

BOM-VL. BEWARING EN ONTSLUITING VAN MULTIMEDIALE DATA IN VLAANDEREN

WERKPAKKET 3 METADATASTANDAARDEN EN UITWISSELINGSFORMATEN

**DEELTAAK 2 GELAAGD METADATAMODEL: BESCHRIJVING MAPPING METADATASTANDAARDEN
NAAR DUBLIN CORE (QUALIFIED)**

Versie 30.06.2009
Auteurs Jan Haspeslagh, mmv. Liesbeth Van Melle
Universiteitsbibliotheek Gent
Sam Coppens
IBBT

Feedback versturen naar:
Inge Van Nieuwerburgh Inge.VanNieuwerburgh@UGent.be
Sam Coppens sam.coppens@ugent.be

BOM-VL WP3: Metadastandaarden en uitwisselingsformaten

Gelaagd metadatamodel: beschrijving mapping van metadastandaarden naar Dublin Core (Qualified)

In WP3, het derde werkpakket binnen het project BOM-VL, wordt een datamodel ontwikkeld voor de langetermijnbewaring van digitale informatie. Voor het beheer van de metadata wordt een gelaagd metadatamodel voorgesteld met een generieke sectoroverschrijdende metadastandaard – (Qualified) Dublin Core – waarnaar de sectorspecifieke metadastandaarden gemapt worden. In WP3 werd gekozen voor de mapping van de volgende metadastandaarden: MARC21 (bibliotheeksector), ISAD(G) en EAD (archieffsector), P/META (audiovisuele sector), CDWA en SPECTRUM (kunstensector en musea).

In dit rapport wordt de generieke standaard (Qualified) Dublin Core beschreven, met zijn voor- en nadelen als overkoepelend metadataschema. Vervolgens worden de sectorspecifieke metadastandaarden bondig toegelicht, met een verwijzing naar de gedetailleerde bronnen voor verdere informatie. Per metadastandaard worden de volgende punten behandeld:

- het volledige mappingschema naar Dublin Core (DC)
- in welke context de velden aan DC gemapt worden en hoe die context in DC getoond wordt
- wat de essentiële (verplichte) velden zijn
- XML-voorbeelden van gemapte records

1. Dublin Core

<http://dublincore.org/>

Dublin Core (DC) voorziet in een eenvoudige gestandaardiseerde set van elementen die 'dingen' zo beschrijven dat ze online terug te vinden zijn. DC wordt zeer veel gebruikt om digitaal materiaal te beschrijven zoals tekst, beelden, video, geluid, websites,... De te beschrijven 'dingen' worden hier verder 'resources' genoemd. Toepassingen van DC maken vaak gebruik van XML- of RDF-structuren.

De DC-standaard bestaat uit twee niveaus:

Simple Dublin Core bevat 15 elementen: contributor, coverage, creator, date, description, format, language, identifier, publisher, relation, rights, source, subject, title en type.

Qualified Dublin Core voegt hier 3 elementen aan toe (audience, provenance en rights holder), en voorziet in een verfijning van de DC-elements: de qualifiers. Het principe bij het specificeren van de DC-elementen (het Dumb-Down Principle) stelt dat een toepassing die de specificatie bij het DC-element niet 'begrijpt', deze qualifier moet kunnen negeren en het DC-element zelf moet kunnen behandelen alsof het een unqualified (breder) element zou zijn.

Qualified Dublin Core bevat bovendien een aantal '**Encoding Schemes**' die een DCElement op een specifieke manier kunnen definiëren. Deze schema's kunnen woordenlijsten zijn, thesauri, notatieregels, lijsten van toegelaten waarden, enzovoort. Een metadatawaarde die in een DCElement van zo'n schema gebruik maakt, kan bijvoorbeeld een trefwoord uit een gecontroleerde woordenlijst zijn, of een formele notatie volgens welbepaalde regels (vb. "2008-12-31" volgens het schema "YYYY-MM-DD"). Als een bepaalde toepassing het schema niet 'begrijpt', blijft de gebruikte metadatawaarde wel nog leesbaar en begrijpelijk voor mensen.

Het onderstaand schema toont de relaties tussen de 15 Simple DC-elementen, de qualifiers en encoding schemes. De website van het Dublin Core Metadata Initiative omvat de gedetailleerde definities van de qualifiers en encoding schemes: <http://dublincore.org/documents/dcmi-terms/>.

DC Element	Element Refinement(s)	Element Encoding Scheme(s)
Title	Alternative	-
Creator	-	-
Subject	-	LCSH MeSH DDC LCC UDC
Description	Table Of Contents Abstract	-
Publisher	-	-
Contributor	-	-
Date	Created Valid Available Issued Modified	DCMI Period W3C-DTF
Type	-	DCMI Type Vocabulary
Format		IMT
	Extent	-
	Medium	
Identifier	-	URI
Source	-	URI
Language	-	ISO 639-2 RFC 1766
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By	URI

	Requires Is Part Of Has Part Is Referenced By References Is Format Of Has Format	
Coverage	Spatial	DCMI Point ISO 3166 DCMI Box TGN
	Temporal	DCMI Period W3C-DTF
Rights		-

De 15 basiselementen en de 3 Qualified DC-elementen worden als volgt gedefinieerd:

15 DC basiselementen

Title: een naam die aan de resource wordt gegeven

Creator: een entiteit die in hoofdzaak verantwoordelijk is voor het creëren van de resource

- Voorbeelden: een persoon, een organisatie, een dienst, ...

Subject: het 'onderwerp', de 'topic' van de resource

- Hierbij wordt vooral gebruik gemaakt van gecontroleerde woordenlijsten, thesauri, classificatiecodes. Indien het onderwerp een context van tijd(sduur) of plaats (geografie) bevat, gebruikt men het element 'Coverage'

Description: een (gedeeltelijke) inhoudelijke/contextuele beschrijving van de resource

- De beschrijving kan bestaan uit (en is niet beperkt tot): een abstract, een inhoudstafel, een grafische voorstelling, vrije tekst, ...

Publisher: een entiteit die verantwoordelijk is voor het beschikbaar stellen van de resource

- Voorbeelden: een persoon, een organisatie, een dienst, ...

Contributor: een entiteit die verantwoordelijk is voor (een) bijdrage(n) aan een resource

- Voorbeelden: een persoon, een organisatie, een dienst, ...

Date: een tijdstip of tijdsduur gelinkt aan een deel van de levenscyclus van de resource

- In dit element kan tijd(sduur) met elke gewenste resolutie uitgedrukt worden. Het gebruik van encoding schemes wordt hier aanbevolen.

Type: de specificatie van het type waartoe de resource behoort

- Het gebruik van encoding schemes wordt hier aanbevolen, met als basisschema het DCMI Type Vocabulary schema (<http://dublincore.org/documents/dcmi-terms/#H7>).
- Voor het beschrijven van bestandsformaten, fysieke media, en dimensies, gebruikt men het element 'Format'.

Format: het beschrijven van bestandsformaat, fysiek medium en/of dimensies van de resource

- Voorbeelden van dimensies zijn afmetingen en tijdsduur. Het gebruik van encoding schemes wordt hier aanbevolen, zoals de lijst van Internet Media Types (MIME): <http://www.iana.org/assignments/media-types/>.

Identificer: een unieke referentie naar de resource in een bepaalde context

- Het wordt aanbevolen om de resource te identificeren met een 'string' volgens een formeel identificatiesysteem (vb. URL, DOI, ISSN, ISBN, ...)

Source: een resource waarvan de beschreven resource afgeleid is

- De beschreven resource kan geheel of gedeeltelijk van de Source afgeleid zijn. Het wordt aanbevolen om ook de Source te identificeren met een 'string' volgens een formeel identificatiesysteem (vb. URL, DOI, ISSN, ISBN, ...)

Language: de taal/talen van de resource

- Het gebruik van encoding schemes wordt hier aanbevolen

Relation: een gerelateerde resource

- Een uitgebreide set van qualifiers beschrijft de verschillende soorten relaties voor dit element: Is Version Of; Has Version; Is Replaced By; Replaces; Is Required By; Requires; Is Part Of; Has Part; Is Referenced By; References; Is Format Of; Has Format.

Voor een volledige definitie van deze qualifiers wordt verwezen naar <http://dublincore.org/documents/dcmi-terms/#H2>.

Coverage: geografische of tijdscontext van de resource, de ruimtelijke toepasselijkheid van de resource, of het gebied waarvoor de resource relevant is

- Bij geografische context en toepasselijkheid kan het gaan om een plaatsnaam of locatie met geocoördinaten. Tijdscontext kan een benoemde (historische) periode zijn, een datum of een date range. Een gebied kan een benoemde administratieve eenheid zijn, of een geografische plaats waarvoor de resource relevant is. Het gebruik van encoding schemes wordt hier aanbevolen, zoals het Universal Transverse Mercator-stelsel (UTM) of de Thesaurus of Geographic Names (http://www.getty.edu/research/conducting_research/vocabularies/tan/index.html)

Rights: informatie over de rechten met betrekking tot de resource

- Dit element bevat de beschrijving van alle mogelijke gebruiks- en toegangsrechten of restricties voor de resource, met inbegrip van intellectueel eigendomsrecht.

3 Qualified DC elementen:

Audience: een entiteit voor wie de resource bedoeld of bruikbaar is

Provenance: een beschrijving van alle wijzigingen in eigendom en/of beheer van de resource sinds zijn creatie, die van belang zijn voor de authenticiteit, integriteit en interpretatie van de resource

- Dit element kan beschrijvingen bevatten over alle veranderingen die opeenvolgende beheerders aan de resource hebben aangebracht.

Rights Holder: een persoon of organisatie die de rechten over de resource bezit en/of beheert.

Opmerkingen bij het gebruik van Dublin Core (Qualified) als overkoepelende metadatastandaard:

- Het element dc:type kan ingezet worden om semantische problemen bij verschillende materiaalsoorten en resource types te voorkomen, door gebruik te maken van specifieke semantische sets bij de types uit de DCMIType-thesaurus (vb. bij dc:type Text, wordt de Contributor role 'auteur'; bij dc:type Image wordt de Contributor role 'schilder', 'fotograaf', ...). Die semantische sets worden dan ontleend aan de velddefinities en bijhorende authority lists van de verschillende gemapte metadatastandaarden.
- Het element dc:relation heeft de nodige qualifiers om de soms complexe relaties in de gemapte metadatastandaarden op te vangen.
- De 'encoding scheme' feature laat toe om sectorspecifieke thesauri, authority files en woordenlijsten te gebruiken, zodat verdere detaillering, standaardisering en uniformiteit van de invulling en doorzoekbaarheid mogelijk zijn.
- Er kunnen bijkomende namespaces gedefinieerd worden waardoor DC (Qualified) met andere metadatastandaarden gecombineerd kan worden (zie <http://dublincore.org/documents/dcmi-namespace/index.shtml>).
- XML en RDF-syntax is mogelijk. DCMI heeft hiervoor uitgebreide guidelines ontwikkeld: <http://dublincore.org/schemas/>.
- Vanwege de algemene definities van elementen en qualifiers bestaat het gevaar voor een verkeerde invulling of foutieve interpretatie van metadata voor verschillende resources binnen een record (bijvoorbeeld een record waarin een originele foto samen met de TIFF-file beschreven wordt: vermenging van dimensiemetadata voor het digitale object (TIFF-file) met dimensiegegevens van het origineel (papieren foto) in het element dc:format). Oplossingen hiervoor liggen opnieuw in het intelligent gebruik van het element dc:type in combinatie met de specifieke encoding schemes.
- Er lijkt wat gebrek te zijn aan specificaties (te weinig attributen, geen mogelijkheid tot annotaties). Indien de invulling van de metadata in het sectorspecifieke metadataschema correct en gedetailleerd genoeg is, wordt de informatie via een solide mapping naar Dublin Core toch opgehaald en getoond.
- Voor het element dc:identificer ontbreekt een encoding scheme voor de toepassingen in het BOM-VI-project. De ingevulde identificer metadata moeten dus nog gekoppeld worden aan een eenduidige, gestandaardiseerde beschrijving van alle sectorspecifieke identificerschema's die in gebruik zijn.

2. Mapping van de sectorspecifieke metadatastandaarden

Algemeen werd gestreefd naar een mapping van alle velden/elementen van de verschillende standaarden naar de DC (Qualified) elementen. De reden hiervoor is dat het binnen elke sector mogelijk moet blijven om metadata in te voeren volgens de eigen specifieke vereisten en granulariteit. Indien men enkel zou opteren voor het mappen van DC-elementen naar specifieke core-velden of elementen binnen iedere sectorstandaard, dan zou er onherroepelijk informatie verloren gaan (onvolledige invulling, verkeerde interpretatie bij de keuze van het invoerveld, ...). De metadata die uiteindelijk via de mapping in DC-records terechtkomen (en getoond worden), zijn weliswaar minimaal maar zullen altijd naar de volledig ingevulde set in het oorspronkelijke record gelinkt kunnen worden. Voor iedere afzonderlijke mapping worden wel een aantal essentiële of core-elementen/velden aangeduid, die samen in het DC-record nog voldoende informatie bevatten om het record te 'begrijpen' en terug te linken naar de volledige oorspronkelijke set.

Waar relevant wordt ook de context van sectorspecifieke elementen toegelicht, en hoe deze context 'vertaald' werd naar het gemapte DC-element. Dit is nodig om begripsverwarring te vermijden bij generieke termen (zoals titel, contributor, format, identifier,) die per sector een zeer verschillende context of invulling kunnen hebben.

a. MARC21: MACHine Readable Cataloging

Volledige versie: <http://www.loc.gov/marc/>

Inleiding

MARC21 is een van de meest gebruikte standaardensets voor het maken van catalogusrecords, die veelal (maar niet uitsluitend) items beschrijven uit bibliotheekcollecties. De standaardenset dient ook als basisschema voor de veld- en relatiestructuur in veel catalogussoftware en databases die literatuurgegevens bevatten. MARC21 kan in XML-schema's weergegeven worden (www.loc.gov/marc/marcdtd/marconv_xml.zip) en kan zo geconverteerd worden naar bijvoorbeeld DC en MODS.

MARC 21 heeft 10 hoofdvelden, met daaronder talloze subvelden, attributen en indicatoren, wat de standaardenset als geheel zeer complex maakt. In de MARC-standaardenset worden 5 types records gedefinieerd: Bibliographic, Holdings, Authority, Classification en Community Information. Elk record bestaat uit een Leader (vaste structuur: 24 karakterposities), een Directory (met de tag, startpositie en lengte van elk veld), en de recordinhoud, verdeeld over Control fields en Data fields. Elke tag bestaat uit een 3-cijfercode. Het zijn deze codes die in het mappingschema gebruikt worden.

Mapping

Alle MARC21-codes werden gemapt aan DC-elementen. De complexiteit zorgt ervoor dat de mapping tot op het niveau van de DC-qualifiers gebeurd is. Sommige MARC21-velden hebben in combinatie met een subveld en/of indicator een andere mapping dan het algemene MARC21-veld.

Vb. MARC21 veld 788 → DC-element dc:subject
MARC21 veld 788 if ind2=2 → DC-element dc:subject + qualifier MeSH

Mappingschema



Opmerkingen

Essentiële velden

MARC21 heeft een zeer fijne resolutie (dus een groot aantal mogelijke MARC21codes) voor alle 18 elementen uit DC Qualified. Enkel voor dc:title zijn er al 21 keuzemogelijkheden in MARC21! Het heeft dus weinig zin om een richtlijn op te stellen met een aantal essentiële MARC-codes die zeker ingevuld moeten zijn.

Omgekeerd kan wel gesteld worden dat op zijn minst voor de volgende DC-elementen een corresponderend MARC21-veld gegevens zou moeten bevatten. Over welk MARC21-veld dit dan gaat, wordt bepaald door de lokale procedures:

- dc:identificer: laat toe het oorspronkelijke MARC-record ten allen tijde te traceren via de unieke ID(s)
- dc:type: gebaseerd op de DCMI Type Vocabulary, laat toe om de onderliggende aanwezige metadata correct te interpreteren (zie het voorbeeld van het record met gegevens over een papieren foto en een TIFF-file)
- dc:date: met gegevens over creation, update, delete status van records

Gebruik van Qualifiers en Encoding schemes

De fijne resolutie van MARC21 records verantwoordt in hoge mate het gebruik van de qualifiers en encoding schemes. Dit geldt in het bijzonder voor:

- dc:format waar de metadata via dc:type (en de bijhorende encoding schemes per type) correct gemapt kunnen worden
- dc:identificer en dc:source: qualifiers en encoding schemes zijn aanwezig/moeten toegevoegd worden aan DC voor DOI, ISBN, ISSN, handles
- dc:relation
- dc:subject waar, naast de reeds aanwezige 'encoding schemes', sectorspecifieke schema's voor bijkomende thesauri, woordenlijsten en classificaties toegevoegd kunnen worden.

b. EAD (Encoded Archival Description)

Volledige versie: <http://www.loc.gov/ead/>

Inleiding

EAD is een datamodel (dtd of schema), gebaseerd op SGML maar omgezet in XML, dat ontwikkeld werd voor het maken, opslaan, publiceren, koppelen en uitwisselen van archiefbeschrijvingen. De standaard wordt internationaal onderhouden en kent verschillende tools:

<http://www.archivists.org/saagroups/ead/tools.html>.

Interessant is de link tussen EAD en ISAD(G) (zie verder): elk ISAD(G) beschrijvingselement heeft 'tags' in EAD, waardoor ISAD (G) aan EAD gemapt kan worden.

Voorbeelden van EAD-records: http://www.loc.gov/ead/tolib/appendix_ca.html.

Mapping: voor deze oefening werden de 146 EAD-elementen rechtstreeks aan DC gemapt, behalve een beperkt aantal *formele elementen* die enkel gebruikt worden voor presentatiedoeleinden in EAD.

Mappingschema



Opmerkingen

Formele elementen

EAD is een XML-uitwisselingsformaat. De onderliggende metadata zijn in tekstformaat (meestal Unicode) opgenomen, zodat de data in veel verschillende vormen gepresenteerd kunnen worden (Word, pdf, html, ...) via bijvoorbeeld stylesheets. Daarom werden de formele elementen van EAD niet gemapt.

Datamodel vs. data content

EAD is een datamodel. Dat betekent dat EAD geen regels voorschrijft over hoe en in welke vorm gegevens in de verschillende velden opgenomen worden. EAD definieert alleen de 146 elementen zelf, en vermeldt welk(e) gegeven(s) in dit element kan (kunnen) voorkomen en wat de mogelijke hiërarchische relaties met andere elementen (niet) kunnen zijn. Vervolgens zetten de archieven zelf het systeem in volgens de eigen noden en procedures.

Een voorbeeld: het EAD-element <archdesc>Archival Description is een 'wrapper' element dat de inhoud, context en omvang van een bepaald archief(onderdeel) beschrijft. De informatie binnen dit element is hiërarchisch georganiseerd, met bijvoorbeeld een beschrijving van het geheel (<archdesc><did><dsc>), en ook beschrijvingen van onderdelen (<archdesc><did><archdescgrp><dsc>). Wat precies in de respectieve <dsc> elementen opgenomen is, wordt o.a. bepaald door de eventuele aanwezigheid van attributen (class, collection, series, fonds, ...). Er zijn heel veel combinaties mogelijk, waardoor het quasi onmogelijk is om een inhoudelijk correcte mapping te voorzien voor alle mogelijke gevallen van metadata-invulling. Een uitgangspunt van het gelaagd metadataamodel is dat de objecten op zich beschreven kunnen worden. In het geval van EAD zal dit omwille van bovenstaande reden tot foute mappings leiden. Het mappen van de EAD-elementen die gebruikt worden om archieven (en grote onderdelen) en collecties als geheel te beschrijven, zal wel tot een goed resultaat kunnen leiden.

EAD en ISAD(G)

In de Vlaamse archiefwereld is EAD op dit moment nog niet echt doorgedrongen. Hier wordt de standaard ISAD(G) gebruikt. Omdat de mapping van EAD naar DC complex is er een ISAD(G)-EAD crosswalk, houdt men best ISAD(G) als standaard voor metadata-invulling aan en, afhankelijk van hoe de partner het organiseert, maakt men gebruik van de crosswalk ISAD(G)-EAD en de mapping naar DC, of van de rechtstreekse ISAD(G)-DC crosswalk (zie onder).

c. ISAD(G): Algemene Internationale Norm voor Archivistisch Beschrijven (General International Standard Archival Description)

Volledige versie: www.ica.org/sites/default/files/isad_g_2e.pdf (Engels)
www.archiefbank.be/doc/isad_NL2000.pdf (Nederlands)

ISAD(G) is een norm die algemene richtlijnen verschaft voor het maken van archivalistische beschrijvingen. Deze ISAD(G)-set van algemene regels zorgt voor consistente archiefrecords die opzoekbaar zijn, uitwisselbaar en integreerbaar binnen een gemeenschappelijk informatiesysteem. De regels worden vastgelegd en beschreven in 26 elementen die gecombineerd kunnen worden om de beschrijving van een archiefbestanddeel te vormen.

Hoewel de nummering/codering van deze elementen niet gestandaardiseerd is, wordt toch aangeraden om de codes 3.1.1. t.e.m. 3.7.3. op te nemen in de toepassingen die ISAD(G) gebruiken.

Mapping: Er werd gekozen voor een rechtstreekse mapping ISAD(G) naar DC. Een alternatief is om de ISAD(G) to EAD crosswalk te gebruiken (http://www.loc.gov/ead/tqlib/appendix_a.html) en vervolgens de EAD to DC mapping.

Mappingschema



Opmerkingen

Essentiële velden

De ISAD(G) standaard definieert zelf 6 essentiële velden:

- 3.1.1.Reference code (dc:identifier)
- 3.1.2.Title (dc:title)
- 3.1.3.Date(s) (dc:date)
- 3.1.4.Level of Description (dc:type)
- 3.1.5.Extent and Medium of the Unit of Description (dc:format)
- 3.2.6.Name of Creator (dc:creator)

Alle 6 de elementen zijn contextueel eenduidig gemapt aan DC-elementen zodat elk correct ingevuld ISAD(G)-record zonder metadata-verlies kan uitgewisseld worden met Dublin Core.

ISAD(G) en EAD

Zoals hoger beschreven kan ISAD(G) als 'tussenstap' gebruikt worden om mappingproblemen tussen EAD en DC te voorkomen.

d. CDWA: CATEGORIES FOR THE DESCRIPTION OF WORKS OF ART

Volledige versie: http://www.getty.edu/research/conducting_research/standards/cdwa/

CDWA is een conceptueel metadataschema voor het beschrijven en ontsluiten van gegevens over kunst, architectuur, andere materiële kunstuitingen, kunst- en beeldcollecties. Het is niet gebonden aan specifieke platformen, databases of software.

CDWA bestaat uit 31 categoriegroepen met 510 categorieën. Hierbinnen zijn een beperkt aantal *core* categorieën geselecteerd, die samen het minimale, essentiële record definiëren. Voor dit core recordschema werd CDWA Lite ontwikkeld

(http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.pdf); CDWA Lite records kunnen ingepast worden in OAI repositories.

Mapping: alle 510 categorieën werden 1 op 1 gemapt. De core categorieën werden met (core) aangeduid, en kunnen dus als aparte groep behandeld worden: CDWA Lite.

Mappingsschema



Opmerkingen

Essentiële velden

CDWA beschrijft zelf de core elementen (CDWA Lite). Het is dus aan te bevelen deze groep elementen (als 'core' aangeduid in de mappingsheet) zeker in te vullen. Contextueel kunnen alle CDWA Lite core elementen eenduidig gemapt worden aan DC Qualified elementen, zodat ook hier geen metadata-verlies zou mogen optreden bij uitwisseling. Een bijkomend argument is de beschikbaarheid van een OAI-compatibel XML-schema, ontwikkeld voor CDWA Lite.

e. P/META Metadata Exchange Scheme

Volledige versie: www.ebu.ch/CMSimages/en/tec_doc_t3295_v0102_tcm6-40957.pdf

P/Meta is een set van definities voor de uitwisseling van informatie over audiovisueel materiaal. Deze set identificeert en beschrijft audiovisuele concepten, onderverdeeld in P/Meta Names en P/Meta Identifiers. Een P/Meta informatieset bestaat uit drie lagen:

- De Definition layer
- De Technology layer
- De Data interchange layer

P/Meta definieert uitsluitend de verschillende informatie-elementen uit de Definition layer. Dit laat de verschillende partijen (producenten, zenders, archieven, klanten) toe om bij de uitwisseling van de media hetzelfde 'woordenboek' te gebruiken.

P/Meta definieert de volgende 5 'exchange concepts', met daarnaast een aantal facultatieve concepten:

- Programma
- Programmagroep
- Item of Programma Item
- Media Object
- Merk ('Brand')
- (Andere concepten: Personen, Organisaties, ...)

Deze concepten worden verder beschreven met attributen (met authority lists voor elk soort attribuut). Verder worden in P/Meta ook P/Meta sets gedefinieerd, bestaande uit P/Meta attributen die met andere P/Meta sets gegroepeerd worden. Deze voorgedefinieerde sets vormen de bouwstenen voor de meest voorkomende informatieuitwisselingen, maar er kunnen ook heel specifieke P/Meta sets voor speciale uitwisselingen geschreven worden.

Mapping: de 'elements' en 'complex types' werden allemaal afzonderlijk gemapt aan DC. Wellicht worden nooit alle 'elements' bij data exchange gebruikt. Voor P/Meta kan gekozen worden om alleen enkele essentiële 'elements' binnen de 5 'exchange concepts' over te houden voor het gelaagd model.

Mappingschema



Opmerkingen

P-META exchange concepts en DC elementen

De 5 genoemde exchange concepts zijn niet eenvoudig te mappen aan een DCElement omdat er weinig overlap is tussen de P-META context van de exchange concepts en de DC context.

Eenzijds kan 'm Element: Programme Group' naar dc:type gemapt worden, maar dit is niet dezelfde conceptuele link als tussen 'm Element: Classification' en dc:type. Anderzijds kan 'm Element: Programme Group' ook naar dc:title gemapt worden, maar dat verengt dan weer de context van het P-META element te veel.

In deze mapping werd gekozen om het 'title'-element voor de 5 P-META exchange concepts als 'essentiële elementen' te beschouwen, en de hiërarchisch hogere exchange concepts zelf te mappen naar dc:description. De richtlijn luidt dan dat in een P-META record altijd de combinatie van het element 'Exchange concept' met het corresponderende element 'title' ingevuld is.

Andere essentiële elementen

P-Meta is bij uitstek een uitwisselingsstandaard waarbij het zeer duidelijk moet zijn wàt precies uitgewisseld wordt. Naast de informatie over de Exchange concepts zelf, zijn de volgende core elementen nodig:

- m Element: Classification (dc:type)
- m Element: ... Title (dc:title) * op de plaats van ... komt één van de exchange concepts

- m Element: Identifier (dc:identifier)
- m Element: OrganisationIdentifier (dc:identifier)
- m Element: ProgrammeGroupidentifier (dc:identifier)
- m Element: Programmeidentifier (dc:identifier)
- m Element: Itemidentifier (dc:identifier)
- m Element: MediaObjectidentifier (dc:identifier)

Het is hierbij niet noodzakelijk dat op elk niveau identifiers ingevuld zijn. Dit is uiteraard afhankelijk van het exchange niveau waarop effectief uitwisseling plaatsvindt.

f. SPECTRUM: UK Museum Documentation Standard

Volledige versie: <http://www.mda.org.uk/spectrum.htm>

SPECTRUM is een door de Britse MDA beheerde standaard voor het professionaliseren van de museale bedrijfsvoering. Alle handelingen met betrekking tot museumstukken, van verwerving tot en met expositie, zijn vervat in 21 procedures. Naast de 21 procedures bevat SPECTRUM ook een overzicht van alle informatie die in het museum vastgelegd moet worden, om de procedures goed te kunnen toepassen. SPECTRUM is dus in de eerste plaats een set richtlijnen.

Elk stukje informatie (informatie-eenheid) wordt in vijf punten toegelicht:

- Definitie van de informatie
- Hoe moet de informatie vastgelegd worden?
- Voorbeelden
- Gebruik van de informatie (bij welke procedures?)
- Tot welke groep van informatie-eenheden behoort het?

SPECTRUM kan toegepast worden in software voor museaal collectiebeheer maar is zelf geen softwarepakket. Het is de bedoeling dat leveranciers van museale software gebruik maken van SPECTRUM, zodat hun software zo goed mogelijk ruimte biedt voor het vastleggen van de benodigde informatie.

Mapping: uit de praktijk blijkt dat een bepaald aantal informatie-eenheden volstaat om de meeste museale handelingen correct te beschrijven. Alleen deze 'selected elements' werden gemapt aan DC. De volledige lijst 'units of information' werd ter informatie bijgevoegd.

Mappingschema



Opmerkingen

Context SPECTRUM vs. context Dublin Core

Aanvankelijk werd Dublin Core ontwikkeld als een schema dat toelaat om digitale online 'resources' via gestructureerde metadata te documenteren en terug te vinden. Alle hierboven behandelde metadatastandaarden hadden te maken met informatie over 'fysieke' resources en hun 'content'. SPECTRUM beheert echter de informatie over handelingen met betrekking tot 'fysieke' resources waarbij de handelingen belangrijker zijn dan de resources (museumstukken) zelf. Er zijn heel veel elementen in SPECTRUM die verwijzen naar begrippen zoals acquisition, entry, disposal, conservation, ownership, exhibition, access, loan, location, movement, handling, place, procedure, request, use, return, enzovoort. Dat die informatie essentieel zal zijn bij het invullen van SPECTRUM-records is evident.

Voor een goede mapping naar Dublin Core volstaat het de elementen uit het SPECTRUM-record die naar de resource (het museumstuk) zelf verwijzen over te nemen. Dit gedeelte van een SPECTRUM-record heeft grotendeels dezelfde context als een Dublin Core elementenset en kan dus perfect gemapt worden. De essentiële elementen voor een correcte mapping (dus niet de essentiële SPECTRUM velden per se) zijn:

- Object number (dc:identifier)
- Title (dc:title)
- Object production date (dc:date)
- Record type (dc:type)
- Unit of information scheme (dc:type)
- Recording date (dc:date)
- Process date (dc:date)

3. Enkele Mapping Voorbeelden:

Deze sectie toont een aantal voorbeelden van records die gemapt zijn naar een Dublin Core RDF beschrijving. Deze mappings zijn gebaseerd op de ontwikkelde mappingschema's. Dergelijke mapping bestaat uit twee stappen. De eerste stap is een mapping van de beschrijving naar RDF. De meeste records zijn nog beschreven in XML of een ander formaat. Deze records dienen eerst gemapt te worden naar RDF. Eens deze beschrijving er is, kan deze worden ingevoerd in een triplestore, wat de tweede stap in het mappingproces is. Hierbij wordt gebruik gemaakt van het Dublin Core schema, waarin de RDF properties gemapt (van b.v. EAD, MARC21, CDWA) worden naar Dublin Core properties teneinde een Dublin Core beschrijving te krijgen van het record. Deze mapping wordt bewerkstelligd door voor elke Dublin Core property een equivalente property te definiëren van de te mappen record beschrijving.

a. MARC21: MACHine Readable Cataloging

Origineel:

```

040      DPL  ^c DPL
099      TMD-808
049      DPLW
100 1    Rucker, Richard A.
245 10 Mountain Training Group trip from Leadville to Aspen, Colorado ^h [graphic]
260      ^c February 1944.
300      1 photoprint ; ^c 20 x 28 cm. (8 x 11 in.)
300      1 phototransparency : ^b Kodachrome slide ; ^c 35mm.
440 0    10th Mountain Division Resource Center Collection ; ^p Slide Collection
500      For further information see 10th Mountain Division Resource Center manuscripts
500      LA1 "M.T.G." trip (31/2 days) Leadville to Aspen (Feb '44) handwritten on original
slide mount.
500      Kodak 828 slide format.
500      Title supplied by cataloger.
520      A line of about twenty skiers climbs up a slope. The men are members of the Tenth
Mountain Division, Mountain Training Group during winter training exercise designed to test
the men's readiness for winter conditions in the mountains by having them ski from Leadville
to Aspen, Colorado. According to the photographer's caption, this was probably taken on the
morning of the second day near Mount Champion.
530      Image file: ^d ZZR700200828
540      Copyright restrictions applying to use or reproduction of this image available from
the Western History/Genealogy Dept., Denver Public Library.
650 7    Soldiers ^x Colorado ^y 1940-1950. ^2 lctgm
650 7    Military training ^z Colorado ^y 1940-1950. ^2 lctgm
650 7    Ski troops ^z Colorado ^y 1940-1950. ^2 lctgm
610 10   United States. ^b Army. ^b Mountain Division, 10th.
655 7    Snapshots. ^2 gmqpc
  
```

Gemapte Dublin Core beschrijving:

```

<?xml version="1.0"?><rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc= "http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://www.example.com/MARC">
  <dc:title> Mountain Training Group trip from Leadville to Aspen, Colorado <\dc:title>
  <dc:description> Kodak 828 slide format.</dc:description>
  <dc:description> Title supplied by cataloger.</dc:description>
  <dc:description> For further information see 10th Mountain Division Resource Center
manuscripts.</dc:description>
  <dc:description> LA1 "M.T.G." trip (31/2 days) Leadville to Aspen (Feb '44) handwritten on
original slide mount.</dc:description>
  <dc:description> A line of about twenty skiers climbs up a slope. The men are members of the
Tenth Mountain Division, Mountain Training Group during winter training exercise designed to
test the men's readiness for winter conditions in the mountains by having them ski from
  
```

Leadville to Aspen, Colorado. According to the photographer's caption, this was probably taken on the morning of the second day near Mount Champion.</dc:description>

```
<dc:description> Image file.</dc:description>
<dc:description> Copyright restrictions applying to use or reproduction of this image
available from the Western History/Genealogy Dept., Denver Public Library.</dc:description>
<dc:subject> Soldiers </dc:subject>
<dc:subject> Military training </dc:subject>
<dc:subject> Ski troops </dc:subject>
<dc:subject> United States</dc:subject>
<dc:subject> Army </dc:subject>
<dc:subject> Mountain Division, 10th.</dc:subject>
<dc:subject> Snapshots</dc:subject>
<dc:contributor> Rucker, Richard A </dc:contributor>
<dc:date> February 1944</dc:date>
<dc:type>Snapshots</dc:type>
<dc:format>20 x 28 cm. (8 x 11 in.)</dc:format>
<dc:format>35mm</dc:format>
</rdf:Description>
</rdf:RDF>
```

b. EAD (Encoded Archival Description)

Origineel:

```
<?xml version="1.0"?>
<?xml-stylesheet href="document.xsl" type="text/xsl"?>
<!DOCTYPE ead SYSTEM "http://ead.lib.virginia.edu/bin/dtd/vivaead/ead.dtd" [
<!ENTITY % notats SYSTEM "http://ead.lib.virginia.edu/bin/dtd/vivaead/eadnotat.ent"> %notats;
<!ENTITY logo SYSTEM "http://ead.lib.virginia.edu/vivaead/logos/uva-sc.jpg" NDATA jpeg>
<!ENTITY conditions SYSTEM "http://www.lib.virginia.edu/speccol/vhp/conditions.html" NDATA
html>
<!ENTITY address SYSTEM "http://ead.lib.virginia.edu/vivaead/add_con/uva-sc_address.xml">
<!ENTITY contact SYSTEM "http://ead.lib.virginia.edu/vivaead/add_con/uva-sc_contact.xml">
]><ead id="viu01401">
<eadheader audience="internal" langencoding="ISO639-2" findaidstatus="edited-full-draft">
<eadid type="SGML catalog">PUBLIC "-//University of Virginia::Library::Special Collections
Dept.//TEXT (US::ViU::viu01401::A Guide a a Thomas Jefferson Letter to Edward Turpin)//EN"
"viu01401.xml"
</eadid>
<filedesc>
<titlestmt>
<titleproper>A Guide to a Thomas Jefferson Letter to Edward A. Turpin<date> 1826 January
10</date>
</titleproper>
<subtitle id="sort">Jefferson, Thomas Letter, 1826 January 10
<num type="collectionnumber">11135
</num>
</subtitle>
<author>Special Collections Staff
</author>
</titlestmt>
<publicationstmt>
<publisher>Special Collections Department, University of Virginia Library
</publisher> &address;
<date type="publication">&copy; 2001 By the Rector and Visitors of the University of Virginia.
All rights reserved.
</date>
<p id="usestatement">
<extref entityref="conditions">Conditions of Use
</extref>
</p>
<p id="filesize">15 kb
</p>
</publicationstmt>
</filedesc>
<profiledesc>
<creation>Machine-readable finding aid derived from MS Word 2000, created by Sue Cuthbertson,
<date> 5/16/2001</date>
</creation>
<language>Description is in
<language>English
</language>
</language>
</profiledesc>
<revisiondesc>
<change>
```



```

<date>2001
</date>
<item>No Changes.
</item>
</change>
</revisiondesc>
</eadheader>
<frontmatter>
<titlepage>
<titleproper>A Guide to a Thomas Jefferson Letter to Edward A. Turpin<date> 1826 January
10</date>
</titleproper>
<subtitle>A Collection in <lb/>The Special Collections Department
<num type="Accession Number">11135
</num>
</subtitle>
<p id="logostmt">
<extptr actuate="auto" show="embed" entityref="logo"/>
</p>
<publisher>Special Collections Department, University of Virginia Library
</publisher>
<date type="publication">2001
</date> &contact;
<list type="deflist">
<defitem>
<label>Processed by:
</label>
<item>Special Collections Department
</item>
</defitem>
<defitem>
<label>Funding:
</label>
<item>Web version of the finding aid funded in part by a grant from the National Endowment for
the Humanities.
</item>
</defitem>
</list>
</titlepage>
</frontmatter>
<archdesc level="collection" langmaterial="eng">
<runner placement="footer">Special Collections, University of Virginia Library
</runner>
<did>
<head>Descriptive Summary
</head>
<unittitle>Thomas Jefferson Letter to Edward A. Turpin
<unitdate type="single" label="Date">1826 January 10
</unitdate>
</unittitle>
<unitid label="Accession number">11135
</unitid>
<physdesc label="Physical Characteristics">One 1 page letter.
</physdesc>
<repository>Special Collections, University of Virginia Library
</repository>
<abstract label="Abstract">Jefferson discusses the students and professors at the University
of Virginia, encloses a copy of the University laws and terms of board and tuition, and closes
with his family's
remembrances and best wishes.
</abstract>
</did>
<admininfo>
<head>ADMINISTRATIVE INFORMATION
</head>
<p></p>
<accessrestrict>
<head>Access
</head>
<p>There are no restrictions.
</p>
</accessrestrict>
<userrestrict>
<head>Use Restrictions
</head>
<p>There are no restrictions.
</p>

```

```

</userrestrict>
<prefercite>
<head>Preferred Citation
</head>
<p>Thomas Jefferson Letter to Edward A. Turpin, 1826 January 10, Accession #11135, Special
Collections Dept., University of Virginia Library, Charlottesville, Va.
</p>
</prefercite>
<acqinfo>
<head>Acquisition Information
</head>
<p>This letter was purchased by the University of Virginia Library from C.G. Sloan &
Company, North Bethesda, Maryland, on January 13, 1994.
</p>
</acqinfo>
<processinfo>
<head>Processing Information
</head>
<p> The letter has been professionally restored but the damaged beginning portion of the
letter is still very difficult to read.
</p>
</processinfo>
</admininfo>
<scopecontent>
<head>Scope and Content Information
</head>
<p>In this one page autograph signed letter, January 10, 1826, Thomas Jefferson, Monticello,
Virginia, writes to Edward A. Turpin, enclosing a copy of the laws and construction of the
University of Virginia and a
notice of the terms of board and tuition, and expressing disappointment at not seeing Turpin
during his return to Albemarle County in the spring.
</p>
<p>Jefferson's comments show the obvious pleasure he took in both professors and students at
his University, "Could you now visit our neighborhood you would see our University in a very
different state from what
it was when you were here. We have been very fortunate in the selection of the Professors
which we received from abroad. They are men of the highest qualifications in point of science,
and in character correct and
amiable. Before the vacation we had 120 students, at first a little unruly, but soon set to
rights and latterly perfect examples of good order and application."
</p>
</scopecontent>
</archdesc>
</ead>

```

Gemapte Dublin Core beschrijving:

```

<?xml version="1.0"?><rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
<rdf:Description rdf:about="http://www.example.com/EAD">
  <dc:title> A Guide to a Thomas Jefferson Letter to Edward A. Turpin </dc:title>
  <dc:title>Jefferson, Thomas Letter, 1826 January 10<\dc:title>
  <dc:subject> Thomas Jefferson Letter to Edward A. Turpin, 1826 January 10, Accession #11135,
Special Collections Dept., University of Virginia Library, Charlottesville, Va.<\dc:subject>
  <dc:description> Special Collections, University of Virginia Library.</dc:description>
  <dc:description> Jefferson discusses the students and professors at the University of
Virginia, encloses a copy of the University laws and terms of board and tuition, and closes
with his family's remembrances and best wishes.</dc:description>
  <dc:description> Machine-readable finding aid derived from MS Word 2000, created by Sue
Cuthbertson </dc:description>
  <dc:description> Jefferson's comments show the obvious pleasure he took in both professors
and students at his University, "Could you now visit our neighborhood you would see our
University in a very different state from what
it was when you were here. We have been very fortunate in the selection of the Professors
which we received from abroad. They are men of the highest qualifications in point of science,
and in character correct and
amiable. Before the vacation we had 120 students, at first a little unruly, but soon set to
rights and latterly perfect examples of good order and application." </dc:description>
  <dc:description> In this one page autograph signed letter, January 10, 1826, Thomas
Jefferson, Monticello, Virginia, writes to Edward A. Turpin, enclosing a copy of the laws and
construction of the University of Virginia and a
notice of the terms of board and tuition, and expressing disappointment at not seeing Turpin
during his return to Albemarle County in the spring. </dc:description>
  <dc:publisher> Special Collections Department, University of Virginia<\dc:publisher>

```



```
<dc:contributor> Special Collections Staff </dc:contributor>  
<dc:date> 5/16/2001</dc:date>  
<dc:identifier> viu01401</dc:identifier>  
<dc:language> English</dc:language>  
<dc:rights> There are no restrictions</dc:rights>  
<dc:format> One 1 page letter.</dc:format>  
</rdf:Description>  
</rdf:RDF>
```

Achtergronddocumentatie:

Digitaal Erfgoed Nederland (2007). Encoded Archival Description (EAD). Versie 1.0.
<http://www.den.nl/docs/20070521104422/>

Digitaal Erfgoed Nederland (2007). Spectrum. Versie 1.
<http://www.den.nl/docs/20060713133105/>

Digitaal Erfgoed Nederland (2008). Dublin Core in samenwerkingsprojecten en publieksgerichte ontsluiting. Versie 1.1. <http://www.den.nl/docs/20050816173630>

Elings, M.W.; Waibel, G. (2007). Metadata for all: Descriptive standards and metadata sharing across libraries, archives and museums. *First Monday*, vol. 12(3)-1628.
<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1628>

Getty Research Institute (2009). Metadata Standards Crosswalk.
http://www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html

Hopper, R. (2000). EBU Project Group P/META Metadata Exchange Standards – Technical Review.
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.6.8540>

Shepherd, E.; Pringle, R. (2002). Mapping descriptive standards across domains: a comparison of ISAD(G) and SPECTRUM. *Journal of the Society of Archivists*, Vol. 23(1): 17-34