsdatabase van de titelbeschrijving van de periodiek waar de afzonderlijke banden en afleveringen bijhoren.		/shipment/*/record/@ID
/deliveredBatch/*/publication/@PPN	KT	Verwijzing naar de titelbeschrijving in de bibliotheekcatalogus.		/shipment/*/record/@PPN
/deliveredBatch/*/publication/@title	KT	Titel van de periodiek waartoe de afzonderlijke banden en afleveringen behoren.		/shipment/*/record/@title
/deliveredBatch/*/publication/@shelfmarkFilm	KT M	Aanvraagnummer van de microfilm.		/shipment/*/record/@shelfmarkFilm
/deliveredBatch/*/publication/@barcodeFilm	KT M	Barcode van de microfilm in de materiaalvoorbereidingsdatabase van de KB.		/shipment/*/record/@barcodeFilm
/deliveredBatch/*/publication/@barcode	KT	Barcode van de band in de materiaalvoorbereidingsdatabase van de KB. Voor convoluten kunnen meerdere barcodes in een band voorkomen. Bij digitale aanlevering geldt een map met scans als een band.		/shipment/*/entity/@barcode
/deliveredBatch/*/publication/@barcodeSourceProvider	KT	Code voor de band in het eigen systeem van de instelling.		/shipment/*/entity/@barcodeSourceProvider
/deliveredBatch/*/publication/@shelfmark	KT	Planknummer van het exemplaar in het magazijn van de instelling.		/shipment/*/entity/@shelfmark
/deliveredBatch/*/publication/@sourceProvider	KT	Bibliotheek of instelling die eigenaar is van het exemplaar.		/shipment/*/entity/@sourceProvider
/deliveredBatch/*/publication/@issueID	KT	ID uit de KB-materiaalvoorbereidingsdatabase van de aflevering van een periodiek.		/shipment/*/entity/issue/@ID
/deliveredBatch/*/publication/@EPN	KT	Verwijzing naar de aflevering in de bibliotheekcatalogus.		/shipment/*/entity/issue/@EPN

Element of attribuut	KT	Beschrijving	Waarde / formaat	Metadatadump / genereren door leverancier
/deliveredBatch/*/publication/@sequenceNo	KT	Volgnummer in de band. Dit is alleen van toepassing op banden waarin meerdere titels zijn opgenomen zoals convoluten en periodieken. In dat geval moet hier worden opgenomen "de hoeveelste" titel of aflevering dit in de band is.		/shipment/*/entity/issue/@sequenceNo
/deliveredBatch/*/publication/@publicationYear	T	Jaar van publicatie van de aflevering (bibliografische notatie).		/shipment/*/entity/issue/@publicationYear
/deliveredBatch/*/publication/@publicationDate	K	Volledige publicatiedatum datum van de aflevering.	yyyy-mm-dd	/shipment/*/entity/issue/@publicationDate
/deliveredBatch/*/publication/@edition	K	Dagdeel waarin het krantennummer is verschenen.	(Ochtend Middag Avond Dag)	/shipment/*/entity/issue/@edition
/deliveredBatch/*/publication/comments	KT	Opmerkingen die van belang kunnen zijn voor het digitaliseren.		/shipment/*/entity/comments

4.5.14 Concordantietabel

Een batch bevat altijd een bijbehorende concordantietabel.

Per batch wordt één concordantietabel geleverd (indien een batch in delen is gesplitst, zijn dit in feite meerdere batches met ieder hun eigen concordantietabel). Deze voldoet aan de volgende specificaties:

- De concordantietabel bevat alle pagina's en/of uitvouwbladen die van een document/object zijn opgenomen. Alle pagina's en uitvouwbladen worden gerelateerd aan het juiste object.
- Alle geleverde bestanden in de batch behalve de checksumbestanden, de concordantietabel zelf en de pakbon moeten in de concordantietabel zijn opgenomen.
- Alle in de concordantietabel genoemde bestanden zijn in de batch geleverd.
- Er mogen geen bestanden in de batch zijn opgenomen, die niet in de concordantietabel staan (uitgezonderd de checksumbestanden, de concordantietabel zelf en de pakbon).
- De concordantietabel heeft de vorm van een csv-bestand.
- Het scheidingsteken is een puntkomma (semicolon ;).
- Rondom de waarden worden aanhalingstekens (quotes) gebruikt. De reden hiervoor is dat zo voorloophnullen in bijvoorbeeld volgnummers niet verdwijnen door datatypering in de verwerkende software. Ook voorkomt dit dat tekens die als scheidingstekens kunnen worden opgevat voor problemen zorgen.
- Als eerste regel worden kopjes voor de kolommen opgenomen. Hierbij wordt gebruik gemaakt van standaard benamingen voor de bestandstypen.
- Voor de volgnummers wordt "volgnr" als kopje gebruikt.
- Voor de technische targets wordt het type target als kop gebruikt.
- Voor tijdschriften die niet op artikel niveau worden ontsloten betreft elke regel van de concordantietabel één opname.
- Voor kranten en tijdschriften die wel op artikelniveau worden ontsloten bevat de concordantietabel één regel voor elke unieke combinatie van een opname en artikel(deel) op die opname (zie voorbeeld aan het einde van dit hoofdstuk). Voor tijdschriftpagina's die geen artikel bevatten (bv. bij voorplat, blanco pagina) hoeft ook geen Tekst-bestand te worden vastgelegd.
- Er mag geen Byte Order Mark (BOM) voorkomen in de concordantietabel.
- De concordantietabel moet de volgende kolommen bevatten:

Kolom	Inhoud	Kolomnaam
Object identifier	Unieke identifier van het digitale object: /shipment/*/entity/@barcode, voorzien van volgnummer in de laatste drie posities Zie PvE 4.1.2 Object Identifiers.	ID
Volgnummer	Volgnummer (vijfcijferig met voorloophnullen). Dit is het logische volgnummer dat de positie van de pagina (of, bij dubbele pagina's, de scan) in het object weergeeft, opeenvolgend en oplopend, beginnend bij 00001 voor elk nieuw object.	volgnr (is gelijk aan /mets:div/@ORDER voor dezelfde pagina/scan)
Master images	Bestandsnaam master image	master
Access images	Bestandsnaam access image	access
Structurele metadata METS	Bestandsnaam METS-bestand	mets
Coördinaten (ALTO)	Bestandsnaam ALTO-bestand	alto
Text (KB text format)	Bestandsnaam text-bestand	text
PDF	Bestandsnaam PDF	pdf

Kolom	Inhoud	Kolomnaam
Targets (1 kolom voor elk type dagtarget; deze kolommen vervallen bij verwerking van door de KB aangeleverde digitale bestanden)	Bestandsnamen targets	<p>“targets<spatie>[targettype]<spatie>[‘master’ of ‘access’]”.</p> <p>Voor [targettype] zijn de mogelijke waarden TON, COL, MTF GRAY, MTF, RGB, GEO.</p> <p>Voorbeeld van kolomnaam: “targets GEO access”</p>

- Op iedere regel van de concordantietabel staan alle bestanden die bij een bepaalde opname horen in de juiste kolom. Bijvoorbeeld: het masterbestand, het accessbestand en het ALTO-bestand van een opname. Zo is te zien welke bestanden bij elkaar horen.
- Bestanden die bij meer dan één opname horen (bijvoorbeeld technische targets en structurele metadata) worden herhaald op iedere regel van het object waar ze bij horen.
- Op iedere regel staat vermeld bij welk object de bestanden horen door middel van de identifier van het object.
Dat wil zeggen dat bij iedere identifier de juiste bestandsnamen in de overige kolommen zijn opgenomen en dat omgekeerd de bestandsnamen aan de juiste object identifier zijn gekoppeld.
- Er worden volgnummers in de concordantietabel opgenomen die aangeven wat de positie van de opname binnen het document/object is. Deze zijn vijfcijferig, waarbij zo nodig voorloophulpen worden gebruikt.
- De object identifier in de concordantietabel komt overeen met de object identifier in het METS-bestand (/mets:mets/@OBJID), die ook gebruikt wordt in de bestandsnamen en in de directories.
- De naamgeving van de bestanden is volgens de specificaties, zie paragraaf 4.1.4.
- Naast alle bestanden die bij objecten horen, worden in de concordantietabel ook de technische targets opgenomen (ieder type target in een aparte kolom)

Voorbeeld van regels (alleen de relevante kolommen) uit een concordantietabel voor een artikel dat over 2 pagina's doorloopt:

ID	Volgnr	...	master	...	text
MMMVC01_000000030	00009	...	MMMVC01_000000030_00009_master.jp2	...	MMMVC01_000000030_00027_text.xml
MMMVC01_000000030	00010	...	MMMVC01_000000030_00010_master.jp2	...	MMMVC01_000000030_00027_text.xml

5 Acceptatiecriteria

De geleverde batches moeten op alle punten voldoen aan de specificaties die in dit Pakket van Eisen zijn opgenomen. De KB controleert de kwaliteit van het geleverde materiaal uitgebreid en op verschillende manieren. Voor bijna alle criteria is 100% correctheid vereist, omdat het anders niet mogelijk is de bestanden adequaat op geautomatiseerde wijze te verwerken. Indien de KB akkoord gaat met iets minder dan 100% correctheid, dan wordt dat vermeld bij de betreffende eis/specificatie.

Geleverde batches die niet door de controle applicatie van de KB verwerkt kunnen worden (bv. doordat er bepaalde mappen/bestanden missen of doordat de naamgeving van bestanden niet kloppen) zullen niet verder gecontroleerd en direct afgekeurd worden. Eventuele fouten die in de volgende versie van deze batch voorkomen, kunnen alsnog tot afkeuring leiden.

Aangenomen wordt dat aan de kant van de opdrachtnemer eveneens sprake is van een automatisch productieproces, waardoor consistentie en kwaliteit worden gewaarborgd.

Het kan voorkomen dat de KB een batch goedkeurt waarin bij nader inzien toch een fout blijkt te zitten waarop de batch afgekeurd had moeten worden. Deze batch blijft goedgekeurd. De KB heeft wel het recht nieuwe (nog te controleren) batches op de betreffende fout af te keuren.

Bijlage A: Omgang met het materiaal ten behoeve van digitalisering

Omvang

Maximaal 750.000 pagina's (volgens transportgegevens KB) mogen bij de Leverancier in de opslag aanwezig zijn.

In overleg met KB mag daarvan afgeweken worden.

Algemeen

1. Nuttig geen voedsel of dranken in de ruimte waar de collectie is (depot, scanruimte).
2. Werk met schone en droge handen, draag geen grote sieraden (ringen, armbanden, kettingen). Gebruik geen handcrème vóór het hanteren van de banden.
3. Lik niet aan de vinger bij het omslaan van een bladzijde. Gebruik ook géén bevochtigers in de vorm van sponsjes. Eventueel kunnen nitriël vingerhoedjes worden gebruikt.
4. Zelfklevende notities (Post-its) mogen nooit op of in het materiaal geplakt worden. Gebruik zuurvrije stroken papier als tijdelijke verwijzing. Verwijder tijdelijke verwijzingen direct na het scannen.
5. Gebruik nooit een pen in de nabijheid van het materiaal. Schrijf met potlood en schrijf en leun niet boven op de banden.

Banden

6. Druk een band nooit met kracht open. Open de band altijd eerst van het midden uit en blader van daar terug naar de 1^{ste} pagina.
7. Ondersteun voor- en achterplat indien nodig. Bijvoorbeeld bij een losse rug die uitstulpt bij het openen van de band zodat de rug de glasplaat niet raakt tijdens het digitaliseren.
8. Vouw uitklapplaten dicht op dezelfde wijze waarop je ze geopend hebt. Maak geen 'nieuwe' vouwen.
9. Mocht een onderdeel van een object loslaten, gooi dit losse deel dan niet weg. Bewaar het in een enveloppe, noteer waar het bij hoort (titel, jaar en bandnummer) en waarschuw de KB.
10. Schadeafhandeling conform Bijlage B

Veiligheid en transport

11. Laat het materiaal nooit onbeheerd achter. Berg het materiaal tijdens korte afwezigheid in een afgesloten ruimte op en aan het einde van een werkdag terug in de krat in het depot.
12. Vervoer de kratten altijd met een karretje. Draag geen (stapels) banden!

Bijlage B: Schade-afhandeling na digitaliseren

Stap	Wie	Omschrijving handeling	Activiteiten	Opmerkingen
1	Leverancier	Scannen	<ul style="list-style-type: none"> Het materiaal wordt gescand 	
1a	Leverancier	Leverancier meldt actuele schade op SharePoint	<ul style="list-style-type: none"> Leverancier meldt onmiddellijk de schade op Sharepoint, met barcode en aard van schade, ook plaatst de Leverancier een target op de plaats van schade in de band 	
1b	Digitalisering: Conserveringsadviseur	SharePoint	<ul style="list-style-type: none"> Conserveringsadviseur constateert schademeldingen 	
1c	Collectiebehoud: Conserveringsadviseur	Overleg bij grote schade met Leverancier en, indien nodig, de instelling	<ul style="list-style-type: none"> In het geval van grote schade vindt er overleg plaats tussen de conserveringsadviseur en de Leverancier en, indien nodig, de instelling 	Let op: laat de productiemangers even weten dat er contact is hierover met de Leverancier en ook melden op sharepoint
2	Digitalisering: Team Materiaalvoorbereiding	Ontvangen geretourneerd materiaal	<ul style="list-style-type: none"> Team Materiaalvoorbereiding ontvangt het materiaal 	
3	Digitalisering: Team Materiaalvoorbereiding	Uitpakken materiaal binnen 10 dagen na ontvangst	<ul style="list-style-type: none"> Het materiaal wordt binnen 10 dagen na ontvangst uitgepakt 	
4	Digitalisering: Team Materiaalvoorbereiding	Terugboeken van het materiaal in de MVB database en het controleren van het materiaal ahv de database en de reeds gemaakte foto's van de schadegevallen	<ul style="list-style-type: none"> Het materiaal wordt op volledigheid gecontroleerd, de schadegevallen gecheckt en daarna teruggeboekt in de database 	
5	Digitalisering: Team Materiaalvoorbereiding	Na terugboeken/controle een mail vanuit team materiaalvoorbereiding dat controle is gedaan en dat: 1. Er géén schades en- of bijzonderheden zijn. 2. Dat er schadegevallen zijn.	e-mail aan alle betrokkenen: Teamleider MVB, Coördinatoren, Conserveringsadviseur en Senior Productiemangers.	
6	Team Digitalisering: Conserveringsadviseur	Direct melden van nieuwe schadegevallen aan Leverancier	Via SharePoint	
7	Collectiebehoud: Conserveringsadviseur	Opstellen schaderapport.	<ul style="list-style-type: none"> Conserveringsadviseur stelt schaderapport op, plaatst deze op de M schijf en op sharepoint en meldt dit bij de Senior Productiemangers Schaderapport wordt opgesteld binnen 5 werkdagen 	
8	Collectiebehoud: Conserveringsadviseur	Zoekt contact met de instelling over het afhandelen van de schadegevallen.	-	

Stap	Wie	Omschrijving handeling	Activiteiten	Opmerkingen
9	Collectiebehoud: Conserveringsadviseur	Conserveringsadviseur handelt schadegevallen af	<ul style="list-style-type: none"> • Hiervoor zijn twee opties: <ol style="list-style-type: none"> 1. Schade wordt hersteld door een extern restauratiebedrijf Leverancier betaalt de kosten 1. Instelling besluit om de schade niet te herstellen 	Indien optie 1 is gekozen: Conserveringsadviseur laat een offerte maken door een extern restauratiebedrijf, stelt de Leverancier hiervan op de hoogte. Restauratie wordt uitgevoerd en factuur wordt doorgestuurd naar de Leverancier. Offerte en factuur worden aan Leverancier gericht. Leverancier betaalt factuur aan restauratiebedrijf.
10	Collectiebehoud: Conserveringsadviseur	Informeert het team MVB hoe het materiaal aan de instelling moet worden terug geleverd.	-	

Bijlage C: JPEG 2000 Formaat parameters

In deze bijlage worden de formaat parameters uit de jpeg2000-specificaties uit paragraaf 4.3.2 nader toelicht.

File format

Defines the format of the produced files (in this case: JP2)

Compression type

Defines the compression used. Within JPEG 2000, both lossless (using the reversible 5-3 wavelet filter) and lossy (using the irreversible 9-7 filter) compression are possible.

Colour transform

Defines the use of a transformation that converts an RGB image to a luminance/chrominance (YC_bC_r) representation prior to compression. This significantly improves the compression efficiency for both lossless and lossy compression⁴.

Number of decomposition levels

This parameter is related to the number of resolution levels (rl) according to $rl = dl + 1$. According to Buckley & Tanner (2009) the first two levels result in better compression, whereas the levels beyond that are mainly added to facilitate the derivation of lower-resolution sub-images when reading the file (e.g. thumbnails).

Progression order

Defines the order in which packets are placed in the code stream. This is mainly important for access optimisation.

Tile size

Defines rectangular regions that are compressed independently. Although it is possible to treat the whole image as one single tile ('no tiles'), this may result in performance problems.

Code block size

Defines the size of the post wavelet-transform code blocks. Large code block values result in a reduction of file size, at the cost of increased memory consumption.

Precinct size

Defines rectangular regions that allow spatial accessibility at the sub-tile level. Precinct size is set separately for all resolution levels.

Number of quality layers

Defines the number of quality layers. The use of layers allows a decoder to access lower quality sub-images using less compressed data.

Target compression ratio

Defines the desired compression ratio (higher values indicate more compression). This is defined separately for all quality layers, where the value of the highest quality layer is the overall compression ratio. Target compression ratios for all lower quality layers follow a logarithmic distribution, as recommended by Buckley & Tanner (2009). Note that for the master images, no target compression ratio is defined for the highest quality layer, as this value corresponds to the overall lossless compression ratio (which depends of the characteristics of the captured image data).

Error resilience

Defines error resilience options. Using these options reduces the impact of any bit corruption of the file. A study by Heydegger (2009) showed that a combination of all of the error resilience options gives results that are more robust than using the options individually. Thus, all options must be used.

Sampling rate

The JP2 file specification defines two sets of header fields for storing grid resolution: “Capture Resolution” and “Default Display Resolution”. The use of the “capture resolution” fields is mandatory for KB projects.

ICC profiles

ICC profiles should always be embedded using the “Restricted ICC method”, as defined by the JP2 format specification. For monochrome images, the ICC profile defines a colour transformation that is described using a gray tone reproduction curve (TRC); for RGB spaces, the profile defines a three-component matrix-based colour transformation (N-component lookup tables are not allowed!)⁵. ICC profiles should always remain unaltered during the JP2 encoding process: for each JP2 the profile should be identical to the one in its TIFF source image.

Capture metadata

Metadata that describe the capture process (XMP TIFF fields in original TIFF image) should be embedded in XMP format in the JP2’s XML box.

Specs reference

Each file should contain a reference to the compression specifications that were used for its creation. This reference takes the form of a text string that is stored as a codestream comment in the JP2 file.