

DIGITALISERING & KWALITEIT

GIDS VOOR KWALITEITSBEHEER EN -CONTROLE IN
EEN PROJECT VOOR DE DIGITALISERING VAN ERFGOED
EN DOCUMENTEN

Schrijven: Chloé Brault

Bijdragers: Florence Gillet, Nicolas Roland

Proeflezen en wetenschappelijke leiding: Anne Chardonnes

Indeling: Chloé Brault

Projectcommissie ADOCHS: Nico Wouters (CegeSoma-AGR), Frédéric Lemmers (KBR), Ann Dooms (VUB), Tan Lu (VUB/KBR), Anne Chardonnes (CegeSoma-AGR/ULB), Nicolas Roland (KBR), Florence Gillet (CegeSoma-AGR), Seth van Hooland (ULB), Chloé Brault (CegeSoma-AGR)

Centre d'Etude Guerre et Société – Studie en Documentatiecentrum Oorlog en Hedendaagse Maatschappij
Square de l'Aviation 29 Luchtvaartsquare 1070 Bruxelles - Brussel

De opstelling van deze gids zou niet mogelijk zijn geweest zonder de financiering van het **Federaal Wetenschapsbeleid** onder contractnummer BR/154/A6/ADOCHS.

Institutionele partners



DIGITALISERING EN KWALITEIT

Gids voor kwaliteitsbeheer en -controle
in een project voor de digitalisering van erfgoed en documenten

Versie 1.1 - September 2021

Door Chloé Brault

Brussel
2021

INHOUDSOPGAVE

	Over deze gids	6
	Achtergrond	6
	Doel en werkingsgebied	7
	Methodologie en structuur	8
01.	Digitalisering: uitdagingen, veranderingen en perspectieven	9
02.	Kwaliteit: rode draad op organisatorisch vlak	14
	2.1 Strategie voor gegevensbeheer en digitaliseringsbeleid: de basis	16
	2.2 Bestek	18
03.	Kwaliteit van de deliverables: beeldbestanden	22
	3.1 Beeld: definitie en kenmerken	22
	Resolutie en definitie	22
	Coderingsdiepte	22
	Modus	23
	Kleurruimte	24
	Formaat	24
	3.2 Criteria voor een kwaliteitsbeeld	25
	Volledigheid	26
	Belichting	28
	Contrast	28
	Kleur	30
	Scherpstelling	30
	Resolutie	31
	Parallellisme	31
	Flare	32
	Vignettering	32
	Ruis	33
	Artefacten	33

04.	Kwaliteit van de deliverables: metadata	35
	4.1 Metadata: definitie en kenmerken	35
	4.2 Rol en functie van metadata	37
	4.3 Voorwaarde voor het gebruik van metadata: interoperabiliteit	37
	4.4 Criteria voor kwaliteitsvolle metadata	39
	4.5 Een gemeenschappelijke visie voor meer interoperabiliteit	40
05.	Beheer van het digitaliseringsatelier	45
	5.1 Behoud van erfgoed- en documentaire collecties	45
	5.2 Lichtbeheer	48
06.	Methodologische fiches om de kwaliteit van een digitaliseringsproject te waarborgen	51
	6.1 Benchmarks bepalen	53
	6.2 Implementatie en uitvoering van een kwaliteitsaanpak	57
	6.3 Kwaliteit van de deliverables bepalen	60
	6.4 Wanneer en hoe kwaliteitscontrole uitvoeren?	62
	6.5 Kwaliteitscontrole bij uitbesteding	65
	6.6 Digitaliserings-apparatuur kiezen	67
	6.7 Lichtbronnen kiezen	69
	6.8 De digitaliseringsketen kalibreren	71
	6.9 Algemene regels voor de beeldvorming	75
	6.10 Aanmaken en invullen van het metadata-referentiebestand	77
	6.11 Opslaan en naamgeving van bestanden	79
	6.12 Retoucheren van beeldbestanden	82
	6.13 OCR van tekstdocumenten	84
	6.14 Inkapselen van metadata	86
	6.15 De laatste kwaliteitscontrole	88
	6.16 Een digitaal databeheersysteem kiezen	89
	Bibliografie	92



OVER DEZE GIDS

Achtergrond

Sinds het midden van de jaren 1990 hebben culturele instellingen zonder enige twijfel hun intrede gemaakt in het digitale tijdperk. Net als overal ter wereld hebben overheden en particuliere bedrijven in België sinds het begin van de jaren 2000 tal van grootschalige digitaliseringsprojecten opgezet¹, en veel grote instellingen beschikken nu over een eigen digitaliseringscel.

Tijdens deze eerste fase was de kwaliteitscontrole een van de grootste struikelblokken. Veel projecten bleken immers de draagwijdte van deze stap, zowel op technisch als op personeelsvlak, in het totale digitaliseringsproces te hebben onderschat. In de meeste gevallen hadden de teams te kampen met een gebrek aan methodologische standaardisatie en aan automatiseringstools om de klus te klaren. Ze moesten daarom vaak handmatig werken, zonder gepaste procedurele richtlijnen voor hun specifieke behoeften.

Kwaliteitscontrole is echter een essentieel onderdeel in elke fase van een digitaliseringsproject om de integriteit en consistentie van de geproduceerde bestanden en gegevens en hun bewaring op lange termijn

te garanderen. Dit geldt zowel voor uitbestede als voor interne digitaliseringsprojecten.

Tegelijkertijd leidde de veelheid van actoren onvermijdelijk tot verschillen in het eindresultaat van de digitalisering. Hoe zorg je ervoor dat een digitaal bestand aan het verwachte kwaliteitsniveau voldoet en dat het een getrouwe en volledige weergave is van het originele document, zodat de gebruiker de integriteit ervan niet in twijfel trekt? Hoe zorg je dat gelijksoortige documenten na digitalisering dezelfde kenmerken hebben, ongeacht de instelling, de operator, de werkplaats en de gebruikte apparatuur?

Uit deze overwegingen werd het onderzoeksproject **ADOCHS** (*Auditing Digitalization Outputs in the Cultural Heritage Sector*) geboren, dat in 2016 van start ging.

Op initiatief van het Rijksarchief – en meer bepaald van CEGESOMA² – in samenwerking met de Koninklijke Bibliotheek van België (KBR), de Vrije Universiteit Brussel (VUB) en de Université libre de Bruxelles (ULB), heeft het ADOCHS-project tot doel de kwaliteitscontrole te verbeteren van de geproduceerde bestanden in het kader van de digitalisering van het erfgoed. Op die manier wil het project het hele digitaliseringsproces

versnellen, de kosten ervan tot een minimum beperken en de waarde van de geproduceerde gegevens in toekomstige projecten verhogen.

Doel en werkingsgebied

Deze gids is in de eerste plaats bestemd voor de digitaliseringsdiensten van de Belgische federale wetenschappelijke instellingen.

Scanmedewerkers vinden er technische aanbevelingen om collecties op een gepaste, homogene manier te digitaliseren. Collectiebeheerders en projectmanagers vinden er een overzicht van digitaliseringsprocessen en componenten, om hen te helpen een digitaliseringsproject secuur uit te werken.

Deze gids wil dus organisatorische en methodologische hulpmiddelen voorstellen om de kwaliteit te waarborgen van de digitale gegevens die in elke fase van het digitaliseringsproces worden geproduceerd.

Op die manier kan iedereen die bij het proces betrokken is, zijn werk zien als onderdeel van een productieketen met een gemeenschappelijk doel.

Het doel van deze gids is tweeledig:

- de criteria voor kwaliteitscontrole standaardiseren
- en zorgen voor het behoud en het (her)gebruik van gedigitaliseerd cultureel erfgoed op lange termijn.

Deze gids wil dus de praktijken van de Belgische federale wetenschappelijke en culturele instellingen harmoniseren. Een harmonisatie die de basis vormt voor een vlotte uitwisseling van informatie binnen en tussen deze instellingen. Ook gebruikers zullen baat hebben bij deze verbetering doordat hun zoekmogelijkheden vergroot en vergemakkelijkt worden.

Kwaliteitscontrole is echter een essentieel onderdeel in elke fase van een digitaliseringsproject om de integriteit en consistentie van de geproduceerde bestanden en gegevens en hun bewaring op lange termijn te garanderen.

1 Zo werd in 2004 het Google-Print-project aangekondigd, in samenwerking met de universiteitsbibliotheken van Harvard, Stanford of Michigan in de VS en Oxford in Groot-Brittannië en de New York Public Library, met als opzet om 15 miljoen boeken te digitaliseren. Het Europeana-project en een aantal digitaliseringsprojecten van particuliere en openbare organisaties zijn een rechtstreeks resultaat van dit initiatief. In zijn artikel schetst Jean-Michel Salaün de geboorte van dit danteske project en de felle reacties die het destijds opriep.

Jean-Michel Salaün. Bibliothèques numériques et Google-Print. Regards sur l'actualité : mensuel de la vie publique en France, La Documentation Française, 2005.
https://archivesic.ccsd.cnrs.fr/sic_00001576/document

2 Het CegeSoma, het Studie- en Documentatiecentrum Oorlog en Hedendaagse Maatschappij, is sinds 2016 de vierde operationele directie van het Belgisch Rijksarchief.

Methodologie en structuur

De gids is onderverdeeld in zes hoofdstukken volgens de onderstaande structuur:

1 Het eerste hoofdstuk schetst in het kort de achtergrond van de digitalisering en haar uitdagingen. Het belicht de meest gangbare doelstellingen van de digitalisering en de ontwikkeling ervan in het licht van de toenemende informatisering van de moderne samenleving. Deze evolutie is in feite het resultaat van institutionele veranderingen die zelf gebaseerd zijn op de veranderende verwachtingen en behoeften van de gebruikers van archieven en bibliotheken.

2 In het tweede hoofdstuk wordt het begrip kwaliteit gedefinieerd volgens de internationale norm ISO-9001³. Deze algemene definitie wordt vervolgens aangepast aan de specifieke kenmerken van het digitaliseringsproces. De strategische en intellectuele ontwikkeling van elk project wordt hier behandeld aan de hand van drie fundamentele documenten: de strategie voor het gegevensbeheer, het digitaliseringsbeleid en ten slotte het bestek.

3 Hoofdstukken drie en vier richten zich op de kwaliteit van de deliverables, met name beelden en bijbehorende metadata. Deze twee begrippen worden op hun beurt toegelicht en gekenschetst om de criteria vast te stellen op grond waarvan zij als kwalitatief goed kunnen worden gedefinieerd.

5 Het vijfde hoofdstuk gaat over de digitaliseringsomgeving en goed werkplaatsbeheer. De digitalisering van erfgoed heeft als bijzonderheid dat het gaat om kostbare materialen – hedendaags of oud – die speciale maatregelen vereisen om beschadiging tijdens de behandeling te voorkomen.

6 Het zesde en laatste hoofdstuk ten slotte bevat een reeks samenvattende fiches voor het waarborgen van de kwaliteit van een digitaliseringsproject. Deze fiches bevatten de kernbegrippen voor elke fase van de kwaliteitscontrole en volgen de chronologie van de digitaliseringsketen. Elke fiche bevat praktische aanbevelingen, hulpmiddelen en tips, en aanvullende bronnen om de beschreven controlefase uit te voeren.

De inhoud van deze gids is gebaseerd op:

- **Gestandaardiseerde referentiesystemen** rond kwaliteit en digitalisering, en meer bepaald rond de ISO-normen.
- **Internationale methodologische richtsnoeren en gidsen voor goede praktijken voor digitalisering.**
- **Algemene monografieën** over digitalisering, kwaliteit en kwaliteitsbeheer, digitale beelden en metadata.
- **Het onderzoek dat het team tijdens het ADOCHS-project** heeft uitgevoerd en meer bepaald het werk van Tan Lu en Anne Chardonrens om de kwaliteit van afbeeldingen en metadata te verbeteren.
- **De vaardigheden en ervaring van de digitaliseringsteams van het Belgisch Rijksarchief en de Koninklijke Bibliotheek van België**, aan de hand van een reeks interviews met technische operatoren en managers van digitaliseringsprojecten in het voorjaar van 2021.

³ Internationale Organisatie voor Standaardisatie (ISO). Quality management systems – Fundamentals and vocabulary, Genève : ISO, ISO 9001 :2015, 2015.

<https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en>

01.

DIGITALISERING: UITDAGINGEN, VERANDERINGEN EN PERSPECTIEVENS

Digitalisering is het omzetten van een analogo medium in een reeks codes die door een machine kunnen worden geïnterpreteerd, zodat elk soort analogo document in digitale vorm kan worden gereproduceerd¹.

Achter deze a priori eenvoudige definitie gaan in werkelijkheid tal van problemen schuil voor de instellingen die deze projecten opzetten en voor hun publiek. Digitaliseringsprojecten moeten culturele, wetenschappelijke, sociale, economische, politieke en technische eisen met elkaar verzoenen.

Het doel van het digitaliseren van documenten hangt sterk af van het geplande gebruik van deze gedigitaliseerde inhoud, en is over het algemeen gericht op de bewaring, verspreiding en valorisatie van erfgoed, archief en documentaire collecties.

Bewaring	Digitalisering voor bewaringsdoeleinden houdt in dat het originele document zo getrouw mogelijk wordt gereproduceerd, zonder het te willen verbeteren of wijzigen. Dit substitueert de behandeling en dus de aantasting van de originele collecties te beperken.
Verspreiding	Digitalisering met het oog op verspreiding heeft tot doel een gedigitaliseerd document online of in een publicatie beschikbaar te maken.
Valorisatie	Digitalisering met het oog op valorisatie is bijvoorbeeld gericht op het creëren van digitale collecties om onderwijs en onderzoek te vergemakkelijken of in de context van een museumtentoonstelling.

¹ Thierry Claerr, Isabelle Westeel, *Manuel de la numérisation*, Electre, Tours, 2011, blz. 18.

De komst van internet en de ontwikkelingen op dat gebied hebben de toegang tot cultuur en de valorisatie ervan mee bepaald. Web 2.0 combineert complexe sociale systemen zoals blogs, wiki's of sociale netwerken en ziet het web als een grote ruimte voor sociale interacties. Het is de samenwerking tussen individuen en platforms die de continue productie van content verzekert.

In die zin gedragen Web 2.0-sites zich meer als gebruikersgerichte portaalsites die de aandacht van de gebruikers willen trekken, eerder dan traditionele websites. Web 2.0 is gebaseerd op aanbevelingsalgoritmen en werkt volgens het principe van matches tussen profielen en content. Deze matches worden met name bepaald aan de hand van gebruiksgegevens van de gebruikers, waardoor het mogelijk wordt groepen te vormen met betrekkelijk homogene kenmerken waaraan en via dewelke informatie kan worden verspreid.

In deze context spelen metadata – die de identificatie en beschrijving van digitale content structureren – uiteraard een centrale rol om een bepaalde verspreidingskring te bereiken en een publiek te vinden. Een accurate beschrijving van content of websites in het kader van zoekmachineoptimalisatie wordt steeds belangrijker, waarbij het structureren en leesbaar maken voor machines een centraal thema is geworden in de benadering van het internet.

Daar ligt de uitdaging van het Web 3.0-project, beter bekend als het semantisch web, dat niet langer tot doel heeft platforms voor gegevensraadpleging te bouwen, maar referentie-informatie toegankelijk en leesbaar te maken voor zowel machines als mensen.

Deze doelstelling impliceert onder meer dat het metadataformaat voldoende flexibel moet zijn om de beschreven digitale content te kunnen raadplegen en verspreiden en tegelijkertijd crossreferencing en transformatie van alle op het internet aanwezige content te bevorderen.

Dit project is vooral zinvol in de context van de digitale verspreiding van erfgoed-, archief- en documentaire collecties: een dergelijke technologie zou het bijvoorbeeld mogelijk maken bekende informatie te verzamelen over corpora die vaak verspreid zijn over verschillende instellingen, zowel nationaal als internationaal.

In deze context spelen metadata een centrale rol om te bereiken en een publiek te vinden

Deze aanpak vereist echter een grote samenhang en een gemeenschappelijke strategie van de culturele en wetenschappelijke sector. De metadata die aan digitale content zijn gekoppeld, moeten voldoende goed geïnformeerd zijn zodat de content door het potentiële publiek kan worden geïdentificeerd².

Met dit doel heeft de Europese Commissie onlangs het ontwikkelingsplan³ voor haar vlaggenschipproject voor de opwaardering van

2 De vindbaarheid van digitale content verwijst naar het gemak waarmee ze al dan niet wordt gevonden door internetgebruikers, ook door wie er niet specifiek naar op zoek was. Dit potentieel hangt af van een aantal factoren, waaronder de metadata en de trefwoorden die eraan zijn gekoppeld, de hyperlinks ernaar, de reclame die ervoor wordt gemaakt, de algoritmen van de gebruikte zoekmachine, en de surfgewoonten van internetgebruikers.

Voor meer details over dit onderwerp zie: José Plamondon. *Êtes-vous repérables ? Documenter vos contenus pour qu'ils soient compris par des moteurs de recherche*, 2019.
http://ipf.ca/wp-content/uploads/sites/2/2019/05/guide-documentation-webseries_20171128_vf-maj_20191405.pdf.

het erfgoed, Europeana, herzien. **Europeana**⁴ is een digitaal platform dat in november 2008 van start ging en toegang biedt tot de digitale bronnen van culturele instellingen in de lidstaten. Het telt nu al bijna 3500 deelnemende instellingen. Let wel: Europeana archiveert geen werken, maar fungeert alleen als zoekinterface. De deelnemende instellingen hebben zich er dan ook toe verbonden hun content te digitaliseren, online toegankelijk te maken en in digitale vorm te bewaren.

De uitdaging van het door de Commissie voorgestelde strategische document **bestaat erin een gemeenschappelijke leidraad te bieden voor digitalisering, en vooral om de deelnemende instellingen aan te moedigen de nodige aandacht te besteden aan metadata om hun zichtbaarheid en de potentiële verspreiding van hun collecties op het internet te bevorderen.** De verschillende maatregelen zijn dus gericht op een betere digitale transitie door aan de volgende punten te werken:

- **Verbeteren van de kwaliteit van metadata en hun interoperabiliteit**
- **Gebruik van nieuwe technologieën, zoals algoritmen voor machinaal leren, om metadataseten automatisch of half-automatisch te verrijken op een snellere en meer schaalbare manier**
- **Versterken van de nationale infrastructuren en ontwikkelen van sterkere interconnecties tussen instellingen**
- **Verbeteren van de technische vaardigheden van actoren in de sector door het opzetten van een kenniscentrum met richtlijnen voor professionals in de erfgoedsector over**

- *onderwerpen als digitalisering, verrijking van metadata, semantische interoperabiliteit, contentcreatie, licentieverlening, hergebruik, bedrijfsmodellen en innovatie.*

Dit perspectief van evolutie in het kader van de verspreiding van digitale content vergt niettemin aanpassingen voor professionals in de erfgoed-, archief- en bibliotheeksector⁵.

De overgang van een beperkte en individuele toegang naar een massacorporus heeft instellingen gedwongen om hun verspreidingsbeleid en de middelen ter ondersteuning ervan te herzien, en zal ze blijven dwingen dit te doen.

Deze paradigmaverschuiving heeft ook de status en de vereiste vaardigheden van de actoren in de sector veranderd: van curatoren van openbaar erfgoed zijn bibliothecarissen, archivarissen en museologen verstrekkers van culturele content geworden. Deze evolutie brengt de ontwikkeling van nieuwe technische expertise met zich mee en vereist bijzondere aandacht voor de veranderende verwachtingen en behoeften van de gebruikers van deze diensten.

3 Europeana, Strategy 2020-2025 Empowering Digital Change, mai 2020, https://pro.europeana.eu/files/Europeana_Professional/Publications/EU2020StrategyDigital_May2020.pdf

4 Europeana Pro, Présentation, [uitliggen], <https://pro.europeana.eu/about-us/mission>

5 Annabelle Boutet et Karine Roudaut, « Les enjeux de la numérisation et de l'ouverture d'archives : le point de vue des professionnels », Terminal <http://journals.openedition.org/terminal/1220>

Een dergelijke studie werd tussen 2015 en 2017 in België uitgevoerd in het kader van het MADDLAIN⁶-onderzoeksproject. Het doel was niet alleen gegevens te verstrekken over de praktijken en behoeften van gebruikers bij de toegang tot digitale informatie, maar ook instellingen te helpen bij het beheer van de methoden en hulpmiddelen om meer inzicht te krijgen in hun publiek.

Dit perspectief van evolutie vergt aanpassingen voor professionals in de erfgoed-, archief- en bibliotheeksector.

Uit deze studie⁷ komen twee fundamentele begrippen naar voren: de behoefte aan toegankelijkheid en de behoefte aan communicatie. Het eindverslag van het project vermeldt in dit verband als « *belangrijkste punten van ontevredenheid die in de enquête naar voren komen: de moeilijkheden die onervaren gebruikers ondervinden om de gezochte informatie te vinden, de leemten in de gepresenteerde informatie – ontoereikende of onvolledige beschrijving van de documenten – en het gebrek aan communicatie over de documenten die al dan niet in de digitale catalogus zijn opgenomen. Tegelijkertijd begrijpen de meeste gebruikers dat niet alle collecties kunnen worden gedigitaliseerd en online beschikbaar kunnen worden gesteld. Wel vragen ze uitleg over de achterliggende technische procedures, over de gemaakte keuzes op vlak van digitalisering en over de mogelijke restricties* » .

Uit deze bevindingen blijkt hoe belangrijk het is

de digitalisering van erfgoed-, archief- en documentaire collecties te benaderen vanuit het oogpunt van de kwaliteit. Kwaliteit van de te leveren diensten, uiteraard, maar ook van de processen en met name van de intellectuele en strategische ontwikkeling van het project. Een gezamenlijke aanpak van deze twee kwaliteitsaspecten maakt het mogelijk alle digitaliseringsprocessen beter te beheersen en in fine betrouwbare en relevante digitale content aan te bieden, zoals we in de volgende hoofdstukken zullen tonen.

6 Het project MADDLAIN (2015-2017) is ontstaan uit een samenwerking tussen de Koninklijke Bibliotheek van België, het Belgisch Rijksarchief en het Centrum voor Oorlogs- en Samenlevingsstudies (Cegesoma). Het had tot doel het gedrag en de behoeften van de verschillende publieken van de instellingen van de documentatieafdeling van het Federaal Wetenschapsbeleid te analyseren op het vlak van de toegang tot digitale informatie.

7 Florence Gillet, Jill Hungenaert, Melissa Hodza, et al., *Identifying Needs to Modernize Access to Digital Data in Libraries and Archives (MADDLAIN) : final report*, Brussels : Belgian scientific Policy, 2018. http://www.belspo.be/belspo/brain-be/projects/FinalReports/MADDLAIN_%20final%20report.pdf

VERDERE INFORMATIE

FONDS INDÉPENDANT DE PRODUCTION, *Êtes-vous repérables? Guide pratique pour documenter vos contenus*, novembre 2017, [Online],

http://ipf.ca/wp-content/uploads/sites/2/2019/05/guide-documentation-webseries_20171128_vf-maj_20191405.pdf

CHAN, L., MOUNIER, P., *Connecting the Knowledge Commons – From Projects to Sustainable Infrastructure : The 22nd International Conference on Electronic Publishing – Revised Selected Papers*. Marseille : OpenEdition Press, 2019.

CANDELA, G., ESCOBAR, P., CARRASCO, R.-C., MARCO-SUCH, M., « A linked open data framework to enhance the discoverability and impact of culture heritage », *In: Journal of Information Science*, 45(6), 2019, blz. 756-766.

FREIRE, N., VOORBURG, R., CORNELISSEN, R., DE VALK, S., MEIJERS, E., ISAAC, A. « Aggregation of Linked Data in the Cultural Heritage Domain: A Case Study in the Europeana Network », *In: Information*, 10(8), 2019.

JONES, E.; SEIKEL, M. *Linked Data for Cultural Heritage*; Facet Publishing: Cambridge, UK, 2016

ACCART, J.-P., *Le Métier de Documentaliste. Avec la collaboration de Réthy Marie-Pierre*. Éditions du Cercle de la Librairie, 2015.

THIAULT, F., « Mutations des métiers de l'information-documentation : vers l'émergence d'une culture de l'information numérique », *In: Les Cahiers de la SFSIC, Société française des sciences de l'information et de la communication*, 2012, blz. 59-62, [Online],

<https://hal.univ-lille.fr/hal-00955279/document>

GILLET, F., HUNGENAERT, J., HODZA, M., en al. *Identifying Needs to Modernize Access to Digital Data in Libraries and Archives (MADDLAIN) : final report*. Brussels : Belgian scientific Policy, 2018, [Online], http://www.belspo.be/belspo/brain-be/projects/FinalReports/MADDLAIN_%20final%20report.pdf



02.

KWALITEIT: RODE DRAAD OP ORGANISATORISCH VLAK

Kwaliteit is een ruim begrip: in de ISO 9000-norm wordt kwaliteit gedefinieerd als « alle kenmerken en eigenschappen van een product, proces of dienst die van invloed zijn op het vermogen ervan te voldoen aan vastgestelde of impliciete behoeften¹ ». In theorie kunnen alle projecten het kwaliteitsconcept en de internationale normen overnemen, ongeacht de doelstellingen, de gebruikte technologieën, de gebruikte middelen of de aard van de organisatie.

Bibliotheken, archieven en musea vertrouwen al deels op deze aanbevelingen bij hun activiteiten, met name in het kader van digitaliseringsprojecten. Kwaliteitscontrole, een integrerend onderdeel van het proces, bevindt zich meestal aan het eind van de keten en betreft de controle van de deliverables: meer bepaald de evaluatie van de kwaliteit van de beelden, de metadata en de bestandsintegriteit.

Kortom, kwaliteit komt er hier op neer dat men zorgt dat de behaalde resultaten

overeenstemmen met de initiële doelstellingen van het project. Maar het waarborgen van de kwaliteit van een digitaliseringsproject mag niet gereduceerd worden tot het a posteriori controleren van deliverables: bij fouten zijn de gemaakte kosten te hoog en is de speelruimte om processen te corrigeren te klein. Deze laatste verificatiestap is uiteraard essentieel voor het digitaliseringsproces, maar moet gepaard gaan met plannings- en controlewerkzaamheden in de gehele digitaliseringsketen.

In die zin moet het begrip kwaliteit niet worden opgevat als een specifieke fase, maar als een instrument voor projectbeheer in alle fasen, een rode draad die de beheersing van alle processen garandeert.

¹ Internationale Organisatie voor Standaardisatie (ISO). Quality management systems – Fundamentals and vocabulary. Genève : ISO, ISO 9000 :2015, 2015.

<https://www.iso.org/obp/ui/#iso:std:iso:9000:ed-4:v1:en>

Op basis van deze definitie impliceert het begrip kwaliteit dus:

- *Technische vaardigheden en een grondige kennis van beelden en metadata, alsook een goed inzicht in digitaliseringsapparatuur.*
- *Een overzicht van de meest voorkomende kwaliteitsproblemen en de oorzaken daarvan, om vóór het project een coherente methodologie te ontwikkelen en zo goed mogelijk in te spelen op onvoorziene gebeurtenissen tijdens het digitaliseringsproces.*
- *Het beheer van de werkomgeving, waaronder de organisatie van ruimtes, de controle van het licht, en de meting en analyse van de atmosferische omstandigheden, om respectievelijk de werkketen te stroomlijnen, optimale scanomstandigheden te garanderen en mogelijke schade aan de erfgoedcollecties tot een minimum te beperken.*
- *Het definiëren en gebruiken van organisatorische hulpmiddelen om de verschillende posten in de digitaliseringsketen op elkaar af te stemmen en de communicatie tussen operatoren te vergemakkelijken.*
- *Meer in het algemeen moet elk digitaliseringsproject gekaderd worden binnen een algemeen beleid voor de bewaring, digitalisering en verspreiding van gedigitaliseerde content, en binnen een strategie voor gegevensbeheer (Digital Data Strategy²).*

Het begrip kwaliteit als leidraad nemen komt eigenlijk neer op « het duurzaam maken van acties, het optimaliseren van fabricageprocessen, het verbeteren van de efficiëntie van een organisatie, het verbeteren van de kwaliteit van haar producten en diensten, het innoveren en het delen van toegevoegde waarde [en ten slotte] een verantwoorde besteding van de geïnvesteerde middelen ten behoeve van de tevredenheid van de eindgebruikers³ ».

Het begrip kwaliteit niet worden opgevat als een specifieke fase, maar als een instrument voor projectbeheer in alle fasen, een rode draad die de beheersing van alle processen garandeert.

Het gaat dus niet enkel om het a posteriori vaststellen van de kwaliteit van een productie, maar om het bepalen van de kwaliteitscriteria van een digitaal product en van de procedures, sequenties en interacties tussen de verschillende fasen van de digitaliseringsketen, het ontwikkelen van een kwaliteitsdienst, en ten slotte om het anticiperen op de verwachtingen en behoeften van de gebruikers

2 Valentina Bachi, Antonella Fresca, Claudia Pierotti, Claudio Prandoni, The digitization age: mass culture is quality culture. challenges for cultural heritage and society. In: Ioannides, M., Magnenat-Thalmann, N., Fink, E., Žarnić, R., Yen, A.-Y., Quak, E. (eds.) EuroMed 2014 en Mike Fleckenstein, Lorraine Fellows, Data Quality. In: Modern Data Strategy. Springer, Cham, 2018.

3 Thierry Claerr, Isabelle Westeel, Manuel de la numérisation, Electre, Tours, 2011, blz. 269.

om kwalitatieve en relevante inhoud aan te bieden. Als elk digitaliseringsproject goed wordt gedocumenteerd, kan het worden geëvalueerd en kunnen alle protocollen en processen worden verbeterd, waardoor de activiteit van een organisatie duurzaam kan worden uitgebouwd. Met dien verstande dat deze aanpak uiteraard moet worden uitgedacht naargelang de technische, personeels- en financiële middelen van de instelling. **Om goed te functioneren moet het systeem worden aangepast aan de feitelijke middelen van de organisatie die het project opzet.**

2.1 Strategie voor gegevensbeheer en digitaliseringsbeleid: de basis

Operationele en strategische besluitvormingsprocessen zijn van essentieel belang voor het welslagen van een project, ongeacht het aantal te digitaliseren documenten of de aard van het project. Het is van belang onderscheid te maken tussen het operationele niveau – dat alle regels omvat voor de verwerking van een bepaald document en voortvloeit uit het digitaliseringsbeleid van de instelling – en de echt strategische kwesties die deel uitmaken van het beheer- en ontwikkelingsbeleid van de culturele en wetenschappelijke instellingen en de strategieën voor gegevensbeheer en gegevenskwaliteit.

Hoewel deze gids niet bedoeld is om in detail in te gaan op de ontwikkeling van een strategie voor gegevensbeheer (Digital Data Strategy) of een digitaliseringsbeleid, is het belangrijk om hun belang en de verschillen ertussen te benadrukken, omdat er operationele kwesties zijn die niet kunnen worden opgelost zonder strategische beoordeling:

De strategie voor gegevensbeheer bestrijkt alle praktijken voor het beheer van gegevens, d.w.z. de informatiemiddelen als een waardevolle hulpbron binnen een organisatie, een instelling of een bedrijf. **Het doel is dan ook om de strategieën, het beleid, de standaarden, architecturen en procedures inzake databeheer te definiëren, goed te keuren en te communiceren, en de opvolging en naleving ervan te verzekeren.** Hoe hoger de kwaliteit van de door de organisatie gecreëerde, geclassificeerde en verwerkte gegevens, des te leesbaarder, relevanter en betrouwbaarder het informatiesysteem is, waardoor de algehele efficiëntie van de organisatie wordt verbeterd.

Het digitaliseringsbeleid vloeit rechtstreeks voort uit de strategie voor gegevensbeheer, maar is meer toegespitst op de digitaliseringsprojecten van de organisatie. Dit document bundelt alle beginselen voor de digitalisering van erfgoed-, documentaire- en archiefcollecties. **Het maakt het dus mogelijk om de uitdagingen en doelstellingen van de instelling te definiëren in termen van de selectie van te digitaliseren documenten, hun bewaring, hun beheer en hun verspreiding in digitale vorm.** Het omvat ook alle procedures voor het ontwikkelen en uitvoeren van projecten, evenals de rol van de teams tijdens het proces.

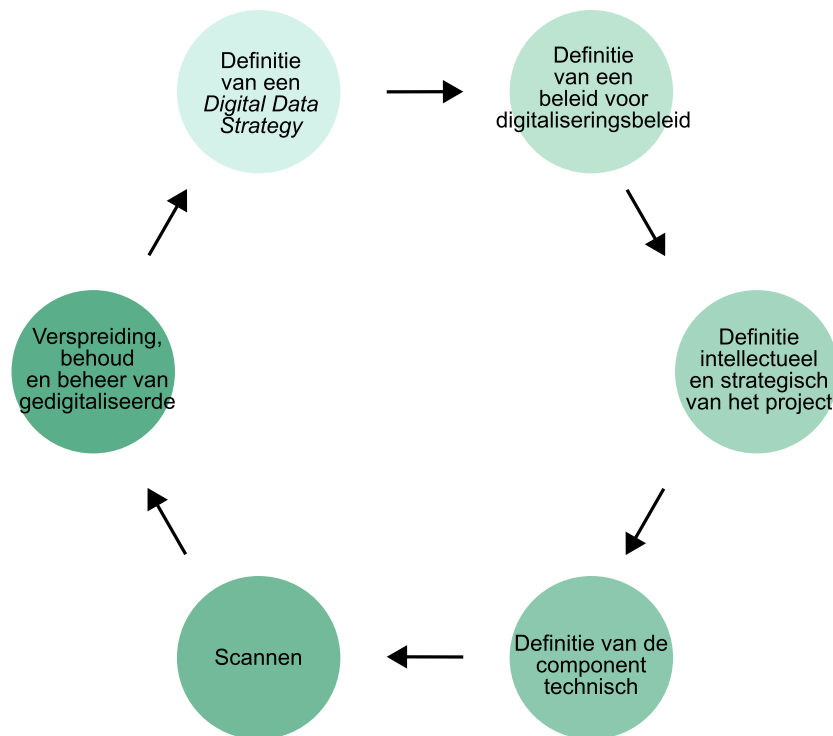


fig. 1 De volledige digitaliseringscyclus

Het institutionele kader van de Digital Data Strategy en van het digitaliseringsbeleid vergemakkelijkt de strategische en intellectuele definitie van digitaliseringsprojecten, met name wat betreft de doelstellingen van deze projecten (bewaring/verspreiding/valorisatie) en de reikwijdte ervan op lange termijn.

Het is van belang dat ervan **bij het begin van het digitaliseringsproject een legitiem en formeel besluitvormingsorgaan wordt aangewezen**, zodat deze kwesties snel kunnen worden aangepakt in geval van een geschil met bijvoorbeeld een externe dienstverlener.

De strategische en intellectuele definitie van het digitaliseringsproject bestaat hoofdzakelijk in het opstellen van een bestek, dat het methodologische kader van het digitaliseringsproject vaststelt en het op zijn beurt mogelijk maakt de technische aspecten

van het project te definiëren (keuze van hardware, software, richtlijnen voor het scannen, transport enzovoort).

Zodra de inhoud gedigitaliseerd is, wordt ze opgeslagen en bewaard in een elektronisch documentbeheersysteem (EDM) of een elektronisch archiveringsstelsel (EAS) met het oog op de verspreiding en/of de bewaring ervan. De implementatie van deze systemen en de doelstellingen en het beheer ervan hangen af van de strategie voor gegevensbeheer. Elke deliverable van een digitaliseringsproject maakt dus deel uit van een veel ruimere cirkel, waarvan alle aspecten in kaart gebracht en geanalyseerd moeten worden.

2.2 Bestek

Het kwaliteitsbeheer van een digitaliseringsproject wordt ontwikkeld zodra het bestek is opgesteld.

Zoals eerder gezegd kadert dit document in de bredere strategie van de organisatie en houdt het rechtstreeks verband met haar doelstellingen op het gebied van de bewaring en verspreiding van haar erfgoed en documentaire collecties. Daarbij moet uiteraard rekening worden gehouden met de juridische haalbaarheid van het project, op basis van de wetgeving inzake intellectuele eigendom. Ten slotte moet het de materiële middelen en organisatorische keuzes weerspiegelen van de organisatie die het project initieert: beschikt de instelling over een digitaliseringsatelier? Wat is zijn rol? Moet het proces geheel of gedeeltelijk worden uitbesteed? Beschikt het team over de nodige vaardigheden en uitrusting voor alle stappen? Wat zijn de doelstellingen van het digitaliseringsproject en waarvoor zal de gedigitaliseerde inhoud worden gebruikt? Deze vragen zijn bepalend voor de te verlenen diensten, de verwachte kwaliteit, de minimumdrempel op kwaliteitsvlak en het beheer van de kwaliteitscontrole.

Om de kwaliteit van een digitaliseringsproject beter te beheersen, moet het volgende in het bestek worden opgenomen:

01 Een nauwkeurige beschrijving van de instelling en haar doelstellingen

Beschrijf de aard en de algemene doelstelling van het project, de soorten documenten, het verwachte resultaat, het beoogde gebruik. Dit is kortom een samenvatting van het digitaliseringsbeleid van de instelling, haar middelen en haar ambities voor het project.

02 Een samenvatting van de te verlenen diensten

Een samenvatting van de fasen van het project. In geval van samenwerking met een derde dient deze clause om het actieterrein van elke partij af te bakenen.

03 Een presentatie van de te digitaliseren documenten

Gedetailleerde beschrijving van de te digitaliseren documenten. In dit luik moet worden vermeld: de aard van de te digitaliseren documenten, hun materiële geschiedenis, staat van bewaring, drager, omvang en hoeveelheid. Ook de specifieke kenmerken van de documenten moeten worden vermeld: heterogene formaten, kwetsbaarheid van de documenten, eventuele bindingen enzovoort. In deze rubriek moet ook worden vermeld of de digitalisering per dubbele pagina of per enkele pagina moet gebeuren, en of alle pagina's van het document moeten worden gedigitaliseerd.

04 De verdeling van de te digitaliseren documenten in batches

Afhankelijk van de heterogeniteit en de hoeveelheid te digitaliseren documenten, is het van essentieel belang de documenten in subsets te groeperen. De omvang en het aantal van deze batches moeten rekening houden met de mogelijkheid om ze binnen een redelijke tijd te verwerken. In deze fase moet de samenstelling van het preproductiemonster worden bepaald: deze batch dient als ijkpunt tijdens de hele productiefase en moet de verscheidenheid van alle gedigitaliseerde documenten weer spiegelen.

05 Workflow en deadlines

Een omschrijving van de workflow van het digitaliseringsproject, het tijdschema voor elke taak en de kwaliteitscontroles voor deze taken. Dit luik moet in detail de preproductiefase omschrijven. Als de preproductiefase niet gevalideerd is, kan de productie zelf niet worden opgestart. Deze tests moeten worden uitgevoerd onder reële productieomstandigheden.

06 Overzicht van de digitaliseringsketen

In dezelfde geest als het vorige punt biedt een grafische voorstelling van de digitaliseringsketen een interessant overzicht om toe te voegen aan het bestek: zo kan eenieder zijn rol binnen het proces en de opeenvolging en interacties tussen de fasen beter begrijpen en de belangrijkste fasen in het kwaliteitsbeheer van het project visualiseren.

07 Algemene uitvoeringsvoorwaarden

Welke basisregels moeten tijdens het hele proces worden gevolgd: digitaliseringsnormen van de instelling, belangrijke beginselen inzake de kwaliteit van de verwachte resultaten (met name wat betreft de leesbaarheid of de esthetiek van de gedigitaliseerde inhoud).

08 Nauwkeurige beschrijving van de voorbereiding van de documenten

Beschrijving van de wijze waarop de documenten moeten worden voorbereid (identificatie en markering van de documenten, indeling naar formaat, verpakking enzovoort).

09 Richtlijnen voor transport en opslag

Het belang van deze clause hangt uiteraard af van de vraag of het digitaliseringsproject geheel of gedeeltelijk wordt uitbesteed. In geval van uitbesteding moeten hier richtlijnen worden beschreven voor het ophalen en vervoeren van de originelen. De eisen inzake bewaring en opslag van documenten moeten duidelijk worden vermeld. Indien de gehele productieketen ter plaatse wordt uitgevoerd, worden in deze sectie de richtlijnen voor tijdelijke opslag vastgesteld, evenals de voorwaarden voor de behandeling van de originele documenten.

10 Kalibratie van de digitaliseringsketen

Beschrijf de procedure die het digitaliseringsatelier moet toepassen om zijn digitaliseringsketen in te stellen

alvorens de productie op te starten – in overleg met de opdrachtgever – en om te controleren of de begininstellingen gedurende de hele dienst gehandhaafd blijven.

De kalibratie van de digitaliseringsketen omvat alle elementen die automatisch kunnen worden aangepast: resolutie, kleur, post-processing, bestandsformaat. De kalibratie wordt onder normale productieomstandigheden uitgevoerd op een reeks documenten die representatief zijn voor de verscheidenheid van de te verwerken documenten, en vervolgens geanalyseerd aan de hand van de indicatoren in het bestek, totdat het kalibratieproces globaal wordt gevalideerd. Naast het kalibreren van de digitaliseringsapparatuur, dient de kalibratie ook als referentie tijdens de operatie om een constante kwaliteit te garanderen.

11 Te volgen procedure in geval van een anomalie

Dit kan een verwijdering van het document uit het productiecircuit inhouden, melding van het geval voor analyse en oplossing, omschrijving van de te ondernemen actie, of opname in het documentatiesysteem voor het systematiseren van deze procedures.

12 Kenmerken van de te verstrekken digitale kopieën

De technische kenmerken (formaat, kleur, bestandsformaat, kadring, vergrotingsfactor, resolutie, colorimetrie, compressietype en -percentage, oriëntatie enzovoort) en de technische metadata van de te leveren beelden.

13 Naamgeving van de bestanden

In dit luik moet de set richtlijnen voor de naamgeving van de bestanden nauwkeurig worden beschreven.

14 Structuur van het referentiebestand van metadata

Oplijsten van alle categorieën die in het metadatabestand moeten worden opgenomen.

15 Uit te voeren indexeringsstakend

Definitie van de uit te voeren indexeringswerkzaamheden en met name van de soorten informatie die moeten worden verzameld. Dit luik specificeert de metadatastandaard die moet worden gevolgd. De concrete vorm van de invoer zal afhangen van de geplande raadplegings- en zoekmethoden: relationele databanken, XML-bestanden enzovoort.

16 Procedures voor de opvolging van het referentiebestand

Bepaal de informatie die de dienstverlener moet verstrekken voor de opvolging van de productie: deadline, vorm waarin het document afgewerkt en overhandigd moet worden enzovoort.

17 Criteria voor de controle van de deliverables

Dit deel beschrijft de laatste fase van de kwaliteitscontrole⁴ aan het eind van de keten, de controle van het geheel van de producten van de dienst. Hier wordt het verwachte kwaliteitsniveau vermeld en de speelruimte van de opdrachtgever om producten te weigeren die niet aan zijn

verwachtingen voldoen. In dit luik wordt ook het aantal controles vermeld die tijdens het digitaliseringsproces zullen worden uitgevoerd, evenals de deadlines voor het controleren van de tussen- en eindresultaten.

18 Levering van digitale beelden

In geval van uitbesteding bepaalt deze clause de wijze van levering van de digitale kopieën, teruggave van de originelen en controle van de verwijderbare dragers waarop de digitale kopieën zijn opgeslagen (USB-stick, externe harde schijf enzovoort).

⁴ In deze gids worden twee soorten kwaliteitscontrole onderscheiden: kwaliteitscontrole als een alomvattende aanpak om het algemene kwaliteitsniveau van de digitaliseringsketen te verbeteren en te bewaken, en kwaliteitscontrole als een op zichzelf staande aanpak.

VERDERE INFORMATIE

Voor een exhaustief overzicht van de wetten, koninklijke besluiten en omzendbrieven inzake het auteursrecht in België, zie bijvoorbeeld [IPNews.be](https://www.ipnews.be), een juridische nieuwssite gewijd aan de nieuwe technologieën.

<https://www.ipnews.be/legislation/legislation-belge/>

Het **Bureau de Coopération Interuniversitaire (BCI) van Quebec** bespreekt in zijn Guide de gestion d'un projet numérique (**Gids voor het beheer van digitale projecten**) een aantal handige evaluatie-, plannings- en controletools voor het goede verloop en het welslagen van een digitaliseringsproject. Hoewel oorspronkelijk bedoeld voor digitaliseringsprojecten van administratieve documenten, leent de digitaliseringsaanpak zich evengoed voor historische of erfgoedarchieven.

<https://www.enssib.fr/bibliotheque-numerique/documents/64628-guide-de-gestion-d-un-projet-de-numerisation.pdf>

Voor concrete voorbeelden van bestekken, zie de **Guide technique de la Direction des Archives de France - Écrire un cahier des charges de numérisation du Patrimoine**.

https://francearchives.fr/file/bf50d8fa5f554586dbf18fdc862d25970a1da0a7/static_4132.pdf

03.

KWALITEIT VAN DE DELIVERABLES: BEELDBESTANDEN

3.1 Beeld: definitie en kenmerken

Een digitaal beeld is elk beeld dat in binaire vorm wordt verkregen, verwerkt en opgeslagen en dat wordt gekenmerkt door de volgende elementen: resolutie, coderingsdiepte, modus, kleuruimte en formaat.

RESOLUTIE EN DEFINITIE

Resolutie wordt vaak geassocieerd met beeldkwaliteit omdat het de fijnheid van de digitalisering uitdrukt en mag niet worden verward met beelddefinitie. De beelddefinitie is het aantal pixels waaruit het beeld bestaat en wordt berekend aan de hand van de resolutie en de grootte van het document. Dit is dus een absolute waarde die het mogelijk maakt een beeld en de gebruikte sensor voor de productie ervan te kwantificeren. Let wel: deze waarde geeft op zich niet de afmetingen van het beeld weer; een pixel heeft immers geen « fysieke » grootte. Zodra het beeld echter wordt weergegeven of afgedrukt, moet men rekening

houden met het weergaveoppervlak, en daarmee de resolutie, de meeteenheid die de fijnheid van digitalisering uitdrukt. Dit wordt uitgedrukt in ppi (*pixels per inch*) of dpi (*dots per inch*).

Algemeen wordt aangenomen dat een minimumresolutie van 300 ppi op 100% van het origineel voldoende is voor de meeste gescande documenten, uitgezonderd voor zeer grote of zeer kleine formaten. De norm lijkt te evolueren naar een resolutie van 400 dpi voor meer leescomfort op het scherm.

CODERINGSDIEPTE

De coderingsdiepte is het aantal bits dat een pixel definieert. De bits bepalen het aantal verschillende waarden die een pixel kan weergeven.

Met 8 bits worden 256 verschillende waarden bereikt. Elk extra bit verdubbelt het aantal mogelijke waarden. Voor een afbeelding in grijstinten is een coderingsdiepte van 8 bits voldoende.

MODUS

Voor een kleurenbeeld wordt een coderingsdiepte van 24 bits aanbevolen. Voor een kleurenbeeld scannen sommige apparaten met een hogere coderingsdiepte, 36 bits of 48 bits (12 of 16 bits per laag).

Een modus is een wiskundig model bedoeld om de onderlinge relaties van kleuren te definiëren aan de hand van het aantal aanwezige lagen in deze modi. Deze lagen bepalen de mate van waarneembare kleurdetails en de bestandsgrootte. De tabel hieronder¹ toont de drie hoofdmodi die worden gebruikt afhankelijk van het type document dat moet worden gescand:

NAAM MODUS	AANTAL LAGEN	DETAILS VAN DE LAGEN	TOEPASSING
CMGZ / CMYK	4	<i>Cyaan - Magenta - Geel - Zwart</i> Wijst aan elke pixel een percentage toe voor elke laag. Hoewel het CMYK-model een standaardkleurenmodel is, kan het exacte kleurengamma variëren afhankelijk van de pers en de drukomstandigheden.	Voorkeursmodus voor offsetdruk. (Het gebruik ervan wordt niet aanbevolen voor scannen.)
RVB / RGB	3	<i>Rood - Groen - Blauw</i> De waarden van elke pixel liggen tussen 0 en 255 en kunnen meer dan 16,7 miljoen verschillende kleuren weergeven.	Voorkeurmodus voor scannen in kleur.
Grijs- tinten / Gray- scale	1	De waarde van elke pixel ligt tussen 0 en 255, zodat 256 helderheidswaarden kunnen worden weergegeven, gaande van zuiver wit (255) tot zuiver zwart (0).	Voorkeurmodus voor scannen in zwart-wit.
Sommige modi, zoals geïndexeerde kleur, bitmap of bitonaal, komen niet in aanmerking voor digitaliseringsprojecten.			

tab. 1 Modi die geschikt zijn voor erfgoeddigitalisering

¹ Informatie uit de digitaliseringsgids van BanQ en de website van Adobe : Marie-chantal Anctil, Michel Legendre, Tristan Müller, Kathleen Brosseau, Louise Renaud, Recueil de règles de numérisation, Bibliothèque et Archives nationales du Québec, Bibliothèque nationale de France et Musée canadien de l'histoire, 2014.
<http://collections.banq.qc.ca/bitstream/52327/2426216/1/4671601.pdf>
<https://helpx.adobe.com/fr/photoshop/using/color-modes.html>

KLEURRUIMTE

Een kleurruimte is een driedimensionaal wiskundig model dat de reeks kleuren weergeeft die binnen een modus waargenomen, gebruikt of weergegeven kunnen worden. Elke kleur wordt dus geassocieerd met coördinaten die een precies punt bepalen en overeenkomen met waarden zoals helderheid, verzadiging en tint.

Er zijn twee soorten kleurruimten:

- **Apparaatafhankelijke kleurruimten**, ook ICC-profielen genoemd
- **Onafhankelijke kleurruimten** die een reeks zichtbare kleuren beschrijven zonder verwijzing naar een apparaat, ook LCC-profielen genoemd

Wat hun type ook is, colorimetrische ruimten worden gedifferentieerd door hun gamma, dat wil zeggen door het aantal toegankelijke en reproduceerbare kleuren binnen deze kleurruimte. Tot de meest voorkomende RGB-kleurruimten behoren, van het kleinste tot het grootste gamma: sRGB, Adobe RGB (1998), of het zichtbare spectrum, zoals weergegeven in de onderstaande figuur.

Let wel: ruimtelijke coördinaten verschillen van ruimte tot ruimte, zodat eenzelfde stel coördinaten met verschillende kleuren overeenkomt, afhankelijk van de kleurruimte. In de bovenstaande illustratie, bijvoorbeeld, komt het groen dat overeenkomt met de RGB-coördinaten [0; 255; 0] overeen met drie verschillende groenen, afhankelijk van het feit of deze coördinaten van toepassing zijn op het zichtbare spectrum, Adobe RGB of sRGB. Voor digitaliseringsdoeleinden wordt in het algemeen Adobe RGB 1998 gebruikt voor een afbeelding in kleur en Gray Gamma 2 voor een afbeelding in grijs tinten².

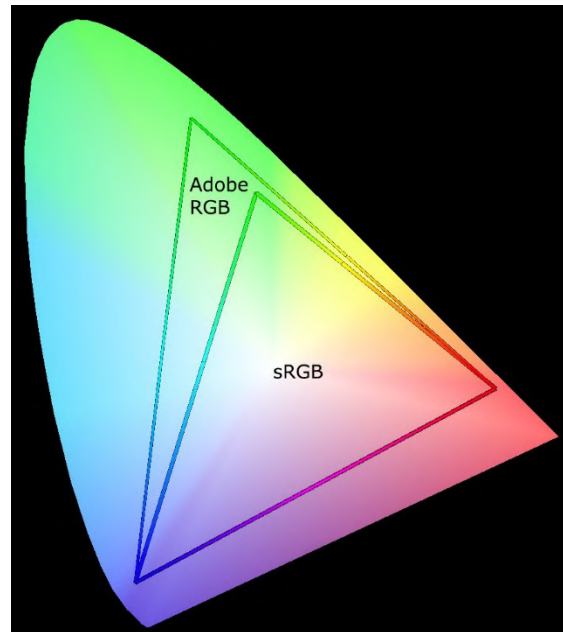


fig 2. Het zichtbare spectrum en de kleurruimtegammata Adobe RGB 1998 en sRGB

FORMAAT

Het eerste beeldbestand in de digitaliseringsfase wordt « RAW » genoemd. Deze algemene benaming betekent dat het verkregen beeldbestand alle ruwe gegevens van de sensor bevat, plus de parameters die nodig zijn om ze om te zetten in een beeldbestand dat op het scherm kan worden bekeken. **Om bruikbaar te zijn, is het nodig het bestand te demasteren, dat wil zeggen naar een ander formaat te converteren.**

Het formaat waarin het gedigitaliseerde documentwordtopgeslagen,verschilt naargelang het doel van het digitaliseringsproject (verspreiding of bewaring).

Zie de onderstaande tabel voor een overzicht van de meest relevante formaten, afhankelijk van het beoogde type operatie:

2 Marie-chantal Anctil, Michel Legendre, Tristan Müller, Kathleen Brosseau, Louise Renaud, Recueil de règles de numérisation, Bibliothèque et Archives nationales du Québec, Bibliothèque nationale de France et Musée canadien de l'histoire, 2014.
<http://collections.banq.qc.ca/bitstream/52327/2426216/1/4671601.pdf>

SOORT DOCUMENT	VERSPREIDING	BEWARING
Tekst	.docx, .odt, .pdf (PDF) .epub .tei	.pdf (PDF/A)
Beelden	.jpg (JPEG), .jpg (JPEG2000). png	.tiff
Kaart/ plattegrond	.dwg, .svg	.dxf .svg

tab 2. Formaten naargelang de toepassing

Bron : Sébastien Soyez, *La numérisation en marche : les étapes de la dématérialisation des processus de travail*, Bruxelles : Archives générales du Royaume et Archives de l'État dans les Provinces, ongedateerd.

http://www.arch.be/docs/brochures/la_numerisation_en_marche.pdf

Voor beeldbestanden wordt vaak de voorkeur gegeven aan het **TIFF-formaat** vanwege de uitstekende kwaliteit. Dit formaat is echter zeer zwaar en niet geschikt om veel te worden bekeken.

Het beste is daarom te beschikken over een eerste versie in TIFF-formaat, die vervolgens op verzoek van de gebruiker of voor intern gebruik kan worden omgezet in een ander formaat, meestal JPEG, JPEG 2000 of PNG, waarbij het verschil ligt in de compressie en dus de bestandsgrootte. JPEG comprimeert het beeld met een zeker kwaliteitsverlies, zodat het gemakkelijk kan worden verspreid, terwijl PNG het beeld comprimeert zonder wijziging en zonder kwaliteitsverlies.

Voor tekstbestanden is het **PDF-formaat de meest aanvaarde norm**, aangezien het erkend is door de ISO (Internationale Organisatie voor Normalisatie). Dit formaat wordt zowel voor

bewarings- als voor verspreidingsdoeleinden gebruikt, omdat het naast een hoge interoperabiliteit ook de mogelijkheid biedt tekst binnen een document te zoeken en/of te selecteren.

OCR (Optical Character Recognition) is het proces waarbij een afbeelding in een tekstdocument wordt omgezet. Het document wordt dan gestructureerd volgens een bepaalde standaard, bijvoorbeeld ePub of TEI (zie kader hieronder).

Merk op dat **geen enkel formaat ontworpen is om permanent te zijn**. Om de duurzaamheid van de gebruikte formaten te garanderen, heeft de Harvard University in samenwerking met Jstor de opensourcesoftware JHOVE³ ontwikkeld, waarmee digitale objecten kunnen worden geïdentificeerd, geanalyseerd en gekarakteriseerd.

De software is dus in staat de eigenschappen van het bestand vast te stellen en na te gaan of het beantwoordt aan bepaalde door de gebruiker gekozen normen en standaarden. Van daaruit is het mogelijk om de relevantie van het digitale document te analyseren, eventuele anomalieën te corrigeren of eventueel naar een ander formaat te converteren in geval van veroudering.

3.2 Criteria voor een kwaliteitsbeeld

Het verkrijgen van beelden die getrouw zijn aan de originele documenten kan een complex proces zijn omdat bij de vastlegging tal van variabelen een rol spelen, zoals de deskundigheid en het beoordelingsvermogen van de digitaliseringsteams. Om instellingen te

3 Open Preservation Foundation, *Getting started with JHOVE*, 2015. <https://jhove.openpreservation.org/getting-started/>

helpen met de technische aspecten van het proces en om een methodologisch kader te bieden dat zowel rigoureuus als realistisch is, bestaat er een aantal normen en documenten inzake goede praktijken die de complexiteit van de beeldkwaliteit aanpakken.

Zo vormen de FADGI⁴ en de Metamorfoze-richtlijnen⁵ die we hieronder weergeven de basis voor de karakterisering van de criteria voor een beeld van goede kwaliteit, waaronder volledigheid, belichting, contrast, kleur, scherpstelling, resolutie, parallelisme, flare, vignettering, ruis en artefacten.

VOLLEDIGHEID

Het eerste criterium voor de kwaliteit van een set beelden is de **volledigheid** ervan. Vooreerst de volledigheid van het geheel: zijn alle pagina's gescand? Maar ook de volledigheid van de elementen van elke afbeelding: is de pagina in zijn geheel gescand? In het laatste geval moet de QA-verantwoordelijke rekening houden met de eisen rond het bijsnijden. De enige manier om de lezer te laten zien dat het hele voorwerp gedigitaliseerd is, is een zwarte rand rond de pagina te laten staan.

Het archiefbestand moet daarom voldoende ruimte rond de pagina hebben om te kunnen bijsnijden, vooral wanneer pagina's waarschijnlijk moeten worden rechtgezet – tools voor het rechtzetten van afbeeldingen snijden min of meer 3% van het aantal pixels bij per inclinatiegraad⁶. Wanneer een bladzijde ontbreekt, bestaat de enige mogelijke correctie erin die bladzijde door middel van een nieuwe scan in de bestaande set in te voegen.

De QA-verantwoordelijke moet ook letten op **de sequentie van de beelden**, d.w.z. de volgorde waarin zij op het scherm verschijnen. De sequentie moet weergeven wat de lezer te zien

krijgt wanneer hij het voorwerp in de leeszaal raadpleegt. Eventuele fouten in de sequentie worden gecorrigeerd door de paginanummering aan te passen. Soms wordt een voorwerp vergezeld van losse blaadjes, bladwijzers of aantekeningen.

De QA-verantwoordelijke moet dan samen met de conservator nagaan of deze voorwerpen in de sequentie moeten worden opgenomen en waar ze moeten worden geplaatst. Gaat het om een noot van de auteur, een bladzijde die per ongeluk uit de band is losgeraakt, of een bibliografische noot? Naar gelang van het geval moeten de beelden worden in- of bijgevoegd en als zodanig worden beschreven.

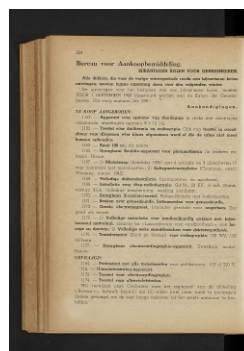


fig.3 Voorbeeld van bijsnijden om de volledigheid van het gedigitaliseerde document te waarborgen.

CegeSoma Collectie

- 4 Thomas Rieger, *Federal Agencies Digital Guidelines Initiative. Technical Guidelines for Digitizing Cultural Heritage Material. Creation of Raster Image File*, 2016, 99 blz.
http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf
- 5 Hans van Dormolen, *Metamorfoze Preservation Imaging Guidelines. Image Quality*, 2012, versie 1.0, 44 blz.
http://www.imagingetc.com/Images/Resources_Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf
- 6 Deze verhouding, die als voorbeeld wordt gegeven, is exponentieel tot 45° en neemt dan omgekeerd af tot 0° bij 90°. Volgens de Metamorfoze-richtlijnen mag er bij ongebonden materialen een inclinatie van maximaal 2° zijn onderaan het beeld. Bij gebonden materialen is het soms onmogelijk om een dergelijke rechteheid in het beeld te garanderen. Wij bevelen daarom aan de ruimte rond het object verder te verbreden.

Bron : Hans van Dormolen, *Metamorfoze Preservation Imaging Guidelines. Image Quality*, 2012, versie 1.0, 44 blz.
http://www.imagingetc.com/Images/Resources_Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

OCR

OPTICAL CHARACTER RECOGNITION

OCR (Optical Character Recognition) is een beeldverwerkingstechniek waarmee tekst wordt geëxtraheerd om ze leesbaar en doorzoekbaar te maken.

Het principe van OCR omvat verschillende stappen:

- Het rechtzetten van de pagina,
- Het binariseren van het document, d.w.z. het omzetten in zwart-wit,
- Het segmenteren van het beeld in tekstgebieden en vervolgens in lijnen, woorden en tekens,
- En ten slotte het classificeren van deze gebieden door het herkennen van tekens en vervolgens woorden.

De meest gebruikte OCR-software is gebaseerd op het principe van deep learning, d.w.z. het verbeteren van de prestaties van de software door te leren van eerder aangetroffen tekens en woorden. Indien een letter of een woord onbekend is bij de software, geeft de persoon die de OCR uitvoert de betekenis ervan aan, die vervolgens wordt onthouden en automatisch wordt opgenomen in het woordenboek van de tool, dat daardoor autonomer en efficiënter wordt.

De moeilijkheden bij OCR zijn van verschillende aard. In de eerste plaats de kwaliteit van het gescande beeld, dat voldoende rechtgezet en gecontrasteerd moet zijn. Ook de staat van het originele document speelt een belangrijke rol: drukfouten, inktdegradatie en eventuele lacunes bemoeilijken de tekenherkenning voor zowel de software als de operator en belemmeren het proces.

Ook de structuur kan problemen geven zijn als ze te complex is – bijvoorbeeld in meerdere kolommen – of als de bladzijde illustraties, zeer kleine lettertypes of handgeschreven toevoegingen bevat. Ten slotte is de software hoofdzakelijk gericht op herkenning van het Latijnse alfabet, waardoor het moeilijk is om bijvoorbeeld het Arabische alfabet te transcriberen, waar de scheiding tussen letters en woorden niet zo duidelijk is.

Op basis van het voortdurende onderzoek omtrent OCR kan men evenwel de volgende perspectieven overwegen:

- Een betere automatisering van de software en een beter beheer van complexe documenten – door de ontwikkeling van methoden, hulpmiddelen en mathematische modellen – dragen bij tot een preciezere beeldsegmentatie in tekstgebieden en tot een beter beheer van documenten die zowel tekst als beelden bevatten⁷.
- Een betere handschriftherkenning. De onderzoeksprojecten HIMANIS⁸ en CLAMM⁹ bijvoorbeeld – onder leiding van het Institut de Recherche et d'Histoire des Textes (IRHT-CNRS) – zijn gewijd aan beeldanalyse en handschriftherkenning voor mediëvisten. Dankzij de ontwikkeling van nieuwe software voor herkenning, classificatie en datering van handschriften konden al corpora van enkele duizenden bladzijden gelezen en geïndexeerd worden.

7 Tan Lu, Homogeneity models for image processing in the cultural heritage sector, [Ongepubliceerde doctoraalscriptie], Vrije Universiteit Brussel, 2020, 245 blz.

8 IRHT - CNRS, Historical MANuscript Indexing for user-controlled (HIMANIS) <https://www.irht.cnrs.fr/fr/ressources/carnet-de-recherche/himanis>

9 IRHT - CNRS, Classification of Latin Medieval Manuscripts (CLAMM) <https://www.irht.cnrs.fr/fr/ressources/sites-web-outils-corpus/clamm>

BELICHTING

Belichting verwijst naar **de inwerking van lichtstralen op de sensor**. In de praktijk is een overbelichte foto een digitale reproductie die te veel licht heeft gekregen en daarom te helder zal lijken, terwijl een onderbelichte foto te weinig licht heeft opgevangen en te donker zal lijken.

Het komt zeer zelden voor dat een belichtingsfout resulteert in **onderbelichting**, hetgeen zich uit in een slechte leesbaarheid van het beeld. **Overbelichting** daarentegen komt vaker voor en is ernstiger, omdat het kan leiden tot gegevensverlies: de RGB-waarden van een digitaal bestand hebben een plafond van 255; als de waarde voor een of meer van de kanalen dit plafond bereikt, wordt het onmogelijk om het beeld te corrigeren, zelfs achteraf¹⁰.

De belichting van een beeld kan worden aangepast door vier elementen in te stellen: de gevoeligheid van de sensor, het diafragma, de sluitertijd of de lichtsterkte. Merk op dat een te grote gevoeligheid de kans op digitale ruis verhoogt.

Met wat ervaring zou de QA-verantwoordelijke in staat moeten zijn om overbelichting op het zicht te detecteren. Gelukkig bestaan er ook tools om hierbij te helpen. Ten eerste **de interne overbelichtingswaarschuwing van de scanner**, die een kleurenmasker toepast op de betrokken gebieden. Deze waarschuwing wordt standaard weergegeven zodra een van de drie RGB-waarden 250 bereikt. Indien geen van deze waarden 255 bereikt, is het nog steeds mogelijk het beeld te recupereren, op voorwaarde dat het opnieuw wordt

gedematriceerd – er wordt dus een nieuw TIFF-bestand op basis van het RAW-bestand gemaakt – door de belichting in het opnamesysteem te corrigeren. Ten tweede kan een tool zoals de « Pipet » in de Photoshop-software bijvoorbeeld helpen om de omvang van het probleem te bepalen en de analyse te verfijnen. Indien blijkt dat een of meer beelden inderdaad overbelicht zijn en er dus informatie verloren is gegaan, moet de QA-verantwoordelijke de beelden verwerpen en het team laten weten dat een nieuwe scan nodig is.

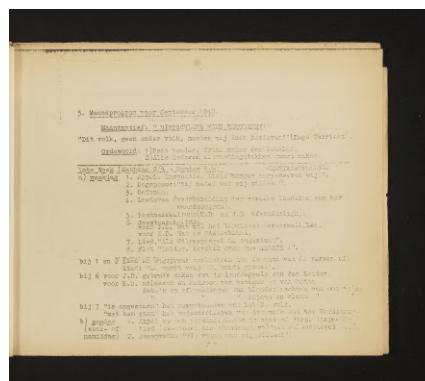


fig. 4 Voorbeeld van een overbelicht beeld

CegeSoma Collection

CONTRAST

Het contrast van een beeld is **het verschil in dichtheid tussen de lichtste en donkerste tinten**. Dit begrip stemt niet overeen met een onderscheid « licht of donker », maar veeleer met de verhouding tussen deze twee soorten tinten.

Het contrast speelt een belangrijke rol bij de leesbaarheid van het beeld en moet daarom zorgvuldig worden ingesteld. **Gain modulation** is in dit opzicht een interessant hulpmiddel, en bestaat erin om – met behulp van specifieke patches die aan de kleurenkaart zijn gekoppeld – het contrastniveau te controleren aan de hand van de grijswaarden en aldus eventuele kleurverzadiging te voorkomen.

¹⁰ Hieruit blijkt hoe belangrijk het is RAW-bestanden te bewaren, ten minste totdat het bewaarbestand wordt gearchiveerd. Er zijn ook scanapparaten die geen RAW genereren. In dat geval moet er bij een fout opnieuw worden gescand. Het is belangrijk hiermee rekening te houden bij de aankoop van apparatuur.

Deze verificatiemethode volgt de **ISO 19264-norm**¹¹, waarin de werking van de methode en de patches en tolerantieniveaus die moeten worden gerespecteerd, in detail zijn beschreven .

De meting van de gain modulation in de hooglichten maakt het dus mogelijk het juiste contrastniveau in het digitale beeld vast te stellen. Er wordt gestreefd naar leesbaarheid zonder verlies van informatie, zachte tinten en vergeelde papierkleuren, en inkten die dicht bij de toon van het papier liggen.

Verkeerd ingestelde contrastniveaus kunnen de inkt onleesbaar maken, vooral wanneer hij in de loop van de tijd bruin is geworden en op de kleur van vergeeld papier gaat lijken. In geval van twijfel moet de QA-verantwoordelijke uiteraard nagaan of de fysieke documenten wel degelijk te ontcijferen zijn. Als blijkt dat het beeld niet overeenkomt met het fysieke voorwerp, moet worden gecontroleerd of de gain modulation juist is.

Om deze waarden te kunnen nagaan, is het van belang **dat elke gemaakte afbeelding een gestandaardiseerde grijschaal van wit naar zwart bevat – type Munsell of Kodak Gray Scale.** Er kan dan snel worden bijgestuurd. Als de gemeten waarden niet goed zijn, moet de set beelden afgekeurd worden.

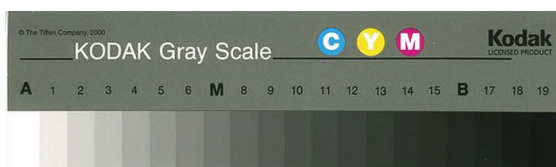


fig. 5 KODAK Gray Scale chart



fig. 6 Munssel Linear Gray Scale chart

De **gain modulation** wordt bijna altijd weergegeven in zijn grafische vorm en wordt soms ook – ten onrechte – gamma genoemd.

Gewoonlijk spreekt men van een **gradatiecurve**. Op digitale spiegelreflexcamera's wordt standaard een **sigmoïd** getekend¹², met name om de donkerste delen te verlichten en de hoge lichten donkerder te maken, zodat de details van de hele opname behouden blijven. In de context van een reproductie, die tot doel heeft een digitale versie aan te bieden die trouw is aan het origineel, moet de gradatiecurve lineair zijn. In tegenstelling tot wat men zou denken, zorgt deze instelling voor een betere leesbaarheid in zachte contrasten, waarbij de sigmoïd zeer weinig groei vertoont in wit en diep zwart.

Het is mogelijk om deze meettechniek in de context van zwart-wit scannen aan te passen aan het hele bereik van neutralen tot de diepste zwarten, met behulp van de Munsell Linear Gray Scale chart – dit wordt opto-elektronische conversiefunctieanalyse of OECF genoemd. Door patches te vergelijken kan de analyse ook een mogelijk flare-effect aan het licht brengen : door bewust voorrang te geven aan de contrastkwaliteit in de hooglichten zullen de beperkingen van het apparaat in de schaduwen naar voren komen.

Als de analyse van de Munsell Linear Grayscale chart ondanks strikte toepassing van de bovenstaande aanbevelingen een niet-lineaire gradatiecurve in de schaduwintinten aan het licht brengt, kan dit alleen liggen aan de kwaliteit van de sensor¹³.

11 Internationale Organisatie voor Standaardisatie (ISO), ISO 19264-1:2021 - Photography – Archiving systems – Imaging systems quality analysis – Part 1: Reflective originals, juni 2021, <https://www.iso.org/obp/ui/#iso:std:iso:19264:-1:ed-1:v1:en>

12 De sigmoïde gradatiecurve is een erfenis van de filmfotografie en, bij televisie, van de kenmerken van de kathodestraalbuis. In de fotografie zijn het de eigenschappen van de film – zilverzouten, filters enzovoort – die de curve bepalen. Deze curve biedt interessante informatie voor de professional die een bepaald beeldresultaat wil bereiken.

13 Om de kwaliteit van de sensor te controleren, moet deze worden blootgesteld aan een lichtbron die – met een IRC van 95% – een reeks flitsen kan produceren waarvan het vermogen een identieke kleurtemperatuur rond 5000 Kelvin moet behouden. De test moet negen keer worden herhaald om de geldigheid van de resultaten te garanderen, onder vooraf bepaalde temperatuur- en vochtigheidsomstandigheden. De analyse van de resultaten, die volgens een strikt schema worden gedocumenteerd, toont het bereik van de door de sensor reproduceerbare dichtheden en de lineariteit van het belichtingstoestel – zie ISO 7589, ISO 14524 en ISO 21550.

KLEUR

De controle van de juiste kleurweergave is waarschijnlijk de meest delicate fase van de kwaliteitscontrole, aangezien de perceptie ervan van individu tot individu kan verschillen¹⁴. Bovendien is de opsporing van een colorimetrische fout veel moeilijker wanneer de set beelden niet klopt of wanneer er van een document slechts één beeld is gemaakt. De fout is gemakkelijker op te sporen als er een plotselinge verandering in kleur is bij het bekijken van een reeks pagina's.

Goed kleurbeheer hangt af van de keuze van de lichtbronnen, de keuze van de scanner, de keuze van de kleurenkaart, de kalibratie van de scanapparatuur en ten slotte de conversie van het RAW-bestand in een ander formaat. In dat geval is het mogelijk foutieve bestanden te corrigeren door het RAW-bestand – mits de oorspronkelijke kwaliteit van de kalibrering voldoende is – opnieuw te dematriceren en te controleren of de ICC- en LCC-profielen en de witbalans van de beelden identiek zijn aan de parameters van het referentiebeeld.



fig. 7 Vergelijking van slechte en goede kleurweergave.

CegeSoma Collectie

14 Er bestaat een aantal kleurwaarnemingsstoornissen, waarvan de ernstigste, achromatopsie, wordt gekenmerkt door een totaal gebrek aan kleurwaarneming. Tussen normaal zien en achromatopsie onderscheiden oogartsen een waaier van aandoeningen die soms zeer moeilijk te diagnosticeren zijn. Daarom kan het interessant zijn het zicht van de scanmedewerkers en QA-verantwoordelijken te testen. De X-Rite Color Challenge bijvoorbeeld is betrekkelijk eenvoudig en snel.

SCHERPSTELLING

Een onscherp beeld is altijd een verloren beeld. Veel beeldbewerkingsprogramma's bieden weliswaar de mogelijkheid om de scherpte te corrigeren, maar de gebruikte algoritmen zijn niet gestandaardiseerd, zodat het niet gegarandeerd is dat het bestand weer perfect leesbaar is. Sommigen vinden dat een leesbaar beeld, zelfs als het enigszins onscherp is, nog steeds een leesbaar beeld is. Maar we moeten ook rekening houden met het feit dat de OCR-software moeite zal hebben met de elektronische interpretatie door het verlies van informatie.

De scherpte van een beeld moet worden gevalideerd door een elektronisch proces. Hiervoor zou elk vastgelegd document moeten worden omgeven door vier QA-62 charts. In de praktijk is dit onhaalbaar, gezien de grootte van sommige documenten in verhouding tot die van de sensoren en de configuratie van sommige boekenwiegens – V-vormig, zuigend enzovoort. De persoon die de controle uitvoert zal het dus op het zicht beoordelen. In geval van twijfel is het belangrijk om het beeld te vergroten tot de maximale weergavegrootte, d.w.z. 100% of schaal 1:1, wat betekent dat elke pixel in het beeld wordt weergegeven door een pixel op het scherm. Vaak vindt de QA-verantwoordelijke een ijkpunt door een als scherp beschouwd beeld met een twijfelachtig beeld te vergelijken.

Wanneer een opnamesysteem herhaaldelijk onregelmatigheden in de beeldscherpte vertoont, moet dit nader onderzocht worden. Dit probleem kan te wijten zijn aan twee factoren: **beweging of onjuiste afstelling**. Bewegingsonscherpte ontstaat bij een te lange sluitertijd. Dit betekent dat zelfs de geringste trilling – bijvoorbeeld alleen al het feit dat de machine zich op een plaats bevindt waar veel mensen passeren – de kwaliteit van het beeld

kan beïnvloeden. Ook de spiegelbeweging van een reflexcamera van slechte kwaliteit kan problemen geven. **De oplossing is dan de sluitertijd te verkorten en de lichtsterkte te verhogen, waarbij het gebruik van een elektronische flitser de beste optie is.**

De tweede factor die terugkerende onscherp-teproblemen kan veroorzaken, is een **verkeerde afstelling van het systeem**, d.w.z. dat er iets gebeurt tussen het instellen en het maken van de opname. Zo hebben sommige vermoeide lenzen moeite met de verticale stand: het mechanisme is niet meer stabiel, de scherpstelling verandert na verloop van tijd en de beelden moeten worden overgedaan. **In dat geval moet degene die de controle uitvoert de hardwareverantwoordelijke verwittigen, zodat de nodige wijzigingen kunnen worden aangebracht.**

RESOLUTIE

Problemen met de resolutie zijn meestal van tweeërlei aard: ofwel is de resolutie ontregeld tijdens de opname, ofwel zijn de instellingen al vóór de eerste opname verkeerd.

In het eerste geval zal de QA-verantwoordelijke een verandering in grootte waarnemen tussen object en beeld, die gepaard gaat met een mindere scherpte. De enige oplossing is de onjuist gekalibreerde beelden opnieuw te scannen, maar dan moet men ook nog bepalen om welke beelden het gaat. Daartoe kan de QA-verantwoordelijke tijdens de kalibratiefase gebruik maken van de kleurenkaart die bij de scanapparatuur hoort. De Munsell-kaart bevat bijvoorbeeld een geïllustreerde lijnaal waarbij een centimeter overeenkomt met ongeveer 118 pixels voor 300 ppi¹⁵.

¹⁵ Deze meting kan niet worden verricht met alle kaarten op de markt: de Kodak Q13/Q14 charts bijvoorbeeld bevatten wel een lijnaal, maar de metingen zijn onnauwkeurig. Dit is ook het geval voor de Color Checker-kaarten. In dergelijke gevallen kan de meting worden verricht met de « lijnaal »-tool in Photoshop of de « bijsnijden »-tool in een ander programma.

De kwaliteit van de scherpte is een extra element om de verandering in resolutie vast te stellen. Merk echter op dat het fenomeen niet altijd gemakkelijk te detecteren is, omdat het zelden het gevolg is van een plotselinge verandering: het is eerder een onderdeel van de scanner – zoals de reproductietafel, het statief enzovoort – dat mogelijk slecht bevestigd is en beetje bij beetje verzakt waardoor de kwaliteit van de afbeeldingen geleidelijk zal verslechteren.

In dergelijke gevallen wordt aangeraden de grootte van de kaart op de scanachtergrond van de eerste en de laatste beelden van een set te vergelijken, om eventuele veranderingen in grootte gemakkelijker te kunnen vaststellen. Dit controlepunt is van essentieel belang, want als het aan het eind van de productiefase niet direct wordt vastgesteld, zal de persoon die belast is met het retoucheren van de beelden het pas tijdens het bijsnijden opmerken en dan wordt de hele digitaliseringsketen vertraagd.

De tweede oorzaak van resolutieproblemen, **een fout bij de eerste instelling**, is moeilijker visueel vast te stellen, aangezien het kan gaan om een slechte verhouding die op een hele reeks beelden, of zelfs op één enkel beeld, werd toegepast. Een ervaren QA-verantwoordelijke zal dit probleem met het blote oog kunnen vaststellen, op voorwaarde dat hij vertrouwd is met het object en met de apparatuur. Ook hier zal de geïllustreerde lijnaal van de Munsell-kaart als controlemiddel worden gebruikt.

PARALLELISME

De QA-verantwoordelijke zal ook tot taak hebben na te gaan of het document tijdens het scannen perfect parallel is aan de fotografische sensor. Dit is gemakkelijk te doen wanneer het gescande object ten minste twee rechte hoeken heeft: het volstaat dan om controlerasters – aanwezig op alle viewers – toe te passen en te

controleren of de paginaranden die aan deze hoeken grenzen, zijn uitgelijnd. Als er verdwijnlijnen in het beeld verschijnen, betekent dit dat er een discrepantie is.

Als het object geen rechte hoeken heeft, moet de QA-verantwoordelijke in geval van twijfel ten minste twee metingen uitvoeren op specifieke punten buiten het midden van het object en ze vergelijken met de waarden op dezelfde punten in het beeld.

Een andere manier om een parallellismefout vast te stellen is te controleren of een deel van het beeld minder scherp is. Daartoe moet de QA-verantwoordelijke het beeld vergroten tot 100% en het naar de vier zijden en het midden verschuiven. Deze methode is echter niet altijd betrouwbaar, omdat sommige groot-formaatsscanners een zodanige scherptediepte hebben dat een verschil in oriëntatie van enkele millimeters onwaarneembaar blijft, vooral omdat de gezichtsscherpte van individu tot individu verschilt.

In het geval van een parallellismeprobleem moeten de gebrekkige beelden afgekeurd worden. Een technicus moet dit dan verder onderzoeken en de beelden moeten opnieuw worden gescand onder gepaste omstandigheden.

FLARE

Flare is **een optische aberratie die optreedt wanneer strooilicht de lens binnendringt**. De lens, die theoretisch alleen de door het object weerkaatste stralen zou moeten opvangen en naar de sensor leiden, capteert dan ook andere stralen, afkomstig van een lambdabron. **Het verschijnsel wordt gekenmerkt door een verzwakking van het contrast over het gehele beeld, soms door chromatische vervorming en, in de ernstigste gevallen, door een halo-effect.**

Nadat het fenomeen is vastgesteld, bestaat de uitdaging erin de oorsprong ervan te vinden om het te verhelpen.

Flare-effecten worden meestal veroorzaakt door een verkeerde plaatsing van de lichtbron – minder dan 30° ten opzichte van het opnamevlak. In dat geval kan het gemakkelijk worden gecompenseerd door een afdekkend zonnescherm te gebruiken. Indien het verschijnsel aanhoudt, is het de moeite waard andere parameters te controleren, zoals de reinheid van de lens – een vingerafdruk op de lens is voldoende om flare te veroorzaken, de vloer van de scanruimte – een glanzende vloer bijvoorbeeld – of de toevallige aanwezigheid van een reflecterend oppervlak in de buurt van het object.

Beelden met een flare-effect moeten afgekeurd worden en het document moet opnieuw worden gescand.

VIGNETTING

Vignettering wordt gekenmerkt door het donkerder worden van de hoeken van een beeld. Bij dit verschijnsel is dus de gelijkmatigheid van het licht op het opnamevlak gewijzigd.

Het is relatief moeilijk te detecteren, omdat het object zelden de gehele scanachtergrond beslaat. Ook is het zo dat de achtergrond zwart is om de leesbaarheid van het beeld te verbeteren en aangezien de paginahoeken van een boek het meest worden blootgesteld aan manipulatie en dus aan slijtage, kunnen deze gebieden al een waarneembare verdonkering op het fysieke object vertonen, waardoor het moeilijker is om vignettering vast te stellen. In geval van twijfel zal de QA-verantwoordelijke de staat van het object dus bijzonder goed moeten controleren.

Vignettering kan verschillende oorzaken hebben. De bekendste is **een incompatibiliteit tussen de optiek en de sensor: de optiek stuurt door zijn vorm een min of meer groot cirkelvormig beeld naar de sensor**, die rechthoekig is. Dit beeld wordt een beeldcirkel genoemd. Als de beeldcirkel niet in staat is de hele sensor gelijkmatig te bedekken, neemt de helderheid in de hoeken van het beeld af. In het kader van de digitalisering van erfgoed wordt de optiek echter geacht om voldoende getest te zijn zodat dit defect niet zou voorkomen. In geval van vignettering is het waarschijnlijker dat er per ongeluk een lamp is verplaatst.

De QA-verantwoordelijke zal dan de persoon inlichten die verantwoordelijk is voor de afstelling van de machines om dit te analyseren. Door de kwaliteiten van het belichtingstoestel te controleren en enkele beelden opnieuw te dematriceren, kan deze laatste de oorzaak van het probleem opsporen.

RUIS

Digitale ruis wordt gedefinieerd als **de aanwezigheid van informatie in het digitale beeld die niet in het originele document voorkomt**. Deze informatie komt voort uit het digitaliseringsproces. Omdat de signaal-ruisverhouding het sterkst is bij weinig licht en het best zichtbaar is op uniforme kleurvlakken, kan de QA-verantwoordelijke het fenomeen het beste detecteren door de zwarte achtergrond van een beeld te observeren.

In de praktijk gaat het om een aantal pixels waarvan de kleuren de leesbaarheid en de scherpte van het beeld beïnvloeden. Meestal is ruis te wijten aan een te hoge sensorgevoeligheid – in ASA of ISO – en is er sprake van 'leesruis'. De eerste reflex van de QA-verantwoordelijke zal daarom zijn om in de technische metadata van het beeld te

controleren of de ISO-waarde op het minimum is ingesteld. Ter indicatie: de meeste sensoren die van de lopende band komen, hebben een optimale gevoeligheid tussen ISO 50 en ISO 200. Nieuwere machines kunnen nog betere resultaten leveren bij waarden tot ISO 400¹⁶.

Aangezien er geen systeem is om dit te corrigeren, moeten beelden met digitale ruis afgewezen en opnieuw gescand worden.

ARTEFACTS

Ten slotte kan het gebeuren dat een artefact per ongeluk de juiste lezing van het beeld verstoort, of gewoon de esthetiek ervan bederft. Het kan stof op een glas zijn, een controlekaart op het bijgesneden beeld, of een object dat bedoeld is om de pagina van een boek open te houden. In dergelijke gevallen hangt alles af van het digitaliseringsbeleid van de instelling. Zo zal een zichtbaar hulpmiddel om de lijnen van een tekst recht te trekken zinvol zijn in het geval van een gedrukte pagina die bestemd is voor OCR, maar moeten dergelijke hulpmiddelen worden uitgesloten voor het maken van een facsimile van een manuscript.



fig. 8 Voorbeeld van een artefact

CegeSoma Collectie

¹⁶ Het is echter niet raadzaam om die waarde te benaderen omdat de herhaalde elektrische ladingen op de fotodiodes en de voortdurende verwerking door de computer het systeem kunnen doen opwarmen en het uitgangssignaal kunnen beïnvloeden – in dat geval spreken we van thermische ruis. Houd er rekening mee dat een sensor die te lang in live view-modus heeft gestaan, warm begint te worden. Het is dus mogelijk dat een reeks beelden wordt beïnvloed door thermische ruis en dat het verschijnsel weer verdwijnt naarmate de sensor afkoelt.

Als richtsnoer kan worden gesteld dat alles wat gemakkelijk uit het beeld kan worden verwijderd en niet bijdraagt tot het lezen, weggenomen moet worden. Zo kan bijvoorbeeld een controlekaart op het object het OCR-proces verstoren, aangezien het tekstgedeelte bevat. Evenzo zullen stof- of papierresten op de boekenwiegen, glazen of achtergronden ertoe leiden dat er een nieuwe scan moet worden gemaakt – afhankelijk van de ernst van het probleem en van de toepassingsdoelstellingen.

De QA-verantwoordelijke kan de aandacht van de operator vestigen op het ondervonden probleem om te voorkomen dat het zich opnieuw voordoet. Is de aantasting van het beeld te wijten aan een materiaalfout, dan brengt hij ook de technische operator op de hoogte, die de nodige maatregelen zal treffen om de machine weer in productie te nemen.

VERDERE INFORMATIE

BIBLIOTHÈQUE NATIONALE DE FRANCE (BnF), *Techniques et formats de conversion en mode texte*, ongedateerd, [online],
<https://www.bnf.fr/fr/techniques-et-formats-de-conversion-en-mode-texte>

CHIRON, G., DOUCET, A., COUSTATY, M., VISANI, M., MOREUX, J.-P., « Impact of OCR Errors on the Use of Digital Libraries: Towards a Better Access to Information », *In: ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, june 2017, Toronto, Canada, 4 blz.
<https://hal.archives-ouvertes.fr/hal-03025508/document>

LU, T., *Homogeneity models for image processing in the cultural heritage sector*, [Ongepubliceerde doctoraalscriptie], Vrije Universiteit Brussel, 2020, 245 blz.

MEMON, J., SAMI, M., KHAN R.-A., UDDIN M., « Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR) », *In: IEEE Access*, vol. 8, blz. 142642-142668, 2020.

RIEGER, T., *Federal Agencies Digital Guidelines Initiative. Technical Guidelines for Digitizing Cultural Heritage Material. Creation of Raster Image File*, 2016, 99 blz.,
[Online],
http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf

VAN DORMOLEN, H., *Metamorphoze Preservation Imaging Guidelines. Image Quality*, 2012, versie 1.0, [Online],
http://www.imagingetc.com/images/Resources/Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

04.

KWALITEIT VAN DE DELIVERABLES: METADATA

4.1 Metadata: definitie en kenmerken

Metadata, in ruime zin, zijn **gegevens over gegevens**. Zij hebben tot doel een ander gegeven te definiëren of te **beschrijven**, ongeacht het medium – papier of digitaal.

De term « metadata », die reeds vóór het digitale tijdperk door bibliotheken en musea werd gebruikt in de context van papieren catalogi en inventarissen, heeft met de toenemende informatisering van de samenleving een ruimere toepassing gekregen. Tegenwoordig komt de term overeen met « **alle informatie die een bron beschrijft, al dan niet digitaal, [en] met als primaire functie de inhoud ervan de bron te beschrijven, zodat het mogelijk is de bron te identificeren, te kwalificeren en te verrijken¹** ».

Net zoals er een breed scala aan erfgoed- en documentaire content is met verschillende

De voornaamste functie is de inhoud van de bron te beschrijven en het mogelijk te maken deze te identificeren, te kwalificeren en te verrijken.

kenmerken, zo zijn er ook verschillende soorten metadata. Deze gegevens worden **intern of extern** ingevoerd in het digitale bestand. Er is sprake van interne metadata wanneer ze ingekapseld zijn in hetzelfde bestand als de gedigitaliseerde bron. Externe metadata bevinden zich dan weer in een bestand of database die losstaat van het gescande document.

¹ Observatoire de la culture et des communications du Québec, État des lieux sur les métadonnées relatives aux contenus culturels, Québec : Institut de la statistique du Québec, 2017, 118 blz.
https://bdso.gouv.qc.ca/docs-ken/multimedia/PB01600FR_MetadonneesCulturel2017H00F00.pdf

Er zijn zeven categorieën te onderscheiden:

CATEGORIE	DEFINITIE	VOORBEELDEN
Identifiers	Identificeren formeel de gedigitaliseerde inhoud, doorgaans als een alfanumerieke tekenreeks.	<ul style="list-style-type: none"> • ISBN • ISNI • URI
Beschrijvende metadata	Beschrijven het gescande document zo nauwkeurig en objectief mogelijk.	<ul style="list-style-type: none"> • Naam van de auteur • Titel van het werk • Materialen • Aanmaakdatum
Administratieve metadata	Geven informatie over de datum waarop het bestand is aangemaakt en de latere wijzigingen.	<ul style="list-style-type: none"> • Aanmaakdatum bestand • Naam van de instelling die het bestand aanmaakt • Identificatie van het brondocument
Wettelijke metadata	Vermelden alle wettelijke voorwaarden waaraan de gedigitaliseerde inhoud is onderworpen: auteursrecht, rechthebbenden, gebruikerslicenties enz.	<ul style="list-style-type: none"> • Naam van de eigenaar of beheerder van het auteursrecht • Bronvermeldingen bij de titel van het werk
Verrijkende metadata	Geven aanvullende informatie over het werk of het document.	<ul style="list-style-type: none"> • Biografie van de artiest • Foto's • Notities • Omslagafbeelding
Technische (of structurele) metadata	Geven informatie over hoe het bestand is gemaakt, het formaat en hoe het kan worden gebruikt.	<ul style="list-style-type: none"> • Software-ID en versie • Scanmodus • Compressieformaat • Bestandsgrootte
Gebruiksmetadata	Informereren over het gebruik van een bepaalde culturele inhoud (meestal automatisch gecreëerd door het platform dat de gedigitaliseerde inhoud verspreidt).	<ul style="list-style-type: none"> • Aantal weergaven • Aanbevelingen van gebruikers

tab. 3 Verschillende soorten metagegevens

Bron : Observatoire de la culture et des communications du Québec, *État des lieux sur les métadonnées relatives aux contenus culturels*, Québec : Institut de la statistique du Québec, 2017, blz. 19
https://bdso.gouv.qc.ca/docs-ken/multimedia/PB01600FR_MetadonneesCulturel2017H00F00.pdf

4.2 Rol en functie van metadata

Metadata worden gebruikt om informatie te ordenen en te sorteren. Dankzij metadata is het mogelijk om catalogi, directory's en databases te doorzoeken om de gezochte informatie te vinden. **De productie van metadata is dus noch een strategie, noch een doel op zich**, maar een middel om tegemoet te komen aan de behoeften en verwachtingen van de verschillende belanghebbenden in de waardeketen van culturele producten. Hoe gestructureerder de informatie, hoe gemakkelijker ze te verwerken en relevant te maken.

In het licht van de steeds toenemende stroom van content, publicaties en informatie in ruimere zin, zorgen metadata er ook voor dat informatie niet simpelweg verloren gaat.

Alvorens gegevens te produceren, is het dus van essentieel belang **te weten in welke context de informatie zal worden gepubliceerd en wat het doel ervan is**. Dit maakt het mogelijk links te leggen die onze informatie verbinden met alle andere relevante informatie. Hoe meer verbanden er zijn tussen verschillende werken, mensen, organisaties, plaatsen en gebeurtenissen, hoe groter het potentieel voor het vinden en (her)gebruiken van al deze informatie.

Om op een relevante manier aan deze behoeften tegemoet te komen, moeten metadata aan bepaalde voorwaarden voldoen: **zij moeten interoperabel zijn, van aanvaardbare kwaliteit, en het resultaat zijn van een visie die door de verschillende betrokken belanghebbenden en gebruikers wordt gedeeld**.

4.3 Voorwaarde voor het gebruik van metadata: interoperabiliteit

De behoefte om informatie te structureren ontstond in de jaren zeventig met het wijdverbreide gebruik van computers in onze samenleving². Een van de belangrijkste elementen in de evolutie van metadata was de behoefte aan **informatie die leesbaar en bruikbaar is voor de verschillende actoren binnen een bepaalde sector, zowel mensen als machines**. Ziedaar de uitdaging van **interoperabiliteit**, gedefinieerd als «*het vermogen van een product of systeem, waarvan de interfaces volledig bekend zijn, om te werken met andere bestaande of toekomstige producten of systemen, zonder beperkingen inzake toegang of implementatie*»³.

Alvorens verder te gaan, is **het belangrijk een onderscheid te maken tussen interoperabiliteit en compatibiliteit**. Compatibiliteit stelt twee systemen in staat met elkaar te communiceren, terwijl interoperabiliteit twee systemen in staat stelt samen te werken. Compatibiliteit situeert zich tussen een formaat A en een formaat B, en bestaat in het feit dat twee elementen elkaar begrijpen: dataformaat A van een bepaalde software is in staat informatie te ontvangen die is vervat in dataformaat B van een andere software, en vice versa. Kortom, compatibiliteit heeft te maken met het vertalen van het ene formaat naar het andere, van geval tot geval, met name wanneer de formaten gesloten zijn.

Interoperabiliteit daarentegen is **algemeen, aangezien zij gebaseerd is op een of meer open formaten (b.v. HTML), onafhankelijk van de**

2 Clarr, Thierry, et Isabelle Westeel, *Numériser et mettre en ligne*. Villeurbanne : Presses de l'enssib, 2010. <http://books.openedition.org/pressesenssib/414>

3 *Définition de l'interopérabilité, geraadpleegd*, <https://aful.org/gdt/interop>

software die wordt gebruikt om de gegevens te creëren en te gebruiken. Deze standaarden hebben een dubbele rol: ze geven aan hoe de dialoog tussen de verschillende elementen zal verlopen en ze maken een communicatiepoort mogelijk die zich aan de veranderende behoeften van de elementen kan aanpassen.

Meer specifiek heeft interoperabiliteit betrekking op drie aspecten van informatie-uitwisseling:

TECHNISCHE INTEROPERABILITEIT

Technische of IT-interoperabiliteit heeft betrekking op de manier waarop systemen aan elkaar worden gekoppeld, de definitie van interfaces, het dataformaat en de protocollen. De term slaat op het vermogen van verschillende technologieën om te communiceren en gegevens uit te wisselen op basis van vooraf gedefinieerde en ondubbelzinnige interfacenormen (bijv.: HTTP, OAI-PMH).

SEMANTISCHE INTEROPERABILITEIT

Semantic interoperability ensures that the exact meaning of the information exchanged is understood by any other application, even if it was not originally designed for that specific purpose. To achieve this, systems need to share a common information exchange model, such as the Dublin Core or MODS metadata models.

SYNTACTISCHE INTEROPERABILITEIT

Tot slot heeft syntactische interoperabiliteit betrekking op de manier waarop de gegevens worden gecodeerd en geformatteerd, met name door de aard, het type en het formaat van de uitgewisselde berichten te definiëren (bijvoorbeeld: XML, RDF).

INTEROPERABILITEIT VAN BEELDEN

HET IIIF-PROJECT

*Hoewel er veel standaarden en protocollen zijn ontwikkeld om aan de behoeften van de uitwisseling van metadata te voldoen, **vielen beelden lange tijd uit de boot**. Het online zetten van kwaliteitsbeelden blijft nog steeds een technisch complex en duur proces. Culturele instellingen hebben weliswaar tools ontwikkeld om digitale inhoud te bekijken, maar deze zijn zelden interoperabel, waardoor de beelden op hun oorspronkelijke site blijven staan en alleen via dit medium beschikbaar zijn, en dus moeilijk kunnen worden gedeeld en hergebruikt.*

*Om dit probleem aan te pakken werd **het IIIF-project** (International Image Interoperability Framework) **opgezet, met als doel een interoperabiliteitskader voor digitale bibliotheken te definiëren**. Deze reeks normen stelt wetenschappelijke en culturele instellingen dus in staat om hun inhoud 'op een gestandaardiseerde manier op het web te promoten en doorzoekbaar, manipuleerbaar en annotateerbaar te maken voor om het even welke compatibele toepassing of software'*

Dit interoperabiliteitskader berust op twee pijlers:

- *Een datamodel, het Shared Canvas⁵, gebaseerd op gelinkte data om gezamenlijk gedigitaliseerde content te beschrijven.*
- *API's (Application Programmable Interface), d.w.z. een reeks functies die toegang geven tot de diensten van een toepassing, via een programmeertaal.*

Deze API's maken het bijvoorbeeld mogelijk om een gestandaardiseerde URL-syntax te definiëren die toelaat om een beeld op afstand te manipuleren of om op gestandaardiseerde wijze informatie over te brengen over de presentatie en structuur van een digitaal object.

Al deze informatie wordt verzameld in één enkel bestand, het « **IIIF-manifest** », dat niet alleen het digitale beeld beschrijft, maar ook de interne structuur van de gedigitaliseerde inhoud. Binnen dit manifest wordt deze informatie gekoppeld aan een canvas, d.w.z. een lege ruimte die op abstracte wijze een bepaalde weergave van de gedigitaliseerde inhoud geeft, bijvoorbeeld een bladzijde, in het geval van een boek. Dit canvas dient als een vergaarbak voor allerlei digitale bronnen die als annotaties op dit canvas zijn ontworpen.

Dit manifest moet vervolgens worden gelezen door een IIIF-compatibele viewer⁴, die de informatie erin dan op afstand kan laden om de digitale inhoud opnieuw samen te stellen in een andere webinterface dan de originele, en ze vervolgens naar believen te kunnen manipuleren en annoteren.

IIIF-technologieën openen dus nieuwe horizons voor de valorisatie van gedigitaliseerd cultureel erfgoed in een van nature gedecentraliseerde omgeving. Voor culturele en wetenschappelijke instellingen bieden dergelijke technologieën tal van voordelen. Financieel weegt het project niet al te zwaar aangezien het door een grote internationale gemeenschap wordt gedragen. Men is minder afhankelijk van een specifieke tool of software en het wordt gemakkelijker om van tool te wisselen en de beste in zijn klasse te kiezen. Meer in het algemeen vergroten de IIIF-normen het potentieel van digitale bibliotheekinhoud.

4 Régis Robineau, Comprendre IIIF et l'interopérabilité des bibliothèques numériques | Insula, 8 november 2016. <https://insula.univ-lille3.fr/2016/11/comprendre-iiif-interoperabilite-bibliothèques-numériques/>

5 Robert Sanderson, Benjamin Albritton, Shared Canvas Data Model, geraadpleegd. <https://iiif.io/model/shared-canvas/1.0/>

6 De IIIF gemeenschap heeft op GitHub alle image viewers geplaatst die compatibel zijn met IIIF technologieën: <https://github.com/IIIF/awesome-iiif#image-viewers>

4.4 Criteria voor kwaliteitsvolle metadata

Net als voor beelden hebben digitaliseringsexperts en wetenschappers criteria vastgesteld voor kwaliteitsvolle metadata. De hieronder gepresenteerde categorisering is rechtstreeks geïnspireerd op het werk van de Task Force Europeana⁷, de onderzoeksgroep FORCE11 en hun FAIR-beginselen⁸, het Observatoire de la Culture et des Communications du Québec⁹, het Quality Assurance Framework for Statistical Data (QAF)¹⁰, ontwikkeld door Statistics Canada (STC), en ten slotte op het werk van Thomas R. Bruce en Diane I. Hillmann¹¹.

Om van goede kwaliteit te zijn, moeten metadata de volgende eigenschappen bezitten:

01 Volledig

Metadata moeten het digitale object zo nauwkeurig mogelijk beschrijven. Deze categorieën moeten zo worden ontworpen dat zij zoveel mogelijk van toepassing zijn op alle digitale bestanden in de collectie.

7 Marie-Claire Dangerfield, Lisette Kalshoven, Report and Recommendations from the Europeana Task Force on Metadata Quality, mei 2015, 54 blz. https://pro.europeana.eu/files/Europeana_Professional/Publications/Metadata%20Quality%20Report.pdf

8 Force 11, Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing versie b1.0 <https://www.force11.org/fairprinciples>

9 Observatoire de la culture et des communications du Québec, État des lieux sur les métadonnées relatives aux contenus culturels, Québec : Institut de la statistique du Québec, 2017, 118 blz. https://bdso.gouv.qc.ca/docs-ken/multimedia/PB01600FR_MetadonneesCulturel2017H00F00.pdf

10 European Statistical System, Quality Assurance Framework, geraadpleegd, versie 2.0, 58 blz. <https://ec.europa.eu/eurostat/documents/64157/4392716/ESS-QAF-V1-2final.pdf/bbf5970c-1adf-46c8-afc3-58ce177a0646>

11 Thomas Bruce, Diane Hillmann, The Continuum of Metadata Quality: Defining, Expressing, Exploiting, 2004, 16 blz. https://www.researchgate.net/publication/247818823_The_Continuum_of_Metadata_Quality_Defining_Expressing_Exploiting

02 Nauwkeurig

Metagegevens moeten correct zijn en gebaseerd op objectieve en feitelijke kennis van het document. Nauwkeurighedsproblemen kunnen zich uiten in typografische fouten, onjuist gebruik van namen of plaatsen of door te veel afkortingen.

03 Betrouwbaar door hun herkomst

De betrouwbaarheid van metadata hangt gedeeltelijk af van de persoon die ze heeft ingevoerd – zijn bekwaamheid, zijn methodologie voor het creëren en verwerken van de metadata en ten slotte zijn ervaring. Soms wijzigt iemand anders de metagegevens nadat het object is gemaakt. Er bestaan echter tools en methodologieën om de betrouwbaarheid van metadata te waarborgen: in de bibliotheeksector bijvoorbeeld ontwikkelden de Amerikanen de Anglo-American Cataloging Rules (AACR2).

04 Conform aan de verwachtingen

Om zinvol te zijn, moeten metadata voldoen aan de verwachtingen en behoeften van de gebruikers van het beschreven digitale object. Hiervoor is het essentieel om de operationele doelstelling van het digitaliseringsproject nauwkeurig te definiëren en de behoeften van de doelgroep van dit project in kaart te brengen.

05 Consistent en coherent

Om de samenhang van de gedigitaliseerde documenten te waarborgen, moeten de gebruikte termen precieze regels volgen om onnauwkeurigheden te voorkomen.

06 Toegankelijk

Metadata moeten leesbaar zijn voor zowel machines als mensen. Het gebruik van een norm die te ingewikkeld of moeilijk leesbaar is, zal de codering en verwerking van het beschreven document bemoeilijken. A fortiori kan dit leiden tot fouten inzake consistentie en nauwkeurigheid.

De toegankelijkheid zit ook in het gemak waarmee het object online kan worden gevonden, en heeft dus te maken met de identifier. Deze laatste moet uniek en persistent zijn. De URI-standaard¹² (Uniform Resource Identifier) voldoet bijvoorbeeld aan deze dubbele eis.

4.5 Een gemeenschappelijke visie voor meer interoperabiliteit

Er bestaat geen universele norm die volledig genoeg is om aan alle bovengenoemde kwaliteitscriteria te voldoen en die tegelijkertijd perfect is afgestemd op de behoeften van culturele en wetenschappelijke instellingen en hun publiek. In de praktijk hebben metadatastandaarden vaak meerdere functies, aangezien elk type culturele instelling – bibliotheek, archief en museum – zijn eigen behoeften heeft voor het identificeren en catalogiseren van zijn collecties.

¹² Een URI is een korte tekenreeks waarmee een fysieke of abstracte bron op een netwerk (bv. een onlinebron) wordt geïdentificeerd en waarvan de syntaxis een standaard volgt die voor het internet is vastgesteld. Het acroniem URI wordt doorgaans gebruikt om zo'n tekenreeks aan te duiden. URI's vormen de basistechnologie van het wereldwijde web omdat alle hyperlinks in deze vorm worden uitgedrukt. Zij vallen in twee categorieën uiteen: URN's en URL's. Voor een overzicht van de door de W3C aanbevolen web-identifiers, zie: <https://www.w3.org/Addressing/>

Bibliotheken hebben een lange traditie in het catalogiseren van hun collecties. Elk werk moet nauwkeurig worden geïdentificeerd via een bibliografisch bestand, maar deze metadata zijn vaak niet erg gedetailleerd of gecontextualiseerd. Eulalia Roel wijst er in dit verband op dat « de kracht van het bibliotheekmodel is dat een minimale hoeveelheid metadata wordt toegekend aan een zeer hoog percentage objecten in de collectie¹³».

Een van de grootste uitdagingen voor bibliotheken was de uitwisseling van bibliografische gegevens tussen bibliotheken en archieven. Dit probleem werd aangepakt met het MARC-formaat van de Library of Congress, dat dankzij zijn architectuur en werking deze uitwisselingen vergemakkelijkt. Uit dit model kwamen verschillende varianten voort: zo ontwikkelde de BnF bijvoorbeeld het InterMARC-formaat, terwijl in Canada het CAN/MARC-formaat werd gecreëerd. In werkelijkheid hebben deze individualiseringsinspanningen de gegevensuitwisseling net complexer gemaakt, met als gevolg dat er conversieprogramma's tussen de metadatastandaarden moeten worden ontwikkeld.

Archieven verschillen van bibliotheken in die zin dat metadata in de eerste plaats de individuele objecten in de collectie moeten contextualiseren door de relaties van een collectie met mensen, plaatsen enzovoort nauwkeurig aan te geven. De grootste moeilijkheid daarbij komt volgens Eulalia Roel voort uit de « grote inconsistentie in de details en variabelen van de beschrijving, en in de sterk uiteenlopende hiërarchische granulariteit van de collectie waaraan metadata worden toegewezen¹⁴». Deze verschillen dragen bij tot misverstanden tussen diverse instellingen wanneer zij metadata over de disciplines heen proberen te gebruiken.

Musea zijn terughoudender in het gebruik van beschrijvende normen voor hun collecties, deels omdat de museumcultuur er geen voorstander van is alle informatie over een collectie vrij toegankelijk te maken voor het publiek. Tegelijkertijd maakt het eclectische karakter van museumcollecties dit vaak tot een uiterst complexe oefening: de rigiditeit van metadata is moeilijk te verenigen met de genuanceerde informatie die voor elk object in een collectie nodig is¹⁵.

Hoewel interoperabiliteit reeds lang wordt beschouwd als het sleutelement voor de uitwisseling van gegevens tussen instellingen en sectoren¹⁶, blijft het onzeker hoe dit niveau van interoperabiliteit kan worden bereikt.

Zoals opgemerkt door Milena Dobrova¹⁷ en Philip Hider¹⁸, vloeit deze moeilijkheid voort uit het feit **dat interoperabiliteit wordt benaderd vanuit twee invalshoeken die niet gemakkelijk te verzoenen zijn.** Enerzijds richten sommige instellingen zich op **interne** interoperabiliteit, terwijl andere de voorkeur geven aan **interinstitutionele ontwikkeling.** Anderzijds zullen sommige instellingen interoperabiliteit in de eerste plaats willen bereiken via de naleving van metadatastandaarden, terwijl andere meer belang hechten aan hoe nieuwe technologieën met uiteenlopende metadataformaten en inhoud kunnen werken¹⁹.

13 Eulalia Roel, « The MOSC project: Using the OAI-PMH to bridge metadata cultural differences across museums, archives, and libraries », In: *Information Technology and Libraries*, 24 (1), 2005, blz. 22-24.

14 *Idem*

15 *Idem*

16 Kalina Sotirova, Juliana Peneva, Stanislav Ivanov, Rositza Doneva, Milena Dobrova, « Digitisation of Cultural Heritage : Standards, Initiatives », In: *Access to Digital Cultural Heritage : Innovative Applications of Automated Metadata Generation*, Plovdiv : Plovdiv University Publishing House, 2012.
<http://www.math.bas.bg/infres/book-ADCH/ADCH-ch1>

17 *Idem*

18 Philip Hider, « Australian digital collections: Metadata standards and interoperability », In: *Australian Academic & Research Libraries*, 35, 2004.

19 *Idem*

Het doel is nu metadatamodellen te ontwikkelen waarmee gedigitaliseerde inhoud tussen de drie sectoren kan worden uitgewisseld. Deze wederzijdse uitwisseling vereist meer flexibiliteit en inclusiviteit in hun interoperabiliteit.

Zonder een uitputtend overzicht te willen geven van de initiatieven die de afgelopen jaren zijn ontwikkeld, stellen we hier enkele modellen voor die in die richting gaan:

Het CIDOC Conceptual Reference Model²⁰ (CIDOC CRM) is een conceptueel model voor de beschrijving van museumobjecten dat sinds 2006 wordt ontwikkeld door het International Committee for Documentation (CIDOC) van de International Council of Museums (ICOM). Het beoogt alle informatie die nodig is voor de documentatie en het beheer van cultureel erfgoed beschikbaar te maken door een gemeenschappelijk en uitbreidbaar semantisch kader te bieden, waarmee alle informatie over cultureel erfgoed in kaart kan worden gebracht. Het is nog in ontwikkeling, maar omvat momenteel 86 klassen en 137 eigenschappen en is gecentreerd rond het begrip gebeurtenis, d.w.z. het soort relatie dat een object verbindt met zijn kenmerken, zoals bijvoorbeeld de aanmaakdatum.

IFLA LRM²¹, in 2017 ontwikkeld door de International Federation of Library Associations and Institutions (IFLA), stelt een nieuw model voor bibliotheken voor. Het is gericht op de behoeften van de gebruikers in die zin dat het de structurering van gegevens voor het zoeken van bibliografische informatie op het internet bevordert zodat gebruikers prioritair bepaalde werken of plaatsen kunnen identificeren alvorens een bron op te zoeken.

Hiertoe is het model gebaseerd op drie concepten: de entiteit, d.w.z. het beschreven object, het attribuut, dat overeenkomt met de

verschillende kenmerken van dit object, en de relaties. Deze relaties betreffen zowel een entiteit met haar eigen attributen als de relaties tussen verschillende entiteiten in een of meer systemen.

Op lange termijn is het de bedoeling bibliografische records te vervangen door een netwerk van relaties tussen entiteiten en zo hun zichtbaarheid op het internet en hun exploitatie door machines te bevorderen.

Records in Contexts (RiC)²² is een project om de archivalistische beschrijvingsnormen te herzien, dat in 2016 werd gestart door de International Council on Archives (ICA). Het doel is de interoperabiliteit van online gepubliceerde archiefbeschrijvingen te verbeteren door een model te creëren dat de bestaande standaarden voor archiefbeschrijving – ISAD(G)²³, ISAAR (CPF)²⁴, ISDF²⁵ en ISDIAH²⁶ – verenigt.

20 Home | CIDOC CRM,
<http://www.cidoc-crm.org/>

21 IFLA, IFLA Library Reference Model (LRM),
<https://www.ifla.org/publications/node/11412>

22 International Council on Archives, Records in Contexts, Conceptual Model, geraadpleegd,
<https://www.ica.org/en/egad-ric-conceptual-model>

23 Een norm die het schrijven van compatibele, relevante en expliciete archiefbeschrijvingen verzekert, en tegelijk het gebruik van gemeenschappelijke authority data mogelijk maakt om de integratie van beschrijvingen uit verschillende repositories mogelijk te maken.

Bron : Observatoire de la culture et des communications du Québec, État des lieux sur les métadonnées relatives aux contenus culturels, Québec : Institut de la statistique du Québec, 2017, blz.98.
https://bdso.gouv.qc.ca/docs-ken/multimedia/PB01600FR_MetadonneesCulturel2017H00F00.pdf

24 International Standard Archival Authority Record for Corporate Bodies, Persons and Families.

Bron : International Council on Archives, ISAAR (CPF), 2011.
https://www.ica.org/sites/default/files/CBPS_Guidelines_ISAAR_Second-edition_EN.pdf

25 International standard for the description of functions and communities associated with the production and management of archives.

Source : International Council on Archives, ISDF, 2011.
https://www.ica.org/sites/default/files/CBPS_2007_Guidelines_ISDF_First-edition_EN.pdf

26 International Standard for Describing Institutions with Archival Holdings.

Bron : International Council on Archives, ISDIAH, 2011.
https://www.ica.org/sites/default/files/CBPS_2008_Guidelines_ISDIAH_First-edition_EN.pdf

Net als de twee hierboven genoemde modellen, stelt RiC ook de implementatie voor van een relationeel systeem tussen de entiteiten en hun kenmerken, evenals tussen deze entiteiten.

Het belang van deze nieuwe metadatamodelen **ligt in hun benadering van digitale objecten, die niet worden gezien als geïsoleerde elementen** waaraan men eenvoudigweg een reeks beschrijvende kenmerken toekent, **maar eerder als een heterogeen geheel van objecten dat pas echt zin heeft door de relaties ertussen.** Door reeds bestaande standaarden samen te brengen, gebaseerd op de RDF-taal (Resource Description Framework) en ten slotte door te vertrouwen op het gebruik van *Linked Open Data* (LOD) sluiten deze nieuwe modellen volledig aan bij de transitie naar het semantische web.

Deze aanpak maakt het immers niet alleen mogelijk een coherenter en zinvoller netwerk van digitale objecten op te bouwen, maar ook om deze objecten meer zichtbaarheid te geven. Voor culturele en wetenschappelijke instellingen opent het vooruitzicht om rijke en complexe corpora te creëren en te bundelen de deur naar een betere valorisatie van hun collecties.

Hoewel deze modellen streven naar meer koppelingen en een betere interoperabiliteit, moet nog worden bekeken hoe de uitwisseling tussen de verschillende sectoren precies zal verlopen. De eerste pogingen tot harmonisatie zijn gedaan met bijvoorbeeld het FRBRoo-project, dat het CRM-model van CIDOC, dat hoofdzakelijk voor musea is ontworpen, en het **FRBR-model**²⁷, dat voor bibliotheken is ontwikkeld, op elkaar probeert af te stemmen of zelfs samen te voegen

Het doel is nu metadatamodelen te ontwikkelen waarmee gedigitaliseerde inhoud tussen de sectoren kan worden uitgewisseld

27 *Introduction à FRBRoo | Documentation, geraadpleegd.*
<https://doc.bibliissima.fr/introduction-a-frbroo>

VERDERE INFORMATIE

CARBONE, F., OOSTERBEEK, L., COSTA, F., « Extending and adapting the concept of quality management for museums and cultural heritage attractions: A comparative study of southern European cultural heritage managers' perceptions », In: *Tourism Management Perspective*, vol. 35, Juli 2020.

FLECKENSTEIN M., FELLOWS, L., « Data Quality », In: *Modern Data Strategy*, Cham: Springer, 2018.

BACHI, V., FRESA, A., PIEROTTI, C., PRANDONI, C., « The digitization age: mass culture is quality culture. challenges for cultural heritage and society », In: Ioannides, M., Magnenat-Thalmann, N., Fink, E., Žarnić, R., Yen, A.-Y., Quak, E. (eds.) *EuroMed 2014*, Limassol, Cyprus, November 3-8, 2014, 17 blz., [Online],

<https://www.digitalmeetsculture.net/wp-content/uploads/2014/12/The-Digitization-Age-Mass-Culture-is-Quality-Culture.pdf>

GUEGUEN, G., MARQUES DA FONSECA, V.-M., D., PITTI, V., SIBILLE, C., *Vers un modèle conceptuel international pour la description archivistique. Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives' Experts Group on Archival Description*, 2013, 17 blz., [Online],

https://www.ica.org/sites/default/files/EGAD_English.pdf

CHARDONNENS, A., *La gestion des données d'autorité archivistiques dans le cadre du Web de données*, [Ongepubliceerde doctoraalscriptie], Université Libre de Bruxelles, 2021, 420 blz., [Online],

<https://difusion.ulb.ac.be/vufind/Record/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/315804/Holdings>

CARON, P., DEFIOLLE, R., LAY, M.-H., *L'enjeu des métadonnées dans les corpus textuels : Un défi pour les sciences humaines*, Presses universitaires de Rennes, 2019.

BRUCE, T., HILLMANN, D., *The Continuum of Metadata Quality: Defining, Expressing, Exploiting*, 2004, 16 blz., [Online],

https://www.researchgate.net/publication/247818823_The_Continuum_of_Metadata_Quality_Defining_Expressing_Exploiting

05.

BEHEER VAN HET DIGITALISERINGS- ATELIER

Hoewel in deze gids tot dusver vooral aandacht is besteed aan digitale documenten, loont het de moeite ook gedigitaliseerde originelen onder de loep te nemen. Deze kostbare en vaak fragiele objecten, zoals oude manuscripten, moeten met bijzondere zorg worden behandeld, maar er moet ook rekening worden gehouden met de omgeving om mogelijke aantasting te beperken.

Hieraan wil de **preventieve conservatie** tegemoetkomen, door ICCROM¹ omschreven als « alle maatregelen en acties die erop gericht zijn toekomstige aantasting of verlies van een kunstvoorwerp of erfgoedcollectie te voorkomen of tot een minimum te beperken²». **Deze maatregelen en acties zijn indirect: zij hebben geen invloed op de materialen en structuren van de objecten en veranderen het uiterlijk ervan niet.**

¹ Het International Centre for the Study of the Preservation and Restoration of Cultural Property (ICCROM) is een intergouvernementele organisatie die in dienst van haar lidstaten ijvert voor het behoud van alle vormen van cultureel erfgoed, overal ter wereld.

² ICCROM, Preventive conservation, geraadpleegd.
<https://www.iccom.org/fr/section/conservation-preventive>

5.1 Behoud van erfgoed- en documentaire collecties

Er zijn tien belangrijke factoren voor de aantasting van erfgoed³:

- Temperatuur,
- Vochtigheid,
- Licht,
- Luchtvervuiling,
- Stof, schimmel,
- Insecten en dieren,
- De menselijke factor – met inbegrip van diefstal en vandalisme – en ten slotte
- Rampen – met inbegrip van branden, overstromingen en gewapende conflicten.

Organische en cellulosematerialen zoals papier, perkament of leer behoren tot de **meest**

³ Canadian Conservation Institute, Agents of deterioration, 2017
<https://www.canada.ca/en/conservation-institute/services/agents-deterioration.html>

gevoelige voor relatieve vochtigheid⁴, omdat zij hygroscoopisch zijn, d.w.z. dat zij vocht kunnen absorberen en desorberen. Met name een te grote variatie in de relatieve vochtigheid kan leiden tot veranderingen in grootte die op hun beurt leiden tot scheuren, breuken, opkrullen, vervormingen⁵, enzovoort.

De tabel hieronder⁶ geeft een overzicht van de omgevingsvariabelen en de toegestane afwijkingen voor de optimale conservatie van de meest voorkomende materialen in archief- en bibliotheekcollecties:

De moeilijkheid van preventieve conservatie ligt in de heterogeniteit van de objecten en documenten die in dezelfde ruimte bijeen zijn gebracht: de optimale omstandigheden voor de conservatie van papier zijn niet dezelfde als voor een foto of voor leer. **Deze cijfers zijn echter slechts indicatief, want meer nog dan de cijfers gaat het erom een zo stabiel mogelijke omgeving te creëren voor de erfgoedcollecties en tegelijk rekening te houden met het comfort van de teams.**

Het uitstekende *Vade-mecum de la conservation préventive*⁷ van het **Centre de Recherche et de Restauration des Musées de France (C2RMF)** bevat een aantal aanbevelingen voor het beheer van de omgevingscondities van collecties en voor het verhelpen van anomalieën, waaronder:

- De relatieve vochtigheid, temperatuur en helderheid van de omgeving meten met behulp van thermohygrografen en lichtmeters
- Een klimaatkaart van het gebouw maken en de meetwaarden op een plattegrond noteren
- Risicogebieden in kaart brengen

- Bronnen van problemen bepalen.
- Variaties in temperatuur en vochtigheid zoveel mogelijk beperken.
- De spreiding van de objecten indien nodig aanpassen volgens klimaatzone.
- De meetapparatuur en specifieke materialen regelmatig controleren en onderhouden.
- De meest stabiele zones identificeren voor het herbergen van de meest fragiele collecties.
- De klimatologische omstandigheden van bepaalde zones indien nodig aanpassen om te weten wanneer en hoe collecties moeten worden verplaatst.
- Een noodplan en een evacuatieplan opstellen in geval van een ramp.

4 Vochtigheid is het gehalte aan waterdamp in de lucht, en het niveau daarvan in de atmosfeer wordt gemeten in relatieve vochtigheid, die « de hoeveelheid waterdamp in een bepaald volume lucht uitdrukt in verhouding tot het maximum dat de lucht bij een gegeven temperatuur en druk kan bevatten ».

Bron : Etienne Fehau, Nathalie Le Dantec, *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 november 2013, blz.10.
https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf

5 Jean-Paul Oddos (dir.), *La conservation : principes et réalités*, Paris : Éd. du Cercle de la Librairie, Coll. Bibliothèques, 1995, 405 blz.

6 In deze tabel wordt een overzicht gegeven van de informatie in de volgende documenten:

Conseil Canadien des Archives, *Manuel de conservation des documents d'archives, chapitre 3 - Environnement*, Conseil canadien des archives, 2003, 18 blz.
http://www.cdncouncilarchives.ca/RBch3_fr.pdf

Pierre Diaz Pedregal, *Climat des magasins d'archives : objectifs, moyens et méthodes - Petit manuel de climatologie appliquée à la conception des bâtiments d'archives*, Direction des Archives de France, 2009, 159 blz.
<https://francearchives.fr/file/47cc986ffbb4b6b02960a1bcbd2b4a6a890cc098/Manuel%20de%20climatologie.pdf>

Etienne Fehau, Nathalie Le Dantec, *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 november 2013, 46 blz.
https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf

7 Etienne Fehau, Nathalie Le Dantec, *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 november 2013, 46 blz.
https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf

SOORT	TEMPERATUUR	RELATIEVE VOCHTIGHEID
Papier	<ul style="list-style-type: none"> • 17 °C tot 18,3 °C Afwijking: 2°C over 24uur 	<ul style="list-style-type: none"> • 45-55% Afwijking: 3% over 24uur
Perkament en leer	<ul style="list-style-type: none"> • 17 °C tot 18,3 °C Afwijking: 2°C over 24uur 	<ul style="list-style-type: none"> • 45-55% Afwijking: 3% over 24uur
Textiel	<ul style="list-style-type: none"> • 17 °C tot 18,3 °C Afwijking: 2°C over 24uur 	<ul style="list-style-type: none"> • 45-55% Afwijking: 3% over 24uur
Foto's Zwart-wit afdrukken Negatieven op glas Zwart-wit gelatinezilvernegatieven op polyesterbasis Celluloseacetaat- en cellulosenitraatnegatieven Zwart-wit gelatinezilvernegatieven op triacetaatbasis Kleurenfilms Kleurproeven	<ul style="list-style-type: none"> • Onder de 18°C • 18°C • 21°C +/- 2° over 24uur • 18°C • 2°C • 2°C • Minder dan 2°C 	<ul style="list-style-type: none"> • 30 - 50 % • 30 - 40 % • 20 - 50 % • 35 - 60 % • 20 - 50 % • 20 - 30 % • 30 - 40 %
Microfilms Zwart-wit gelatinezilvernegatieven op polyesterbasis Warmtebehandelde zilverfilms Vesiculaire films op polyesterbasis Zilverbeeld behandeld door bleken op polyesterbasis	<ul style="list-style-type: none"> • Maximaal 21°C 	<ul style="list-style-type: none"> • 20-50%

tab. 4 Aanbevolen temperatuur en relatieve vochtigheid voor de preventieve conservering van papier, textiel en leder

5.2 Lichtbeheer

Licht is het kleine deel van elektromagnetische golven dat door ons gezichtsvermogen wordt waargenomen. Het wordt zichtbare straling genoemd om aan te geven dat er andere soorten straling zijn die onzichtbaar zijn voor ons visuele systeem, maar die in wisselende verhoudingen samengaan met zichtbare straling: ultraviolette en infrarode straling. Alle lichtbronnen, van de zon tot TL-buizen en halogeenlampen, zenden een zeer grote hoeveelheid straling uit, zowel zichtbare als onzichtbare. Jean-Jacques Ezrati heeft de spectrale samenstelling van de meest voorkomende lichtbronnen op een rijtje gezet:

Het erom een zo stabiel mogelijke omgeving te creëren voor de erfgoedcollecties en tegelijk rekening te houden met het comfort van de teams.

LICHT-BRONNEN	ULTRAVIOLETTE STRALING	ZICHTBARE STRALING	INFRARODE STRALING	Globale STRALING
Gemiddeld daglicht	6%	44%	50%	100%
Wolfram/halogeenlamp	1%	9.5%	90%	100%
TL-buis 3000 K	1%	89%	10%	100%
TL-buis 5000 K	2%	88%	10%	100%
LED	0%	100%	0%	100%

tab.5 De verschillende soorten lichtbronnen en hun straling

Bron : Jean-Jacques Ezrati, *Éclairage d'exposition, musée et autres espaces*, Paris : Eyrolles, 2015, blz.97

Deze stralingen hebben fotochemische en/of thermische effecten op organische materialen, hetgeen leidt tot een onomkeerbare aantasting van het object.

Ultraviolet licht heeft een fotochemische werking en is verantwoordelijk voor kleurveranderingen, het bruin worden van het papier en een algemene mechanische verzwakking van de werken. Warmteproducerende infraroodstralen hebben

rechtstreeks een verzachtend, uitdrogend of ontvlammend effect op organische materialen. Meer in het algemeen versnellen ze de chemische aantasting ervan. Lichtgevoelige werken worden aangetast telkens wanneer ze aan licht worden blootgesteld.

Deze aantasting is cumulatief en onomkeerbaar. **Het komt er dus op aan ultraviolette straling te onderdrukken, infrarode straling te verminderen en zichtbare straling te beheersen.**

Interessante oplossingen zijn de installatie van zonweringen en UV-filters op vensters, het gebruik van lichtmeters en de installatie van ledverlichting. Kortom, de bescherming van erfgoedobjecten tegen **licht bestaat erin de hoeveelheid licht te beperken en de intensiteit of duur van de blootstelling te verminderen.**

Tegelijkertijd is het lichtbeheer een bijzonder gevoelige kwestie voor digitaliseringsateliers. Ter herinnering: de uitdaging van digitalisering bestaat erin om het originele document zo getrouw mogelijk te reproduceren. Te veel licht doet echter afbreuk aan de waarneming van kleuren en daarmee aan de betrouwbaarheid van de digitale weergave.

Lichtbeheer houdt in dat geen enkele ander bron de belichting verstoort, zodat het van

essentieel belang is te werken in een omgeving die ondoordringbaar is voor daglicht.

Zo niet zal een machine die in de winter is gekalibreerd, alle beelden in heldere perioden overbelichten, met een onomkeerbaar verlies van informatie als gevolg.

Andere lichtbronnen, zoals hulpverlichting voor het lezen van administratieve documenten of om ongelukken te voorkomen en computerschermen, kunnen in de scanruimte worden geplaatst, mits ze geen invloed hebben op het door het object weerkaatste licht. Het is daarom van essentieel belang dat deze bronnen niet op het object zijn gericht en dat hun vermogen zeer laag is – maximaal 32 lux op het object.



*fig. 9 Digitaliseringsatelier
Nationaal Archief van België*

VERDERE INFORMATIE

ICCROM, een intergouvernementele organisatie ten dienste van haar lidstaten, bevordert het behoud van het cultureel erfgoed in al zijn vormen over de gehele wereld. Hierbij stelt ze verschillende hulpmiddelen ter beschikking van culturele instellingen om hen in dit proces te ondersteunen:

- Gids voor risicobeheer voor cultureel erfgoed:
<https://www.iccrom.org/publication/guide-risk-management>
- De ABC-methode, een methodologie voor het identificeren, evalueren en corrigeren van de risico's en factoren die bijdragen tot de aantasting van erfgoedcollecties:
<https://www.iccrom.org/publication/abc-method-risk-management-approach-preservation-cultural-heritagf>

Het **Canadian Conservation Institute (CCI)** biedt erfgoedprofessionals een bondig naslagwerk met praktische tips voor het beschermen van collecties tegen aantastende factoren. Deze aanbevelingen zijn onderverdeeld in drie delen: gebouw en faciliteiten, uitrusting en materiaal, en procedures.

<https://www.canada.ca/en/conservation-institute/services/preventive-conservation/framework-preserving-heritage-collections.html>

Het **Centre de Recherche et de Restauration des Musées de France (C2RMF)** verricht en publiceert onderzoek op het gebied van behoud en restauratie van het erfgoed, en op het gebied van het archiefbeheer en het gebruik van nieuwe technologieën voor dit doel. Op zijn website stelt het centrum een methodologie voor om musea en culturele instellingen te helpen bij het opstellen van hun conservatie- en beschermingsplan.

<https://c2rmf.fr/conserver/fiches-techniques>

In België beschikt het Koninklijk Instituut voor het Kunstpatrimonium (KIK) over een cel die zich bezighoudt met onderzoek inzake preventieve conservatie. Het instituut biedt advies en analyses aan culturele en wetenschappelijke instellingen die hun aanpak van preventieve conservatie willen verbeteren.

<https://www.kikirpa.be/nl/conservatie-restauratie/cel-preventieve-conservatie>

EZRATI, J.-J. *Éclairage d'exposition, musée et autres espaces*, Paris : Eyrolles, 2015.

CONSEIL CANADIEN DES ARCHIVES, *Manuel de conservation des documents d'archives. Chapitre 3 – Environnement*, 2003.

http://www.cdncouncilarchives.ca/RBch3_fr.pdf

FEHAU E., LE DANTEC N., *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 november 2013, 46 blz.

https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf



06.

METHODOLOGISCHE FICHES OM DE KWALITEIT VAN EEN DIGITALISERINGS- PROJECT TE WAARBORGEN

*De onderstaande fiches volgen de chronologie van de digitaliseringsketen en **geven een overzicht van de elementen waarmee rekening moet worden gehouden tijdens de verschillende fasen van digitalisering.***

*Zij bestaan elk uit drie categorieën: **kernbegrippen**, dat wil zeggen de theoretische elementen die men moet kennen over een bepaalde fase, **aanbevelingen** in de vorm van een lijst, en ten slotte de **tools** om deze aanbevelingen uit te voeren.*

VÓÓR

Uw digitale gegevensstrategie bepalen

Vaststelling van uw digitaliseringsbeleid

INTELLECTUELE EN STRATEGISCHE OMSCHRIJVING VAN HET PROJECT

Benchmarks bepalen

Implementatie en uitvoering van een kwaliteitsaanpak

Kwaliteit van de deliverables bepalen

Wanneer en hoe kwaliteitscontrole uitvoeren?

Kwaliteitscontrole bij uitbesteding

DEFINITIE VAN DE TECHNISCHE ASPECTEN

Digitaliserings-apparatuur kiezen

Lichtbronnen kiezen

De digitalisering sketen kalibreren

DIGITALISERING

Algemene regels voor de beeldvorming

Aanmaken en invullen van het metadatareferentiebestand

Opslaan en naamgeving van bestanden

POST-PRODUCTIE

Retoucheren van beeldbestanden

OCR van tekstdocumenten

Inkapselen van metadata

De laatste kwaliteitscontrole

VERSPREIDING BEHOUD

Een digitaal databeheersysteem kiezen

01/ BENCHMARKS BEPALEN

KERNBEGRIPPEN

Een benchmark is een samenhangend geheel van gegevens of informatie die door een gemeenschap van actoren worden gedeeld.

Belangrijke kenmerken zijn stabiliteit, betrouwbaarheid, eenheid van betekenis en interoperabiliteit. Dit soort document kadert in de massale ontwikkeling van informatie-uitwisseling en communicatie en vloeit voort uit de behoefte aan een gemeenschappelijke taal voor deze uitwisselingen. Het gebruik van benchmarks situeert zich binnen een dynamisch perspectief, bijvoorbeeld om processen en producten te identificeren, analyseren, classificeren, vergelijken of bij te werken.

Volgens deze definitie kan de benchmark verschillende vormen aannemen. In het kader van een digitaliseringsproject zijn er drie soorten benchmarks:

- Normen.
- Interne documentaire benchmarks.
- Methodologische benchmarks.

Normen

Normen zijn gestandaardiseerde en gedocumenteerde definities van goede praktijken en processen die op verschillende gebieden worden toegepast¹.

Op internationaal niveau is de ISO (Internationale Organisatie voor Normalisatie) de referentie op dit gebied. Er zijn ook normalisatie-instellingen die zich op specifieke gebieden toelagen, zoals het World Wide Web Consortium (W3C) of de Organization for the Advancement of Structured Information Standards (OASIS), die zich richten op normen voor de IT-sector.

¹ Wanneer de specificaties van een norm volledig door een organisatie worden overgenomen, is het mogelijk een certificering aan te vragen die de technische bekwaamheid en betrouwbaarheid van de organisatie bevestigt. Hoewel deze aanpak in vele opzichten interessant is, blijft hij facultatief. In België is BELAC sinds 2006 de enige organisatie die met deze certificeringen is belast.

Interne documentaire benchmarks

Interne documentaire benchmarks nemen meestal de vorm aan van een audit. Het gaat om min of meer uitgewerkte documenten die specifiek zijn voor een organisatie en waarin de processen, middelen en belangrijke informatie die over een bepaald onderwerp gekend moeten zijn, worden verzameld. Dit soort documenten wordt steeds meer gestandaardiseerd, met name dankzij de ISO 19001-norm, die een strikt en volledig kader biedt voor het opstellen van deze documenten. De BnF biedt op haar website ook een [methodologisch overzicht voor de voorbereiding van audits](#).

Methodologische benchmarks

Methodologische benchmarks ten slotte zijn, zoals de naam al zegt, corpora van methoden en goede praktijken met betrekking tot een bepaald onderwerp.

Deze documenten evolueren per versie en kunnen van om het even welke aard zijn: algemeen of specifiek voor een domein of proces, en al dan niet intern. In de meeste gevallen moeten dergelijke documenten worden aangepast aan de specifieke kenmerken van de organisatie of instelling die ze gebruikt.

AANBEVELINGEN

- ➔ ANCTIL, M.-C, LEGENDRE, M., MÜLLER, T., MAILLET, D., BROUSSEAU, K., RENAUD, L., *Recueil de règles de numérisation*, Bibliothèque et Archives nationales du Québec, Bibliothèque nationale de France et Musée canadien de l'histoire, 2014.
<https://numerique.banq.qc.ca/patrimoine/details/52327/2426216>
- ➔ BIZINGRE, J., PAUMIER, J., RIVIÈRE, P., *Les référentiels du système d'information*, Paris : Dunod, 2013.
- ➔ BURESI, C., et CÉDELLE-JOUBERT, L., *Conduire un projet de numérisation*, Paris : Tec et Doc, 2002, 326 blz.
- ➔ CLAERR, T., WESTEEL, I., et MELOT, M., *Manuel de la numérisation*, Paris : Éd. du Cercle de la Librairie, 2011. 317 blz.
- ➔ CLAERR, T., et WESTEEL, I., *Numériser et mettre en ligne*, Villeurbanne : Presses de l'enssib, 2010.
<http://books.openedition.org/pressesenssib/414>

- ➔ **CLAERR, T., WESTEEL, I., ORY-LAVOLÉE, B., GEORGES, N.,** *Manuel de constitution de bibliothèques numériques*, Paris : Éd. du Cercle de la Librairie, 2013. 407 blz.
- ➔ **DIRECTION DES ARCHIVES DE FRANCE. MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION - COMITÉ DE PILOTAGE NUMÉRISATION - MISSION DE LA RECHERCHE ET DE LA TECHNOLOGIE (DDAI),** *Écrire un cahier des charges de numérisation du patrimoine. Guide technique. Documents reliés- manuscrits- plans- dessins- photographies- microformes*, Paris : Archives de France, 2008, 62 blz.
https://francearchives.fr/file/bf50d8fa5f554586dbf18fdc862d25970a1da0a7/static_4132.pdf
- ➔ **DIRECTION INTERMINISTÉRIELLE DU NUMÉRIQUE ET DU SYSTÈME D'INFORMATION ET DE COMMUNICATION DE L'ETAT,** *Référentiel général d'interopérabilité*, 2015, versie 2, 84 blz.
https://www.numerique.gouv.fr/uploads/Referentiel_General_Interoperabilite_V2.pdf
- ➔ **FEHAU E., LE DANTEC N.,** *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 november 2013, 46 blz.
https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf
- ➔ **FORCE 11,** *Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0*, ongedateerd.
<https://www.force11.org/group/fairgroup/fairprinciples>
- ➔ **PLAN D'EXPLOITATION ET DE PRÉSERVATION DES PATRIMOINES (PEP'S),** *Normes et lignes directrices techniques et organisationnelles pour les initiatives de numérisation des patrimoines culturels soutenues par la Communauté française, ministère de la communauté française*, Bruxelles : Secrétariat général – Préservation et exploitation des patrimoines, 2009, 61 blz.
http://www.peps.cfwb.be/index.php?eID=tx_nawsecuredl&u=0&file=fileadmin/sites/numpat/upload/numpat_super_editor/numpat_editor/documents/CFWB/Normes.pdf&hash=a2d0e740e46e34b8c3475c604c3e96edbb8a8712
- ➔ **RIEGER, T.,** *Federal Agencies Digital Guidelines Initiative. Technical Guidelines for Digitizing Cultural Heritage Material. Creation of Raster Image File*, 2016, 99 blz.
http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf
- ➔ **VAN DORMOLEN, H.,** *Metamorphoze Preservation Imaging Guidelines. Image Quality*, 2012, 44 blz.
http://www.imagingetc.com/images/Resources/Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

ISO 9000 :2015	Systemen voor kwaliteitsbeheer <i>Basisbeginselen en begrippenlijst</i>
ISO 9001 :2015	Systemen voor kwaliteitsbeheer <i>Eisen</i>
ISO 9004 :2018	Kwaliteit van een organisatie <i>Richtsnoeren voor goede prestaties</i>
ISO 19011 :2018	Richtsnoeren voor de audit van beheersystemen
ISO 31000	Risicobeheer
ISO 15489-1 :2016	Informatie en documentatie - Documentatiebeheer <i>Deel 1: Begrippen en beginselen</i>
ISO/IEC 27001:2013	Informatietechnologie - Beveiligingstechnieken - Beheersystemen voor informatiebeveiliging <i>Eisen</i>
NF Z42-026	Definitie en specificaties van diensten voor getrouwe digitalisering van documenten op papier en controle van deze diensten
ISO 2859-1	Bemonsteringsvoorschriften voor testen per attribuut <i>Deel 1: Bemonsteringsprocedures voor testen per partij, geïndexeerd volgens aanvaardbaar kwaliteitsniveau</i>
FD Z42-017: 2009	Elektronische beeldvorming - Digitalisering van documenten - Uitbesteding van digitaliseringsprocessen <i>Handleiding voor het opstellen van technische specificaties</i>
ISO 12641-1 :2016	Grafische technologie - Uitwisseling digitale preprintgegevens <i>Kleurdoelen voor kalibratie bij scannerinvoer</i>
ISO 19264-1 :2021	Fotografie - Archiveringssystemen - Kwaliteitsanalyse van beeldvormingssystemen <i>Deel 1: Reflecterende originelen</i>
ISO 7589 :2002	Fotografie - Lichtbronnen voor sensitometrie <i>Specificaties voor daglicht, gloeilampen en printers</i>
ISO 14524 :2009	Fotografie - Elektronische fotocamera's <i>Methoden voor het meten van opto-elektronische conversiefuncties (OECF's)</i>
ISO 21550 : 2004	Fotografie - Elektronische scanners voor fotografische beelden <i>Dynamisch-bereikmetingen</i>
ISO 25964-2:2013	Informatie en documentatie <i>Thesaurus en interoperabiliteit met andere thesauri</i>

02 / IMPLEMENTATIE EN UITVOERING VAN EEN KWALITEITSAANPAK

KERNBEGRIPPEN

Kwaliteitsbenadering of kwaliteitsbeheer is een beheersconcept om de werkprocessen en de resultaten ervan beter te begrijpen en te analyseren.

Dit concept is gebaseerd op de PCDA-methode voor kwaliteitsbewaking (Plan - Do - Change - Act), die een virtueuze cirkel wil tot stand brengen waarin voortdurend wordt gestreefd naar het verbeteren van de kwaliteit van projecten en instellingen.

De ISO 9000 normenfamilie is de erfgenaam van deze methode en biedt een algemeen methodologisch kader voor het verbeteren van werkprocessen en resultaten.

Doelstellingen en voordelen

- ➔ De individuele prestaties van het digitaliseringsatelier versterken door de organisatie ervan te verbeteren.
- ➔ Verbetering van de kwaliteit van de deliverables.
- ➔ Beter rekening houden met de vraag van gebruikers en content aanbieden die aan hun verwachtingen voldoet.
- ➔ De instelling betrekken bij de Europese digitale transitie en een wereldwijde organisatiecultuur delen.

AANBEVELINGEN

- ➔ Een **strategie voor gegevensbeheer** ontwikkelen.
- ➔ Een **digitaliseringsbeleid uitwerken**.
- ➔ Een **bestek opstellen** voor elk digitaliseringsproject.
- ➔ Een **workflow** creëren voor het digitaliseringsproces.
- ➔ De **rol en taken van elk teamlid** nauwkeurig omschrijven.
- ➔ Een gemeenschappelijke tool voor **projectmonitoring gebruiken** in het team om de communicatie te vergemakkelijken.
- ➔ In de loop van het project fouten, **anomalieën en storingen registreren**, evenals de aanpak ervan.
- ➔ Aan het einde van het digitaliseringsproject **met het hele team het project evalueren en vergelijken met eerdere digitaliseringsprojecten**: waren er terugkerende fouten? Is er adequaat gereageerd op deze fouten? Zijn de verwachte resultaten bereikt?
- ➔ **Strategische documenten en workflow waar nodig aanpassen om volgende projecten te verbeteren.**

TOOLS

- **ISO 9000 :2015 - Kwaliteitsbeheersysteem**

- Basisbeginselen**

- Begrippenlijst**

- **ISO 9004 :2018 - Kwaliteit van een organisatie**

- Richtsnoeren voor goede prestaties**

- **Project monitoring tools : Projector et Redmine**

- Projector** en **Redmine** zijn opensourcetools voor projectopvolging. Zij bieden talrijke functionaliteiten voor het opvolgen van taken, het visualiseren van de voor een project beschikbare middelen of het beheren van de planning binnen een team.*

VERDERE INFORMATIE

BAROUCH, G., *Révolution des pratiques managériales : le changement par la qualité*, La Plaine Saint-Denis : Afnor Éditions, 2013.

GIESEN, E., *Démarche qualité et norme ISO 9001 : Une culture managériale appliquée à la recherche*, Marseille : IRD Éditions, 2018, 180 blz.

CARBONE F., OOSTERBEEK L., COSTA C., FERREIRA AM., « Extending and adapting the concept of quality management for museums and cultural heritage attractions: A comparative study of southern European cultural heritage managers' perceptions », *In: Tourism Management Perspective*, vol.35, Juli 2020.

03 / KWALITEIT VAN DE DELIVERABLES BEPALEN

KERNBEGRIPPEN

Kwaliteit: Alle eigenschappen en kenmerken van een product, proces of dienst waardoor het in staat is om aan uitdrukkelijke of impliciete behoeften te voldoen.

Onderkwaliteit of non-kwaliteit: Niet-conformiteit van het product met de minimumvereisten die van een gedigitaliseerd document worden verwacht en niet-naleving van de leesbaarheid of esthetische integratie, waardoor het gebruik ervan de facto in het gedrang komt.

Overkwaliteit: Overmatige optimalisatie van een product, waardoor buitensporig veel middelen en kosten worden gebruikt om het product te verbeteren, zonder echte toegevoegde waarde.

Minimale of aanvaardbare kwaliteitsdrempel: Toegestane foutenmarge bij het uitvoeren van een digitaliseringsproject. Deze fouten moeten gering zijn en mogen de leesbaarheid, de esthetische integriteit en de bruikbaarheid van het document niet in het gedrang brengen.

Doelstellingen

- ⇒ Kwaliteit wordt gedefinieerd in relatie tot het beoogde gebruik van de gedigitaliseerde documenten (bewaring/verspreiding/valorisatie).
- ⇒ Tijdens de ontwikkeling van het project moeten drie mijlpalen worden vastgesteld: non-kwaliteit, overkwaliteit en de aanvaardbare kwaliteitsdrempel.
- ⇒ Voor elk criterium dat als sleutelement bij de analyse van de kwaliteit van een productie wordt gedefinieerd, moeten drempelwaarden worden vastgesteld waarboven of waaronder beelden al dan niet conform worden verklaard.
- ⇒ Deze drempels kunnen worden gebaseerd op een of meer waarden voor elk criterium.

AANBEVELINGEN

VOORBEELDEN VAN ERNSTIGE FOUTEN

- ➔ Ontbreken van het referentiebestand van de dienstverlener.
- ➔ Niet-naleving van dragers en digitale bestandsformaten.
- ➔ Niet-naleving van de naamgeving voor dragers, mappen, bestanden en afbeeldingen.
- ➔ Niet-naleving van mappen en bestandsstructuren.
- ➔ Inconsistentie in de inhoud van de belangrijkste velden van het referentiebestand.
- ➔ Ontbreken van de gevraagde indexering.
- ➔ Niet-naleving van de belangrijkste indexering volgens de verstrekte regels.
- ➔ Onvolledige digitalisering (sommige documenten werden niet gereproduceerd).

VOORBEELDEN VAN LICHT FOUTEN

- ➔ Inconsistentie in de inhoud van de secundaire velden van het referentiebestand.
- ➔ Niet-conformiteit van de ingevoerde indexering.
- ➔ Niet-naleving van de opnamevolgorde.
- ➔ Slechte kadrering die geen invloed heeft op de lezing en niet leidt tot verlies van informatie.
- ➔ Kleine afwijking in kleur en/of densiteit ten opzichte van de referentiebestanden...).
- ➔ Onjuiste oriëntatie van opname.
- ➔ Veranderde getrouwheid aan het originele document (problemen met lichtverdeling, vlakheid, perspectief, scherpte, strepen...).
- ➔ Gebruik van vuile glazen plaat om een document plat te drukken bijvoorbeeld.

04 / WANNEER EN HOE KWALITEITSCONTROLE UITVOEREN?

KERNBEGRIPPEN

Tijdens de digitaliseringsfase zijn er drie soorten kwaliteitscontroles:

- **Visuele controle**, waarbij de leesbaarheid en de esthetiek van het beeld worden geverifieerd.
B.v.: marges, achtergrondkleur, onscherpte enzovoort.
- **Technische controle**, die bestaat uit het controleren van de technische elementen van de gedigitaliseerde inhoud.
B.v.: resolutie/definitie, bestandsformaten, metadata (EXIF, IPTC XMP-model of andere), aanwezigheid van kleurprofiel, analyse van testpatronen enzovoort.
- **Integriteits- of consistentiecontrole**, om na te gaan of de modellen werden gevolgd voor de integratie van beelden in het digitale documentbeheersysteem.
B.v.: boomstructuur, bestandsnaam, mapnaam, volledigheid enzovoort.

Theoretisch kan elke stap van het digitaliseringsproces aan een kwaliteitscontrole worden onderworpen: afhankelijk van het resultaat kan het proces worden voortgezet of geheel of gedeeltelijk worden overgedaan.

Vanuit pragmatisch oogpunt is het zo dat hoe vroeger een niet-naleving wordt ontdekt, hoe lager de **herstelkosten zullen zijn**. Een niet-conform product dat door de eindgebruiker wordt ontdekt, kost in principe minstens tweemaal zoveel als een goed vervaardigd product.

De herstelkosten kunnen in werkelijkheid veel hoger zijn. Een niet-conformiteit die tijdens de productie wordt ontdekt, betekent een aanzienlijke extra kostenpost, maar minder dan een globale herneming. Een goede kwaliteitsbewaking heeft zo vanzelf een financiële rechtvaardigingsgrond.

AANBEVELINGEN

FASE VAN DE DIGITALISERINGSKETEN	PARAMETERS DIE MOETEN WORDEN GECONTROLEERD	AANBEVOLEN HULPMIDDELEN
Vorbereiding van de omgeving	<ul style="list-style-type: none"> • Temperatuur • Relatieve vochtigheid • Licht 	<ul style="list-style-type: none"> • Thermohygrometers • Luxmeters
Keuze van lichtbronnen	<ul style="list-style-type: none"> • Plaatsing en inclinatie van de lichtbronnen rond de opnameapparatuur 	
Selectie en voorbereiding van documenten	<ul style="list-style-type: none"> • Staat van de originele documenten • Externe voorwerpen die voor het scannen moeten worden verwijderd (paperclips, losse papieren, bladwijzers enz.) 	
Aanmaken en invullen van het referentiebestand met beschrijvende metadata	<ul style="list-style-type: none"> • Consistentie tussen de ingevoerde gegevens en de gebruikte standaard • Volledigheid en juistheid van de ingevoerde metadata • Consistentie van de gebruikte terminologie 	<ul style="list-style-type: none"> • OpenRefine
Kalibratie van de keten	<ul style="list-style-type: none"> • Staat van de apparatuur: stabiliteit, zuiverheid van de lens • Invloed van licht op de scanapparatuur • Gebruikte kleurruimte • Consistentie van ICC-profielen toegepast op apparaten 	<ul style="list-style-type: none"> • AutoSFR • OpenDice • Universal Test Target (UTT)
Preproductie & Productie	<ul style="list-style-type: none"> • Volledigheid • Belichting • Contrast • Kleur • Scherpstelling • Resolutie • Parallellisme • Flare • Vignettering • Ruis • Artefacten 	
Opslag en naamgeving bestanden	<ul style="list-style-type: none"> • Consistentie tussen bestandsnamen en het naamgevingsschema van de instelling 	<ul style="list-style-type: none"> • File Renamer Basic • BulkRename

FASE VAN DE DIGITALISERINGSKETEN	PARAMETERS DIE MOETEN WORDEN GECONTROLEERD	AANBEVOLEN HULPMIDDELEN
Retoucheren van beelddocumenten	<ul style="list-style-type: none"> • Consistentie tussen de ICC-profielen van de scanapparatuur en het programma • Aanwezigheid en juistheid van technische metadata • Technische kenmerken van de beelden 	<ul style="list-style-type: none"> • Photoshop • Lightroom • Gimp
OCR van tekstdocumenten	<ul style="list-style-type: none"> • Weglatingen • Classificatieproblemen OCR-software • Segmentatieproblemen 	<ul style="list-style-type: none"> • ABBY Fine Reader • Tesseract • Layout Evaluation
Integratie van technische en administratieve metadata	<ul style="list-style-type: none"> • Volledigheid van de aanwezige metadata: beschrijvend / administratief / technisch • Juistheid van de paden tussen verschillende metadatabestanden 	<ul style="list-style-type: none"> • Standard METS
Eindcontrole van de kwaliteit	<ul style="list-style-type: none"> • Consistentie tussen de in het bestek beschreven criteria en de verkregen resultaten: • Conformiteit van de formaten • Conformiteit van de bestandsnamen • Volledigheid van de beelden • Technische kenmerken van de beelden • Aanwezigheid en juistheid van de beschrijvende, technische en administratieve metadata • Volledigheid en juistheid van het METS-bestand 	

05 / KWALITEITS- CONTROLE BIJ UITBESTEDING

KERNBEGRIPPEN

Uitbesteding van productie verwijst naar de mogelijkheid voor een organisatie om het digitaliseringsproces geheel of gedeeltelijk toe te vertrouwen aan een externe dienstverlener. Uitbesteding kan worden gebruikt om een gebrek aan technische middelen op te vangen of om een grote stroom van te digitaliseren documenten te beheren. Het kan ook een strategie zijn om de kosten van een bepaalde operatie te beheersen of om meer deskundigheid te verwerven. Uitbesteding vereist echter extra stappen om de controle te bewaren over de digitaliseringsketen en de kwaliteit van de te leveren prestaties te beschermen.

De omvang en de aard van de taken die aan de dienstverlener worden toevertrouwd, worden al van bij het opstellen van het bestek gedefinieerd. De controlemethoden moeten ook nauwkeurig worden omschreven.

AANBEVELINGEN

- ➔ **Het actiegebied van alle werknemers bepalen:** soort taken, verloop van de prestaties, deadlines.
- ➔ **De algemene uitvoeringsvoorwaarden bepalen** (te volgen normen, methodologische normen enzovoort).
- ➔ **Bepalen hoe originele documenten moeten worden vervoerd en behandeld.**

- ➔ **De manier van opvolging tijdens de productie bepalen**, met inbegrip van het soort rapporten en de frequentie ervan.
- ➔ **Bepalen welke stappen moeten worden ondernomen in geval van anomalieën tijdens de productie.**
- ➔ **De dienstverlener instructies verstrekken in verband met het naamgevingsschema, de metadatanorm en de indexering die hij moet toepassen.**
- ➔ **De verantwoordelijkheid delen voor de kwaliteitscontrole van de deliverables.**
- ➔ **Eenvoudige, betrouwbare en relevante controle-indicatoren bepalen.**
- ➔ **De toegestane foutenmarge bepalen.**

Voorbeeld: De Nationale Bibliotheek van Frankrijk (BnF)

Sinds haar eerste grootschalige digitaliseringsprojecten koos het BnF ervoor een deel van haar digitaliseringsprojecten uit te besteden en dus haar kwaliteitscontrolebeleid te herzien met het oog op een vlotte en transparante samenwerking met de externe dienstverleners.

De gekozen oplossing bestond erin **een kwaliteitsaanpak te ontwikkelen gebaseerd op fluïditeit**, met de volgende kernbeginselen:

Betrokkenheid van de dienstverlener vanaf het moment dat het bestek wordt opgesteld, met name om overeenstemming te bereiken over de werkmethoden, de kwaliteit van het digitale document, de processen van de digitaliseringsketen, de indicatoren die moeten worden gecontroleerd enzovoort.

Een permanente dialoog verzekeren, met name in de vorm van door de dienstverlener opgestelde wekelijkse rapporten waarin het aantal documenten en pagina's wordt vermeld dat elke werkplaats binnenkomt en verlaat.

Deze rapporten maken het mogelijk om problematische fasen te identificeren, de oorzaken ervan te analyseren en de digitaliseringsketen tijdens de productie bij te sturen.

Gedeelde verantwoordelijkheid voor kwaliteitscontrole.

Voor meer details zie:

BELLIER, L., « Numérisation, pour une nouvelle approche de la qualité », *In: Bulletin des bibliothèques de France*, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB), 2014.

<https://bbf.enssib.fr/contributions/numerisation-pour-une-nouvelle-approche-de-la-qualite>

06 / DIGITALISERINGS- APPARATUUR KIEZEN

AANBEVELINGEN

Het aanbod van scanners voor de digitalisering van erfgoed en documenten is nog nooit zo groot geweest. Door de wedijver tussen fabrikanten op het gebied van technische innovaties is het soms moeilijk om de fundamentele selectiecriteria voor digitaliseringsapparatuur vast te stellen.

Bij het kiezen van de juiste scanapparatuur moet men in de eerste plaats kijken naar:

- ➔ Het soort documenten dat moet worden gescand.
- ➔ Hun formaten.
- ➔ De hoeveelheid documenten en hun kostbaarheid.
- ➔ Hun staat en hoe ze tijdens het proces kunnen worden behandeld.
- ➔ Het beschikbare budget.
- ➔ De kwalificatiegraad van de teams.
- ➔ De mate van complexiteit bij het gebruik van de apparatuur en de vraag of de teams zijn opgeleid om de apparatuur te gebruiken.
- ➔ De ruimte om de machines op te slaan.
- ➔ De omgevingsomstandigheden in de scanruimte.

Zodra deze elementen zijn vastgesteld, wordt de keuze van de scanapparatuur gebaseerd op zuiver technische kenmerken:

- ➔ **Resolutie**, of de nauwkeurigheid van de beeldopname.
- ➔ **De diepte van de analyse**, of het aantal kleuren en grijstinten dat het toestel kan produceren.
- ➔ **De interface** en de aanwezigheid van USB-poorten.
- ➔ **Scansnelheid**.

Geschikte en aanbevolen toestellen voor het digitaliseren van erfgoed en documenten zijn:

- ➔ **Camera's**.
- ➔ **Drumscanners**.
- ➔ **Reprostatieven**.
- ➔ **Gespecialiseerde scanners voor het scannen van gebonden materialen**.

Opgelet

- ➔ Na langdurig gebruik, of door een fout in de behandeling, **kunnen systeemelementen achteruitgaan of buiten werking raken** zonder dat dit per se met het blote oog waarneembaar is.
- ➔ Daarom is **het belangrijk regelmatig kwaliteitscontroles van de geproduceerde beelden uit te voeren**. De frequentie van dit soort controles hangt af van de intensiteit waarmee de machines worden gebruikt en van de aanbevelingen van de fabrikant. Deze controle maakt het ook mogelijk de geproduceerde beelden tussen twee tests te valideren en draagt bij aan de optimalisatie van een eventuele printopdracht.

07 / LICHTBRONNEN KIEZEN

KERNBEGRIPPEN

Zoals gezegd tast natuurlijk licht de kwaliteit van erfgoedobjecten aan. De keuze van de lichtbronnen en hun sterkte en plaatsing is dus van essentieel belang om de kwaliteit van het digitaliseringsproces te optimaliseren.

De kwaliteit van verlichting wordt beoordeeld door het vermogen om het hele kleurenspectrum waarneembaar te maken, onderscheidend vermogen genoemd. Dit vermogen wordt gemeten in Colour Rendering Index (CRI) met behulp van een thermocolorimeter speciaal voor digitale beelden. De maximale CRI-waarde is 100 en komt overeen met ideaal daglicht. De minimumwaarde is 0 en komt overeen met een licht waarbij het onmogelijk is kleuren te onderscheiden. **De minimale CRI-waarde voor een goede kleurweergave is 95.**

Om deze drempelwaarde te bereiken en een gelijkmatige verlichting te garanderen, **moeten de lichtbronnen van dezelfde kwaliteit en sterkte zijn en symmetrisch aan weerszijden van het beeldvlak geplaatst zijn, onder een hoek van 30-45°** naar het midden van het beeldvlak. Daarbij moet hun hoogte minimaal gelijk zijn aan die van de lens, die is uitgerust met een zonnekap, met name om flare te voorkomen.

AANBEVELINGEN

- ➔ **De elektronische flitser** is de beste oplossing om een object correct weer te geven. Met een CRI van ongeveer 98% heeft flitslicht een lichtopbrengst die de contrastwaarden verbetert.
- ➔ **Metaalhalidelampen** – HMI, MSR, GEMI enzovoort – zijn een interessant alternatief. Zij waren oorspronkelijk bestemd voor film en televisie en produceren een stabiel licht met een breed en compleet spectrum. Gekalibreerd tussen 5000 en 5500 Kelvin, is hun CRI ongeveer 95.

- ➔ De soorten verlichting die niet worden aanbevolen voor scannen zijn fluorescentielampen - waarvan de CRI te laag is - en leds - die geen optimale kleurweergave bieden.
- ➔ De twee gebruikte lichtbronnen moeten aan weerszijden van de scanapparatuur worden geplaatst. In het geval van een groot document opteert men beter voor vier lichtbronnen.
- ➔ Het licht moet **gedempt en constant** zijn. Aanbevolen maximale helderheid: **32 lux**.
- ➔ Het scherm moet worden beschermd tegen **daglicht** of directe reflecties van de lichtbronnen om te voorkomen dat licht weerkaatst wordt door de scanapparatuur.

08 / DE DIGITALISERING SKETEN KALIBREREN

KERNBEGRIPPEN

Kalibreren bestaat uit het instellen van de parameters van alle scanapparatuur en randapparatuur volgens de technische specificaties die in het bestek staan vermeld. De volgende parameters kunnen worden ingesteld:

- ➔ Resolutie.
- ➔ Witbalans.
- ➔ Belichting.
- ➔ Scherpstelling.
- ➔ Contrast.
- ➔ Kleuren en kleurweergave.

Consistent en nauwkeurig kleurbeheer vereist **het gebruik van betrouwbare ICC-profielen voor alle kleurapparatuur**. Zonder controle van deze profielen kan een gescande afbeelding kleurverschillen vertonen van apparaat tot apparaat, door een eenvoudig verschil tussen de scanner en de grafische software. Met een betrouwbaar profiel kan het programma dat de afbeelding importeert verschillen tussen apparaten corrigeren en zo de ware kleuren van een gescande afbeelding weergeven. Het is daarom van essentieel belang – ongeacht het gekozen kleurprofiel – **dat deze profielen tussen de verschillende apparaten op elkaar worden afgestemd en op dezelfde wijze worden gekalibreerd**.

Opgelet



Elke scanner heeft zijn zwakke punten wat betreft kleurweergave. Daarom is het van essentieel belang de scanapparatuur regelmatig te kalibreren door een document te scannen en de kleuren van het digitale bestand en het origineel te vergelijken om de betrouwbaarheid van de weergave te waarborgen. Een kleurenkaart – of kleurentestkaart – die naast het gescande object wordt geplaatst, geeft een beter inzicht in de kleurweergave.

TOOLS

Normen en benchmarks

- **Standard ISO 12641-1 :2016 - Grafische technologie**

Uitwisseling digitale preprintgegevens - Kleurdoelen voor kalibratie bij scannerinvoer

- **Standard ISO 7589 :2002 - Fotografie**

Lichtbronnen voor sensitometrie - Specificaties voor daglicht, gloeilampen en printers

- **Standard ISO 14524 :2009 - Fotografie**

Elektronische fotocamera's - Methoden voor het meten van opto-elektronische conversiefuncties (OECF's)

- **Standard ISO 21550 : 2004 - Fotografie**

Elektronische scanners voor fotografische beelden - Dynamisch-bereikmetingen

- **Metamorfoze gids**

http://www.imagingetc.com/images/Resources_Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

- **FADGI gids**

http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf

Kleurruimten en ICC-profielen

- **Adobe RGB 1998**

Aanbevolen kleurruimte voor kleurencannen bestaande uit drie niveaus (Rood - Groen - Blauw).

- **Adobe sRGB**

Alternatieve kleurruimte voor kleurencannen bestaande uit drie niveaus (Rood - Groen - Blauw).

- **Gray Gamma 2.2**

Aanbevolen ICC-profiel voor grijs tinten. Wordt gebruikt om de wit- en grijsbalans in te stellen.

- **Kodac Gray Scale**

Alternatief ICC-profiel voor grijs tinten. Wordt gebruikt om de wit- en grijsbalans in te stellen.

- **Digital Color Checker**

Aanbevolen ICC-profiel voor kleurencans.

Evaluatietools

- **Open DICE**

Gratis software voor het meten en analyseren van de technische criteria van scanners. Deze software voor automatische controle maakt gebruik van verschillende ISO-normen om de kwaliteit van de door scanners geproduceerde beelden en de technische onderdelen van de scanapparatuur te analyseren.

<http://www.digitizationguidelines.gov/guidelines/digitize-OpenDice.html>

- **Auto SFR**

Gratis programma dat is ontwikkeld om beeldvormingsprofessionals de werkelijke resolutie van afbeeldingen te helpen bepalen en de juiste resolutie in te stellen voor de te scannen documenten, afhankelijk van het type document.

http://www.digitizationguidelines.gov/guidelines/OpenDICE/AutoSFR_manual.pdf

● UTT

Een methode om alle parameters van een opnamesysteem te controleren, ontwikkeld door de Nederlandse Koninklijke Bibliotheek. Deze gestandaardiseerde kaart – beschikbaar in formaten gaande van DIN A4 tot DIN A0 – kan worden gebruikt om parameters van scanapparatuur te valideren zoals: resolutie, contrast, witbalans, gain modulation, uniformiteit van het licht op het object, ruis, kleurweergave, geometrische vervorming en parallelisme.

<http://universaltesttarget.com/about.php>

AANBEVELINGEN

- ➔ De afstelling van de lichtbronnen moet vóór de kalibratie worden uitgevoerd, om de kalibratie niet te verstoren. Maximaal toegestane helderheid: 32 lux.
- ➔ Aanbevolen resolutie:
 - 300 ppi voor documenten tussen DIN 5 en DIN A2.
 - 400 ppi voor andere formaten.
- ➔ Voorkeur geven aan **Adobe RGB 1998** voor kleurencans. Voor afbeeldingen in grijstinten geniet **Gray Gamma 2.2** de voorkeur.
- ➔ Om de kwaliteit van de reproductie van de tonen en tinten van het originele document te garanderen, is het **noodzakelijk de referentiekleurenkaart onder dezelfde voorwaarden te scannen als de gereproduceerde documenten.**
- ➔ Deze **testkaart moet elke dag worden gescand** om ervoor te zorgen dat de machine-instellingen conform blijven.

09 / ALGEMENE REGELS VOOR DE BEELDVORMING

AANBEVELINGEN

- ➔ De documenten moeten volledig worden gescand: **het beeld moet het volledige document weergeven en alle bladzijden omvatten** in het geval van een boek.
- ➔ Om zeker te zijn dat het document volledig is, **laat men het beste een marge van 0,25 tot 1 cm rondom het document**. Hoe groter het document, hoe groter de te nemen marge.
- ➔ Indien de omtrek van het te reproduceren document onregelmatig is, laat dan een kader open dat breed genoeg is om het gehele document aan alle zijden te omsluiten.
- ➔ Het wordt aanbevolen om documenten in **kleur** te scannen.
- ➔ **Documenten moeten zoveel mogelijk platgelegd worden**, afhankelijk van de aard van het document, de staat waarin het verkeert en de kostbaarheid ervan.
- ➔ Het platdrukken **moet gelijkmatig gebeuren** om vervorming van het document te voorkomen.
- ➔ De oriëntatie van het document moet overeenkomen **met de leesrichting**. Indien het formaat van het document dit niet toelaat, zal het tijdens de postproductie in deze richting moeten worden georiënteerd.
- ➔ Voor documenten met gedrukte **tekst moet de tekstregel horizontaal lopen, met een toegestane afwijking van 1,30°**.

- ➔ Bij het rechtzetten richt men zich **het beste op de tekst eerder dan op de pagina**, met het oog op de OCR-fase.
- ➔ **Gebruik een neutrale scanachtergrond – grijs, zwart of wit** – om het contrast en dus de leesbaarheid van het document te vergroten.
- ➔ Gebruik voor ingebonden documenten **met tussenruimten een mat kartonnen afdekblad in een neutrale kleur** om de leesbaarheid te vergroten en de OCR te vergemakkelijken.
- ➔ Om fouten te corrigeren die tijdens de procedure opduiken, **is het aanbevolen het RAW-bestand** te bewaren zodat het opnieuw kan worden gedematriceerd en de anomalie kan worden gecorrigeerd.

10 / AANMAKEN EN INVULLEN VAN HET METADATA-REFERENTIEBESTAND

KERNBEGRIPPEN

Het referentiebestand is een dossier met de beschrijvende metadata van het gedigitaliseerde document, zoals auteur, volume, afmetingen, conserveringskenmerken en, meer in het algemeen, alle essentiële elementen die de scanmedewerkers moeten kennen met het oog op de verspreiding en de catalogisering van het document.

De definitie van deze categorieën hangt opnieuw af van het beoogde gebruik van de gedigitaliseerde documenten, maar ook van de vraag of al dan niet een gestandaardiseerd metadatamodel wordt gebruikt – waarvan de categorieën reeds zijn gedefinieerd.

In geval van uitbesteding moet het referentiebestand vooraf worden opgesteld en aan de dienstverlener worden overgemaakt. Zodra dit bestand is ingevuld en ingediend, moet de integriteit ervan worden gecontroleerd om de juistheid van de informatie te garanderen.

AANBEVELINGEN

Het referentiebestand moet ten minste de volgende gegevens bevatten:

- ➔ De **beschrijvende categorieën** van de gedigitaliseerde objecten (unieke identifier, auteur, titel, aanmaakdatum, onderwerp, afmetingen, beschrijving enz.)
- ➔ De **definitie** van deze categorieën.

- ➔ De **mate van verplichting** van de **metadatacategorieën** (verplicht/optieel/aanbevolen).
- ➔ Het **soort gegevens** – hoeveelheid, datum, gecontroleerde terminologie, vrije tekst enz.
- ➔ De **invoernormen voor vrije tekst** en het referentiedocument voor de metadatastandaard die door de instelling wordt gevolgd.
- ➔ De **naam van de invoerder** en de aangebrachte wijzigingen om deze te kunnen traceren en mogelijke betrouwbaarheidsproblemen te kunnen opsporen.

TOOL

- **Open Refine**

Open Refine is een opensourceprogramma voor het opschonen en corrigeren van een set gegevens, zodat ze consistenter zijn. Het dient voor:

- *Importeren van gegevens in verschillende formaten*
- *Toepassen van celtransformatie*
- *Verwerken van cellen die meerdere waarden bevatten*
- *Maken van koppelingen tussen gegevensverzamelingen*

<https://openrefine.org/documentation.html>



11 / OPSLAAN EN NAAMGEVING VAN BESTANDEN

KERNBEGRIPPEN

Bij elk digitaliseringsproject rijst de vraag naar de naamgeving van de bestanden die zullen worden geproduceerd. Deze kwestie sluit aan bij het semantische web, waarvan een van de basisprincipes het gebruik van persistente identifiers is.

Van een identifier wordt gezegd dat hij persistent is wanneer hij een bron en alle bijbehorende metadata op unieke wijze aanduidt, waardoor deze bron gemakkelijker te vinden is op het internet. Deze identifier – zoals bijvoorbeeld de URI, ARK of DOI – garandeert ook het hergebruik van digitale bronnen en hun zichtbaarheid doordat hij toelaat ze aan andere bronnen te koppelen en zo de lees- of toegangspaden te vermenigvuldigen.

Gezien het grote aantal documenten dat in een scanproject moet worden beheerd, is het niet haalbaar de bestanden afzonderlijk te benoemen. Het gebruik van halfautomatische programma's maakt een massaal beheer van bestandsidentificatoren mogelijk, ongeacht hun type, gewicht of locatie.

AANBEVELINGEN

- ➔ Ongeacht het gekozen beheerprogramma is het essentieel een conventie voor de naamgeving te definiëren om het digitaliseringsproces te vergemakkelijken en het risico op informatieverlies te verminderen.

➔ Er is geen perfecte manier om de elementen in een bestandsnaam te structureren en te ordenen. Wel is het belangrijk om consequent te zijn. Hieronder volgen enkele van de meest voorkomende categorieën om documenten in culturele en wetenschappelijke instellingen te identificeren:

- Alfabetieke identificatiecode.
- Type document.
- Datum van aanmaak van het document (datumformaat JJJJMMDD).
- Onderwerp van het document.
- Ontvanger in geval van een aanvraag tot reproductie.
- Versie.

➔ **Vormregel die moeten worden nageleefd:**

- Scheiding van woorden in een veld met hoofdletters of onderstrepingstekens.

➔ **Verboden tekens:**

- Tekens met accenten.
- Leestekens.
- Punt.
- Spatie.

➔ **Het is aanbevolen het in de instelling geldende naamgevingsschema te volgen of, indien dit nog niet bestaat, een schema op te stellen dat specifiek is voor digitaliseringsoperaties.**

➔ **Voor unieke identifiers zorgen.**

➔ **In geval van verwijdering of niet-publicatie informatie geven over de omstandigheden en indien nodig naar een vervangende bron verwijzen.**

➔ **Het ID-beheersbeleid van de instelling bekendmaken**

TOOLS

● **File Renamer Basic**

File Renamer Basic is een programma voor het hernoemen van bestanden voor Windows. Het kan een naamgevingsformaat toepassen op alle bestanden in een map en kan de technische metadata van beeldbestanden (EXIF) batchgewijs bewerken.

https://download.cnet.com/File-Renamer-Basic/3000-2248_4-10306538.html

- **Bulk Rename Utility**

Bulk Rename Utility is ook een programma voor het hernoemen van bestanden voor Windows. Het biedt dezelfde functionaliteit als File Renamer Basic, maar is speciaal ontworpen voor het verwerken van mappen met enkele duizenden bestanden.

<https://www.bulkrenameutility.co.uk/>

- **Name Changer**

Name Changer is een uitstekend alternatief voor Mac.

<https://www.macupdate.com/app/mac/21516/namechanger>

12 / RETOUCHEREN VAN BEELDBESTANDEN

KERNBEGRIPPEN

Ondanks alle voorzorgsmaatregelen die tijdens de opname zijn genomen, is het niet ongevoel dat na de productie retouches moeten worden uitgevoerd om het geproduceerde beeld te perfectioneren of kleine anomalieën te corrigeren. Niet alle hieronder vermelde handelingen zijn noodzakelijk. Het is aan de met de controle belaste persoon om – op basis van het bestek of een referentiedocument – te bepalen welke elementen moeten worden gecorrigeerd.

Dit referentiedocument moet **alle parameters bevatten** die in die fase moeten worden gecontroleerd, en de resultaten die voor elk soort document moeten worden bereikt. Deze informatie wordt aan het proces gekoppeld en geïntegreerd in de metadata bij de beelden, waardoor het mogelijk wordt de historiek ervan te reconstrueren.

Het retoucheren van beeldbestanden omvat:

- ➔ Draaien van het beeld om het in de leesrichting te plaatsen. Dit is bijzonder belangrijk in het geval van een tekstdocument waarop OCR wordt toegepast. De tekstregel moet horizontaal lopen, ook al is de band of het document niet helemaal horizontaal.
- ➔ Clipping, dat bestaat uit het elimineren van pixels buiten het gebied van het gescande object.
- ➔ Verzagingscorrectie.
- ➔ Ruisonderdrukking.
- ➔ Verscherpen of verzachten van het beeld.
- ➔ Contrastcorrectie.
- ➔ Gammacorrectie of wijziging van de luminantiewaarden van de pixels.

AANBEVELINGEN

- ➔ Vóór het retoucheren nagaan of het ICC-profiel van de beeldverwerkingssoftware hetzelfde is als dat van de scanner.
- ➔ Met een kopie van het gescande bestand werken: het is niet gemakkelijk om alle bewerkingen ongedaan te maken.

TOOLS

- **Photoshop**

Photoshop is wellicht de bekendste software voor het bewerken van beelden. Deze software, ontwikkeld door Adobe, heeft veel functies en is ongetwijfeld de meest complete tool om een gerasterd beeld te bewerken.

https://www.adobe.com/be_nl/products/photoshop.html

- **Lightroom**

Lightroom is een interessant alternatief omdat het speciaal is ontwikkeld voor het bewerken van foto's. Het biedt ook een collaboratieve bewerkingsmodus, zodat meerdere mensen aan dezelfde afbeelding kunnen werken indien nodig.

https://www.adobe.com/be_nl/products/photoshop-lightroom.html

- **Gimp**

GIMP is een platformonafhankelijk beeldbewerkingsprogramma dat beschikbaar is voor GNU/Linux, OSX, Windows en andere besturingssystemen. Het is minder geavanceerd dan de twee bovengenoemde programma's, maar heeft het voordeel dat het gratis is en compatibel is met alle platforms.

<https://www.gimp.org/>

13 / OCR VAN TEKSTDOCUMENTEN

KERNBEGRIPPEN

OCR (*Optical Character Recognition*) of optische tekenherkenning is een beeldverwerkingstechniek waarmee tekst wordt geëxtraheerd om ze leesbaar en doorzoekbaar te maken.

Het OCR-proces omvat verschillende fasen, waaronder:

- ➔ **Het rechtzetten** van de pagina.
- ➔ **Het binariseren van het document**, d.w.z. het omzetten in zwart-wit.
- ➔ **Het segmenteren** van het beeld in tekstgebieden en vervolgens in lijnen, woorden en tekens.
- ➔ En ten slotte **het classificeren** van deze gebieden door het herkennen van tekens en vervolgens woorden.

AANBEVELINGEN

- ➔ Om de kwaliteit van de scan te controleren, is het aanbevolen gebruik te maken van een referentietekst, de zogenaamde grondwaarheid.
- ➔ Om een grondwaarheid te verkrijgen, moet een 'handmatige' scan van een beelddocument worden uitgevoerd:
 - Binariseren van het beeld.
 - Automatisch verwijderen van ruis en handmatig corrigeren van eventuele fouten of onnauwkeurigheden.

- Verwijderen van randen.
- Handmatig of semiautomatisch aanmaken van tekstzones.
- Classificeren van de zones volgens type (afbeelding, tekst, handgeschreven tekst, tekening enz.).
- Handmatig bepalen van de leesvolgorde tussen de tekstblokken in het document.
- Definiëren van de tekstregels en handmatig invoeren van een representatieve zin, d.w.z. het meest heterogene geheel van normale en speciale tekens.

➔ Op basis van dit referentiedocument wordt vervolgens aan de hand van feitelijke criteria beoordeeld in hoeverre de software in staat is tekstzones, regels en tekens te herkennen en met name:

- Weglatingen.
- Classificatieproblemen.
- Segmentatieproblemen, d.w.z. het opdelen van het hele beeld in subdelen – regels, woorden, tekens.

➔ De OCR-controle kan worden geautomatiseerd met behulp van het LayoutEval-softwarepakket, dat kwaliteitsverschillen detecteert op basis van het geladen referentiebestand.

TOOLS

- **ABBY FineReader**

ABBY FineReader – Eigen software

https://www.abbyy.com/media/6652/guide_english.pdf

- **Tesseract**

Tesseract – Open source software

<https://ichi.pro/fr/guide-du-debutant-sur-tesseract-ocr-18608940018749>

- **Layout Evaluation**

Layout Evaluation – Eigen software

<https://www.primaresearch.org/tools/PerformanceEvaluation>

14 / INKAPSELEN VAN METADATA

KERNBEGRIPPEN

Om bruikbaar te zijn en verspreid te kunnen worden, moet elk digitaal bestand een reeks metadata bevatten die de digitale inhoud zo nauwkeurig mogelijk beschrijven. Deze metadata kunnen extern of intern zijn, maar moeten in de postproductiefase binnen hetzelfde document worden verzameld, met het oog op een optimale bewaring en exploitatie.

- ➔ Beschrijvende metadata omvatten alle gegevens die de digitale inhoud beschrijven – auteur, aanmaakdatum van het document, onderwerp enzovoort – en worden vastgelegd in een apart document.
- ➔ Technische en administratieve metadata geven alle informatie weer met betrekking tot het digitaliseringsproces van het object: datum van vastlegging, model van de gebruikte scanner of camera, kleuruimte van het bestand, ICC-profiel, resolutie, grootte in pixels, coderingsdiepte enzovoort. Deze metadata zijn doorgaans intern en worden automatisch gegenereerd door de scanapparatuur voor het vastleggen van beelden.

AANBEVELINGEN

De METS-standaard (Metadata Encoding and Transmission Standard), ontwikkeld door de Digital Library Federation, laat toe om alle informatie met betrekking tot de interne structuur van documenten gedetailleerd weer te geven. Dit document, dat gebaseerd is op het XML-schema, maakt het mogelijk alle elementen met betrekking tot de beschrijving van het object en de voorwaarden voor het creëren en exploiteren van de gedigitaliseerde inhoud te groeperen en aan elkaar te koppelen. Het METS-bestand moet daarom worden gezien als een verzameling metadata die naar verschillende gelinkte bestanden verwijst. Het heeft het voordeel dat het compatibel is met andere metadatastandaarden zoals MARC, EAD of MODS.

Een METS-bestand omvat zeven categorieën:

- 1** De kopregel, die het METS-document zelf beschrijft.
- 2** Beschrijvende metadata of het pad naar het bestand dat deze informatie bevat.
- 3** Administratieve metadata, die informatie verschaffen over de wijze waarop de bestanden zijn gecreëerd en opgeslagen en over de intellectuele-eigendomsrechten van het oorspronkelijke object.
- 4** De sectie bestanden, die informatie geeft over de elektronische versies van het digitale object.
- 5** De structuurkaart, die, zoals de naam aangeeft, een hiërarchische structuur tekent voor het digitale object, en elk element van die structuur koppelt aan inhoudsbestanden en metadata.
- 6** Structurele links, die de bestaande hyperlinks beschrijven tussen elementen binnen de structuurkaart.
- 7** En tot slot de gedragingen, die de uitvoeringsmodaliteiten bepalen van de verschillende bestanden die aan het METS-document zijn gekoppeld.

TOOLS

- **Algemene documentatie en gebruikersgids voor de METS-standaard**
http://www.loc.gov/standards/mets/METSOverview.v3_en.html
- **Huidig schema van de METS-standaard**
<http://www.loc.gov/standards/mets/mets.xsd>
- **Programma's voor het maken en beheren van METS-bestanden**
<http://www.loc.gov/standards/mets/mets-tools.html>
- **Complete gids voor het aanmaken en bewerken van METS-bestanden, ontwikkeld door de Bibliothèque nationale de France**
https://www.bnf.fr/sites/default/files/2021-01/ref_num_metadonnees_mets.pdf

15 / DE LAATSTE KWALITEITSCONTROLE

AANBEVELINGEN

In dit stadium wordt de kwaliteitscontrole niet alleen toegepast op de digitale beelden, maar ook op alle producten van de dienst:

- De conformiteit van dragers en bestandsformaten.
 - De naamgeving van dragers, mappen, bestanden en beelden.
 - De volledigheid van de beeldbestanden in verhouding tot het aantal te reproduceren originelen en/of de bijbehorende indexering.
 - De integriteit van de gevraagde indexering op basis van de geformuleerde inputregels.
 - De geschiktheid van de technische kenmerken van de beeldbestanden voor de verwachte doelstellingen, die objectief moeten kunnen worden geëvalueerd.
- ➔ Een volledige controle van de verkregen resultaten is alleen mogelijk wanneer het aantal te scannen documenten beperkt is. **In de meeste gevallen wordt geopteerd voor steekproeven.** De norm NF ISO 2859-1 voorziet in een methodologie voor dit soort controle.
- ➔ Bemonstering vindt altijd plaats **aan de hand van de minimale kwaliteitsdrempel**, d.w.z. een maximumpercentage aan gebrekkige objecten in een partij om die partij als bevredigend te kunnen beschouwen.
- ➔ De verificatie van de indicatoren gebeurt in verhouding tot **de fouten die als ernstig of licht worden beschouwd.**



16 / EEN DIGITAAL DATABEHEERSYSTEEM KIEZEN

KERNBEGRIPPEN

Langetermijnbewaring van elektronische inhoud is de reeks acties, instrumenten en methoden die worden toegepast om elektronische inhoud te verzamelen, te identificeren, te selecteren, te rubriceren, te vernietigen en te bewaren op een veilig medium, met de bedoeling deze inhoud te exploiteren en na verloop van tijd toegankelijk te maken. Bij elektronische archivering gaat het niet alleen om het opslaan en delen van informatie, maar ook om het waarborgen van de integriteit en vertrouwelijkheid van gedigitaliseerde inhoud.

De keuze van het systeem voor het beheer en/of de bewaring van de gedigitaliseerde documenten is dus van fundamenteel belang. In de eerste plaats moet het meest geschikte type systeem worden gekozen naargelang van de documenten die worden gearhiveerd:

- Is het systeem bedoeld om levende en bewerkbare documenten te beheren, dan kiest men het beste voor een elektronisch documentbeheersysteem (EDM).
- Als de documenten daarentegen vastliggen en gevalideerd zijn, is het beter een beroep te doen op een elektronisch archiveringssysteem (EAS).

De elektronische kluis is een aanvullend element op het EAS. In het Engels wordt gesproken van een « Digital Safe-Deposit Box Component », een term die wordt gebruikt in de norm NF Z42-020. De twee begrippen mogen niet worden verward: de digitale kluiscomponent verzekert de integriteit en traceerbaarheid van de digitale objecten die erin worden gedeponeerd, terwijl het digitale documentbeheersysteem specifieke archiveringsfuncties biedt: bestendiging van formaten, documenttypologie, beheer van de levenscyclus, archiveringsprofielen enzovoort.

EDMS

ELECTRONIC DOCUMENT MANAGEMENT SYSTEM

DOELSTELLINGEN:

- Centraliseren van alle documenten in één punt, ongeacht hun oorsprong en formaat.
- Harmoniseren van processen om het documentbeheer binnen een instelling te vergemakkelijken.
- Dagelijks beheer van documenten binnen een instelling.

FUNCTIES:

- Maakt wijziging en/of vernietiging van documenten mogelijk.
- Maakt het beheer van bewaartermijnen mogelijk.
- Kan een georganiseerde opslagstructuur omvatten, beheerd door de gebruikers.
- Scannen van documenten (indien op papier), vastleggen, indexeren, archiveren, opslaan, toegankelijk maken en verspreiden.

EAS

ELECTRONIC ARCHIVING SYSTEM

DOELSTELLINGEN:

- Zorgen voor toegang, vertrouwelijkheid, integriteit, duurzaamheid en traceerbaarheid van elk document.
- Garanderen van archivering met bewijskracht, d.w.z. archivering die de bewaring van een document op lange termijn garandeert, met inachtneming van de wettelijke voorschriften en de archiveringsnormen.

FUNCTIES:

- Voorkomt wijziging en/of vernietiging van documenten.
- Verplichte controle van de bewaartermijnen.
- Strikte archiveringsstructuur.

AANBEVELINGEN

- ➔ Bij het implementeren van een elektronisch archiveringssysteem is het aanbevolen om een archiveringsbeleid op te zetten dat de wettelijke, operationele, technische en functionele beperkingen definieert, evenals de toegangsrechten. Hierdoor wordt de beveiligingsfunctie van het EAS nog versterkt.
- ➔ EAS en EDM zijn complementaire oplossingen bij het beheer van erfgoedarchieven: EDM vergemakkelijkt bijvoorbeeld het beheer van actuele documenten en de samenwerking in het kader van interne digitaliseringsprojecten, terwijl een EAS permanente conserveringsoplossingen biedt voor gedigitaliseerde inhoud.

- ➔ Opteren voor een EAS dat voldoet aan de norm NF Z42-013. Het EAS garandeert de authenticiteit en conformiteit van het document via een exhaustief bewijsdossier (traceerbaarheid van de ondernomen acties, integriteitsbewijzen, bestaan van het document op een bepaalde datum).

TOOLS

- **Norm AFNOR NF Z 42-013**

De norm NF Z42-013, die in 1999 door AFNOR werd gepubliceerd en vervolgens in 2001, 2009 en recentelijk in 2020 is herzien, bevat een reeks functionele, organisatorische en infrastructurele eisen en aanbevelingen voor het ontwerp en de werking van een elektronisch archiefsysteem (EAS).

De tekst behandelt het organiseren van de functies van elektronische archivering, overdracht, bewaring, toegankelijkheid, communicatie en afvoerbeleid, interoperabiliteit van het EAS, restitutie en omkeerbaarheid.

- **Norm AFNOR NF Z 42-020**

De norm Z42-020, die in 2012 door AFNOR werd gepubliceerd, heet 'Functionele specificaties van een digitale kluiscomponent bedoeld voor het behoud van digitale informatie in condities die de integriteit ervan op lange termijn garanderen'. Deze norm richt zich enkel op de opslaglaag – de digitale kluis – en gaat niet in op kwesties als authenticatie, vertrouwelijkheid, metadata of interfacing met andere specifieke lagen.

- **Norm ISO 14641-1**

De internationale norm ISO 14641-1 is volledig gebaseerd op de norm NF Z42-013 en bevat ook de specificaties voor de implementatie van een elektronisch archiveringssysteem. De ontwikkeling ervan staat echter los van de AFNOR-norm.

- **Norm ISO 14721**

ISO 14721, oorspronkelijk gepubliceerd in 2003 en herzien in 2012, is een omzetting van het OAIS (Open Archive Information System), een conceptueel model voor het beheer, de archivering en de langetermijnbewaring van digitale documenten.

- **Norm ISO 15489-1**

Deze norm, die in 2016 is herzien, beschrijft de concepten en beginselen voor het beheer van documenten met bewijskracht van natuurlijke of rechtspersonen, privaat of publiek, en definieert de kenmerken die een document zijn gezaghebbende waarde of bewijskracht geven: authenticiteit, integriteit, betrouwbaarheid en bruikbaarheid.

ANCTIL, M.-C, LEGENDRE, M., MÜLLER, T., MAILLET, D., BROUSSEAU, K., RENAUD, L., *Recueil de règles de numérisation*, Bibliothèque et Archives nationales du Québec, Bibliothèque nationale de France et Musée canadien de l'histoire, 2014, [Online],

<http://collections.banq.qc.ca/bitstream/52327/2426216/1/4671601.pdf>

ACCART, J.-P., *Le Métier de Documentaliste*. Avec la collaboration de Réthy Marie-Pierre. Éditions du Cercle de la Librairie, 2015.

BACHI, V., FRESA, A., PIEROTTI, C., PRANDONI, C., « The digitization age: mass culture is quality culture. challenges for cultural heritage and society », In: Ioannides, M., Magnenat-Thalmann, N., Fink, E., Žarnić, R., Yen, A.-Y., Quak, E. (eds.) EuroMed 2014.

BAROUCH, G., *Révolution des pratiques managériales : le changement par la qualité*, La Plaine Saint-Denis : Afnor Éditions, 2013.

BELLIER, L., « Numérisation, pour une nouvelle approche de la qualité », In: *Bulletin des bibliothèques de France*, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB), 2014, [Online],

<https://bbf.enssib.fr/contributions/numerisation-pour-une-nouvelle-approche-de-la-qualite>

BIBLIOTHÈQUE ET ARCHIVES NATIONALES DU QUÉBEC., *La numérisation des documents administratifs : Méthodes et recommandations : version revue et corrigée*, Montréal : Bibliothèque et Archives nationales du Québec, mai 2012, 35 blz., [Online],

<https://docplayer.fr/79033783-La-numerisation-des-documents-administratifs.html>

BIBLIOTHÈQUE NATIONALE DE FRANCE, *Techniques et formats de conversion en mode texte*, BnF - Site institutionnel, ongedateerd, [Online],

<https://www.bnf.fr/fr/techniques-et-formats-de-conversion-en-mode-texte>

BIZINGRE, J., PAUMIER, J., RIVIÈRE, P., *Les référentiels du système d'information*, Parijs : Dunod, 2013.

BOUTET, A. and ROUDAUT, K., « Les enjeux de la numérisation et de l'ouverture d'archives : le point de vue des professionnels » In: *Terminal*, 110 | 2012, blz. 27-37.

BRUCE, T., HILLMANN, D., *The Continuum of Metadata Quality: Defining, Expressing, Exploiting*, 2004, [Online],

https://www.researchgate.net/publication/247818823_The_Continuum_of_Metadata_Quality_Defining_Expressing_Exploiting

BURESI, C., CÉDELLE-JOUBERT, L., *Conduire un projet de numérisation*, Parijs : Tec et Doc, 2002, 326 blz.

CARBONE, F., OOSTERBEEK, L., COSTA, F., « Extending and adapting the concept of quality management for museums and cultural heritage attractions: A comparative study of southern European cultural heritage managers' perceptions », In: *Tour Manage Perspect*, 2020, 10 blz.

CARON, P., DEFIOLE, R., LAY, M.-H., *L'enjeu des métadonnées dans les corpus textuels : Un défi pour les sciences humaines*, Rennes : Presses universitaires de Rennes, 2019, 206 blz.

CHARDONNENS, A., *La gestion des données d'autorité archivistiques dans le cadre du Web de données*, [Ongepubliceerde doctoraalscriptie], Université Libre de Brussel, 2021, 420 blz., [Online],
<https://difusion.ulb.ac.be/vufind/Record/ULB-DIPOT:oai:dipot.ulb.ac.be:2013/315804/Holdings>

CHAUMIER, J., *Document et numérisation. Enjeux techniques, économiques, culturels et sociaux*, Parijs : AdBS éditions, 2006, 119 blz.

CHIRON, G., DOUCET, A., COUSTATY, M., VISANI, M., and MOREUX, J.-P., « Impact of OCR Errors on the Use of Digital Libraries: Towards a Better Access to Information », In: *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 2017, blz. 1-4.

CLAERR, T., WESTEEL, I., and MELOT, M., *Manuel de la numérisation*, Parijs : Éd. du Cercle de la, 2011. 317 blz.

CLAERR, T., WESTEEL, I., *Numériser et mettre en ligne*, Villeurbanne : Presses de l'Enssib, 2010, [Online],
<http://books.openedition.org/pressesenssib/414>.

CLAERR, T., WESTEEL, I., ORY-LAVOLÉE, B., GEORGES, N., *Manuel de constitution de bibliothèques numériques*, Parijs : Éd. du Cercle de la librairie, 2013, 407 blz.

CHEVRY, E., *Stratégies numériques*, Parijs : Hermès Science publications-Lavoisier, 2011, 272 blz.

COMMISSION NATIONALE INFORMATIQUE ET LIBERTÉS, *Comment permettre à l'homme de garder la main? Les enjeux éthiques des algorithmes et de l'intelligence artificielle. Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion éthique confiée par la Loi pour une république numérique*, December 2017, [Online],
https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf

CONSEIL CANADIEN DES ARCHIVES. *Manuel de conservation des documents d'archive, Chapitre 3 – Environnement*, 2003, [Online],
http://www.cdncouncilarchives.ca/RBch3_fr.pdf

DANGERFIELD, M.-C., KALSHOVEN, L., *Report and Recommendations from the Task Force on Metadata Quality*, Mei 2015, 54 blz., [Online],
http://pro.europeana.eu/files/Europeana_Professional/Publications/Metadata%20Quality%20Report.pdf

DIRECTION DES ARCHIVES DE FRANCE. MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION - COMITÉ DE PILOTAGE NUMÉRISATION - MISSION DE LA RECHERCHE ET DE LA TECHNOLOGIE (DDAI), *Écrire un cahier des charges de numérisation du patrimoine. Guide technique. Documents reliés- manuscrits- plans- dessins- photographies- microformes*, Parijs : Archives de France, 2008, 62 blz., [Online],
https://francearchives.fr/file/bf50d8fa5f554586dbf18fdc862d25970a1da0a7/static_4132.pdf

DIRECTION INTERMINISTÉRIELLE DU NUMÉRIQUE ET DU SYSTÈME D'INFORMATION ET DE COMMUNICATION DE L'ÉTAT. *Référentiel général d'interopérabilité*, 2015, versie 2, [Online],
https://www.numerique.gouv.fr/uploads/Referentiel_General_Interoperabilite_V2.pdf

DUFRÈNE, B., IHADJADENE, M., DENIS, B., BRUCKMANN, D., BARBIER, B., et RACINE, B., *Numérisation du patrimoine : quelles médiations ? quels accès ? quelles cultures ?*, Parijs : Hermann, 2013, 311 blz.

D'HOORE, M., *La numérisation du patrimoine écrit : du virtuel à la réalité*, Brussel : Archives et bibliothèques de Belgique, 2011, 280 blz.

EAD, *EAD en bibliothèque, guide des bonnes pratiques, Origine - origination*, 2020, [Online],
<https://www.ead-bibliotheque.fr/guide/donnees-du-did/origination/>

EZRATI, J.-J., *Éclairage d'exposition, musée et autres espaces*, Parijs : Eyrolles, 2015.

FEHAU E., LE DANTEC N., *Vade-mecum de la conservation préventive*, C2RMF, versie van 18 November 2013, [Online],
https://c2rmf.fr/sites/c2rmf.fr/files/vademecum_cc.pdf

FEYLER, F., « De la compatibilité à l'interopérabilité en matière de repérage d'information pertinente. La problématique et l'exemple d'OTAREN », In : *Documentaliste-Sciences de l'Information*, vol. 44, no. 1, 2007, blz. 84-92.

FLECKENSTEIN, M., FELLOWS, L., « Data Quality », In: *Modern Data Strategy*, Cham: Springer, 2018.

FONDS INDÉPENDANT DE PRODUCTION. *Êtes-vous repérables? Guide pratique pour documenter vos contenus*, November 2017, [Online],
<http://ipf.ca/FIP/ressources/etes-vous-reperables/>

FONDS DES MÉDIAS DU CANADA. *Découvrabilité : vers un cadre de référence commun*, 2016, [Online],
<https://trends.cmf-fmc.ca/fr/rapports-de-recherche/decouvrabilite-vers-un-cadre-de-reference-commun>

FORCE 11, *The Fair Data Principles*, ongedateerd, [Online],
<https://www.force11.org/group/fairgroup/fairprinciples>

GIESEN, E., *Démarche qualité et norme ISO 9001 : Une culture managériale appliquée à la recherche*, Marseille : IRD Éditions, 2018, [Online],
<http://books.openedition.org/irdeditions/20609>

GILLET, F., HUNGENAERT, J., HODZA, M., en al., *Identifying Needs to Modernize Access to Digital Data in Libraries and Archives (MADDLAIN) : final report*, Brussel : Belgian scientific Policy, 2018, [Online],
http://www.belspo.be/belspo/brain-be/projects/FinalReports/MADDLAIN_%20final%20report.pdf

GUEGUEN, G., MARQUES DA FONSECA, V.-M., D., PITTI, V. , SIBILLE, C., *Vers un modèle conceptuel international pour la description archivistique. Toward an International Conceptual Model for Archival Description: A Preliminary Report from the International Council on Archives' Experts Group on Archival Description*, 2013, [Online],
https://www.ica.org/sites/default/files/EGAD_English.pdf

INTERNATIONAL CENTRE FOR THE STUDY OF THE PRESERVATION AND RESTORATION OF CULTURAL PROPERTY (ICCROM), *A Guide to Risk Management of Cultural Heritage*, 2016, [Online],
<https://www.iccrom.org/publication/guide-risk-management>

INTERNATIONAL CENTRE FOR THE STUDY OF THE PRESERVATION AND RESTORATION OF CULTURAL PROPERTY (ICCROM), *The ABC Method. A risk management approach to the preservation of cultural heritage*, 2016, [Online],
<https://www.iccrom.org/publication/abc-method-risk-management-approach-preservation-cultural-heritagf>

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Information and Documentation: Part 1: Guidelines*, Geneva: ISO, ISO 15489-1: 2001, 2001.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Information and Documentation: Part 2: Practical Guide*, Geneva: ISO, ISO 15489-2: 2001, 2001.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Photography - Sensitometric Illuminants - Specifications for Daylight, Artificial Light and Shutter*, Geneva: ISO, ISO 7589:2002, 2002.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE(ISO), *Photography - Electronic scanners for photographic images - Dynamic interval measurements*, Geneva: ISO, ISO 21550: 2004, 2004.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Document Management : Electronic Document File Format for Long-term Preservation: Part 1: Use of PDF 1.4 (PDF/A-1)*, Geneva: ISO, ISO 19005-1, 2005.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Photography - Electronic Cameras - Methods of Measurement of Optoelectronic Conversion Functions*, Geneva: ISO, ISO 14524: 2009, 2009.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Space data and information transfer systems - Open archival information system (OAIS) - Reference model*, Geneva: ISO, ISO 14721: 2012, 2012.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Quality management systems - Fundamentals and vocabulary*, Geneva: ISO, ISO 9001:2015, 2015.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Information and documentation - Records management - Part 1: Concepts and principles*, Geneva: ISO, ISO 15489:2016, 2016.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Quality management systems - Requirements*, Geneva: ISO, ISO 9004:2018, 2018.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Quality management - Quality of an organization - Guidance to achieve sustained success*, Geneva: ISO, ISO 9004:2018, 2018.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Electronic document management - Design and operation of an information system for the preservation of electronic documents - Specifications*, Geneva: ISO, ISO 114641:2018, 2018.

INTERNATIONALE ORGANISATIE VOOR STANDAARDISATIE (ISO), *Information and documentation - The Dublin Core set of metadata elements - Part 2*, Geneva: ISO, ISO 15836-2:2019, 2019.

INSTITUT DE CONSERVATION DU CANADA (ICC), *Plan de préservation des collections patrimoniales : Stratégies pour éviter ou limiter les dommages*, ongedateerd, [Online],
<https://www.canada.ca/fr/institut-conservation/services/conservation-preventive/plan-preservation-collections-patrimoniales.html>

INSTITUT ROYAL DU PATRIMOINE ARTISTIQUE (IRPA), *Cel preventieve conservatie*, KIK-IRPA, ongedateerd, [Online],
<https://www.kikirpa.be/nl/wie-is-wie/preventieve-conservatie?lang=nl>

KULESZ, O., *Culture, machines et plateformes : L'intelligence artificielle et son impact sur la diversité des expressions culturelles*, Comité intergouvernemental pour la protection et la promotion de la diversité des expressions culturelles, Douzième session, Parijs, UNESCO, 11-14 December 2018, DCE/18/12.IGC/INF.4, 13 November 2018, [Online],
https://en.unesco.org/creativity/sites/creativity/files/12igc_inf4_fr.pdf

LU, T., *Homogeneity models for image processing in the cultural heritage sector*, [Ongepubliceerde doctoraalscriptie], Vrije Universiteit Brussel, 2020, 245 blz.

MANDERLKERN, D., *Diffusion des données publiques et révolution numérique*, Rapport rédigé à la suite des travaux de l'Atelier présidé par D. Mandelkern au Commissariat au Plan, Parijs : La Documentation Française, 1999, 124 blz.

MEMON, J., SAMI, M., KHAN R.-A., UDDIN M., « Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR) », In : *IEEE Access*, vol. 8, blz. 142642-142668, 2020.

MICKELVEY, F., et HUNT, R., *Responsabilité algorithmique et découverte de contenus numériques, document de réflexion, Rencontre internationale sur la diversité des contenus à l'ère numérique*, Patrimoine canadien, February 2019, [Online],
<https://fr.readkong.com/page/responsabilite-algorithmique-et-decouverte-de-contenus-2696781>

MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION, *Feuille de route stratégique : Métadonnées culturelles et transition Web 3.0*, Januari 2014, 38 blz. [Online],
<https://www.enssib.fr/bibliotheque-numerique/documents/64776-feuille-de-route-strategique-metadonnees-culturelles-et-transition-web-3-0.pdf>

NAPOLI, P., *Diversité de contenus à l'ère numérique : découvrabilité de contenu diversifié aux échelons local, régional et national, document de réflexion, Rencontre internationale sur la diversité des contenus à l'ère numérique*, Patrimoine canadien, Februari 2019, [Online],
<https://culturenumerique.qcnum.com/wp-content/uploads/2019/03/Napoli-De%CC%81couvrabilite%CC%81-de-contenu-diversifie%CC%81-aux-e%CC%81chelons-local-re%CC%81gional-et-national.pdf>

PLAN D'EXPLOITATION ET DE PRÉSERVATION DES PATRIMOINES (PEP'S), *Normes et lignes directrices techniques et organisationnelles pour les initiatives de numérisation des patrimoines culturels soutenues par la Communauté française*, ministère de la communauté française Brussel : Secrétariat général – Préservation et exploitation des patrimoines, 2009, 61 blz., [Online],
http://www.peps.cfwb.be/index.php?eID=tx_nawsecuredl&u=0&file=fileadmin/sites/numpat/upload/numpat_super_editor/numpat_editor/documents/CFWB/Normes.pdf&hash=a2d0e740e46e34b8c3475c604c3e96edbb8a8712

OBSERVATOIRE DE LA CULTURE ET DES COMMUNICATIONS DU QUÉBEC.

État des lieux sur les métadonnées relatives aux contenus culturels, Québec :

Institut de la statistique du Québec, Observatoire de la culture et des communications du Québec, 2017, 118 blz., [Online],

<https://statistique.quebec.ca/fr/fichier/etat-des-lieux-sur-metadonnees-relatives-aux-contenus-culturels.pdf>

PLAMONDON, J., *Bien documenter pour favoriser la découverte en ligne. Travailler avec des données*, 2019, [Online],

<https://espaceschoregraphiques2.com/wp-content/uploads/2019/12/Guide-me%CC%81adonne%CC%81es-FJPP.pdf>

PLAMONDON, J., *Données ouvertes et liées : le Web comme base de données*, Maart 2018, [Online],

<https://joseeplamondon.com/donnees-ouvertes-liees-web-base-de-donnees/>

PLAMONDON, J., *Produire des données : entre outils de marketing et bases de connaissances*, 21 Augustus 2019, [Online],

<https://joseeplamondon.com/produire-donnees-culture-marketing-ou-bases-connaissances>

RIEGER, T., *Federal Agencies Digital Guidelines Initiative. Technical Guidelines for Digitizing Cultural Heritage Material. Creation of Raster Image File*, 2016, 99 blz. [Online],

http://www.digitizationguidelines.gov/guidelines/FADGI%20Federal%20%20Agencies%20Digital%20Guidelines%20Initiative-2016%20Final_rev1.pdf

SALAÛN, J.-M., « Bibliothèques numériques et Google-Print », *In: Regards sur l'actualité : mensuel de la vie publique en France*, La Documentation Française, 2005, [Online],

https://archivesic.ccsd.cnrs.fr/sic_00001576/document

SIBILLE, C., « Avantages et inconvénients d'un encodage en EAD », *In: La Gazette des archives*, n°220, 2010-4, Les instruments de recherche : évolutions, publics et stratégies, blz. 179-194.

SOYEZ, S., *La numérisation en marche : les étapes de la dématérialisation des processus de travail*, Brussel : Archives générales du Royaume et Archives de l'État dans les Provinces, ongedateerd, [Online],

http://www.arch.be/docs/brochures/la_numerisation_en_marche.pdf

TCHÉHOUALI, D., « Note d'orientation pour une étude sur les enjeux, défis et opportunités de la découvrabilité en ligne des contenus culturels francophones », *In: Dispositif d'observation des dynamiques culturelles et linguistiques*, Comité d'orientation des Dynamiques culturelles et linguistiques, April 2018, [Online],

https://www.francophonie.org/sites/default/files/2020-01/perspectives_de_recherche_web-2.pdf

TESSIER, M., FERRY, L., RACINE, B., JEANNENEY, J.-N., *La révolution du livre numérique : état des lieux, débats, enjeux*, Paris : Odile Jacob, 2011. 208 blz.

THIAULT, F., « Mutations des métiers de l'information-documentation : vers l'émergence d'une culture de l'information numérique », *In: Les Cahiers de la SFSIC*, Société française des sciences de l'information et de la communication, 2012, blz. 59-62.

VAN DORMOLEN, H., *Metamorphoze Preservation Imaging Guidelines. Image Quality*, 2012, versie 1.0, [Online],
http://www.imagingetc.com/images/Resources/Images/PDFs_DownloadFiles/Metamorfoze_Preservation_Imaging_Guidelines_1.0.pdf

VILLANUEVA, E., SHIRI, A., « Methodological Diversity in the Evaluation of Cultural Heritage Digital Libraries and Archives: An Analysis of Frameworks and Methods / Diversité méthodologique dans l'évaluation des bibliothèques et les archives numériques du patrimoine culturel : Une analyse des cadres et des méthodes », *In: Canadian Journal of Information and Library Science*, vol. 43 no. 3, 2021, blz. 316-342.

