



APEF – DTR cooperation

3 October 2016



What is the DTR project?

The project Digitale Taken Rijk (DTR) is a program launched by the Dutch Ministry of Education, Culture and Science to build a digital infrastructure and service organisation for the Dutch National Archives and 11 Dutch Regional Historical Centers (RHC's), former State Archives in the Dutch provinces, during the period: 2013 – 2016

- WP1 – infrastructure
- WP2 – service organisation
- WP3 – digitisation
- WP4 – sustainability



What is the DTR project?

Various projects, for instance for realising a collective:

- Trusted Digital Repository (e-Depot, based on Preservica)
- Collection Management System (MAIS-Flexis)
- Authority file for record creators (Actorenregister)

and ...

... aggregated search through the collections ...



What is the APEF – DTR connection?

For the infrastructure DTR wants to make use of the aggregator function of the Archives Portal Europe.

At first it was the idea to just re-use the code of the Archives Portal Europe and to further develop this “branch“ separately, adding extra functionality needed for DTR purposes.

But, after comparing the DTR and APEF development roadmaps and feature request lists, it was decided to further develop the Archives Portal Europe together, concentrating on three tasks which are important for both parties.



Main tasks of APEF – DTR cooperation

1. Building an **API** to open up the functionality of the Archives Portal Europe to third parties and to implement the possibility for the content providers to switch this option on or off in the Archives Portal Europe's back-end
2. Improving the uploading and processing of **EAC-CPF records** in the Archives Portal Europe and to improve the way in which this records creators information can be linked with other files and can be searched and found via the Archives Portal Europe's front-end and via the API
3. Enabling the uploading/processing of “**Additional Finding Aids** (indexes)” in the Archives Portal Europe and make this additional information available via the Archives portal Europe's front-end and the API



Ad 1: why developing an API?

APEF:

- to be able to easily share data and a search functionality with multiple third parties (CENDARI, EHRI, Europeana Cloud, etc.)
- to foster the use of Open Data

DTR:

- to be able to present (Dutch) Archives Portal Europe data in multiple (12) websites with their own look & feel

<https://www.youtube.com/watch?v=s7wmiS2mSXY>



Ad 2: why improve EAC-CPF handling?

APEF:

- to be able to go beyond institution based linking of records creators to finding aids → *better relations management*
- to be able to eventually create a central European archives records creators authority file system

DTR:

- to be able to upload Dutch government records creators data from a central system (Actorenregister) and make them searchable via the Archives Portal Europe API



Ad 3: why get “additional“ FA’s in?

APEF:

- to be able to add more and meaningful metadata to descriptive units and make them searchable, which will make archival material more accessible to a wider audience, which can eventually lead to better quality metadata and Linked Open Data

DTR:

- to be able to add more and meaningful metadata to descriptive units and make them searchable, which will make archival material more accessible to a wider audience, which can eventually lead to better quality metadata and Linked Open Data



What has been achieved? → API

October 2016: nine APE services on production:

- **for EAD:**
 - a simple search facility to provide basic information of a search result from finding aids and their descriptive units
 - a possibility to fetch more detailed information of a specific finding aid
 - a possibility to fetch more detailed information of a specific descriptive unit
- **for EAC-CPF:**
 - a simple search facility to provide basic information of a search result from archival records creators, including filtering information
 - a possibility to fetch all details of a specific archival records creator, including filtering information



What has been achieved? → API

October 2016: nine APE services on production:

- **Extra:**
 - a possibility to fetch a list of all archival institutions that have enabled Open Data in the Archives Portal Europe's Dashboard and thus give access to their data-sets via the API
 - a possibility to fetch a list of all finding aids of one of these archival institutions
 - a possibility to download finding aids of these archival institutions as raw data (EAD/XML)
 - a possibility to download archives records creators information of these archival institutions as raw data (EAC-CPF/XML)

http://wiki.archivesportaleurope.net/index.php/APE_API_documentation

http://wiki.archivesportaleurope.net/index.php/Using_My_Pages



APE API Explorer
showing how it
technically works

What has been achieved? → API

TOKEN

API REFERENCE

- content
- download
- institute
- search

Search for EAC-CPF

Search for EAD

version

Archives Portal Europe

This is a list of usable API services available from Archives Portal Europe. Here you can test these services, but to do so you need API key. Please register at [APE server](#) to get your API key.

Search for EAD

POST /search/ead

Parameters

body (required)

Content type: application/vnd.ape-v2+json

SearchRequest Search EAD units Count should not be more than 50

Test this endpoint

TRY

Response Type: application/vnd.ape-v2+json

Response Messages

RESPONSE SAMPLE

RESPONSE SCHEMA

BODY SAMPLE

```
{
  "count": "5",
  "startIndex": 0,
  "query": "*",
  "filters": [
    {
      "facetFieldName": "string",
      "facetFieldIds": [
        "string"
      ]
    }
  ],
  "dateFilters": [
    {
      "dateFieldName": "string",
      "dateFieldId": "string"
    }
  ]
}
```

BODY SCHEMA

<https://www.archivesportaleurope.net/ApeApi/#!/search/search>



Enable/Disable
use of the API in the
APE's Dashboard

What has been achieved? → Dashboard



Content provider information

[View/modify EAG file](#)

[Upload EAG file](#)

[Download EAG file](#)

[Manage open data for API](#)

[Change institutions name](#)

[Set feedback e-mail address](#)

Content information

[Preferences for uploaded files](#)

[Create automatic harvesting function](#)

[Upload content](#)

[Create EAC-CPF](#)

[Content manager](#)

[Manage topic mappings](#)

Manage open data for API

Open data flag is enabled.

This means that at the moment your data will not only be published via the interface of the Archives Portal Europe, but also made available as "open data" under CC0 license to be used by anyone via the Archives Portal Europe's API

In case you want to change that: check mark the option "Disable open data" below and click on the Submit-button. Then all your current data and all data which you will ingest in the Archives Portal Europe in the future will not be made available as "open data" anymore via the Archives Portal Europe's API.

Disable open data

Submit



What lies ahead – API

- Improving the existing nine API services
- Adding new services for providing detailed information from Additional Finding Aids, in particular on persons
- Deploying iterative new releases of the API services after the improvements and the implementation of the services for Additional Finding Aids



What lies ahead – EAC-CPF handling

Currently in Archives Portal Europe it is only possible to:

- upload/harvest EAC-CPF records per institution
- manually connect EAC-CPF records to EAD files of more than one institution (*and only when EAC-CPF files are created within the Dashboard*)

Currently it is not possible to:

- recognise double records for a specific CPF entity
- establish relations between double records and their linked EAD files



What lies ahead – EAC-CPF handling

- We want to deal with the “upload per institution only“-concept for all files (so for EAC-CPF as well as EAD files):
 - more than one OAI-PMH endpoint
 - more than one OAI-PMH set
 - adapt the current Dashboard workflow if necessary
- We don't want to create unique apeEAC-CPF “umbrella“ records (yet), but we will start improving the relationship between records first (-> *decision of the APEF Standards Working Group*)



What lies ahead – Additional Finding Aids

What are additional finding aids exactly?

Most of the time the descriptive units of a finding aid describe the archival material very generally, although there is a wealth of information available in them, f.i. on events, persons, places, subjects, etc.

Often this information is available, but in a variety of formats: handwritten lists, indexes, card systems, word processor or spreadsheet documents, databases, etc., either made when the material was created or later on by archivists and/or volunteers.

The variety of this kind of information and the variety of the ways in which it is stored, makes it hard to determine a metadata format for it.



What lies ahead – Additional Finding Aids

It's worthwhile to explore how to use this rich information.

The DTR project is going to pilot with two types of additional finding aids: Notary records (from the 11 RHC's) and VOC administrative records and ships logs (from the Nationaal Archief).

Since these mainly contain data on people, the DTR project thought it could make use of the metadata format A2A/XML which is created for a Dutch genealogical portal (<http://www.den.nl/standaard/386/>).

However, then the Archives Portal Europe should “learn” how to deal with this new (Dutch) metadata format and this might not be a suitable format for all sorts of other additional finding aids.



What lies ahead – Additional Finding Aids

Why not use EAD/XML for additional finding aids?

After all: it *is* “finding aid“-information and it can be put in the original finding aid, f.i. as an index within the descriptive unit.

But that will result in huge EAD/XML files which will probably generate performance problems in the EAD/XML files maintenance tools.

But we dealt with this before: when we tried to give access to all scans of digitised archival material via a finding aid in EAD/XML → then we decided to link EAD/XML files to METS/XML files per descriptive unit.

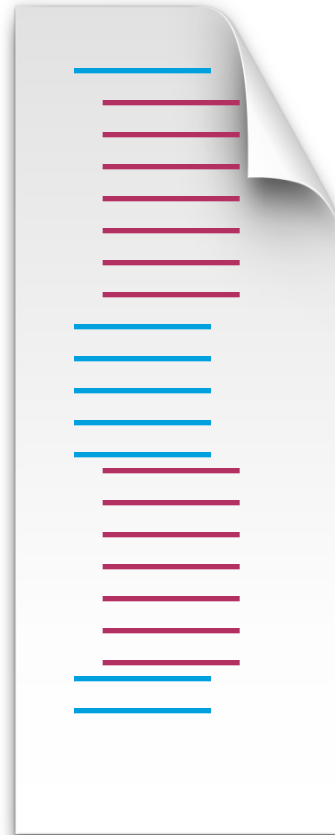
We can do the same with Additional Finding Aids: link the descriptive units of the original EAD finding aid to another (additional) EAD finding aid.



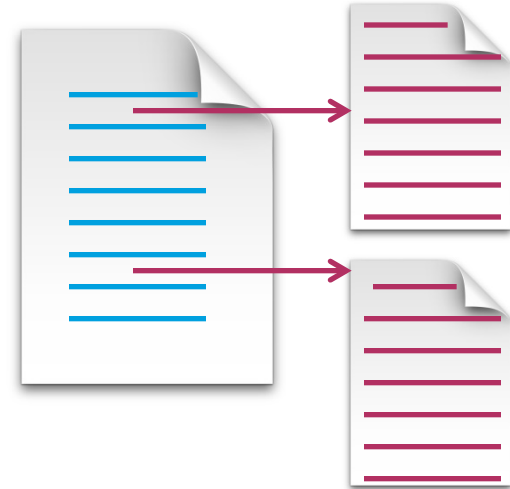
EAD no links:



EAD incl. <dao/>



EAD + METS





EAD

no additional information



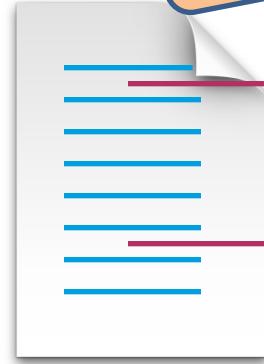
EAD

incl additional information



EAD

+ additional finding aids





What lies ahead – additional finding aids

We can already do that now in apeEAD ...

- But only very basically in a table in <scopecontent>, because apeEAD doesn't accept specific EAD2002 elements which we could use for this, like: <chronlist>, which allows tagging of events

```
<scopecontent>
  <table>
    <tgroup cols="2">
      <tbody>
        <row>
          <entry>Date</entry>
          <entry>1820, Dec. 20</entry>
        </row>
        <row>
          <entry>Event</entry>
          <entry>Born, Jacob son of Taylor Jones and Dicey Duke; Little River, Orange Co., N.C.</entry>
        </row>
        <row>
          <entry>Date</entry>
          <entry>1842</entry>
        </row>
        <row>
```



What lies ahead – additional finding aids

... even better with improved apeEAD ...

- in addition: apeEAD doesn't accept inline tagging of <geogname>, <corpname>, <famname>, <persname>, <subject>, etc., nor does the portal take care of this in providing search results yet ...

```
<chronlist>
  <chronitem>
    <date>1820, Dec. 20</date>
    <event>Born <persname>Jacob</persname>, son of <persname>Taylor Jones</persname> and <persname>Dicey Duke</persname>;
    <geogname>Little River, Orange Co., N.C.</geogname></event>
  </chronitem>
  <chronitem>
    <date>1842</date>
    <event>Married <persname>Mary Caroline Clinton</persname></event>
  </chronitem>
  <chronitem>
    <date>1844</date>
    <event>Born <persname>Sidney Jones</persname></event>
  </chronitem>
</chronlist>
```



```

<scopecontent>
  <chronlist>
    <chronitem>
      <datesingle>1820, Dec. 20</datesingle>
      <chronitemset>
        <event>Born
        <persname>
          <part localtype="firstname">Jacob</part>
          <part localtype="role">child</part>
        </persname>
        <persname>
          <part localtype="firstname">Taylor</part>
          <part localtype="lastname">Jones</part>
          <part localtype="role">mother</part>
        </persname>
        <persname>
          <part localtype="firstname">Dicey</part>
          <part localtype="lastname">Duke</part>
          <part localtype="role">father</part>
        </persname>
        <geogname>
          <part>Little River</part>
          <part>Orange Co.</part>
          <part>N.C.</part>
          <geographiccoordinates coordinatesystem="UTM">17S 682871m
        </geogname>
        <event>
      </chronitemset>
    </chronitem>
  </chronlist>
</scopecontent>

```

... but much better in EAD3 ...

So we have 3 possibilities:

- current apeEAD -> table
- improved apeEAD -> extra elements plus 'inline' tagging
- move to EAD3, at least for Additional Finding Aids

... the winner is: option 3

because:

- it is more suitable for tagging Additional Finding Aids in more detail
- we already are familiar with the basics of this metadata format (and the same goes for "Europe")
- we (APEF) have to start on moving to EAD3 at some point anyway



thank you for your attention

more information:

<http://www.archivesportaleuropefoundation.eu>

<http://www.archivesportaleurope.net>

<http://wiki.archivesportaleurope.net>

wim.van.dongen@nationaalarchief.nl