

Jailbreaking Governmental Data: PDF becomes RDF

mro.name/33c3

Overview

- there is an "OpenData Initiative"¹ of the bavarian government
- administrations love to publish PDFs
- e.g. the „Monument Lists“ (non-personal, non-controversial)
- I turn them 1:1 to RDF plus a simple HTML view

¹ <http://www.ldbv.bayern.de/aktuell/archiv/613.html>

Chieming

Baudenkmäler

D-1-89-114-1

Am Venusberg 2. Kapelle, Satteldachbau mit Putzgliederung, um 1842.
nachqualifiziert

D-1-89-114-12

Außerlohen 4. Zugehöriges Zuhause, zweigeschossiger Satteldachbau, biedermeierlich, 1.
Hälfte 19. Jh.
nachqualifiziert

The Start

- monuments are well managed with a neat file number
- the lists² are a bit hidden but online – yippie!
- but PDF³, so no deep linking to individual monuments, they don't have a URI⁴

2 <http://www.blfd.bayern.de/denkmalerschaffung/denkmalliste/denkmalliste/>

3 http://geodaten.bayern.de/denkmal_static_data/externe_denkmalliste/pdf/denkmalliste_merge_189114.pdf

4 <https://www.w3.org/QA/Tips/uri-choose>

mro.name/33c3

A kingdom for a Bot!

- daily iterates all bavarian „Official Municipality Numbers“⁵
- get the PDF monument list per municipality
- turn to XML via pdftohtml⁶
- turn to TTL via denkmaeler-xml2ttl-cmd⁷ (300 LOC of Go)
- turn to RDF via rapper⁸

5 https://de.wikipedia.org/wiki/Amtlicher_Gemeindeschl%C3%BCssel#Aufbau

6 <https://poppler.freedesktop.org/>

7 <https://github.com/mro/Denkmaeler/tree/master/denkmaeler-xml2ttl-cmd>

8 <http://librdf.org/raptor/rapper.html>

```
?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE pdf2xml SYSTEM "pdf2xml.dtd">

<pdf2xml producer="poppler" version="0.41.0">
<page number="1" position="absolute" top="0" left="0" height="1263" width="892">
  <fontspec id="0" size="16" family="Helvetica" color="#000000"/>
  <fontspec id="1" size="19" family="Times" color="#000000"/>
  <fontspec id="2" size="25" family="Times" color="#000000"/>
  <fontspec id="3" size="16" family="Times" color="#000000"/>
  <fontspec id="4" size="16" family="Times" color="#000000"/>
  <text top="51" left="602" width="253" height="19" font="1">Regierungsbezirk Oberbayern</text>
  <text top="87" left="765" width="90" height="19" font="1">Traunstein</text>
  <text top="123" left="772" width="83" height="19" font="1">Chieming</text>
  <text top="202" left="38" width="113" height="24" font="2"><b>Chieming</b></text>
  <text top="202" left="684" width="171" height="24" font="2"><b>Baudenkmäler</b></text>
  <text top="260" left="38" width="100" height="16" font="3"><b>D-1-89-114-1</b></text>
  <text top="260" left="204" width="527" height="16" font="3"><b>Am Venusberg 2. </b> Kapelle, Satteldachbau mit Putzgliederung, um 1842.</text>
  <text top="281" left="204" width="122" height="16" font="3"><b>nachqualifiziert</b></text>
  <text top="329" left="38" width="109" height="16" font="3"><b>D-1-89-114-12</b></text>
  <text top="329" left="204" width="650" height="16" font="3"><b>Außerlohen 4. </b> Zugehöriges Zuhaus, zweigeschossiger Satteldachbau, biedermeierlich, 1.</text>
  <text top="350" left="204" width="97" height="16" font="4">Hälfte 19. Jh.</text>
  <text top="370" left="204" width="122" height="16" font="3"><b>nachqualifiziert</b></text>
  <text top="419" left="38" width="109" height="16" font="3"><b>D-1-89-114-11</b></text>
  <text top="419" left="204" width="562" height="16" font="3"><b>Bachfeld. </b> Privatkapelle, massiver Satteldachbau, bez. 1883, mit Ausstattung.</text>
  <text top="440" left="204" width="122" height="16" font="3"><b>nachqualifiziert</b></text>
  <text top="488" left="38" width="109" height="16" font="3"><b>D-1-89-114-45</b></text>
  <text top="488" left="204" width="625" height="16" font="3"><b>Dorfstraße 1. </b> Ehem. Bauernhaus, Wohnteil verputzt mit Kniestock und zwei Lauben,</text>
  <text top="509" left="204" width="254" height="16" font="4">geschnitzte Türrahmung bez. 1838.</text>
  <text top="529" left="204" width="122" height="16" font="3"><b>nachqualifiziert</b></text>
  <text top="578" left="38" width="109" height="16" font="3"><b>D-1-89-114-46</b></text>
  <text top="578" left="204" width="614" height="16" font="3"><b>Dorfstraße 5; Dorfstraße 7. </b> Bauernhaus, Mittertennbau, Wohn- und ehem. Stallteil</text>
```

mro.name/33c3

```

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix dbp: <http://dbpedia.org/property/> .
@prefix cc: <http://creativecommons.org/ns#> .

<http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf>
  cc:attributionName "Baudenkmäler" <http://www.w3.org/2001/XMLSchema#string> ;
  cc:attributionURL <http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf> ;
  cc:license <http://creativecommons.org/licenses/by-sa/3.0/> ;
  dct:format <http://purl.org/NET/mediatypes/application/rdf+xml> ;
  dct:hasFormat <http://geodaten.bayern.de/denkmal_static_data/externe_denkmalliste/pdf/denkmalliste_merge_189114.pdf> ;
  dct:source <http://geodaten.bayern.de/denkmal_static_data/externe_denkmalliste/pdf/denkmalliste_merge_189114.pdf> ;
  dct:spatial <http://sws.geonames.org/6556396/> ;
  dct:relation <http://purl.mro.name/denkmaeler> ;
  dct:created "2016-08-31" <http://www.w3.org/2001/XMLSchema#date> ;
  a foaf:Document ;
  foaf:primaryTopic <http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/> .

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .
@prefix gn: <http://www.geonames.org/ontology#> .
@prefix c: <http://nwalsh.com/rdf/contacts#> .

<http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#D-1-89-114-1>
  a geo:SpatialThing ;
  rdfs:label 1 ;
  dct:identifier "D-1-89-114-1" ;
  dct:isPartOf <http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#Baudenkmäler> ;
  dct:type <http://www.geodaten.bayern.de/denkmaltyp#Baudenkmäler> ;

```

mro.name/33c3

```

< dct:identifizier >D-1-89-114-19</ dct:identifizier >
< dct:isPartOf rdf:resource="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#Baudenkmäler" />
< dct:type rdf:resource="http://www.geodaten.bayern.de/denkmaltyp#Baudenkmäler" />
< rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">45</ rdfs:label >
</ geo:SpatialThing >
< geo:SpatialThing rdf:about="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#D-1-89-114-2">
  < c:street >Hauptstraße 1</ c:street >
  < dct:description xml:lang="de">Kath. Pfarrkirche Mariä Himmelfahrt, neuromanische Anlage mit einschiffigem Langhaus und Westturm, nach Plänen von Johann Marggraff, 1882/83; mit Ausstattung; im Friedhof Kriegerdenkmal, Gusseisen, nach 1871.</ dct:description >
  < dct:identifizier >D-1-89-114-2</ dct:identifizier >
  < dct:isPartOf rdf:resource="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#Baudenkmäler" />
  < dct:type rdf:resource="http://www.geodaten.bayern.de/denkmaltyp#Baudenkmäler" />
  < rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">9</ rdfs:label >
</ geo:SpatialThing >
< geo:SpatialThing rdf:about="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#D-1-89-114-20">
  < c:street >Tabinger Straße 9</ c:street >
  < dct:description xml:lang="de">Bauernhof, Einfirstanlage, Wohnteil aus Tuff-Bruchsteinmauerwerk, mit hölzerner Hochlaube und biedermeierlicher Haustür, um 1830/40.</ dct:description >
  < dct:identifizier >D-1-89-114-20</ dct:identifizier >
  < dct:isPartOf rdf:resource="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#Baudenkmäler" />
  < dct:type rdf:resource="http://www.geodaten.bayern.de/denkmaltyp#Baudenkmäler" />
  < rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">46</ rdfs:label >
</ geo:SpatialThing >
< geo:SpatialThing rdf:about="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#D-1-89-114-21">
  < c:street >Tabinger Straße 11</ c:street >
  < dct:description xml:lang="de">Ehem. Gutshaus, zweigeschossiger Putzbau mit Kniestock und Flachsatteldach, Kern wohl 18. Jh., im 19. Jh. ausgebaut, Portal bez. 1838, Hochlaube bez. 1895.</ dct:description >
  < dct:identifizier >D-1-89-114-21</ dct:identifizier >
  < dct:isPartOf rdf:resource="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#Baudenkmäler" />
  < dct:type rdf:resource="http://www.geodaten.bayern.de/denkmaltyp#Baudenkmäler" />
  < rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">47</ rdfs:label >
</ geo:SpatialThing >
< geo:SpatialThing rdf:about="http://linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf#D-1-89-114-22">
  < c:street >Tabinger Straße 24</ c:street >
  < dct:description xml:lang="de">Kleinbäuerliches Anwesen mit Mittertenne. Wohnteil massiv gemauert. wohl 1. Viertel 19. Jh.</ dct:description >

```

mro.name/33c3

Denkmalliste Chieming x +

linkeddata.mro.name/open/country/DE/AGS/09/1/89/114/denkmal.rdf

Suchen

Shaare link Meistbesucht Getting Started

Denkmalliste Chieming

Quelle, Stand: 2016-08-13, [DBPedia](#),

[Bayernweite Übersicht](#),
[GitHub](#),

Baudenkmäler

[D-1-89-114-1](#)
Am Venusberg 2; Kapelle, Satteldachbau mit Putzgliederung, um 1842.

[D-1-89-114-12](#)
Außerlohen 4; Zugehöriges Zuhaus, zweigeschossiger Satteldachbau, biedermeierlich, 1. Hälfte 19. Jh.

[D-1-89-114-11](#)
Bachfeld; Privatkapelle, massiver Satteldachbau, bez. 1883, mit Ausstattung.

[D-1-89-114-45](#)
Dorfstraße 1; Ehem. Bauernhaus, Wohnteil verputzt mit Kniestock und zwei Lauben, geschnitzte Türrahmung bez. 1838.

[D-1-89-114-46](#)
Dorfstraße 5; Dorfstraße 7; Bauernhaus, Mittertennbau, Wohn- und ehem. Stallteil giebelgeteilt, Obergeschoss südseitig Blockbau mit Bundwerk, nordseitig massiv, bez. 1826; Backofenhäuschen, 1. Hälfte 19. Jh.

[D-1-89-114-55](#)
Dorfstraße 9 a; Ehem. Getreidekasten, wohl 16./17. Jh., Neuaufstellung vor Dorfstraße 9a.

[D-1-89-114-47](#)

[mro.name/33c3](https://linkeddata.mro.name/33c3)

Result

- nicely (deep-)linkable RDF/XML files
- deployed to a webserver (no DB, no CGI, no JS)
- git versioning as a bonus

mro.name/33c3

Todos

- geo-code the addresses
- make a nice map
- tersen the documentation

mro.name/33c3

Thank You

for your attention

<Marcus Rohrmoser> 33c3@mro.name

mro.name/33c3