

Preserveringsplan

2025-2028

Mei, 2025



KB } nationale
bibliotheek

Colofon

Uitgever:
KB, 2025

Redactie:
Sam Alloing, Daniel Steinmeier

Beeld:
KB, 2025

Met dank aan:
Inge Hofsink, Fiona Rombouts,
Geertrui Verbraak, Johan van der Knijff,
Meta van der Waal - Gentenaar

Meer informatie:
Jeffrey.vanderHoeven@kb.nl

© KB, Den Haag.

Inhoud

1.1	Inleiding	3
2	Integriteit	5
2.1	Versie-integriteit	5
2.2	Informatie-integriteit	6
2.3	Collectie Integriteit	7
3	Authenticiteit	8
3.1	Bron	8
3.2	Event geschiedenis	9
3.3	Intentie	11
3.4	Significant Properties	12
4	Duurzame toegankelijkheid	13
4.1	Leesbaar	13
4.2	Interpreteerbaar	14
4.3	Betrouwbaar	16
4.4	Beschikbaar	18
4.5	Designated community	19
4.6	Hergebruik	20
4.7	Bescherming van de collectie	22
4.8	Documentatie	24
4.9	Deskundigheid	26
4.10	Netwerken	28
4.11	Infrastructuur	29
4.12	Veiligheid	30
4.13	Recht	31

1.1 Inleiding

Dit document is bedoeld om uiting te geven aan de principes zoals vastgelegd in het [digitaal preserveringsbeleid](#). Het preserveringsplan is de tactische uitwerking van de strategische uitgangspunten die daar staan genoemd. Het houdt dezelfde volgorde en onderwerpen aan als het strategisch beleid om gemakkelijker de koppeling tussen de twee documenten te kunnen zien. Het plan bevat een aantal gebieden waar we de komende jaren aan willen werken en bundelt een aantal basisvragen die gebruikers kunnen stellen over onze digitale collecties en waar preservering probeert een oplossing voor te bieden. Per vraagstuk wordt benoemd welk deelaspect of begrip ontleend aan de standaarden rondom preservering is betrokken en hoe we vooruitgang op dit onderdeel willen aantonen de komende jaren.

Omdat het tactisch beleid is zijn de strategieën tot op zekere hoogte wat praktischer uitgewerkt dan in het strategisch beleid maar nog steeds generiek genoeg om te kunnen gelden voor meerdere collecties. Per vraagstuk wordt ter verduidelijking ook een voorbeeld gegeven van een item uit de collectie dat het vraagstuk illustreert. Zo wordt duidelijk voor de lezer hoe de theorie van preservering gekoppeld kan worden aan casussen die ook daadwerkelijk voorkomen in de praktijk. Ook kan aan de hand hiervan dan nagevolgd worden hoe de voorgestelde oplossing het vraagstuk zou oplossen of verbeteren en aan welk aspect van de gebruikersbeleving dit ten goede komt.

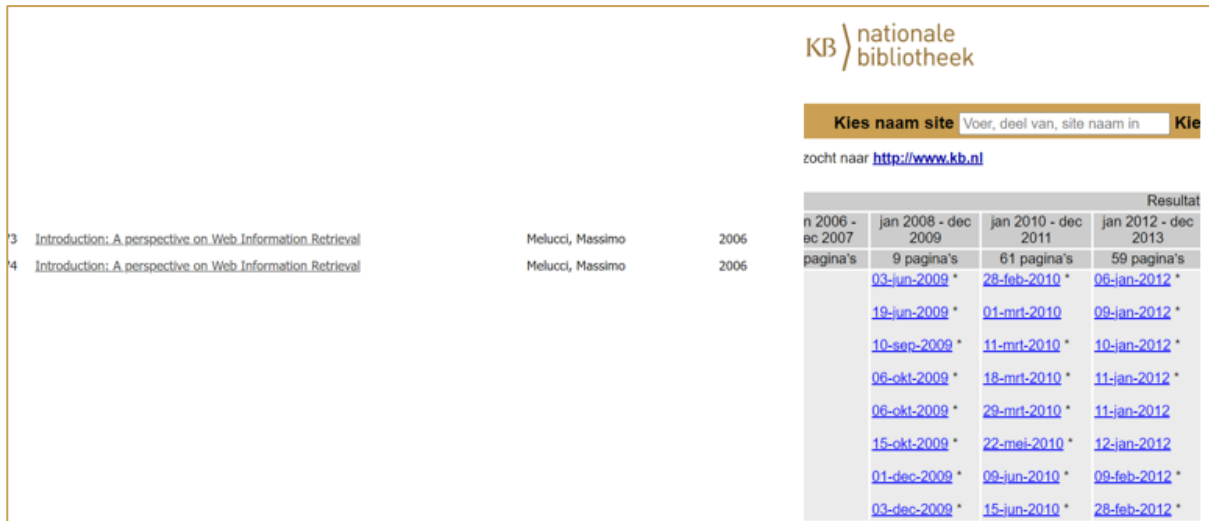
Niet alle onderwerpen uit het strategisch beleid worden behandeld maar vooral die onderwerpen die wat ons betreft meer prioriteit moeten krijgen de komende jaren. Uiteraard betekent dit niet dat we de andere dingen niet gaan doen maar eerder dat we deze voortzetten omdat deze al zijn geïmplementeerd. Een voorbeeld is het aspect van bit-integriteit. Dit is geïmplementeerd in de huidige systemen in de vorm van checksum-controles. Dit begrip is ook ingeburgerd geraakt als controlemechanisme voor migraties en behoeft in dit plan daarom geen speciale aandacht. Een ander voorbeeld is IP-integriteit. Het controleren van de volledigheid van de levering is inmiddels een standaard onderdeel van de workflows binnen de waardeketen.

Hieronder volgen in het kort de actiepunten voor de komende drie jaar.

Actiepunten	Prioriteit	Planning
Versiebeleid implementeren	1	2025
Beter onderscheid maken tussen foutgevallen	1	Iteratief per collectie
Migratie versnellen	1	2025
Lerend werken in praktijk brengen	1	2025
Risicomitigatie implementeren	1	2025
Foutgevallen opsporen met formaatvalidatie	1	Iteratief per collectie
Corrupte bestanden opnieuw laten leveren	1	Iteratief per collectie
Meedoen aan werkgroepen rondom maatschappelijke thema's	1	2025-2027
Kennis uitwisselen in netwerkverband	1	2025-2027
Representation information opstellen en duurzaam bewaren	1	Iteratief per collectie
Rechteninformatie duurzaam bewaren	1	Iteratief per collectie
Bewijs rondom collectie-integriteit opslaan	2	Iteratief per collectie
Contextinformatie rondom authenticiteit opslaan	1	Iteratief per collectie
Experimenteren met emulatie	2	2026
Informatie over benodigde software vastleggen	2	2026
Beperkingen voor vormen van hergebruik in kaart brengen	2	2026
Context informatie toegankelijk maken	2	2027
Procedures voor conserveringsprocessen vastleggen	2	2027
Beleidsevaluatie implementeren	2	2027
Preserveringsstrategieën opstellen per collectie	2	2026
Informatie over significant properties vastleggen	2	2026
Beleidsuitwerking op collectieniveau	2	Iteratief per collectie
Impact van sociale en technologische ontwikkelingen in kaart brengen op hoofdlijnen	3	2025-2027
Impact van securitymaatregelen vastleggen	3	2027

2 Integriteit

2.1 Versie-integriteit



KB nationale bibliotheek

Kies naam site Voer, deel van, site naam in Kies

zoekt naar <http://www.kb.nl>

	n 2006 - dec 2007	jan 2008 - dec 2009	jan 2010 - dec 2011	jan 2012 - dec 2013	Resultat
3 Introduction: A perspective on Web Information Retrieval	Melucci, Massimo	2006	9 pagina's	61 pagina's	59 pagina's
4 Introduction: A perspective on Web Information Retrieval	Melucci, Massimo	2006	9 pagina's	61 pagina's	59 pagina's
			03-jun-2009 *	28-feb-2010 *	06-jan-2012 *
			19-jun-2009 *	01-mrt-2010	09-jan-2012 *
			10-sep-2009 *	11-mrt-2010 *	10-jan-2012 *
			06-okt-2009 *	18-mrt-2010 *	11-jan-2012 *
			06-okt-2009 *	29-mrt-2010 *	11-jan-2012
			15-okt-2009 *	22-mei-2010 *	12-jan-2012
			01-dec-2009 *	09-jun-2010 *	09-feb-2012 *
			03-dec-2009 *	15-jun-2010 *	28-feb-2012 *

Figuur 1 Screenshot website ibl.kb.nl, KB, 2025

Welke soorten versies hebben we?

Een belangrijk aspect van integriteit is versie-integriteit. Informatie over versies maakt dat gebruikers kunnen bepalen naar welke versie van iets ze kijken en welke andere versies nog beschikbaar zijn. Op het plaatje links is te zien hoe deze informatie niet altijd beschikbaar is. Op basis van een zoekopdracht in onze portal IBL komen twee ogenschijnlijk dezelfde artikelen naar boven. Hoe weten we wat het verschil is tussen de twee?

Het plaatje rechts geeft een andere situatie weer die ook is te relateren aan versies: binnen webarchivering worden periodiek snapshots gemaakt van een bepaalde website. Hierdoor ontstaan er versies van de website die van elkaar te onderscheiden zijn op basis van creatiedatum. Alle snapshots zijn bij elkaar verzameld in de interface en hebben een overkoepelende identifier omdat het snapshots zijn van dezelfde website. Dit is een voorbeeld waar gebruikers daadwerkelijk kunnen zien welke versies beschikbaar zijn en hoe ze samenhangen.

Het doel is om ook voor andere collecties verder te komen met dit vraagstuk. Het doel voor de komende jaren is om versiebeleid te implementeren waarin tenminste bepaald wordt welke versies opgeslagen worden, hoe versies te herkennen en te onderscheiden zijn en op welk niveau een persistent identifier wordt toegekend. Er is versiebeleid nodig op het niveau van bestanden, metadata en publicaties. Het uiteindelijke doel is dat in het preserveringssysteem op een consistente manier links te leggen zijn tussen versies. Zo blijft deze informatie voor de lange termijn toegankelijk.

2.2 Informatie-integriteit



Figuur 2 Screenshot website dds.nl, KB, 2025

Wat betekent het?

Een ander belangrijk aspect van integriteit is informatie-integriteit: alle informatie die nodig is om het object technisch en inhoudelijk te begrijpen. Het is de bedoeling dat deze informatie gelinkt is aan het object zodat gebruikers weten waar de benodigde informatie te vinden is. Dit om te zorgen dat ze uiteindelijk het object kunnen raadplegen en duiden.

Het plaatje hierboven laat een concreet geval zien. Links zien we het openingsscherm van De Digitale Stad. Zonder contextinformatie weten gebruikers van nu wellicht niet het verschil tussen bewoner en 'tourist'. Daardoor kunnen ze ook geen goede keuze maken in de navigatie van de website. Wanneer ze zouden klikken op 'bezoek de stad als bewoner' zouden ze bij een foutmelding uitkomen omdat de inlog-functionaliteit niet meer werkt. Dit is technische informatie die gebruikers van nu niet kunnen weten wanneer het niet verstrekt wordt samen met het object zelf.

Tegelijkertijd is er ook de inhoudelijke informatie: wat was het verschil tussen bewoners en toeristen? In de Digitale Stad werd veel gebruik gemaakt van metaforen: zo waren er huizen, maar ook flats. Deze begrippen spreken niet voor zich maar we kunnen wel verwijzen naar informatie waar dit wordt uitgelegd. Op deze manier kunnen gebruikers zelfstandig het object raadplegen zonder afhankelijk te zijn van instructie van experts. Om verder te komen met het aanbieden van context-informatie willen we de komende jaren inzetten op het opstellen van toelichting, zoals de collectiebeschrijving van De Digitale Stad. Daarnaast willen we deze informatie ook opnemen en linken in het preservingsysteem zodat deze duurzaam beschikbaar blijft.

2.3 Collectie Integriteit

Titel, Vol Iss met ontbrekende artikel(en)	Springer	Dig Magazijn	Ontbrekend (neg. getal)
Acta mathematica Hungarica Vol:111Iss:1-2	12	12	0
Acta mathematica Hungarica Vol:111Iss:3	5	2	-3
Analog integrated circuits and signal proce Vol:48Iss:2	10	0	-10
Apoptosis Vol:11Iss:3	14	14	0
Apoptosis Vol:11Iss:8	18	18	0
Artificial life and robotics Vol:10Iss:1	16	16	0
Astrophysics and space science Vol:300Iss:4	9	8	-1
Brain tumor pathology Vol:23Iss:1	10	10	0
Bulletin of experimental biology and medic Vol:140Iss:6	31	30	-1
Bulletin of experimental biology and medic Vol:141Iss:1	37	41	4

Figuur 3 Screenshot intranet plein.kb.nl, KB, 2025

Hoe weten we dat we niks kwijt zijn?

Het e-depot is al meer dan 20 jaar oud en de collecties zijn in die tijd een aantal keer gemigreerd. In sommige gevallen is bekend dat materiaal hierdoor kwijt is geraakt of nog op storage staat buiten het huidige e-depot. Het kan ook zijn dat bepaalde issues nooit zijn aangeleverd.

Bovenstaande screenshot is een voorbeeld van een inventarisatie die is gedaan waaruit blijkt dat er in sommige gevallen issues van tijdschriften missen. In het kader van collectie-integriteit willen we in ieder geval duidelijke informatie kunnen bieden aan gebruikers over welk materiaal we wel en niet beschikbaar hebben en waarom. Nu dat we opnieuw migreren naar een nieuw preserveringssysteem is er ook de kans om compleetheid van de collectie te inventariseren. Zo worden er wellicht ook andere gevallen bekend die nu nog niet in beeld zijn.

We kunnen voortgang tonen de komende jaren op het gebied van collectie-integriteit wanneer er meer informatie beschikbaar is over de compleetheid van de collecties die we migreren naar het preserveringssysteem, inclusief redenen voor het eventuele ontbreken van materiaal. Een ander voorbeeld van concreet bewijs van collectie-integriteit kunnen log-files zijn waaruit blijkt dat alle bestanden van de ene omgeving naar de andere zijn verhuisd met behoud van bit-integriteit.

3 Authenticiteit

3.1 Bron



Figuur 4 Screenshot internal storage environment, KB, 2025

Hoe zag het er vroeger uit?

De standaarden op het gebied van digitale preserving gaan ervan uit dat materiaal waarschijnlijk niet door de tijd heen toegankelijk kan blijven in het oorspronkelijke formaat. Niet voor alle collecties geldt dat dit nu al een urgent probleem is. Er zijn echter al collecties aanwijsbaar waar we dit zien.

Een concreet voorbeeld is de collectie van optische dragers. De oorspronkelijke cd's worden nu overgezet naar een iso-image of WAV-files met bijbehorende logfiles en metadata. Het resultaat is te zien aan de linkerkant van de afbeelding. Maar we kunnen hier ook zien hoe het origineel eruit zag. Bij het overzetten van de dragers wordt er ook een afbeelding van de hoes gemaakt zodat ontwerp en tekst van het origineel inzichtelijk blijft zelfs al is de oorspronkelijke drager alleen nog toegankelijk als een data-bestand op disk. Deze hoestekst bevat vaak belangrijke informatie over de inhoud van het object en wat er technisch vereist is om het af te spelen. Deze informatie is niet altijd beschikbaar als metadata.

Om authenticiteit van deze collectie voor gebruikers inzichtelijk te maken is het nodig dat deze informatie beschikbaar is. Daarnaast is het ook nodig dat er context informatie wordt opgesteld die gebruikers duidelijk maakt hoe de huidige versies van deze objecten tot stand zijn gekomen en hoe ze zich verhouden tot de originele bestanden.

3.2 Event geschiedenis



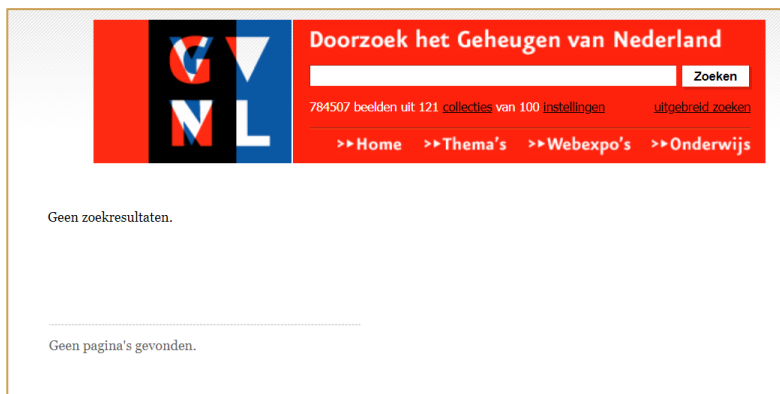
Figuur 5 Screenshot kbresearch.nl / Internet Archive, KB, 2025

Hoe weten we dat het origineel is?

Een ander belangrijk aspect van authenticiteit is de geschiedenis van het object. Vanaf het moment dat een object wordt opgenomen in de collectie worden er acties op uitgevoerd. Met event-geschiedenis willen we voor gebruikers inzichtelijk maken wat er is gebeurd met het object zodat ze zichzelf kunnen verzekeren dat de objecten die ze raadplegen ook daadwerkelijk de herkomst hebben die ze lijken te hebben.

Bijvoorbeeld wanneer objecten worden gemigreerd is het belangrijk dat deze actie wordt vastgelegd en ook dat de documentatie van eerdere acties wordt meegenomen bij de migratie. Om hier verder in te komen willen we de komende jaren documenteren welke events bewaard zijn gebleven uit eerdere generaties van het e-depot, welke events worden gegenereerd bij de migratie en welke events worden bijgehouden in het nieuwe preservingssysteem. Daarnaast willen we ook vastleggen op welke preservingsstrategieën we in willen zetten voor een bepaalde collectie.

Bovenstaande afbeelding is een voorbeeld van hoe verschillende preservingsstrategieën een ander resultaat kunnen opleveren. Links is te zien hoe een reconstructie van de website NL-menu heeft geleid tot een bepaalde lay-out, rechts is te zien hoe dezelfde website er uit ziet op basis van webharvesting. De websites verschillen totaal en de versie aan de rechterkant bevat elementen die niet authentiek zijn. Op basis van een dergelijke afweging kan verantwoord worden waarom in dit geval reconstructie een authentiekere resultaat oplevert dan het geval is bij webharvesting. Door preservingsstrategieën vast te leggen met uitleg waarom voor een bepaalde strategie is gekozen, kunnen we gebruikers meer inzicht geven in de authenticiteit van onze collecties.



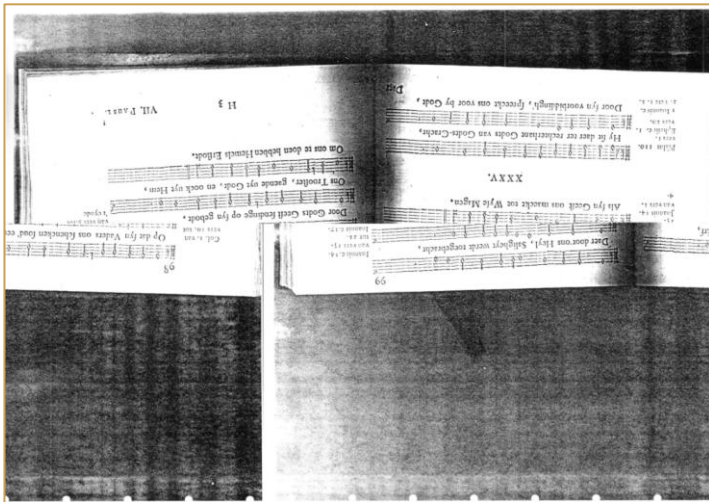
Figuur 6 Screenshot KB webarchive / Geheugen van Nederland, KB, 2025

Hoe is het tot stand gekomen?

Om de geschiedenis van onze objecten beter te snappen is het ook nodig dat de event-geschiedenis van het ingestproces duurzaam wordt bewaard, dat afspraken over het verkrijgen van de content inzichtelijk zijn en dat er context-informatie beschikbaar is over hoe de collectie tot stand is gekomen.

Een voorbeeld van waarom dit belangrijk kan zijn is hierboven zichtbaar. Het betreft de website van Het geheugen van Nederland zoals aanwezig in ons webarchief. Voor een gebruiker die niet bekend is met de criteria die we stellen aan websites en de technische mogelijkheden en beperkingen van de tools waarmee we websites binnenhalen, zal het niet duidelijk zijn waarom juist deze website geen enkele content toont. In dit geval is daar bewust voor gekozen omdat het niet de bedoeling was alle objecten van deze website te dupliceren in het webarchief. Wanneer we meer informatie kunnen geven aan gebruikers omtrent dergelijke keuzes kunnen ze ook zelfstandig beschouwen of een object authentiek genoeg is voor het doeleinde dat ze voor ogen hebben.

3.3 Intentie



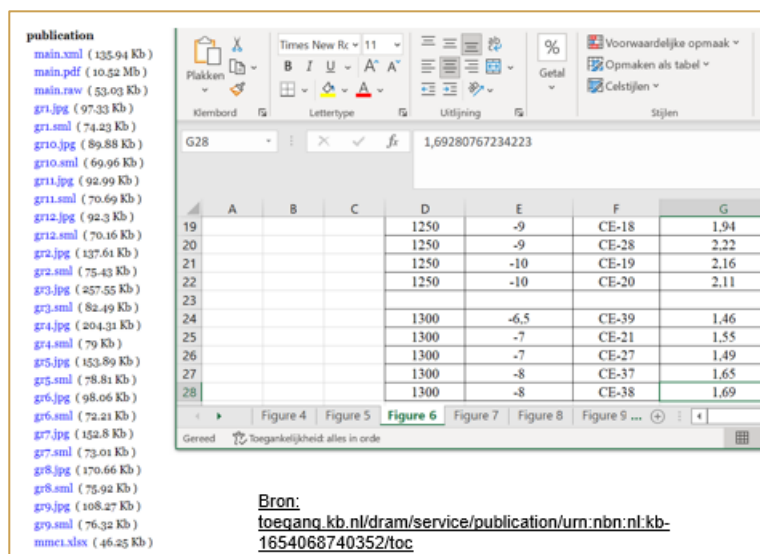
Figuur 7 Screenshot brug061chri01_0028.jpg, KB, 2025

Hoort dit zo?

Misschien wel het meest filosofische aspect van authenticiteit is intentie: het idee dat het object op een juiste manier de oorspronkelijke betekenis representeert. We kunnen dit uiteraard niet voor alle objecten op een absolute manier zekerstellen maar we kunnen wel veel voorkomende fouten proberen af te vangen.

De bovenstaande afbeelding geeft dit idee weer: wat ook de oorspronkelijke bedoeling geweest kan zijn, dit is overduidelijk niet de bedoeling. Deze fout is ontdekt bij het valideren van gedigitaliseerde bestanden. Omdat dit bestand niet valide was kon op basis van steekproeven aangetoond worden dat de fouten in het bestand ook daadwerkelijk leiden tot weergaveproblemen. Het migreren van collecties geeft ons de kans bestandsformaten te identificeren en te valideren. Op deze manier kunnen we objecten die niet de oorspronkelijke intentie representeren opsporen en vervangen voor betere exemplaren.

3.4 Significant Properties



Figuur 8 Screenshot bestanden en databestand van urn:nbn:nl:kb-1654068740352, KB, 2025

Wat moet bewaard blijven voor de lange termijn?

Het concept significant properties in de preserveringsliteratuur is gerelateerd aan het idee van betekenis zoals vervat in digitale objecten. Het idee is dat een bestand eigenschappen heeft waarvan sommige essentieel zijn voor de betekenis terwijl andere eigenschappen niet cruciaal zijn om de inhoud te begrijpen, bijvoorbeeld omdat ze slechts gekoppeld zijn aan een specifiek formaat.

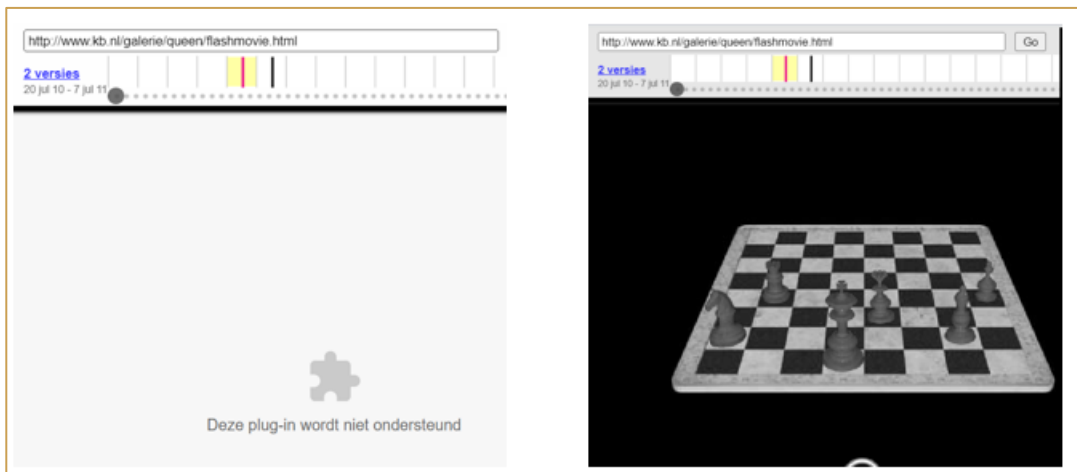
Dit onderscheid is met name belangrijk wanneer het object niet meer toegankelijk is in het oorspronkelijke formaat. Het bestand kan dan gemigreerd worden naar een nieuw formaat. Het kan echter zo zijn dat niet alle eigenschappen bewaard blijven in het nieuwe formaat. Zolang het niet-essentiële eigenschappen betreft is dit geen probleem. Als een simpel voorbeeld zou gedacht kunnen worden aan een tekstbestand: de tekst is essentieel maar als deze bewaard blijft in een nieuw formaat en de bestandsgrootte verandert daarbij dan is dit geen probleem voor de betekenis van het object.

Echter zijn er ook complexere voorbeelden waarbij er moeilijke keuzes gemaakt moeten worden. Op de afbeelding is een excel te zien. Stel dat we deze zouden omzetten naar PDF dan zouden de precieze waarden verloren gaan. En als de waarden het resultaat zijn van één of meer onderliggende formules of macro's ben je die ook kwijt.

Daarnaast zouden de waarden ook niet meer afzonderlijk toegankelijk zijn. Gaat hiermee de oorspronkelijke betekenis verloren? Dit ligt eraan of de precieze waarden en de functionaliteit om deze afzonderlijk te kunnen bewerken essentieel is voor hergebruik. In dit geval weten we dat wetenschappers dit belangrijk vinden. Door dergelijke informatie vast te leggen kunnen we zeker stellen dat dit bij toekomstige migraties niet over het hoofd wordt gezien.

4 Duurzame toegankelijkheid

4.1 Leesbaar



Figuur 9 Screenshot KB Webarchief / kb.nl, KB, 2025

Wat als het niet werkt?

Onder duurzame toegankelijkheid verstaan we alle criteria die nodig zijn om te zorgen dat gebruikers zelfstandig kennis kunnen nemen van de collecties en zowel technisch als inhoudelijk begrijpen waar ze naar kijken. Het idee van leesbaarheid binnen preservering gaat ervan uit dat bestandsformaten verouderd raken en dat gebruikers speciale kennis nodig hebben om te weten hoe ze iets moeten openen. Het doel is dat ze een versie van de content kunnen bekijken in een courante omgeving.

Op het plaatje is een voorbeeld te zien van een flash-animatie in een webtentoonstelling van de KB. Wanneer gebruikers weten welke bestandsformaten in de collectie zitten en met welke software deze te openen is dan kunnen ze zelfstandig deze content bekijken. Hiervoor is het nodig dat we meer bestandsformaten analyseren. De technische informatie - waaronder informatie over de omgeving die nodig is om de content te bekijken – slaan we op bij de objecten.

Een andere optie om verouderde formaten leesbaar te houden is om niet alleen informatie te bieden maar ook de software die nodig is om formaten weer toegankelijk te maken. Bijvoorbeeld op computers in de leeszaal. Dit is recent gebeurd voor Flash-materiaal in het webarchief met behulp van de plugin Ruffle.

Een doel voor de komende jaren is dat we meer gaan experimenteren met emulatie om te zien hoe dit op een toegankelijke manier aangeboden kan worden aan gebruikers. Daarnaast willen we ook vastleggen welke collecties het meest geëigend zijn om duurzaam toegankelijk te houden middels de strategie van emulatie.

4.2 Interpreteerbaar

Titel:	The Utrecht Psalter : picturing the psalms of David / realisation: Koert van der Horst, Frits Ankersmit Bijbel ; O.T. Ps. (lat)
Uitgave:	Utrecht : Utrecht University Library
Jaar:	cop. 1996
Collatie:	1 cd-rom. : facs. ; 4 3/4 in + boekje ([8] p., Engels)
Annotatie:	Facsimile van Utrecht University Library , Ms. 32 met transcriptie van de Latijnse tekst, alsmede vertalingen in het Engels, Nederlands en Frans Systeemeisen: IBM PC of compatible; MS-DOS 3.3 (of hoger); Windows 3.11 of Windows for Workgroups 3.11; MSCDEX 2.2 (of hoger), CD-ROM speler, SVGA monitor, [8]
Samenvatting:	150 psalmen en 16 teksten van de Utrecht Psalter
Auteur:	Koert van der Horst (ISNI 0000 0001 0925 8866) ; Frits Ankersmit (ISNI 0000 0000 4269 2317)
Organisatie:	Universiteitsbibliotheek, Utrecht
Onderwerp:	handschriften; religieuze kunst; middeleeuwen
UNESCO rubriek:	Algemeen
Brinkmannummer:	B9641482
Jaar:	cop. 1996
Links:	WorldCat

Figuur 10 Screenshot KB-catalogus, KB, 2025

Hoe moet ik het openen?

Het bestandsformaat kennen geeft nog geen garantie dat de gebruiker ook weet hoe het geopend moet worden. Sommige bestandsformaten zijn immers verouderd geraakt en werken niet meer met moderne software. Dit is een van de oorspronkelijke redenen om aandacht te hebben voor digitale preservatie.

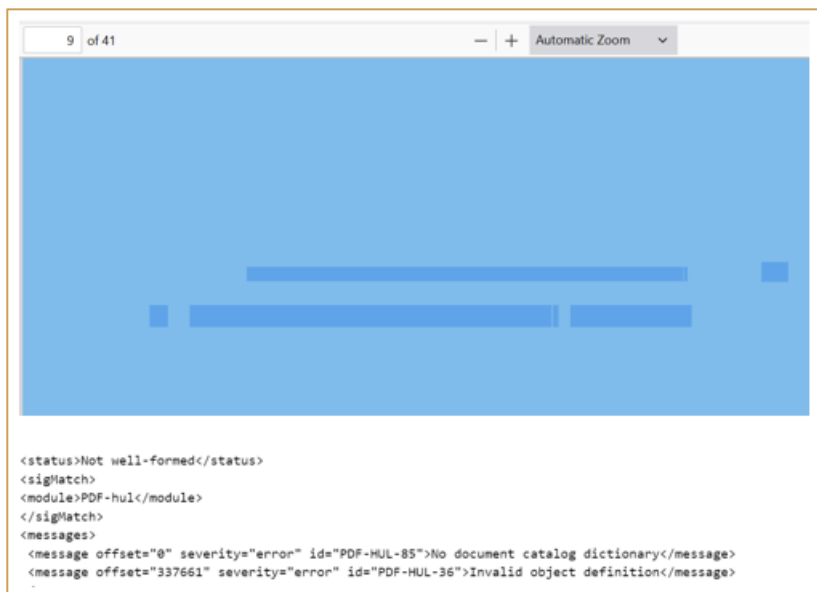
De oplossing vanuit preservatie is het hebben van *representatie informatie*. Dit kan uitleg zijn over hoe het formaat technisch in elkaar zit maar kan ook semantische informatie zijn. Bijvoorbeeld over de betekenis van kolommen in een bepaald excel-template. Software telt ook als representatie-informatie omdat de standaarden rondom preservatie ervanuit gaan dat dit de meest gebruikersvriendelijke manier is om bestanden leesbaar te houden. Het aanbieden van de juiste software in de leeszaal is dus een manier om representatie informatie aan te bieden tezamen met het object. Er kan echter ook verwezen worden naar documentatie wanneer we denken dat gebruikers niet voldoende kennis hebben om te weten hoe ze iets moeten openen.

Een voorbeeld is de cd-rom van het Utrecht Psalter die we in de collectie hebben. Deze werkt alleen op een oude versie van Windows 3.11 en wordt door de universiteitsbibliotheek van Utrecht als verloren beschouwd omdat deze niet meer te openen is in moderne versies van Windows. Om dit toch toegankelijk te maken zou er gekeken moeten worden naar emulatie. Daarnaast is er in onze collectie nog context-informatie te vinden over deze cd-rom, inclusief informatie over de totstandkoming en beperkingen die deze bevatte.¹ Dit is belangrijk om aan te kunnen bieden aan gebruikers die deze versie willen bekijken.

¹ Zie: https://www.dbnl.org/tekst/_jaa008201501_01/_jaa008201501_01_0007.php#376

We kunnen vooruitgang tonen met betrekking tot dit vraagstuk door in kaart te brengen welke formaten speciale software nodig hebben om geopend te kunnen worden en ook door semantische context informatie op te slaan of te linken.

4.3 Betrouwbaar



Figuur 11 Screenshot corrupte PDF / JHOVE error, KB, 2025

Wat als het echt stuk is?

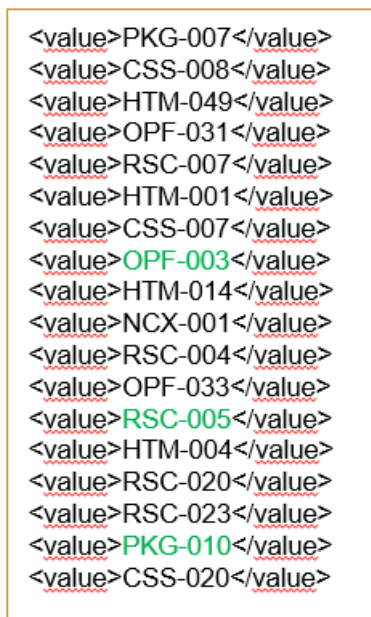
In sommige gevallen helpt het niet meer om informatie of software aan te bieden omdat de content die we hebben opgeslagen simpelweg corrupt is. Vanuit preserving is het belangrijk dat een digitaal archief betrouwbaar is. Dit betekent niet alleen dat materiaal integer en authentiek is zoals hierboven beschreven, maar ook dat gebruikers zelf het bewijsmateriaal hiervoor kunnen inzien en beoordelen. Wanneer een gebruiker verrast wordt met een bestand dat niet te openen is maakt dit geen betrouwbare indruk. Dit willen we voorkomen, maar hoe herkennen we zulke bestanden?

Hierboven zie je een voorbeeld van een corrupte PDF waarvan de tekst niet leesbaar is. Voor dit soort situaties zijn er tools beschikbaar die kunnen controleren of een bestand voldoet aan de specificatie. Dit heet formaatvalidatie. Uit formaatvalidatie bleek dat dit bestand fouten bevatte. Bij inspectie blijkt dat deze fouten zo serieus zijn dat de content helemaal niet meer zichtbaar is. Niet alle fouten die uit formaatvalidatie komen zijn echter belangrijk voor correcte weergave. Sommige fouten zijn afwijkingen van de standaard maar zonder dat dit gevolgen heeft voor het tonen van de inhoud. Het omgekeerde kan ook, bijvoorbeeld een PDF bestand waarvan de inhoud niet getoond kan worden door het gebruik van encryptie. Ondanks dat het bestand volledig conform de standaard is kan deze kenmerken bevatten die het bestand minder geschikt maken voor duurzame toegankelijkheid.

De komende jaren willen we meer bestanden onderzoeken en beter zicht krijgen op welke fouten leiden tot corrupte weergave. Dit betekent dat we niet alleen kijken naar validatie tegen de standaard maar ook naar andere risicogeveallen. Op deze manier kunnen we onderzoeken welke corrupte bestanden aanwezig zijn in de collectie en proberen de juiste versies opnieuw te verkrijgen. Hiermee kunnen we ook zorgen dat voor bepaalde formaten het kennisniveau stijgt waardoor we er meer

controle over hebben in het kader van functionele conservering. Het is hierbij belangrijk dat we beter zicht krijgen op welke risico's voor welke formaten spelen en welke tools geschikt zijn voor het detecteren hiervan.

4.4 Beschikbaar



```
<value>PKG-007</value>
<value>CSS-008</value>
<value>HTM-049</value>
<value>OPF-031</value>
<value>RSC-007</value>
<value>HTM-001</value>
<value>CSS-007</value>
<value>OPF-003</value>
<value>HTM-014</value>
<value>NCX-001</value>
<value>RSC-004</value>
<value>OPF-033</value>
<value>RSC-005</value>
<value>HTM-004</value>
<value>RSC-020</value>
<value>RSC-023</value>
<value>PKG-010</value>
<value>CSS-020</value>
```

Figuur 12 Screenshot interne documentatie, KB, 2025

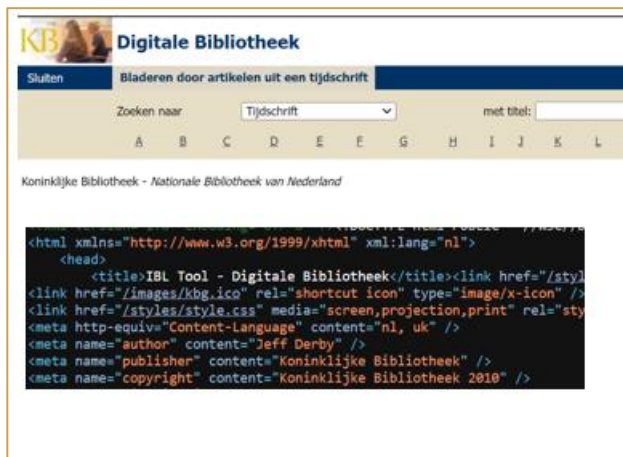
Wat als het alleen maar stuk lijkt?

Een van de uitgangspunten van het strategisch beleid is het principe van minimal ingest: het idee dat materiaal eerst veilig wordt opgeslagen voordat er bestandsvalidatie wordt toegepast. Dit maakt ook dat materiaal sneller beschikbaar is voor gebruikers. Een van de redenen dat materiaal niet beschikbaar komt is dat bestandsvalidatie een fout oplevert die ingest van het materiaal verhindert.

Zoals op bovenstaande afbeelding te zien zijn er verschillende foutcodes die optreden in een dergelijk proces. Zoals eerder genoemd zijn niet alle fouten problematisch voor weergave. Dit geldt in het voorbeeld voor de groen gekleurde waardes. Het is dan ook belangrijk dat deze niet in de weg staan van ingest.

Bij de migratie naar een nieuw systeem hebben we de kans om de huidige logica van het validatieproces aan te passen zodat er onderscheid gemaakt kan worden tussen foutmeldingen die corrupte bestanden opleveren en foutmeldingen die alleen afwijkingen zijn van de standaard zonder dat dit gevolgen heeft voor de weergave.

4.5 Designated community



Figuur 13 Screenshot ibl.kb.nl, KB, 2025

Hoe blijven we bij?

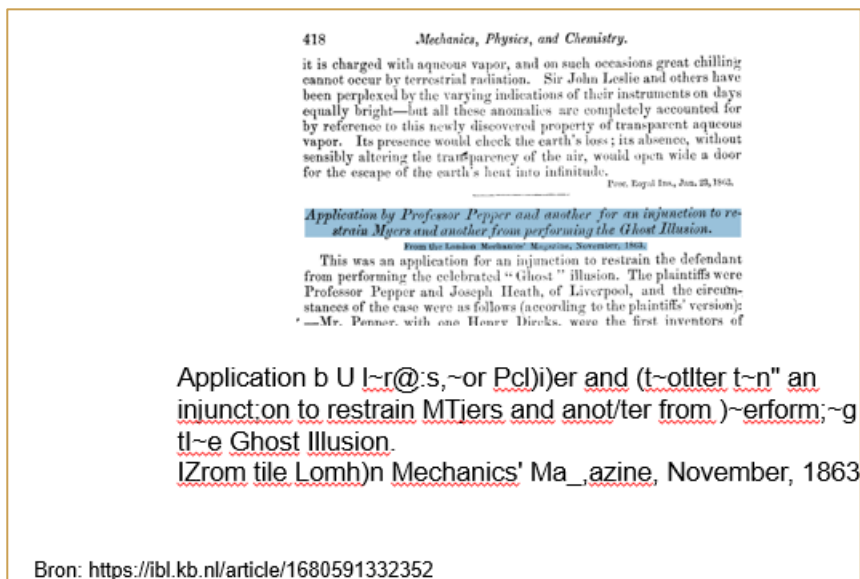
We kunnen niet voorzien wat er in de toekomst gaat gebeuren daarom is het belangrijk om in ieder geval bij te blijven bij sociale en technologische ontwikkelingen. Deze zijn ook vaak niet los van elkaar te zien. Het hierboven genoemde voorbeeld van Flash laat dit zien. Dit formaat is verouderd geraakt omdat de fabrikant ervoor heeft gekozen dit formaat niet langer te ondersteunen. Aan de andere kant was het formaat al omstreden omdat er zorgen waren over de veiligheidsaspecten ervan. Daarnaast was er ook concurrentie van andere technieken zoals HTML5 dat door veel mensen werd geadopteerd als een alternatief.

Door ontwikkelingen omtrent gebruik en ondersteuning van bestandsformaten in de gaten te houden kunnen we zorgen dat we bij blijven. Het is daarbij ook belangrijk dat we als organisatie in staat zijn om sneller in te spelen op veranderingen. Een voorbeeld is de migratie naar een nieuw preserveringssysteem. Wanneer we in staat zijn om het proces van collecties aansluiten te versnellen dan kunnen we ook sneller profiteren van de nieuwe functionaliteit van het preserveringssysteem, zoals formaatanalyse en extractie van technische metadata.

Ook in de diensten die we aanbieden moet uiteindelijk aan te tonen zijn dat we rekening houden met verandering in gebruikerswensen en technologische ontwikkelingen. Op de afbeelding hierboven is de website te zien die nu toegang biedt tot de collectie van wetenschappelijke artikelen. Zoals blijkt uit de broncode is deze website al 14 jaar geleden gemaakt en is de navigatie en toegankelijkheid ervan niet naar de meest recente maatstaven.

Tegelijkertijd is het moeilijk om een dergelijke website te vernieuwen omdat het gebruik ervan laag is. Dit maakt dat een vernieuwingstraject in de praktijk minder prioriteit krijgt. Om hier verder mee te komen willen we de komende jaren in ieder geval in kaart brengen welke sociale en technische ontwikkelingen hebben plaatsgevonden die relevant zijn voor de diensten die we bieden zodat dit meegenomen kan worden in de besluitvorming rondom vernieuwingstrajecten.

4.6 Hergebruik



Figuur 14 Screenshot ibl.kb.nl/article/1680591332352, KB, 2025

Is het duurzaam toegankelijk?

Direct gerelateerd aan bovengenoemde punt is het idee van hergebruik. Om te bepalen in hoeverre ons materiaal herbruikbaar is moeten we inzichtelijk hebben wat de designated community verwacht van de diensten en de content die we bieden. Dit is ook gerelateerd aan wat er technisch mogelijk is en dit wordt ten dele bepaald door keuzes die in het verleden zijn gemaakt. Daarnaast zijn al deze dingen ook voortdurend aan verandering onderhevig.

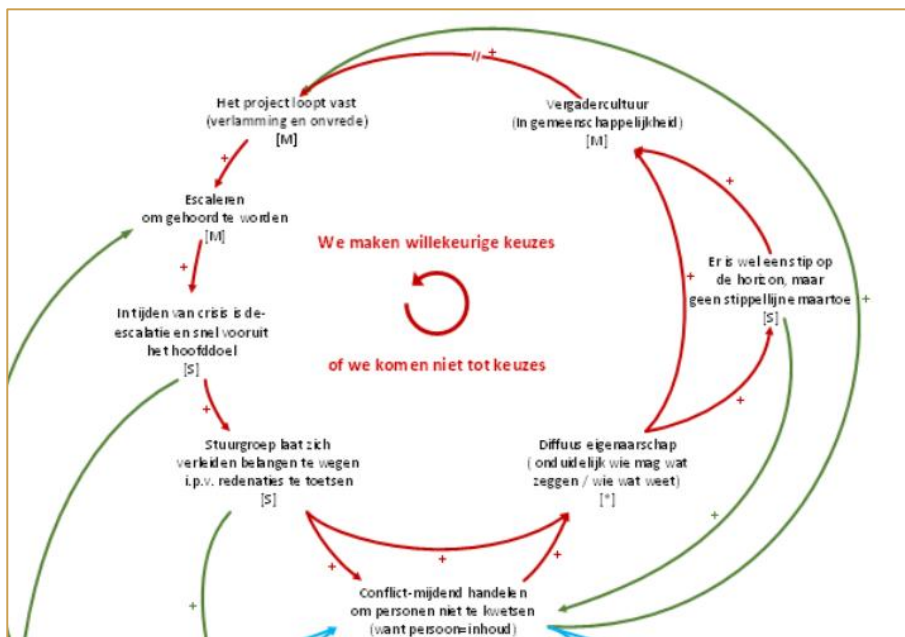
Een voorbeeld is het wetenschappelijke artikel dat in bovenstaande afbeelding te zien is. Dit is een voorbeeld van een van de oudste artikelen die we hebben. De tekst in de PDF is afkomstig van een gedigitaliseerde afbeelding waar OCR op is toegepast. De OCR maakt dat de losse woorden te selecteren en te kopiëren zijn. Wanneer we dit echter proberen is te zien dat de tekst niet goed herkend is. Het is dus eigenlijk niet mogelijk om de tekst op een betrouwbare manier te herbruiken door middel van kopiëren. We zien hier verschillende factoren samenkomen: er zijn technische beperkingen aan OCR en als instituut vinden we het in eerste instantie belangrijk dat de tekst in ieder geval leesbaar is. Aan de andere kant verwachten gebruikers van nu dat materiaal gekopieerd kan worden en begrijpen ze wellicht niet waarom dit niet werkt.

Daarnaast zijn er ook sociale ontwikkelingen die maken dat toegankelijkheid van de tekst belangrijker is geworden. Er is nu meer aandacht voor het feit dat niet iedereen de tekst kan lezen en dat tekst automatisch voorleesbaar moet zijn. Wanneer de tekst niet betrouwbaar herkend is kan deze ook niet voorgelezen worden voor mensen met een visuele beperking.

We kunnen deze technische onvolkomenheden niet zomaar oplossen maar het is wel belangrijk dat we informatie kunnen verstrekken aan gebruikers in hoeverre materiaal herbruikbaar is. Op Delpher

zien we bijvoorbeeld dat er een waarschuwing aanwezig is over de kwaliteit van de OCR maar dit geldt niet voor alle platforms. Daarnaast kunnen uiteraard ook andere aspecten van hergebruik een rol spelen. In de komende jaren willen we beter in kaart brengen welke vormen van hergebruik we kunnen garanderen voor welk materiaal, bijvoorbeeld door dit vast te leggen in collectieprofielen.

4.7 Bescherming van de collectie



Figuur 15 Screenshot interne documentatie op intranet plein.kb.nl, KB, 2025

Hoe maken we keuzes?

Om in te kunnen spelen op ontwikkelingen is het ook goed om te kijken naar hoe besluitvorming werkt binnen de organisatie. We kunnen in kaart brengen welke veranderingen impact hebben op de duurzame toegankelijkheid van de collectie. Dit is een kwestie die voor de hele organisatie van belang is. Het is daarom belangrijk om grip te hebben op hoe argumenten rondom duurzame toegankelijkheid meegenomen kunnen worden in besluitvormingstrajecten rondom technische vernieuwing en collectieontwikkeling. Daarnaast is het ook van belang dat we het preservingsbeleid zoveel mogelijk toetsbaar maken en bijstellen op basis van nieuwe inzichten.

Het huidige document is hiertoe een aanzet. We willen nog beter uitleggen aan de organisatie welke aspecten van duurzaamheid aandacht behoeven en op basis waarvan we denken dat dit zo is. Daarnaast willen we binnen de organisatie het gesprek aangaan over mogelijke oplossingen, waarbij de aandachtspunten in het huidige plan als een startpunt kunnen dienen. Wanneer beleid niet wordt gevolgd willen we beter inzicht krijgen in de waardeprioriteiten die in de praktijk worden gesteld en hoe we hier in toekomstig beleid beter rekening mee kunnen houden.

Door ons beleid uit te leggen en toetsbaar te maken nodigen we mensen uit om met ons mee te denken en mee te helpen de collecties duurzaam toegankelijk te houden. Daarnaast biedt het voor onszelf de mogelijkheid om het beleid stapsgewijs realistischer en uitvoerbaarder te maken. Bovengenoemde afbeelding toont een voorbeeld van evaluatie zoals dit binnen de organisatie heeft plaatsgevonden. Wanneer we patronen herkennen hebben we ook handvaten om deze te doorbreken. Dit plan is alweer het tweede preservingsplan dat wordt opgesteld. De afgelopen vijf

jaar hebben we de kans gehad om te zien wat is gelukt en wat niet. Op basis daarvan willen we in dit plan weer gericht bepalen welke gebieden nog aandacht behoeven en dit ook tussentijds opvolgen.

4.8 Documentatie

```
2016-01-05 05:02:34 I : Computing delivery #8469511 for customer profile e-Depot Netherlands -  
2016-01-05 05:02:34 I : Starting with units updated between 2016-01-04 05:00:00 and 2016-01-05  
2016-01-05 05:02:39 I : ... 4039 base unit(s)  
2016-01-05 05:02:39 I : Applying SUB_ON_MATCH filter 'BMC_BiomedCentral'  
2016-01-05 05:02:39 I : ... subtracted 336 unit(s)  
2016-01-05 05:02:39 I : Applying SUB_ON_MATCH filter 'subEcomed'  
2016-01-05 05:02:39 I : ... subtracted 0 unit(s)  
2016-01-05 05:02:39 I : Applying SUB_ON_MATCH filter 'BMC_PhysMathCentral_Pub_Sub'  
2016-01-05 05:02:39 I : ... subtracted 0 unit(s)  
2016-01-05 05:02:39 I : Applying SUB_ON_MATCH filter 'BMC_ChemistryCentral'  
2016-01-05 05:02:39 I : ... subtracted 0 unit(s)  
2016-01-05 05:02:39 I : Applying SUB_ON_MATCH filter 'CoverDateHistorical_SUB'  
2016-01-05 05:02:39 I : ... subtracted 42 unit(s)  
2016-01-05 05:02:39 I : Applying global SUB_ON_MATCH filter 'Global_Professional_SUB'  
2016-01-05 05:02:39 I : ... subtracted 22 unit(s)
```

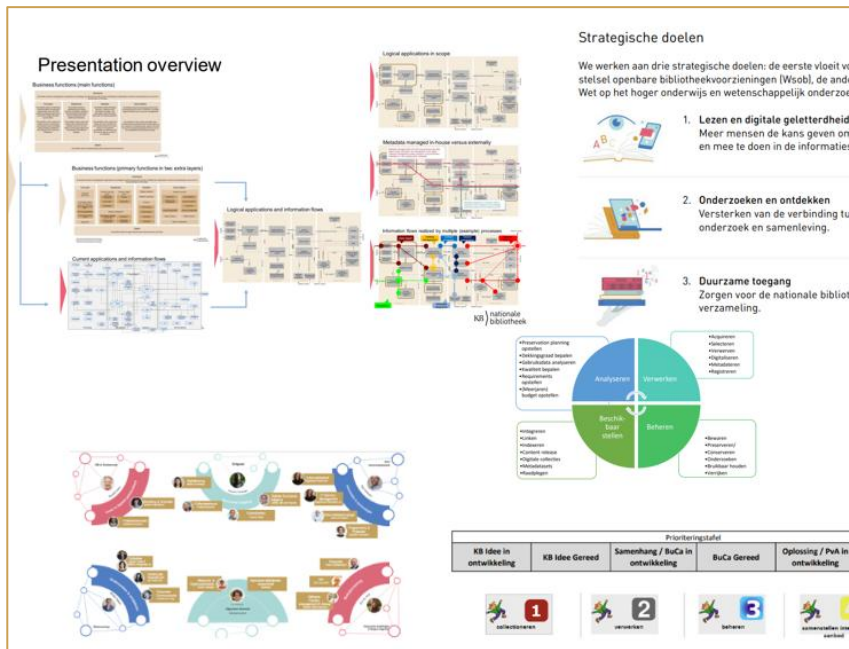
Figuur 16 Screenshot logfile ingest proces, KB, 2025

Welke ideeën zitten er achter?

Om het preserveringsbeleid effectiever te maken is het ook belangrijk dat bepaalde beleidsgebieden zoals genoemd in het strategisch plan beter uitgewerkt worden naar de praktijk. Een voorbeeld is het begrip van authenticiteit. We hebben dit uitgewerkt aan de hand van drie factoren: intentie, geschiedenis en herkomst. Echter wat betekent dit in de praktijk voor de informatie die wordt opgeslagen per collectie?

Een voorbeeld, zoals te zien hierboven, zijn de ontvangstbevestigingen die we krijgen bij sommige collecties zoals wetenschappelijke artikelen. Zijn deze belangrijk voor authenticiteit of niet? En als we zeggen dat deze belangrijk zijn, hoe zijn deze dan door gebruikers in te zien? En hoe moeten ze deze informatie gebruiken om meer inzicht te krijgen in de geschiedenis van het materiaal? Voor sommige collecties zijn deze ideeën nog niet tot in detail uitgewerkt.

De komende jaren willen we beter inzichtelijk maken hoe het beleid daadwerkelijk per collectie geïmplementeerd is. In sommige gevallen zal misschien blijken dat er een verschil is tussen het ideaal en de praktijk. Juist door deze verschillen inzichtelijk te maken kunnen we betrouwbare informatie verstrekken aan gebruikers.



Figuur 17 Interne documentatie op intranet plein.kb.nl, KB, 2025

Hoe werkt het geheel?

Een uitgangspunt van preservering is dat je dingen niet alleen goed moet doen maar ook dat je bewijs hiervan moet kunnen tonen. In hoeverre voldoen de praktische handelingen die zijn uitgevoerd op de collecties aan de procedures die hiervoor zijn opgesteld en hoe kunnen gebruikers dit voor zichzelf nagaan? Zoals aan bovenstaande collage te zien is worden er veel modellen gebruikt in de organisatie om het werk in te richten. Er is natuurlijk het OAIS-model zelf dat voor preserveringsexperts het leidende model is maar daarnaast zijn er in de organisatie ook vele andere modellen in omloop zoals bedrijfsarchitectuur, domeinmanagement en Agile-werkprocessen.

Welke impact hebben deze modellen op preservering en hoe kunnen deze modellen elkaar versterken? Dit is iets dat we de komende jaren beter in kaart willen brengen.

4.9 Deskundigheid

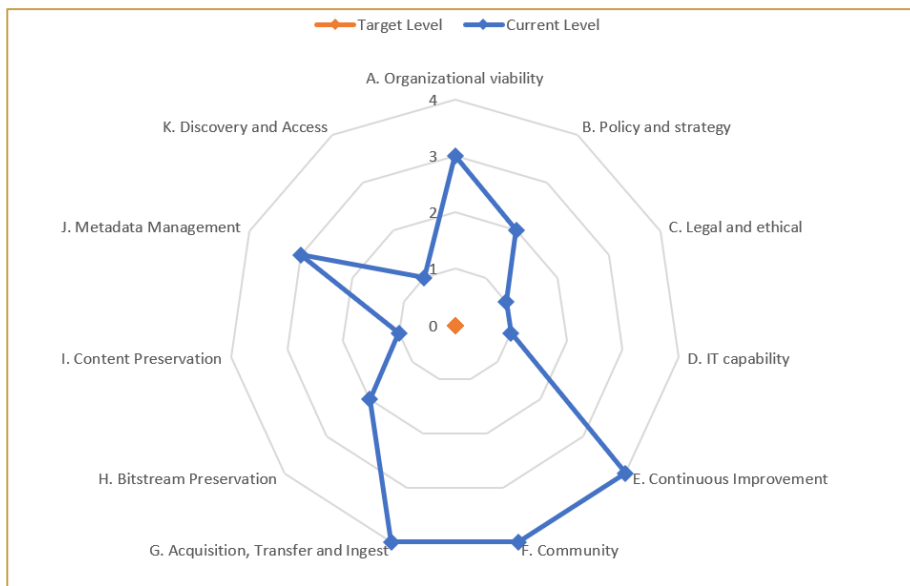


Figuur 18 Interne documentatie op intranet plein.kb.nl, KB, 2025

Hoe weten we dat het geheel werkt?

Het doel van preserving is om materiaal duurzaam toegankelijk te houden met betrouwbare processen. Maar hiervoor is het ook nodig om veranderingen bij te houden. Betrouwbare processen suggereren stabiliteit maar meegaan met verandering veronderstelt flexibiliteit. Dit zijn tegengestelde waarden die in balans moeten zijn wanneer we aan beide waarden aandacht willen geven. Dit heeft ook impact op de manier waarop we werken. We moeten zowel dingen vastleggen als ook experimenteren. Het is belangrijk dat we kennis en deskundigheid inzetten maar ook dat we nieuwe kennis opdoen door middel van meer creatieve werkvormen. Binnen de organisatie is er veel ervaring met dingen vastleggen in richtlijnen, procedures en beleid. Daarom is het voor preserving van belang om juist ook aandacht te besteden aan de andere kant van het spectrum: flexibiliteit.

Op de afbeelding is een rapport te zien waarbij het programma NDM is geëvalueerd. De lessen die hieruit zijn gekomen willen we ook toepassen in nieuwe ontwikkeltrajecten. Hoe weten we dat we als organisatie flexibel genoeg zijn om in te spelen op nieuwe ontwikkelingen? De komende jaren willen we experimenteren met methodes om hier beter antwoord op te kunnen geven. We willen verder komen met onderzoekend werken. Daarnaast willen we bekijken hoe we aantoonbaar kunnen maken waar verbetering mogelijk is op het gebied van organisatorische flexibiliteit.



Figuur 19 DPC RAM visualisering toegepast op de situatie van de KB, KB, 2025

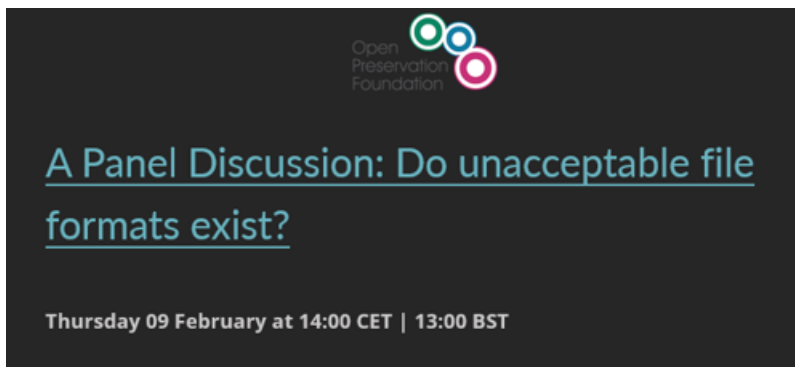
Hoe blijven we relevant?

Flexibiliteit is ook van belang in het kader van maatschappelijke thema's die invloed hebben op het gebied van preserving. Er is al een aantal thema's aan te wijzen die de afgelopen jaren steeds meer een rol spelen in het denken rondom preserving. Een van die thema's is inclusiviteit. Hoe zorgen we dat er ruimte is voor meerstemmigheid in de collectie? Een ander thema is Green IT. Wat we opslaan en beschikbaar stellen verbruikt stroom. Hoe kunnen we tegengestelde waarden zoals veiligheid van de collectie en belasting voor het milieu tegen elkaar afwegen? Als culturele instelling staan we ook voor publieke waarden. Hoe hebben deze thema's impact op het preservingbeleid?

Op het plaatje zie je een grafiek van het Rapid Assessment Model van de Digital Preservation Coalition zoals ingevuld voor de KB. Dit model is een manier om de volwassenheid van preservingoplossingen in de organisatie te bepalen aan de hand van een aantal onderwerpen. Meer nog dan in de officiële standaarden vanuit preserving zijn in dit model ook maatschappelijke vraagstukken meegenomen zoals ethiek en inclusiviteit. Zelfs al hebben we CoreTrustSeal behaald dan kunnen we aan de score volgens dit model zien dat er nog meer te doen is op thema's zoals toegankelijkheid en ethiek.

Werken aan thema's met maatschappelijke relevantie is ook voor preserving een belangrijke manier om in te spelen op ontwikkelingen en flexibel te blijven. We willen hier op inzetten door deel te nemen aan werkgroepen om zelf te blijven leren en om het perspectief vanuit preserving te kunnen brengen.

4.10 Netwerken



Figuur 20 Screenshot OPF website, KB, 2025

Waar halen we de kennis vandaan?

Bijdragen aan (internationale) samenwerkingsverbanden is belangrijk om de digital preservation community te ondersteunen en een professioneel netwerk op te bouwen. Dit laat zien dat we niet in isolatie bezig zijn met dit onderwerp maar dat we gebruik maken van de kennis en hulpmiddelen die beschikbaar zijn in de verschillende netwerken. Wanneer we bijvoorbeeld meewerken aan het verbeteren van analysesoftware dan leren we zelf meer over bestandsformaten en kunnen we zorgen dat de tools die we gebruiken beter voldoen aan onze eisen.

Maar ook op het gebied van kennisdeling kunnen we van elkaar leren. Op de afbeelding is hier een voorbeeld van te zien. In het strategisch beleid wordt ons uitgangspunt rondom voorkeursformaten verder uitgewerkt. We willen geen formaten verbieden maar meer kennis opdoen over wat we in huis hebben. Maar hoe wordt hierover gedacht in de buitenwereld? Dit hebben we getoetst aan de hand van een paneldiscussie. Hieruit blijkt ook hoe moeilijk het is om dergelijke onderwerpen bespreekbaar te maken. Binnen preserving is het gebruikelijker om richtlijnen te volgen dan om dingen ter discussie te stellen. Toch is dit belangrijk wanneer we verder willen komen binnen dit vakgebied.

Daarom zetten we de komende jaren in op samenwerking in netwerkverband: om onze ideeën te toetsen, om inspiratie op te doen en om kennis te delen.

4.11 Infrastructuur



Figuur 21 Incident bij de KB (afbeelding van intranet KB), KB, 2023

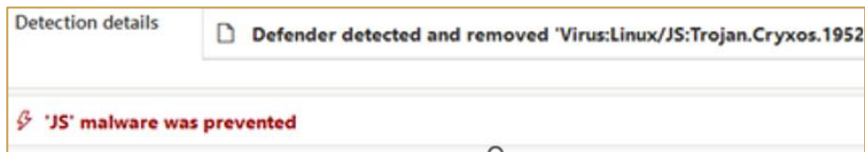
Wat als er iets misgaat?

Niet alles is te voorzien maar voor voorspelbare dingen kunnen we maatregelen treffen. Waar het bijvoorbeeld gaat om technische infrastructuur die faalt of incidenten die plaatsvinden op fysieke locaties zijn er noodmaatregelen gedefinieerd. Maar hoe weten we of deze procedures ook werken?

Op de afbeelding is een voorbeeld te zien waarbij er rookontwikkeling was ontstaan. In dit geval zijn procedures gevolgd om de schade in te perken en in kaart te brengen. Vanuit preservering is het belangrijk dat ook impact op de digitale collectie op deze manier in noodplannen en procedures terecht komt.

Daarnaast is het belangrijk dat er backup- en recoverytesten gebeuren zodat we zeker kunnen stellen dat wanneer er een incident plaatsvindt de collectie met succes hersteld kan worden. Ook willen we de komende jaren beter in kaart brengen welke andere soorten risico's voorzien kunnen worden zodat we hier beter op voorbereid zijn.

4.12 Veiligheid



Figuur 22 Screenshot Windows Defender notificatie, KB, 2025

Is het veilig?

Een duidelijk risico dat al geïdentificeerd is, is het geval van virussen in de collectie. Zoals op de afbeelding te zien is, zijn er gevallen bekend waar de huidige virussoftware malware heeft ontdekt in bepaalde collecties. Tegelijkertijd blijken niet alle meldingen ook echt betrouwbare informatie te geven. Het gaat soms om valse positieven die worden aangemerkt als virus op basis van algemene kenmerken.

We willen de komende jaren dus beter in kaart brengen welke virussen daadwerkelijk aanwezig zijn en ook meer inzicht in welke gevaren dit oplevert voor gebruikers. Wanneer we deze bestanden simpelweg zouden verwijderen dan zouden er problemen optreden met de integriteit en authenticiteit van de collecties. Tegelijkertijd is het natuurlijk niet de bedoeling dat gebruikers daadwerkelijk risico lopen in het gebruik van de collectie. Om deze reden willen we beter in kaart brengen welke maatregelen de juiste balans treffen tussen veiligheid en integriteit.

4.13 Recht

NBN	: urn:nbn:nl-kb:edepot:ewtij:1145529936376 (permalink: http://resolver)
Titel	: Introduction: A perspective on Web Information Retrieval
Rechten	: restricted
NBN	: urn:nbn:nl-kb-1704369916416 (permalink: http://resolver)
Titel	: Introduction: A perspective on Web Information Retrieval
Rechten	: free

Figuur 23 Screenshot ibl.kb.nl, KB, 2025

Wat mogen we ermee?

Tenslotte is het belangrijk dat ook juridische randvoorwaarden meegenomen worden als onderdeel van preservering. Er zijn overeenkomsten waarin afspraken zijn vastgelegd omtrent preservering van het materiaal. Het doel is om deze duurzaam te bewaren als onderdeel van de collectie waar de overeenkomst betrekking op heeft.

Daarnaast is het ook belangrijk dat gebruikers inzage hebben in de rechten die berusten op het materiaal. Zoals te zien is op de afbeelding zijn er soms problemen met de consistentie van de rechteninformatie die wordt getoond. Hetzelfde artikel is in twee vormen te vinden waarbij de status bij het ene item 'restricted' is terwijl de status bij de andere versie 'free' is. Hoe weten we welke informatie juist is? Versiebeleid implementeren zal hierbij ook helpen. De komende jaren willen we vastleggen voor welke collecties de informatievoorziening rondom rechten verbeterd kan worden.

