



## Technische uitleg dataset EDBO

Deze pagina geeft uitleg over en voorbeelden van de

- Beschrijvende metadata (dcx),
- Structurele metadata (mpeg21-didl),
- Full-text objecten (pdf, ocr-xml, alto-xml, jpg, jpf)
- Harvest-API o.b.v. OAI -PMH
- Zoek-API o.b.v. SRU

voor de set EDBO (<http://www.kb.nl/banners-apis-en-meer/dataservices-apis/early-dutch-books-online>) die de Koninklijke Bibliotheek als open data aanbiedt.

Om snel aan de slag te kunnen met EDBO bieden we een [Snelstart EDBO \(pdf\)](#) aan. Ook kan het nuttig zijn de [Snelstart SGD \(pdf\)](#) te raadplegen om meer inzicht te krijgen hoe SRU-queries binnen de KB-infrastructuur opgebouwd moeten worden.

### Voorwaarden hergebruik & bronvermelding EDBO

Op de website van EDBO zijn de [algemene gebruiksvoorwaarden](#) gemeld.

Samengevat: individuele items en niet-substantiële delen van de gehele website zijn vrij, het databankrecht wordt voorbehouden voor substantiële delen of de gehele website.

Verder gelden de onderstaande bepalingen:

1. De metadata van de EDBO-set zijn beschikbaar onder een [CC0-licentie](#).
2. Individuele full-text items en een niet-substantieel deel van deze set zijn te gebruiken onder een [Public Domain Mark](#) ;
3. Voor de gehele set full-text items of substantiële delen ervan geldt:
  - voor niet-commerciële doelen mag alles worden hergebruikt.
  - voor commerciële doelen mogen ook de gehele deelcollecties van de Universiteitsbibliotheek Amsterdam en de KB worden gebruikt;
  - de Leidse deelcollectie of substantiële delen daarvan mogen alleen voor niet-commerciële doelen worden hergebruikt.

### *Bronvermelding*

Bronvermelding is wenselijk: Vermeld bij (her)gebruik van de hele set, of substantiële delen daarvan, de naam "Early Dutch Books Online", gevolgd door de namen van de deelnemende bibliotheken.

Ook wordt het op prijs gesteld als de naam van de bibliotheek waar het object uit afkomstig is, vermeld wordt (metadata: `<dc:identifier xsi:type="shelfmark">`):

- Amsterdam, Universiteitsbibliotheek

- Den Haag, Koninklijke Bibliotheek
- Leiden, Universiteitsbibliotheek

## Beschrijvende metadata

Iedere boektitel in EDBO is van een aantal beschrijvende metagegevens voorzien, de tabel hieronder beschrijft de velden

Let op: de naam van de set binnen de KB-infrastructuur is DPO (dus niet EDBO)

Voorbeeldrecords	
<ul style="list-style-type: none"> <li>• <a href="http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:141982977">http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:141982977</a></li> <li>• <a href="http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:182873757">http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:182873757</a> (meerdere titel)</li> <li>• <a href="http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:141105593">http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:141105593</a></li> </ul>	
<dc:title>	Titel van het boek (of de serie boeken)
<dcx:annotation>	Annotatie van het boek (of de serie boeken)
<dc:language>	Taal van het boek (of de serie boeken)
<dcterms:extent>	Aantal delen in deze titel
<dc:date>	Jaar van uitgave
<dc:creator>	Auteur
<dc:contributor>	Andere auteur(s)
<dc:subject>	Onderwerpstrefwoord
<dc:publisher>	Drukker/Uitgever, plaats, jaren actief
<dcx:metadataSource>	Bron van de beschrijvende metadata STCN – Short Title Catalogue Netherlands – <a href="http://www.kb.nl/stcn/index-en.html">http://www.kb.nl/stcn/index-en.html</a>
<dc:identifier xsi:type="shelfmark">	Exemplaar. dit is de plaats waar het origineel te vinden is. Dit wordt aangeduid met de bezittende instelling en het nummer waaronder het origineel bij die instelling opgevraagd kan worden. Dit is bijvoorbeeld een catalogusnummer of inventarisnummer
<dcx:recordIdentifier dcx:type="PPN">	PPN – Pica Productie Nummer: dit is het identificatienummer van de object in de algemene catalogus (Gemeenschappelijk Geautomatiseerd Catalogussysteem, GGC)
Met deze PPN kan het beschrijvende metadatarecord aangeropen worden via de volgende syntax <a href="http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:&lt;PPN&gt;">http://services.kb.nl/mdo/oai?verb=GetRecord&amp;identifier=DPO:DPO:&lt;PPN&gt;</a> , waarbij <PPN> dan vervangen wordt door bv 141982977	

## Structurele metadata & objecten

De beschrijvende metadata van een object met *identifier* DPO:DPO:141105593 is op te vragen via

<http://services.kb.nl/mdo/oai?verb=GetRecord&identifier=DPO:DPO:141105593> (het laatste getal is het unieke PPN van deze titel)

In het record staat in het veld <dcterms:hasFormat xsi:type="dcterms:URI" dcx:anchorText="Full text EDBO"> een link naar <http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21>

Met behulp van de identifier *dpo:3686* kunnen we nu de webpagina's, structurele metadata en objecten afleiden

#### Webpagina's

- *Webpagina – beeldweergave document, voorkant*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21> (persistent) of  
<http://www.earlydutchbooksonline.nl/nl/view/image/id/dpo:3686:mpeg21> (niet-persistent)
- *Webpagina – beeldweergave document, 7e pagina*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21:0007> (persistent) of  
[www.earlydutchbooksonline.nl/nl/view/image/id/dpo:3686:mpeg21:0007](http://www.earlydutchbooksonline.nl/nl/view/image/id/dpo:3686:mpeg21:0007) (niet-persistent)
- *Webpagina – tekstweergave van document, 7e pagina*  
<http://www.earlydutchbooksonline.nl/nl/view/text/id/dpo:3686:mpeg21:0007> (niet-persistent)

#### Structurele metadata (*mpeg21-didl*)

- <http://services.kb.nl/mdo/oai?verb=GetRecord&identificer=DPO:dpo:3686:mpeg21&metadataPrefix=didl>  
In dit structurele metadata record staat in het veld <dcx:recordIdentifier xsi:type="dcx:PPN"> ook weer de PPN "141105593" vermeld. Hiermee is via de syntax <http://services.kb.nl/mdo/oai?verb=GetRecord&identificer=DPO:DPO:<PPN>> dan weer het beschrijvende record op te vragen

#### Afbeeldingen

- *JPEG2000 (.jpf) van pagina 7 van document*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21:0007:image>
- *Mid-size JPG van pagina 7 van document*  
<http://imageviewer.kb.nl/ImagingService/imagingService?&id=dpo:3686:mpeg21:0007:image>
- *Thumb JPG van pagina 7 (300px hoog, 180 graden geroteerd)*  
<http://imageviewer.kb.nl/ImagingService/imagingService?r=180&h=300&id=dpo:3686:mpeg21:0007:image>

#### Tekstobjecten

- *PDF hele document*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21:pdf>
- *OCR/TXT van pagina 7 van document*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21:0007:ocr>
- *ALTO woordcoördinaten van pagina 7 van document*  
<http://resolver.kb.nl/resolve?urn=dpo:3686:mpeg21:0007:alto>

De URLs die beginnen met *http://resolver.kb.nl* zijn persistent en moeten bij voorkeur gebruikt worden in verwijzingen.

#### Metadata harvesten via de download-API (OAI-PMH)

De EDBO-dataset heeft een download-API o.b.v. het [OAI-PMH-protocol](#). Hiermee kun je zowel de beschrijvende als de structurele metadata harvesten m.b.v. een OAI-harvester

Let op: de naam van de set binnen de KB-infrastructuur is DPO (dus niet EDBO)

1. De OAI-PMH-syntax is hoofdlettergevoelig!
2. Beschrijving KB OAI-repository : <http://services.kb.nl/mdo/oai?verb=Identify>
3. Voor het harvesten van de eerste 400 records beschrijvende metadata:  
<http://services.kb.nl/mdo/oai?verb=ListRecords&set=DPO&metadataPrefix=dcx> (let op: DPO dus met hoofdletters!)
4. *ResumptionToken*: met de laatste XML-tag in deze respons (de ResumptionToken) - DPO!2011-01-28T15:08:08.209Z!!dcx!0 - kan de volgende batch van 400 records opgevraagd worden:  
<http://services.kb.nl/mdo/oai?verb=ListRecords&resumptionToken=DPO!2011-01-28T15:08:08.209Z!!dcx!0> (metadataPrefix=dcx hoeft dan niet meer in de query string aangegeven te worden.) Op die manier kan iteratief de hele set beschrijvende metadata overgehaald worden.
5. *Structurele metadata* : kan op vergelijkbare manier geharvest worden door metadataPrefix=didl te kiezen:  
<http://services.kb.nl/mdo/oai?verb=ListRecords&set=DPO&metadataPrefix=didl>
6. *Identifiers*: met <http://services.kb.nl/mdo/oai?verb=ListIdentifiers&set=DPO&metadataPrefix=dcx> (of metadataPrefix=didl voor structurele metadata) kan een lijst van de identifiers worden opgevraagd.

## **Dataset doorzoeken via de zoek-API (SRU)**

De EDBO-dataset heeft ook een zoek-API o.b.v. het SRU-protocol. Hiermee kun je de beschrijvende metadata doorzoeken. SRU is gebaseerd op CQL.

Let op: de naam van de set binnen de KB-infrastructuur is DPO (dus niet EDBO)

Enkele voorbeelden van SRU-queries voor EDBO zijn:

- Zoek naar alle publicaties uitgegeven tussen 1760 en 1791. Let op de URL-encoding!  
[http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&recordSchema=ddd&x-collection=DPO\\_boekdeel&query=\(date%20within%20%221760%201791%22\)](http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&recordSchema=ddd&x-collection=DPO_boekdeel&query=(date%20within%20%221760%201791%22))
- Er is ook een dcx recordschema beschikbaar. Dit dcx recordschema eist wat meer van de SRU-server, dus vriendelijk verzoek dit alleen te gebruiken wanneer het ddd recordschema niet voldoet.  
[http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&recordSchema=dcx&x-collection=DPO\\_boekdeel&query=\(date within "1760 1791"\)](http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&recordSchema=dcx&x-collection=DPO_boekdeel&query=(date%20within%20%221760%201791%22))
- Zoek in de abstracts van de 'DPO\_pagina' index voor de term 'water'.  
[http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO\\_pagina&query=water](http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO_pagina&query=water)
- Zoek naar publicaties die de term 'water' in de titel hebben  
[http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO\\_boekdeel&query=dc.title ANY "water"](http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO_boekdeel&query=dc.title ANY )
- Als voorgaande, maar nu beperkt tot publicaties uit 1760 t/m 1791.  
[http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO\\_boekdeel&query=dc.title ANY "water" AND \(date within "1760 1791"\)](http://jsru.kb.nl/sru/sru?version=1.2&maximumRecords=50&operation=searchRetrieve&startRecord=1&x-collection=DPO_boekdeel&query=dc.title ANY )

Zie tevens [Snelstart EDBO \(pdf\)](#) en de [Snelstart SGD \(pdf\)](#) voor meer uitleg over & meer voorbeelden van SRU-queries binnen de KB-infrastructuur

## **Vragen of opmerkingen over de EDBO-set?**

Mail naar [dataservices@kb.nl](mailto:dataservices@kb.nl)

*Laatste update : 21-11-2012*