OGUZHAN-BURAK BOZKURT

KUNST IM ÖFFENTLICHEN RAUM (STUTTGART)



DATASET

- 60 sculptures from the Stuttgart Kunstmuseum which are located in public spaces. The datasets contains longitude and latitude values for the locations as well as a reference image. No license issues to be worried about, since the "freedom of panorama" applies here. Images from the dataset can be used and distributed.
- Data Formats: CSV & XML
- Example XML Extraction of one single art object

```
<Object>
 <ID>69414</ID>
 <Titel DE>Marsyas I</Titel DE>
 <K_nstler>Hrdlicka, Alfred</K_nstler>
 <Entstehungsjahr>1959 - 1962</Entstehungsjahr>
 <Technik_DE>Bronze</Technik_DE>
 <Lizenzhinweis>Foto: Matter Of</Lizenzhinweis>
 <Inventarnummer>P-278</Inventarnummer>
 <Lebensdaten>1928 - 2009</Lebensdaten>
 <Datensatztyp>Plastik und Skulptur</Datensatztyp>
 <Jahr_bis>1962</Jahr_bis>
 <Jahr von>1959</Jahr von>
 <KoordinateB>48.776699</KoordinateB>
 <KoordinateL>9.180017</KoordinateL>
 <Digital_AssetsList>
   <Digital_Assets>
     <Dateiname>P-278_v2.jpg</Dateiname>
   </Digital_Assets>
 </Digital_AssetsList>
</Object>
```

OPEN REFINE WORKFLOW

- I tried to normalize / clean the dataset in openrefine using some transformation techniques, adding and removing columns and getting rid of null values
- Separated comma separated names into first and last name
- Used RegEx for date types, so that the dataset always has a consistent value for the year, instead of multiple different values
- Reconciled the dataset with wikidata, using common properties for some values



