

An aerial photograph of a Dutch landscape. A river winds through the center, surrounded by green fields and a town with dense housing. The image is partially covered by a dark green overlay on the left and bottom.

Aanlevervoorwaarden

Digitaal archief overheden

e-Depot Midden-Holland

Colofon

Uitgave: © 2024 Streekarchief Midden-Holland (SAMH) te Gouda

Datum: 12 juni 2024

Versie: 1.0

Afbeelding voorzijde:

Luchtfoto met delen van de gemeenten Gouda, Krimpenerwaard, Waddinxveen en Zuidplas (2008).

Bron: NL-GdSAMH 1548 Collectie digitale foto's gemeente Gouda (zonder inv.nr.).

Afbeelding achterzijde:

Schilderijkaart (ca. 1520) met o.a. de plaatsen Gouda, Schoonhoven, Haastrecht, Stolwijk en Berkenwoude.

Bron: NL-GdSAMH Collectie Kartografische documenten (0054.2021).

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand en/of openbaar gemaakt in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of op enige andere manier zonder voorafgaande schriftelijke toestemming van de uitgever.

Actualisatie

Door veranderende wet- en regelgeving en andere inzichten kunnen deze aanlevervoorwaarden aangepast worden. De coördinator Informatiebeheer & Toezicht actualiseert periodiek dit document. Een nieuwe versie wordt gedeeld met de deelnemende gemeenten. Grote wijzigingen worden eerst voorgelegd aan het dagelijks bestuur van het streekarchief.

Versiebeheer

Versie	Opmerkingen	Datum
0.6	Conceptversie	13-11-2023
0.8	Conceptversie na reviewronde	01-12-2023
0.9	Conceptversie na bespreking met het streekarchief	04-12-2023
0.91	Conceptversie na toevoegingen Van Kaliber	21-01-2024
0.92	Schone versie, zonder opmerkingen	01-02-2024
0.93	Schone versie t.b.v. het DB	31-05-2024
1.0	Vastgesteld door de het dagelijks bestuur van het streekarchief	12-06-2024

Begrippen

Afkorting of begrip	
Archiefvormer	Persoon, groep personen of organisatie die zelfstandige archiefvorming als een van zijn of haar activiteiten heeft..
Beheer e-depotvoorziening	Technisch- en databeheer: verantwoordelijkheid van de leverancier (het betreft een SaaS oplossing).
	Functioneel beheer: verantwoordelijkheid van het streekarchief.
	Inhoudelijk beheer op de informatieobjecten, metadata, openbaarheid is de verantwoordelijkheid van het streekarchief. Uitzondering het pre-depot: hier ligt de verantwoordelijkheid van het inhoudelijk beheer bij de archiefvormer.
Digitaal archief	Een samenhangende verzameling van digitale bestanden digitale informatieobjecten, digitale archiefbescheiden, of digitale documenten en metadata, bijvoorbeeld bouwdoSSIers of ruimtelijke plannen.
e-Depot	Het geheel van organisatie, beleid, processen en procedures, financieel beheer, personeel, databeheer, databeveiliging en aanwezige hard- en software, dat duurzaam beheren en raadplegen van te bewaren digitale archiefbescheiden mogelijk maakt.
Hybride archief	Archief dat zowel uit papieren als digitale documenten bestaat.
Ingest	Ingesten betreft het importeren van het digitale archief in het e-depot.
MDTO	MDTO is een afkorting van Metadatagegevens voor duurzaam toegankelijke overheidsinformatie. MDTO is een landelijke standaard, waarin is vastgelegd welke set van kenmerken (metadata) lokale en rijksoverheden tenminste aan informatieobjecten moeten meegeven, zodat informatie in het e-depot straks makkelijk te vinden en eenduidig te interpreteren is. De MDTO-standaard voorziet informatie op eenzelfde manier van metadata, met als doel om informatieprocessen en -systemen op elkaar aan te laten sluiten. Het draagt bij aan goede vindbaarheid en duurzame toegankelijkheid.
Metadata	Metadata zijn gegevens die de context, inhoud, structuur en vorm van informatie en het beheer ervan door de tijd heen beschrijven.
Ongestructureerde data	Data zonder vastgestelde structuur en die niet geordend is op een vastgestelde, vooraf vastgelegde manier. Bijvoorbeeld: e-mail, mappenstructuur, etc.
OPEX	Open Exchange Format (OPEX) is een uitwisselingsformaat in gebruik door Preservica. Digitale archieven die worden opgenomen in het e-depot dienen metadata aan te bieden in het OPEX-formaat om een soepele ingest mogelijk te maken.
Pre-depot	Een pre-depot is bestemd voor uitgeplaatste archieven: dit zijn archieven die niet of nog niet hoeven te worden overgebracht naar een archiefbewaarplaats. Voor de zorgdrager betekent uitplaatsing dat hij verantwoordelijk voor en eigenaar is van archiefinformatie in een voorziening die hij niet zelf beheert. Het inhoudelijk beheer is de verantwoording van de zorgdrager. Het functioneel beheer is de verantwoording van het streekarchief. Het technisch beheer is uitbesteed aan de leverancier van de e-depotvoorziening.
Preservica	Leverancier van de e-depotvoorziening van het streekarchief. Het streekarchief gebruikt Preservica CE als e-depotvoorziening.
Sidecarstructuur	Een bestandsindeling waarbij elk digitaal object (zoals een document, afbeelding, of video) vergezeld gaat van een afzonderlijk metadata-bestand. Dit metadata-bestand wordt het "sidecar-bestand" genoemd en is in een standaard MDTO formaat XML. Het bevat essentiële informatie over het digitale object, zoals de auteur, datum van creatie, formaatdetails, en andere beschrijvende gegevens.

3.

Archief-Kast.

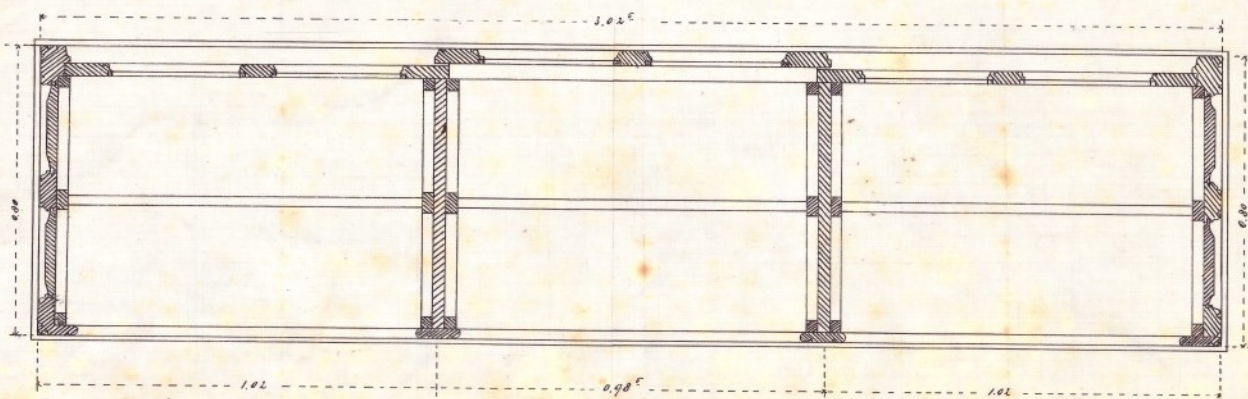
Opstand

61.

505-11-32-1



Plattegrond.



Gouda den 25. Dec 1891
de Opvoeder - Bouwmeester.
P. J. J. J. J.

Schaal 1 à 10

33 1/2

Inhoudsopgave

1.	Inleiding	7
2.	Inhoudelijke voorwaarden	9
2.1	Archief is volledig	10
2.2	Ordering is juist en volledig	10
2.3	Dossiers zijn juist en volledig	10
2.4	Dossiers zijn geschoond	10
2.5	Dossiers zijn afgesloten	10
2.6	Naamgeving is juist, volledig en vrij van persoons- of bedrijfsgegevens	11
2.7	Metadagegevens zijn juist en volledig	11
2.8	Metadata voor 'beperking gebruik' zijn juist en volledig	12
3.	Technische voorwaarden aan bestanden	13
3.1	Virusvrij	14
3.2	Versleutelde bestanden	14
3.3	Beschadigde bestanden	14
3.4	Preserveerbaar bestandsformaat	14
4.	Technische voorwaarden aan metadata	17
4.1	Metadata mapping	18
4.2	MDTO in OPEX	18
4.3	Sectorspecifieke metadata	18
4.4	Aanvullende eisen	19
5.	Technische voorwaarden aan sidecarstructuur	20
5.1	Export uit bronsysteem	21
5.2	Sidecarstructuur	21
5.3	Aanvullende eisen	22

6.	Technische voorwaarden aan de aanlevering.....	23
6.1	Aanlevering van de sidecarstructuur	24
6.2	Aanlevering van de controlegegevens	24
6.3	Aanvullende eisen	24
7.	Uitzonderingen.....	25
7.1	Hybride archief	26
7.2	Ongestructureerde archieven	26
7.3	Pre-depot	26
8.	Begrippenlijst.....	27
8.1	Aggregatieniveaus	28
8.2	EventTypeLijst	28
8.3	Waarderingen	30
8.4	Relatietypen (betrokkene)	30
8.5	Relatietypen (informatieobject)	31
8.6	BeperkingGebruikTypeLijst	31
8.7	ChecksumAlgoritme	31

The background image shows a study room in a chocolate factory archive. In the foreground, there is a wooden desk with a computer monitor, a magazine rack filled with various publications, and a small informational sign. The sign lists opening hours: 'Dinsdag 12.00 - 16.00 uur' and 'Vrijdag 10.00 - 17.00 uur'. Behind the desk, there are tall wooden bookshelves filled with books and documents. A wooden cabinet with the text 'Streekarchief Midden-Holland' is visible in the background. The overall atmosphere is that of a well-organized research space.

I. Inleiding

Hoofdstuk I: Inleiding

Dit document heeft als doel om de vereisten te beschrijven waaraan een digitaal archief moet voldoen wanneer het door een overheidsorganisatie wordt overgebracht of uitgeplaatst naar het e-Depot Midden-Holland, beheerd door het Streekarchief Midden-Holland (SAMH).

Het biedt een overzicht van de criteria en normen die moeten worden gevolgd om ervoor te zorgen dat het digitale archief correct en duurzaam kan worden beheerd in het *e-Depot Midden-Holland*. Er worden voorwaarden gesteld aan de inhoudelijke kwaliteit en aan de technische kwaliteit van het digitaal archief.

Dit document is bedoeld voor de bij het Streekarchief Midden-Holland aangesloten overheidsorganisaties, voor digitaal archief van particulieren/bedrijven gelden andere aanlevervoorwaarden.

De archiefvormer is zelf verantwoordelijk voor een correcte aanlevering van het archief op basis van onderstaande eisen.

Het streekarchief

Het Streekarchief Midden-Holland is het digitale en papieren geheugen van de regio Midden-Holland en bewaart in het Gouwedepot 10 kilometer archieven uit deze regio. Digitale archieven worden duurzaam toegankelijk gemaakt in het e-depot. In de studiezaal in de Chocoladefabriek in Gouda zijn op aanvraag originele bronnen bekijken. Samen met onderzoekers en vrijwilligers maakt het streekarchief bronnen toegankelijk en werkt het aan erfgoedprojecten. Tot slot organiseren we cursussen en activiteiten en maken samen met onze samenwerkingspartners het verleden zichtbaar.

De gemeenten

Het Streekarchief Midden-Holland is een samenwerkingsverband tussen de gemeenten Gouda, Krimpenerwaard, Krimpen aan den IJssel, Waddinxveen en Zuidplas. De Omgevingsdienst Midden-Holland (ODMH), de Regio Midden-Holland (RMH) en Promen maken eveneens deel uit van de samenwerking.

Archieven en collecties

Naast de gemeentelijke archieven bewaren we ook particuliere archieven die veel over de geschiedenis van dit gebied vertellen. Denk aan bedrijfsarchieven, kerkelijke archieven, archieven van scholen, maatschappelijke organisaties en verenigingen. Ook kunt u het verleden ontdekken via onze collecties foto's en prentbriefkaarten, kaarten en plattegronden oorlogsdocumentatie, kranten en tijdschriften. Binnen onze bibliotheekcollectie neemt de Goudse stadslibrje een bijzondere plaats in. Deze historische wetenschappelijke boekencollectie vindt haar oorsprong in het jaar 1594.

2. Inhoudelijke voorwaarden

Hoofdstuk 2: inhoudelijke voorwaarden

De inhoudelijke voorwaarden verwijzen naar de mate waarin de inhoud van het aangeleverde archief, zoals een document, een presentatie, een rapport, of een ander soort, voldoet aan normen of criteria. Bij het aanleveren van archief aan het e-depot dient te worden voldaan het onderstaande. Dit geldt ook voor eventuele latere aanvullingen aan een al in het e-depot opgenomen archief(deel).

2.1 Archief is volledig

Het archief moet volledig zijn. Het is belangrijk ervoor te zorgen dat het archief zo uitgebreid en informatief mogelijk is voor toekomstige raadplegers, zoals burger en historische onderzoekers. Het belang van volledigheid kan variëren afhankelijk van het specifieke doel van het archief en de context waarin deze wordt beheerd.

2.2 Ordening is juist en volledig

Ordening en structuur: archiefstukken moeten georganiseerd zijn volgens een bepaalde structuur of classificatie, zoals zaak, onderwerpenclassificatie of een ander relevant schema. Dit is nodig voor het beheren en het toegankelijk maken. Het MDTO bevat en beschrijft deze elementen.

2.3 Dossiers zijn juist en volledig

Het is van belang dat dossiers juist en volledig zijn. De dossiers kunnen getoetst worden op juistheid en volledigheid.

Juistheid:

Een juist dossier bevat correcte en nauwkeurige informatie. Onjuiste gegevens kunnen leiden tot verkeerde beslissingen.

Volledigheid:

Dossiers moeten alle relevante informatie bevatten die noodzakelijk is/was om op basis van de dossiers de benodigde werkzaamheden uit te voeren. Het dossier kan gebruikt worden als bewijs of controle van dit (overheids-)handelen.

2.4 Dossiers zijn geschoond

Het is van belang dat dossiers geschoond zijn conform de geldende *Selectielijst gemeenten en intergemeentelijke organen* van de Vereniging van Nederlandse Gemeenten (VNG).

2.5 Dossiers zijn afgesloten

Alleen dossiers die zijn afgesloten en waar de besluitvorming is afgerond mogen in het e-depot worden opgenomen.

2.6 Naamgeving is juist en volledig en vrij van persoons- of bedrijfsgegevens

- Beschrijf mapnamen en bestandsnamen betekenisvol en goed.
- Indien mapnamen persoons- of bedrijfsgegevens bevatten is dat afgestemd met de Functionaris gegevensbescherming van zowel zorgdrager als het streekarchief.
- Kies een vaste volgorde voor de opbouw van de bestandsnamen.
- Gebruik de originele extensie en gebruik die termen nooit in de naam.
- Schrijf de datum volgens het stramien JJJJMMDD.
- De tekens < > : " / | \ ? * # & ~ % { } mogen niet in de bestandsnaam voorkomen.
- De letter- en tekencombinaties CON, PRN, AUX, NUL, COM1, COM2, COM3, COM4, COM5, COM6, COM7, COM8, COM9, LTP1, LTP2, LTP3, LTP4, LTP5, LTP6, LTP7, LTP8, LTP9, .lock, _vti_ en desktop.ini mogen niet als bestandsnamen voorkomen.
- Bestands- en padnamen gezamenlijk mogen niet langer zijn dan 256 tekens.

2.7 Metadatagegevens zijn juist en volledig

Voor elk archiefstuk moet gedetailleerde informatie worden verstrekt, zoals de titel, datum van creatie, auteur, onderwerp, en andere relevante gegevens zoals vereist door het MDTO. Metadata spelen een essentiële rol bij het identificeren en beheren van archiefstukken in het e-depot. Het is belangrijk dat de metadata nauwkeurig en duidelijk zijn ingevuld, zonder ongebruikelijke of organisatiespecifieke afkortingen.

Sommige metadata moeten verplicht worden verstrekt volgens de voorschriften van het MDTO en moeten dus altijd aanwezig zijn. Er zijn ook metadata-elementen die "verplicht zijn indien van toepassing". Dit betekent dat ze verplicht zijn wanneer redelijkerwijs kan worden verwacht dat een dossier of bestand deze metadata bevat. Als deze metadata niet beschikbaar zijn in het bronsysteem, moet mogelijk aanvullende informatie uit andere bronnen worden verkregen. Dit is met name relevant voor systemen die content en metadata in verschillende systemen opslaan.

Het kan voorkomen dat niet alle benodigde metadata in het MDTO kunnen worden opgenomen. Dit betreft sectorspecifieke metadata die de algemene structuur van het MDTO overstijgen. In dergelijke gevallen moet in overleg met het streekarchief worden besloten hoe deze metadata in een apart metadata-schema kunnen worden opgenomen.

Voor bepaalde metadata zijn lijstwaarden beschikbaar (voor elk van de velden die een type aanduiden) die verplicht moeten worden gebruikt; deze waarden zijn vastgesteld in het MDTO.

Aanvullende "open" lijsten worden in Hoofdstuk 8, Begrippenlijsten, verder uitgewerkt op basis van de pilotervaringen. Het betreft de onderstaande lijsten:

- Aggregatieniveaus
- Eventtype
- Relatietypen (betrokkenen)
- Relatietype (informatieobject)
- Beperking gebruik
- Checksum Algoritme

2.8 Metadata voor 'Beperking gebruik' zijn juist en volledig

Privacybescherming en intellectueel eigendom van informatie zijn belangrijke onderwerpen die van invloed zijn op de toegang. Dat geldt ook na aanlevering aan het e-depot. De beperking op de openbaarheid die wordt gezien als een tijdelijke uitzondering wordt vastgelegd in een akte die wordt opgemaakt tussen de aanbieder en het streekarchief.

Daarbij is het van belang dat het MDTO metadata element 'Beperking gebruik' in de metadatamapping is opgenomen voor zowel het intellectueel eigendom (auteurs- en portretrecht), privacybescherming, en andere mogelijke beperkingen. Hiermee wordt zeker gesteld dat het bronsysteem deze waarden kan aanleveren.

Een controle op de juistheid en volledigheid van deze gegevens, in het bronsysteem, wordt uitgevoerd.

3. Technische voorwaarden aan bestanden



Hoofdstuk 3: technische voorwaarden aan bestanden

3.1 Virusvrij

Het is van belang dat het aangeleverde archief vrij is van computervirussen.

3.2 Versleutelde bestanden

Hoewel het gebruik van versleuteling binnen de bedrijfsvoering van de zorgdragers essentieel is voor de beveiliging van gevoelige informatie, zoals persoonlijke gegevens en financiële informatie, beperkt dit na aanlevering aan het e-depot de duurzame toegankelijkheid. Daarom mogen bestanden die worden aangeleverd niet versleuteld zijn, ook niet als de encryptiesleutel samen met de gegevens kan worden verstrekt.

3.3 Beschadigde bestanden

Alle over te dragen bestanden moeten kunnen worden geopend en gelezen in een daarvoor juiste applicatie. Bestanden met bestandsgrote 0bytes mogen niet in de collectie aanwezig zijn, omdat dit vaak een indicatie is van een corrupt gemaakt digitaal bestand.

3.4 Preserveerbaar bestandsformaat

Het streekarchief kan alle bestandsformaten die in het PRONOM register staan opslaan in het e-depot. Het preserveeringsbeleid leidt echter tot een kleinere set van voorkeursformaten en geaccepteerde formaten. Hierbij wordt van de volgende uitgangspunten gebruik gemaakt:

- Open bestandsformaten hebben de voorkeur. Dit zijn bestandsformaten die niet van een bepaalde software afhankelijk zijn.
- Het is alleen toegestaan gebruik te maken van een losless compressietechniek op de bestanden in het over te brengen digitaal archief. Losless bestandscompressie is een techniek om de bestandsgrootte van gegevens te verminderen zonder verlies van kwaliteit.
- Duurzame bestandsformaten hebben de voorkeur. Dit zijn bestandsformaten waarvan wordt verwacht dat deze in de toekomst ook nog te openen zijn.
- Indien een bestandsformaat niet open is, maar wel de *de facto* standaard, wordt deze in veel gevallen geaccepteerd. Hierbij is ook in acht genomen of het in Preservica mogelijk is het bestandsformaat te normaliseren of migreren naar een open bestandsformaat.
- Het streekarchief accepteert geen bestanden die in een container (bijvoorbeeld zip) worden aangeleverd.
- Bestandsformaten die niet in het PRONOM-register staan worden niet geaccepteerd.
- Indien een bestandsformaat als 'endangered' is bestempeld in de BitList van de Digital Preservation Coalition wordt deze niet opgenomen in het e-depot, tenzij hierop een normalisatie of migratie kan worden uitgevoerd.

Toepassings-gebied	Voorkeurs-formaat	Geaccepteerd for-maat	Toegestane Pronom ID's
Afbeelding	TIFF, JPEG2000	PNG, JPEG, GIF, DNG	fmt/353 , fmt/153 , fmt/154 , x-fmt/392 , fmt/11 , fmt/12 , fmt/13 , fmt/41 , fmt/42 , fmt/43 , fmt/44 , fmt/3 , fmt/4 , fmt/436 , fmt/152 , fmt/730
Vector afbeelding	SVG		fmt/91 , fmt/92
Tekst	ODF (.odt en .ott), PDF/A-1, PDF/A-2	DOC, DOCX, PDF 1.0-1.7, PDF/X, TXT, EPUB	fmt/136 , fmt/290 , fmt/291 , fmt/39 , x-fmt/275 , x-fmt/274 , fmt/412 , fmt/37 , fmt/38 , fmt/95 , fmt/354 , fmt/476 , fmt/477 , fmt/14 , fmt/15 , fmt/16 , fmt/17 , fmt/18 , fmt/19 , fmt/20 , fmt/276 , fmt/478 , fmt/157 , fmt/158 , fmt/488 , x-fmt/111 , fmt/483
Audio	WAV, WAVE, RF64	MP3, MP1	fmt/1 , fmt/703 , fmt/2 , fmt/707 , fmt/704 , fmt/527 , fmt/712 , fmt/713 , fmt/951 , fmt/6 , fmt/141 , fmt/142 , fmt/143 , fmt/134 , fmt/347
Video	MXF	MKV, mp4 (MPEG-4 Part 10/H.264)	fmt/200 , fmt/786 , fmt/569 , fmt/199 met codec H.264 LIC ID fdd000081
Email	EML	MSG	fmt/278 , x-fmt/430
Website	WARC		fmt/1355 , fmt/1281
Spreadsheet	ODF (.ods, .ots), CSV	XLS, XLSX	fmt/137 , fmt/294 , fmt/295 , x-fmt/18 , fmt/55 , fmt/56 , fmt/57 , fmt/58 , fmt/59 , fmt/61 , fmt/214
Presentatie	ODF (.odp, otp)	PPT, PPTX, PUB, VSD	fmt/138 , fmt/292 , fmt/293 , fmt/126 , fmt/215 , x-fmt/257 , x-fmt/258 , fmt/443
Database	ODF (.odb), SQL, SIARD	ACCDB, MDB	fmt/140 , fmt/444 , fmt/424 , fmt/206 , fmt/161 , fmt/995 , fmt/1196 , fmt/275 , x-fmt/66 , x-fmt/240 , x-fmt/241 , x-fmt/238 , x-fmt/239
Software en code	XML, CSS, JSON, DTD, XSL, XSD, XML, HTML		x-fmt/224 , x-fmt/315 , fmt/817 , x-fmt/280 , x-fmt/281 , fmt/101 , fmt/96 , fmt/97 , fmt/98 , fmt/99 , fmt/100 , fmt/102
3D	GLB, GLTF	X3D, OBJ, E57, 3DM, PLY	fmt/1316 , fmt/1314 , fmt/1315 , fmt/579 , fmt/580 , fmt/581 , fmt/582 , fmt/1210 , fmt/643 , x-fmt/432 , x-fmt/433 , x-fmt/434 , x-fmt/435 , fmt/831
CAD en Geo	GML, KML, Geo-TIFF	DWG, APR, PRJ, ADF, AUX	x-fmt/227 , fmt/244 , fmt/155 , fmt/21 , fmt/22 , fmt/23 , fmt/24 , fmt/25 , fmt/26 , fmt/27 , fmt/28 , fmt/29 , fmt/30 , fmt/31 , fmt/32 , fmt/33 , fmt/34 , fmt/35 , fmt/36 , x-fmt/455 , fmt/434 , fmt/531 , fmt/1395 , x-fmt/317 , fmt/332 , fmt/320 , x-fmt/218 , fmt/1215

* Bestandsformaten die rood zijn kunnen in Preservica worden opgenomen, maar niet worden getoond in de viewer. Dikgedrukte bestandsformaten kunnen opgeslagen worden, maar enkel na migratie in een ander bestandsformaat worden getoond.

Van de meeste bestandsformaten bestaan verschillende versies. Niet elke versie wordt geaccepteerd. Controleer daarom in de tabel welke Pronom ID's aangeboden kunnen worden.

Als een bestandsformaat niet in de bovenstaande tabel is opgenomen, maar wel aanwezig is in het te archiveren archief, is het essentieel dat de archiefvormer contact opneemt met het streekarchief om de mogelijke oplossingen en procedures te bespreken. Het is van groot belang om eventuele onzekerheden en vragen met betrekking tot bestandsformaten proactief aan te pakken om een soepele en succesvolle overdracht naar het e-depot te waarborgen.

Het PRONOM-register en aanvullende informatie zijn beschikbaar op de volgende locatie: www.nationalarchives.gov.uk/PRONOM/Default.aspx. Hier zijn meer details te vinden over bestandsformaten en gerelateerde informatie die nuttig kan zijn bij het archiveren en overdragen van digitale materialen naar het e-depot.

The background of the slide is a faded, grayscale photograph of the Gemeentehuis Waddinxveen. The building is a modern, multi-story structure with large glass windows and a flat roof. In the foreground, there is a paved area, some greenery, and a street lamp. The sky is overcast with clouds.

4. Technische voorwaarden aan metadata

Hoofdstuk 4: technische voorwaarden aan metadata

4.1 Metadata mapping

Voordat digitaal archief naar het e-depot kan worden overgebracht, is het verplicht om een vergelijking/mapping te maken tussen de "MDTO in OPEX"-structuur en de metadata-velden van de archiefdocumenten en dossiers in het bronsysteem. Deze mapping is uniek voor elke collectie en bronsysteem.

Naast de standaard MDTO-velden kunnen ook specifieke metadata van het digitale archief worden opgenomen, zelfs als deze niet in het MDTO zijn gedefinieerd. Het is essentieel dat deze mapping wordt gemaakt en verstrekt als onderdeel van de aanlevering van het digitale archief aan het e-depot.

4.2 MDTO in OPEX

Het streekarchief maakt gebruik van het MDTO (Metagegevens voor Duurzaam Toegankelijke Overheidsinformatie, de opvolger van TMLO) en OPEX (Open Exchange Format) om metadata te importeren in het e-depot. In de OPEX-metadata worden belangrijke gegevens zoals de beveiligingstag, integriteit en ID opgenomen, die essentieel zijn voor een correcte opname in het e-depot. De mapping voor MDTO en OPEX vormt het metadataschema voor het digitale archief.

Meer informatie over MDTO:

- MDTO XML-schema specificatie, versie informatie en xsd: <https://www.nationaalarchief.nl/archiveren/mdto>
- XML-voorbeelden per aggregatieniveau: <https://www.nationaalarchief.nl/archiveren/mdto/xml-schema>

Meer informatie over OPEX:

- Guide to OPEX ingest en Monitoring: [Using OPEX and PAX for Ingesting... | Preservica Developers website](#)
- Bijlage 2 - OPEX xsd: <cdn2.assets-servd.host/preservica-core/production/resources/OPEX-Metadata.xsd>

4.3 Sectorspecifieke metadata

Het kan voorkomen dat niet alle benodigde metadata in het MDTO in te passen zijn. Dit gaat om sectorspecifieke metadata die het generieke karakter van het MDTO overstijgen. In overleg met het streekarchief dienen deze metadata in een separaat metadataschema geplaatst te worden door de aanleverende overheidsorganisatie.

Ook deze metadata moeten in hetzelfde XML-bestand als de OPEX en MDTO metadata geplaatst worden (binnen het element OPEX.descriptiveMetadata).

4.4 Aanvullende eisen

- Bij de aanlevering van archiefbescheiden aan het streekarchief dient van elk archiefstuk een hashcode (fixity) meegegeven te worden in het SHA-256 algoritme, in het OPEX schema én in het MDTO schema.
- Het systeem is hoofdlettergevoelig. De naam van zowel het record/de map als het bijbehorende metadatabestand moeten op exact de wijze zijn gespeld inclusief hoofd- en kleine letters. Dit geldt ook voor de lijstwaarden in het metadataschema.
- De XML sidecarbestanden moeten zijn opgesteld in de tekenset UTF8. Let op dat sommige databases standaard exporteren naar de set “western” en dat bijzondere lettertekens (zoals: â, of ë) dan verkeerd worden weergegeven.
- Het datumformaat moet zijn JJJJ-MM-DD (bijvoorbeeld: 2024-12-23) dit is conform ISO 8601: JJJ-MM-DDThh:mm:ss. Dit is ook gelijk het formaat voor Tijd. Het formaat voor jaar is JJJJ.
- Indien het element <naam> binnen MDTO veel karakters bevat moet bij de mapping de totale lengte van alle mapnamen en bestandsnaam niet de lengte van 256 karakter overschrijden. Dit is een beperking in Windows, die overigens in sommige gevallen, afhankelijk van de versie en andere programmatuur die deze beperking nodig heeft, kan worden uitgezet in de RegEdit van Windows.

The background image shows the modern architecture of the Gemeentehuis Krimpen aan den IJssel. Three flags are visible: the Dutch flag, the Krimpen aan den IJssel municipal flag, and the European Union flag. A wide ramp with glass railings leads up to the building entrance. In the foreground, there is a paved plaza with a bench and some greenery.

5. Technische voorwaarden aan sidecarstructuur

Hoofdstuk 5: technische voorwaarden aan sidecarstructuur

5.1 Export uit het bronsysteem

Bij uitplaatsing en overbrenging moet de aanleverende overheidsorganisatie ervoor zorgen dat het in een verplichte structuur wordt aangeleverd; de "MDTO in OPEX" sidecar-structuur.

Er zijn verschillende manieren waarop dit kan gebeuren. In overleg met het streekarchief worden de mogelijkheden onderzocht en gekozen.

- Het bronsysteem heeft een exportfunctie die het archief geautomatiseerd exporteert in de juiste "MDTO in OPEX" sidecarstructuur.
- Het bronsysteem heeft een exportfunctie die het archief exporteert in het "MDTO" sidecarstructuur. In dit geval zijn er extra stappen nodig om de benodigde "OPEX-schil" te creëren.
- Het bronsysteem heeft geen exportfunctie die het archief in de juiste sidecarstructuur kan exporteren. In dit geval is het meestal mogelijk om de bestanden en de bijbehorende metadata apart aan te leveren. Dit betekent dat de bestanden op een netwerkschijf worden geplaatst en dat de metadata op verschillende niveaus beschikbaar wordt gesteld in Excel-bestanden.

Het is de verantwoordelijkheid van de zorgdrager om het archief “in goede geordende en toegankelijke staat” aan het streekarchief aan te leveren. De aanlevering moet voldoen aan de aanlevervoorwaarden beschreven in dit document. De verplichte sidecarstructuur wordt hieronder toegelicht.

5.2 Sidecarstructuur

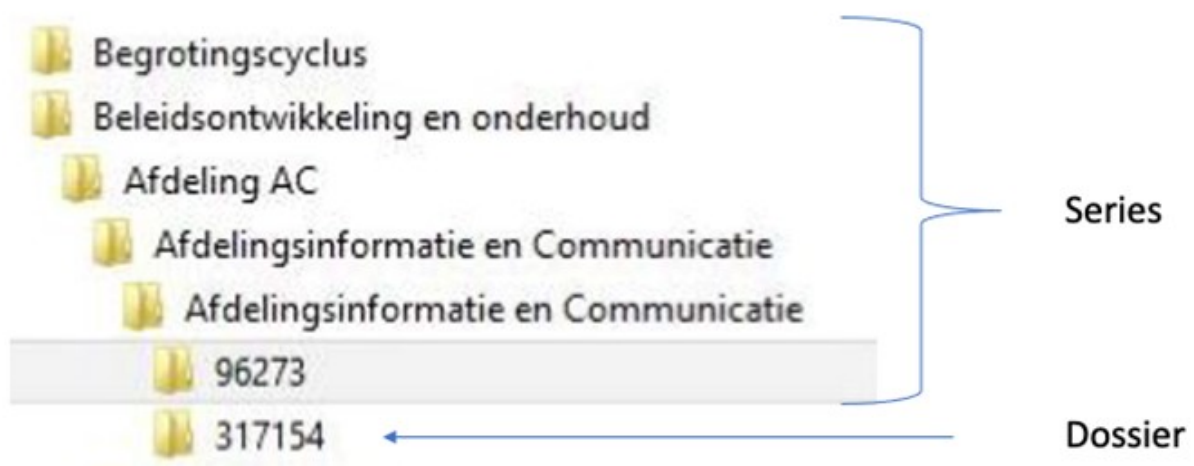
In de sidecarstructuur wordt de inhoud van het digitale archief georganiseerd als een reeks mappen en bestanden, vergelijkbaar met een mappenstructuur. Elke map vertegenwoordigt een niveau in de archiefhiërarchie, zoals een archief, serie, dossier (ook wel zaak genoemd), sub dossier en document. Elke versie van een document wordt als een afzonderlijk bestand behandeld. Binnen deze mappen en bestanden is er een apart metadatabestand, genaamd "sidecar". Het metadatabestand heeft dezelfde naam als de map of het bestand waarbij het behoort, maar het heeft de extensie ".opex" toegevoegd.

Het is belangrijk dat de extensie van het metadatabestand moet overeenkomen met die van het bijbehorende bestand. Bijvoorbeeld, als er een bestand is met de naam "bestand.docx," dan zal het metadatabestand voor dat bestand "bestand.docx.opex" worden genoemd.

Waar het metadatabestand wordt geplaatst, hangt af van of het bij een map of een bestand hoort. Voor mappen bevindt het metadatabestand zich binnen die map, op een lager niveau. Voor bestanden bevindt het metadatabestand zich naast dat bestand, op hetzelfde niveau in dezelfde map.

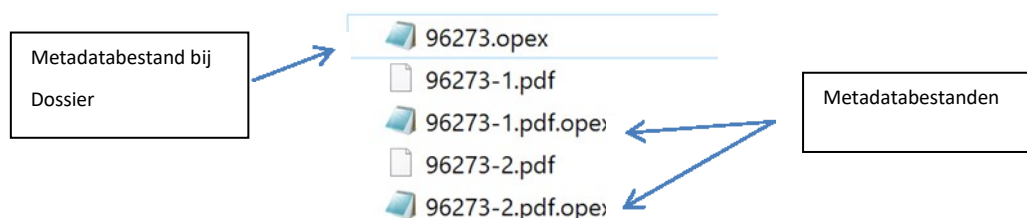
In de Nederlandse archiefwereld worden meestal vijf niveaus van aggregatie gebruikt: archief, serie, dossier en document. Het hoogste niveau is het archief, gevolgd door één of meerdere niveaus van series. Dan komt minstens één niveau van dossiers (zaken), en vervolgens het niveau van records zijn, hoewel dit laatste niveau niet altijd verplicht is.

Dit is hoe de inhoud van een dossier (of map) eruit kan zien in de sidecar-structuur:



Het laagste niveau is het "bestand", wat betekent dat het verwijst naar de technische details van een registratie. Er moeten geen aparte metadatabestanden voor dit niveau aangeleverd te worden, omdat Preservica deze gegevens automatisch maakt wanneer het archief wordt geïmporteerd.

Hier is een voorbeeld van de inhoud van een dossier of map met de naam '96273', zoals te zien in de vorige afbeelding, inclusief bestanden en metadata:



In het "record" (de map met het nummer 96273) staat het volgende:

- Een metadatabestand genaamd '96273.opex' voor het dossier '96273'
- Twee bestanden genaamd '96273-1.pdf' en '96273-2.pdf'
- Bij deze twee bestanden horen metadatabestanden met de namen '96273-1.pdf.opex' en '96273-2.pdf.opex'

5.3 Aanvullende eisen

- Het systeem is hoofdlettergevoelig. De naam van zowel het record/de map als het bijbehorende metadatabestand moeten op exact de wijze zijn gespeld inclusief hoofd- en kleine letters.
- De naam van een metadatabestand is identiek aan die van de map of bestand waar het bij hoort, aangevuld met de extensie '.opex'. De extensie van het bestand moet opgenomen worden in de bestandsnaam van het metadatabestand. Bijvoorbeeld bestand.docx.opex bij het bestand 'bestand.docx'

The background image shows the Gemeentehuis Zuidplas, a modern building with a facade of vertical metal slats and large glass windows. The building is multi-storied and has a contemporary architectural style. A small tree is visible in the foreground to the right of the building.

6. Technische voorwaarden aan de aanlevering

Hoofdstuk 6: technische voorwaarden aan de aanlevering

6.1 Aanlevering van de sidecarstructuur

De archiefvormende overheidsorganisatie is verantwoordelijk voor het veilig versturen van het conform de aanlevervoorwaarden opgestelde digitaal archief naar de IT-omgeving van het streekarchief. De keuze van deze software gebeurt in overleg met het streekarchief, om zo een veilige en betrouwbare verzending te waarborgen. Het streekarchief voorziet in een gebruikersnaam en wachtwoord om een beveiligde verbinding op te zetten.

Toelichting:

De uitwisseling met de transfer-NAS van het e-depot is beveiligd door middel van een SSL certificaat. SSL staat voor Secure Socket Layer en is een beveiligingsprotocol dat zorgt voor een versleutelde verbinding tussen een webserver en een internetbrowser. Dit verzekert dat alle gegevens, die tussen de webserver en browser worden verstuurd, onleesbaar zijn en dus geheim blijven voor tussenliggende partijen. Daarnaast beschikt het systeem over een geo-fence. Alleen Nederlands verkeer wordt toegelaten en daarbij geldt een controle op te veel foutieve logins. Het onjuist aanmelden zal leiden tot een tijdelijke blokkade die alleen opgelost kan worden door de beheerder. Ook is de NAS voorzien van two-factor-authenticatie op de login-credentials. Mocht een wachtwoord en account uitlekken, dan is het niet zomaar mogelijk dat men toch het systeem inkomt. Tot slot zal de informatie die wordt verstrekt in de NAS gecontroleerd worden op virussen alsook een consistentie van de bestandsindeling op het moment van binnenkomst. De verstrekte login-accounts zijn persoonlijk en bieden alleen toegang aan de map(pen) die gekoppeld zijn aan het account. Het is dus niet mogelijk dat andere partijen inzage hebben in de aangeleverde gegevens.

6.2 Aanlevering van de controlegegevens

Bij de aanlevering moeten de volgende, generieke, controle gegevens worden aangeleverd:

- Totale omvang in Bytes
- Aantal mappen, inclusief de hoofdmap
- Aantallen bestanden
- Aantal metadatabestanden (met extensie .opex)
- Aantal, indien toegestaan, op het laagste niveau leeg gelaten mappen (dossiers zonder documenten)
- Metadata-mapping
- Concept acte van overbrenging (of uitplaatsing) en eventueel aanvullende gegevens over de autorisatiestructuur (openbaar, niet openbaar)

6.3 Aanvullende eisen

- Afhankelijk van de omvang van de sidecar is een snelle en stabiele internetverbinding is een vereiste. (Om 250GB per uur te kunnen verzenden in een verbinding nodig van minimaal 560Mbps)
- Sommige besturingssystemen maken verborgen bestanden aan die onbedoeld in de sidecar kunnen staan en worden meegezonden. Deze bestanden maken de aanlevering ongeldig en moeten voor aanlevering worden verwijderd.

7. Uitzonderingen

Hoofdstuk 7: uitzonderingen

7.1 Hybride archief

Een hybride archief is een archiefsysteem dat zowel fysieke (papieren) documenten als digitale bestanden bevat. Het wordt gekoppeld aan het archiefbeheersysteem om het beheer en de toegankelijkheid van archiefmaterialen te vergemakkelijken.

Binnen een hybride archief is het belangrijk om een koppeling te maken tussen de digitale en fysieke dossiers. Dit kan bijvoorbeeld worden gedaan door altijd de archiefcode en het inventarisnummer op te nemen de digitale dossiers. Op deze manier kunnen de fysieke en digitale documenten met elkaar worden verbonden en geïdentificeerd.

Deze verwijzing naar het fysieke archief moet opgenomen worden in het metadataschema (MDTO). Dit kan worden gedaan in het element "Raadpleeglocatie" om gebruikers te informeren over waar het fysieke archief zich bevindt en hoe het kan worden geraadpleegd.

De aanlevering van archiefmaterialen in een hybride archief moet altijd in overleg moet plaatsvinden. Er moeten afspraken worden gemaakt tussen de betrokken partijen over de wijze van aanlevering en de vereiste metadata.

Door een hybride archief op een gestructureerde manier in te richten en rekening te houden met de koppeling tussen fysieke en digitale dossiers, kan een efficiënt beheer en raadpleging van archiefmaterialen worden gewaarborgd.

7.2 Ongestructureerde archieven

Bij ongestructureerde archieven ontbreekt het aan een vastgestelde structuur. Deze archieven zijn niet geordend op een vastgestelde, vooraf vastgelegde manier.

Het aanleveren van archieven in het pre-depot is alleen mogelijk na overleg met het streekarchief. Hier worden aanvullende voorwaarden voor opgesteld.

7.3 Pre-depot

Het is mogelijk om archieven en collecties uit te plaatsen in het e-depot van het streekarchief. Het gaat in dit geval om archieven die niet of nog niet hoeven te worden overgebracht naar een archiefbewaarplaats. Het inhoudelijk beheer is de verantwoording van de zorgdrager. Het functioneel beheer is de verantwoording van het streekarchief. Het technisch beheer is uitbesteed aan de leverancier van de e-depotvoorziening.. In het geval van uitplaatsing wordt gesproken over een *pre-depot*.

Het aanleveren van archieven in het pre-depot is alleen mogelijk na overleg met het streekarchief. Hier worden aanvullende voorwaarden voor opgesteld.

8. Begrippenlijsten

Hoofdstuk 8: Begrippenlijsten

De volgende begrippenlijsten moeten worden gebruikt voor de “labels” binnen de aan te leveren MDTO xml.

Mochten de labels niet voldoen, dan kan in overleg met SAMH een label worden toegevoegd. Dat geldt alleen als de begrippenlijst van het type “open” is.

8.1 Aggregatieniveau

Type begrippenlijst: Open

Het label “Archiefstuk” wordt niet gebruikt.

Label	Definitie
Archief	Geheel van informatieobjecten, ontvangen of opgemaakt door een archiefvormer.
Serie	Verzameling van dossiers, fysieke archiefbestanddelen en/of stukken, numeriek, alfabetisch, chronologische of logisch geordend, ontstaan vanuit een identieke "handeling", dan wel een identieke vorm hebbend dan wel verwante inhoud bevattend.
Dossier	Geheel van fysieke of virtueel gekoppelde informatieobjecten die op één onderwerp betrekking hebben.
Archiefstuk	Enkelvoudig informatieobject of informatie-eenheid. Enkelvoudig wil zeggen dat het stuk niet meer dan één component bevat.

8.2 EventType

Type begrippenlijst: Open

De labels creatie en bevriezing moeten worden aangeleverd, tenzij deze informatie niet beschikbaar is en ook niet uit andere MDTO elementen, eventueel bij benadering, kan worden afgeleid.

Label	Definitie
Creatie	Creatie van een informatieobject door een auteur.
Ontvangst	Ontvangst van een informatieobject door de archiefvormer.
Verzending	Verzending van een informatieobject door de archiefvormer.
Opname	Opname van een informatieobject in een applicatie met de bijbehorende metagegevens. De opname vindt bijvoorbeeld plaats na creatie of ontvangst en wordt gerealiseerd door registratie en opslaan.
Digitalisering	Het scannen van een fysiek informatieobject, waardoor een digitaal informatieobject ontstaat.
Vervanging	Formele vervanging van een informatieobject door een ander informatieobject, waarbij het vervangende informatieobject de plaats en archiefwetelijke status overneemt en het originele informatieobject die plaats en status verliest.
Bevriezing	'Bevriezen' van het Informatieobject, waarna geen wijzigingen meer zijn toegestaan. Voorbeelden zijn afsluiten van dossier of afronden van een tekstdocument.
Conversie	Omzetting van het Informatieobject van het ene naar het andere formaat.
Export	Exporteren van een informatieobject en de metagegevens uit de applicatie.
Import	Importeren van een informatieobject met de metagegevens uit de applicatie.
Kopie	Kopiëren van een informatieobject met de metagegevens binnen de applicatie zodat een nieuw informatieobject bestaat. NB. Unieke metagegevens worden bij een kopie wel gewijzigd.
Migratie	Verplaatsen van het Informatieobject van de ene hard- en/of softwareconfiguratie naar een andere, zonder het formaat te wijzigen.
Vernietigen	Vernietigen van informatie is het blijvend ontoegankelijk maken van die informatie, waardoor deze niet meer vindbaar, beschikbaar, leesbaar, interpreterbaar en betrouwbaar is.
Overbrenging	Formeel overbrengen van het Informatieobject en bijbehorende metagegevens naar een archiefbewaarplaats, waarbij het zorgdragerschap ook wordt overgedragen.
Uitplaatsing	Uitplaatsen van een informatieobject en bijbehorende metagegevens naar een beheeromgeving buiten de eigen organisatie, zonder daarbij het zorgdragerschap over te dragen.
Wijziging	Wijzigen van het Informatieobject of de bijbehorende metagegevens.
Publicatie	Publiceren van het Informatieobject en bijbehorende metagegevens, bijvoorbeeld op een openbare webpagina.
Accordering	Het accorderen van een informatieobject. Dit kan bijvoorbeeld door een digitale handtekening.
Validatie Handtekening	Controle of de digitale handtekening daadwerkelijk door de desbetreffende Actor is gezet.

8.3 Waardering

Type begrippenlijst: Gesloten

Dit is een gesloten lijst aanvullingen zijn niet mogelijk.

Code	Label	Definitie
B	Blijvend te bewaren	Het Informatieobject dient blijvend bewaard te worden.
V	Tijdelijk te bewaren	Het Informatieobject dient tijdelijk bewaard te worden en na afloop van de bewaartermijn vernietigd te worden.
N	Nader te bepalen	Er is mogelijk een waardering. Maar de aard daarvan is niet vastgelegd als type waardering en niet vastgelegd in de meta-gegevens.

8.4 Relatietype (betrokkene)

Type begrippenlijst: Open

8.5 Relatietypen (informatieobject)

Label	Definitie
Belanghebbende	De persoon of organisatie die een belang heeft bij de inhoud van het Informatieobject. Bijvoorbeeld als iemand het onderwerp is van een Informatieobject.
Indiener	Indiener van een verzoek of aanvraag waar het Informatieobject betrekking op heeft.
Rechthebbende	Degene die een wettelijk recht op het Informatieobject kan laten gelden. Zoals auteurs- of portretrecht.

Label	Betekenis
Heeft Versie	Een gerelateerd object dat een versie, editie of een aanpassing is van het beschreven object.
Wordt aangehaald door	Een gerelateerd object dat refereert aan, citeert of op een andere wijze verwijst naar het beschreven object.
Is vervangen door	Een gerelateerd object dat het beschreven object vervangt, verdringt of opvolgt.
Is vereist door	Een gerelateerd object dat het beschreven object nodig heeft ter ondersteuning van de functie, levering of coherentie.
Is een versie van	Een gerelateerd object waarvan het beschreven object een versie, editie of een aanpassing is.
Refereert aan	Een gerelateerd object waarnaar wordt gerefereerd, uit wordt geciteerd of op een andere wijze naar wordt verwezen door het beschreven object.
Vervangt	Een gerelateerd object dat wordt vervangen, verdrongen of opgevolgd door het beschreven object.
Heeft nodig	Een gerelateerd object wat het beschreven object nodig heeft ter ondersteuning van de functie, levering of coherentie.

8.6 BeperkingGebruikType

Type begrippenlijst: Open

Als de beperking wel bekend is maar er is daarvoor geen type gedefinieerd, dan wordt het type 'Overig' gebruikt. In dat geval moet de aard van de beperking volledig in <beperkingGebruikNadereBeschrijving> beschreven worden.

Label	Definitie
Geen beperking	Er rust geen beperking op het gebruik van het informatieobject.
Nader te bepalen	Er is mogelijk een beperking. Maar de aard daarvan is niet vastgelegd als type beperking en niet vastgelegd in de metagegevens.
Overig	Niet nader gespecificeerde beperking. Deze waarde wordt gebruikt als er geen andere geschikt type gedefinieerd is en er in de documentatie wel een beperking is omschreven.

8.7 ChecksumAlgoritme

Type begrippenlijst: Open

Alleen SHA-256 wordt geaccepteerd

Label	Definitie
SHA-256	Cryptografisch hashalgoritme ten behoeve van authenticatie en integriteitscontrole. Zie https://www.forumstandaardisatie.nl/open-standaarden/sha-2



Streekarchief Midden-Holland

Klein Amerika 20 | 2806 CA Gouda

info@samh.nl | <https://www.samh.nl>