

Google Cloud Vision AI for text extraction

Presenter: Steve Baskauf
steve.baskauf@vanderbilt.edu




Art Gallery Wikidata project (spring 2020)

April 2021: all 6884 works had items in Wikidata

September 2022: images of 1300 public domain works in Wikimedia Commons

Statements

instance of print edit
↳ 1 reference
+ add value

image edit

A Ghost Painting Coming to Life in the Studio of the Painter Okyō, from the series Yoshitoshi ryakuga (Sketches by Yoshitoshi) - Vanderbilt Fine Arts Gallery - 1992.083.tif
2,890 × 2,215; 18.31 MB
↳ 0 references
+ add reference
+ add value

inception 1882 edit
↳ 1 reference
+ add value



"Prints" are the largest category of work in the Vanderbilt Art Gallery

Top groupings (Minimum 1 items)			
Name	Count	image (P18)	inception (P571)
print (Q11060274)	2918	26.25% (766)	78.51% (2291)
photograph (Q125191)	564	0.18% (1)	92.73% (523)
painting (Q3305213)	521	22.46% (117)	51.82% (270)
image (Q478798)	311	3.54% (11)	62.38% (194)
artist's book (Q1062404)	278	41.37% (115)	75.18% (209)
drawing (Q93184)	273	14.29% (39)	72.53% (198)
figurine (Q1066288)	177	25.99% (46)	22.6% (40)
sculpture (Q860861)	163	42.94% (70)	50.31% (82)

"Print": poster or art print?

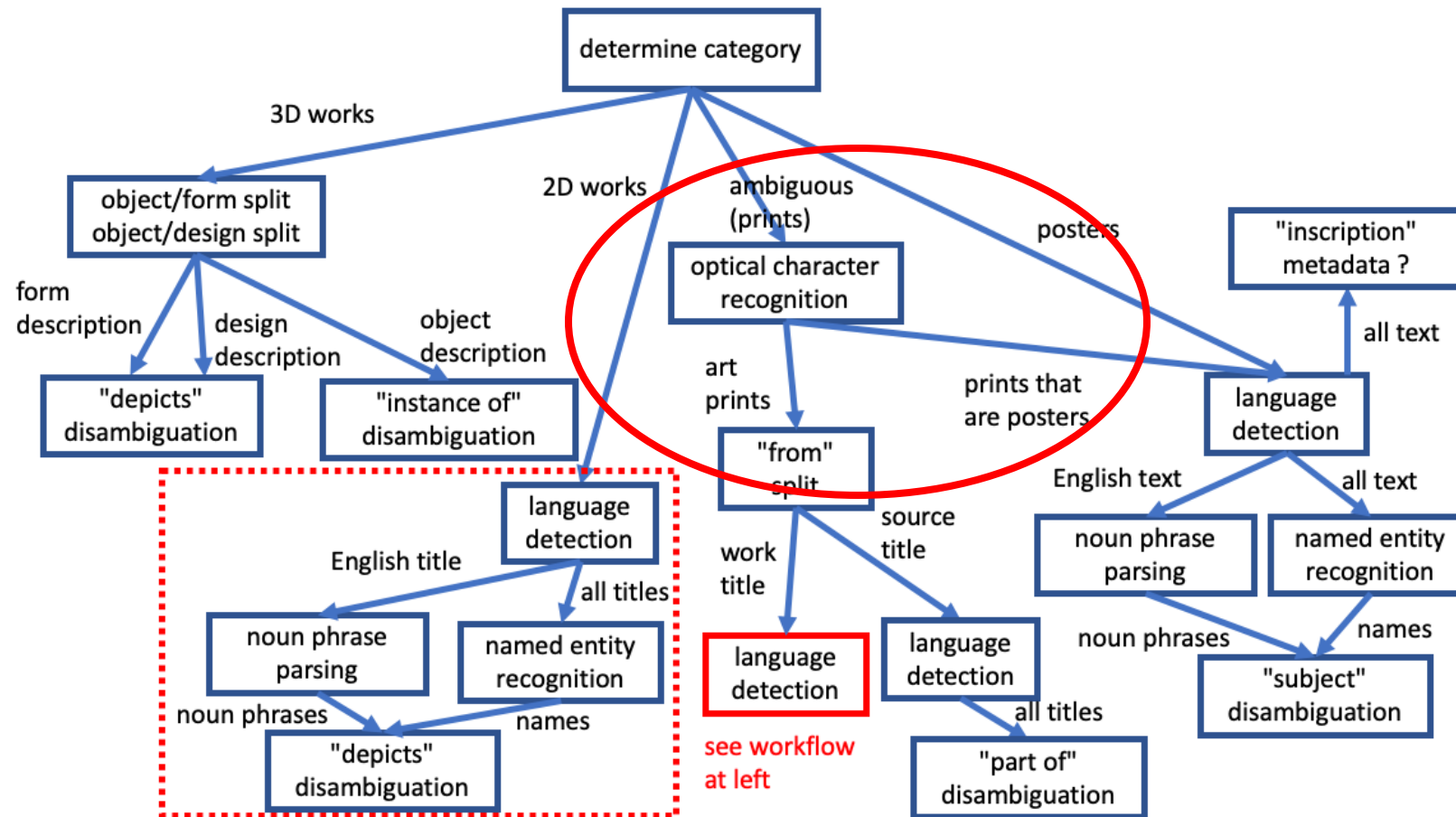


"An Idyll", 1905, print by Charles Haslewood Shannon. Vanderbilt Art Gallery 2003.022



"This Device on Hat or Helmet Means U.S. Marines", 1918, print by Charles Buckles Falls. Vanderbilt Art Gallery 1979.1218P

OCR as a potential method to split prints



Experiments with Keras OCR - 2021



Extremely simple Python script.
Generally poor performance with atypical fonts.
Detected text given as individual tokens (words) with coordinates.

Project proposal by Aiden Layer in 2021

Suggested using Google Vision AI and Clarafai to detect and tag objects depicted in Vanderbilt Art Gallery works.

Project Proposal

Using **AI Assisted Tagging**
to Improve Findability of
the Vanderbilt University
Art Collection

Author: Aiden Layer, Vanderbilt University Fine Arts Gallery



Google Vision AI object analysis (left) and label analysis (right).

Lorna Simpson, Backdrops Circa 1940s, 1998, Screen print on two felt panels, Dr. and Mrs. E. William Ewers Gift for Fine Arts Fund Purchase, Vanderbilt University Fine Arts Gallery, 1999.143

Google Cloud Vision AI online console

<https://cloud.google.com/vision/>

DEMO

Try the API

Try the API

Drag image file here or
Browse from your computer

KEY FEATURES

Three computer vision offerings to meet you where you are

Vertex AI Vision

[Vertex AI Vision](#) is a fully managed end to end application development environment that lets you easily build, deploy and manage computer vision applications for your unique business needs. Vertex AI Vision includes [Streams](#) to ingest real-time video data, [Applications](#) that lets you create an application by combining various components and [Vision warehouse](#) to store model output and streaming data.

Navigation: A red arrow points to the 'Demo' link in the left sidebar under the 'Vision AI' section.

Try the API

Objects Labels Logos **Text** Properties Safe Search

+Page 1

+Block 1

THIS DEVICE ON HAT OR HELMET MEANS U.S. MARINES

+Block 2

ENLIST TODAY AT

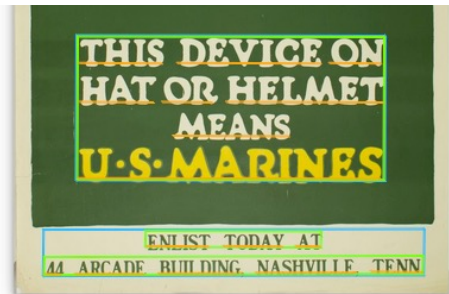
44 ARCADE BUILDING, NASHVILLE, TENN.

This_Device_on_Hat_or_Helmet_Means_U.S._Marines_-_Vand...

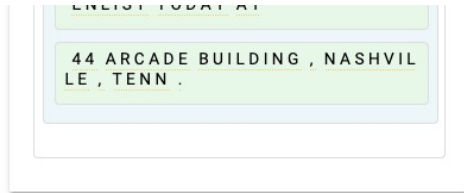
Show JSON ▾

RESET NEW FILE

Automation: JSON to request and for response



This_Device_on_Hat_or_Helmet_Means_U.S._Marines_-_Vand...



```
{
  "requests": [
    {
      "features": [
        {
          "maxResults": 50,
          "type": "LANDMARK_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "FACE_DETECTION"
        },
        {
          "maxResults": 50,
          "type": "OBJECT_LOCALIZATION"
        },
        {
          "maxResults": 50,
          "type": "LOGO_DETECTION"
        }
      ]
    }
  ]
}
```

Hide JSON ^

RESET

NEW FILE

```
3360     "y": 997
3361   }
3362 }
3363 },
3364 "description": "THIS DEVICE ON\nHAT OR HELMET\nMEANS\nU.S.MARINES\nENLIST TODAY AT\n44 ARCADE BUILDING, NASHVILLE, TENN.",
3365 "locale": "en"
3366 },
3367 {
3368   "boundingPoly": {
3369     "vertices": [
3370       {
3371         "x": 105,
3372         "y": 625
3373       },
3374       {
3375         "x": 235,
3376         "y": 625
3377       },
3378       {
3379         "x": 235,
3380         "y": 666
3381       },
3382       {
3383         "x": 105,
3384         "y": 666
3385       }
3386     ]
3387   },
3388   "description": "THIS"
3389 },
3390 {
3391   "boundingPoly": {
3392     "vertices": [
3393       {
3394         "x": 258
```

entire text

position of each word

Python script to automate process

1. Generate image at suitable size and format
2. Upload images to Google Cloud bucket
3. Loop through the list of images and analyze
4. Extract data from results



1. Generating images at suitable size and format using IIIF image request

Used JPEG images 1000 pixels on longest dimension.

2.1. Image Request URI Syntax

The IIIF Image API URI for requesting an image **MUST** conform to the following URI Template:

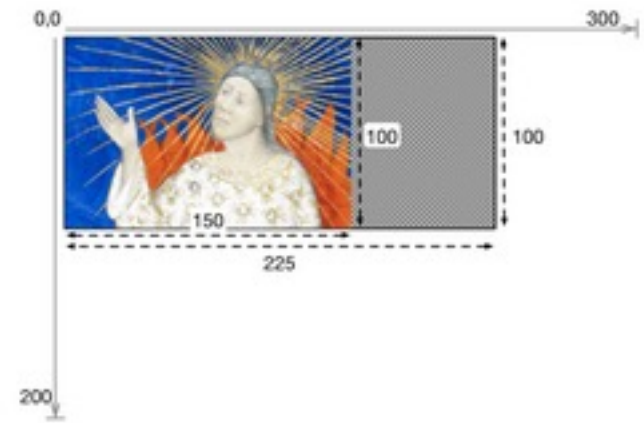
```
{scheme}://{server}/{prefix}/{identifier}/{region}/{size}/{rotation}
/{quality}.{format}
```

For example:

```
http://www.example.org/image-service/abcd1234/full/full/0/default.jpg
```

The parameters of the Image Request URI include region, size, rotation, quality and format, which define the characteristics of the returned image. These are described in detail in [Image Request Parameters](#).

[From https://iiif.io/api/image/2.1/](https://iiif.io/api/image/2.1/)



6 size=!225,100

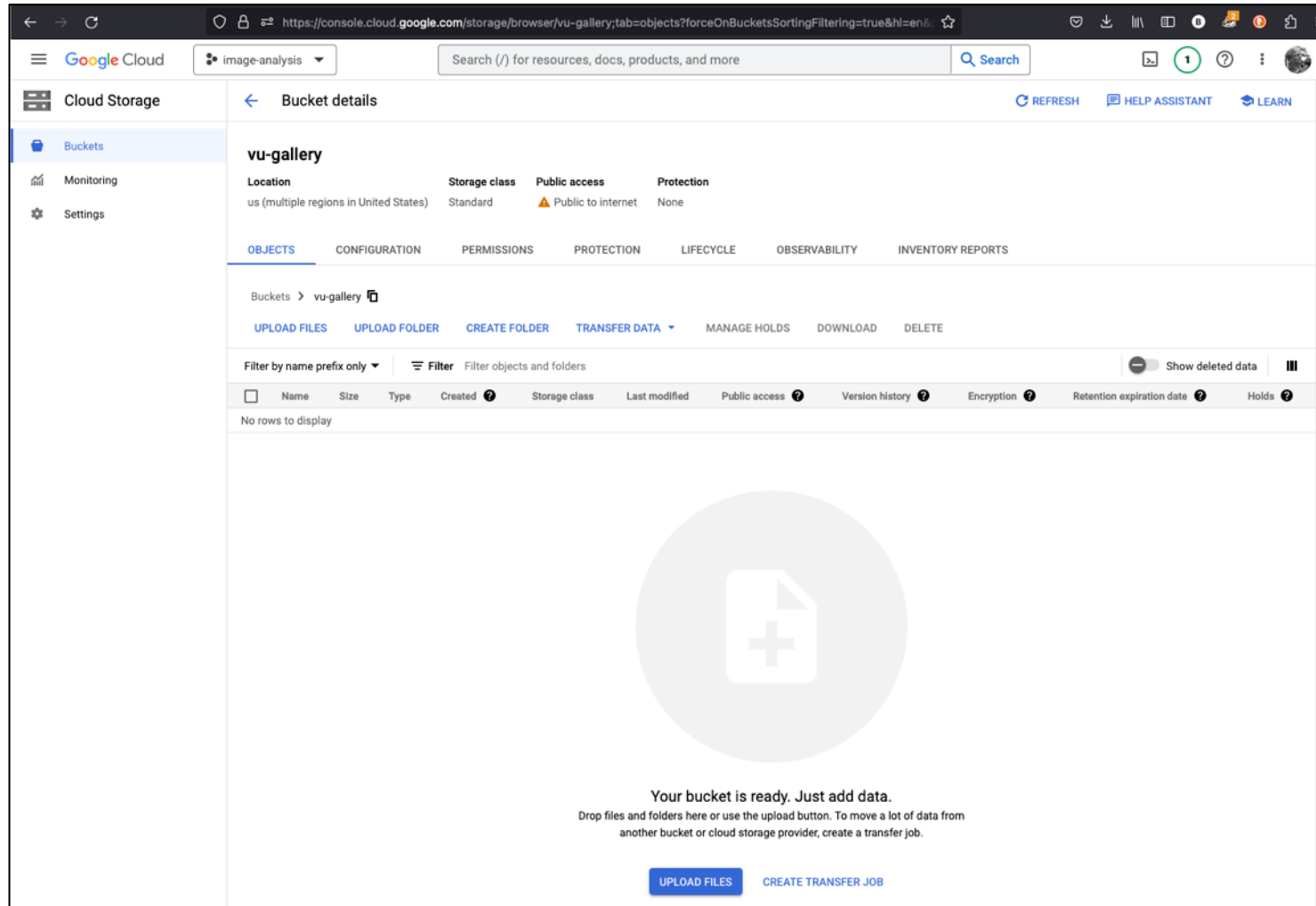
.../full/!225,100
/0/default.jpg

N.B. Returned image is 150,100 px

2. Upload images to Google Cloud bucket

Note: direct URL access required too much work to prepare metadata.

Need to create a Google Cloud account at cloud.google.com. Bucket access must be public.



The screenshot shows the Google Cloud Storage console interface. The browser address bar displays the URL: <https://console.cloud.google.com/storage/browser/vu-gallery;tab=objects?forceOnBucketsSortingFiltering=true&hl=en&...>. The page title is "Bucket details" for the bucket named "vu-gallery".

Key details for the bucket "vu-gallery":

- Location:** us (multiple regions in United States)
- Storage class:** Standard
- Public access:** Public to internet
- Protection:** None

The "OBJECTS" tab is selected, showing a table with columns: Name, Size, Type, Created, Storage class, Last modified, Public access, Version history, Encryption, Retention expiration date, and Holds. The table is currently empty, displaying "No rows to display".

Below the table, there is a large circular icon with a document and a plus sign, indicating that the bucket is ready for data. The text below the icon reads: "Your bucket is ready. Just add data. Drop files and folders here or use the upload button. To move a lot of data from another bucket or cloud storage provider, create a transfer job." At the bottom, there are two buttons: "UPLOAD FILES" and "CREATE TRANSFER JOB".

3. Analyze images using API

For reference, see <https://cloud.google.com/vision/docs/detect-labels-image-client-libraries>

Script at https://github.com/HeardLibrary/linked-data/blob/master/image_analysis/google_cloud_vision/google_cloud_vision.ipynb

- Need to create a project and get an access key
- There is a Python client library (google-cloud-vision) that handles API interactions. Code examples were helpful.
- Used the JSON request examples from the online console to figure out how to send the request.
- Used the JSON response examples to figure out how to extract the returned text data

4. Extract data

Script turns JSON to CSV spreadsheet

Data consist of text (all and individual words) and position within image.

Output includes language detection.

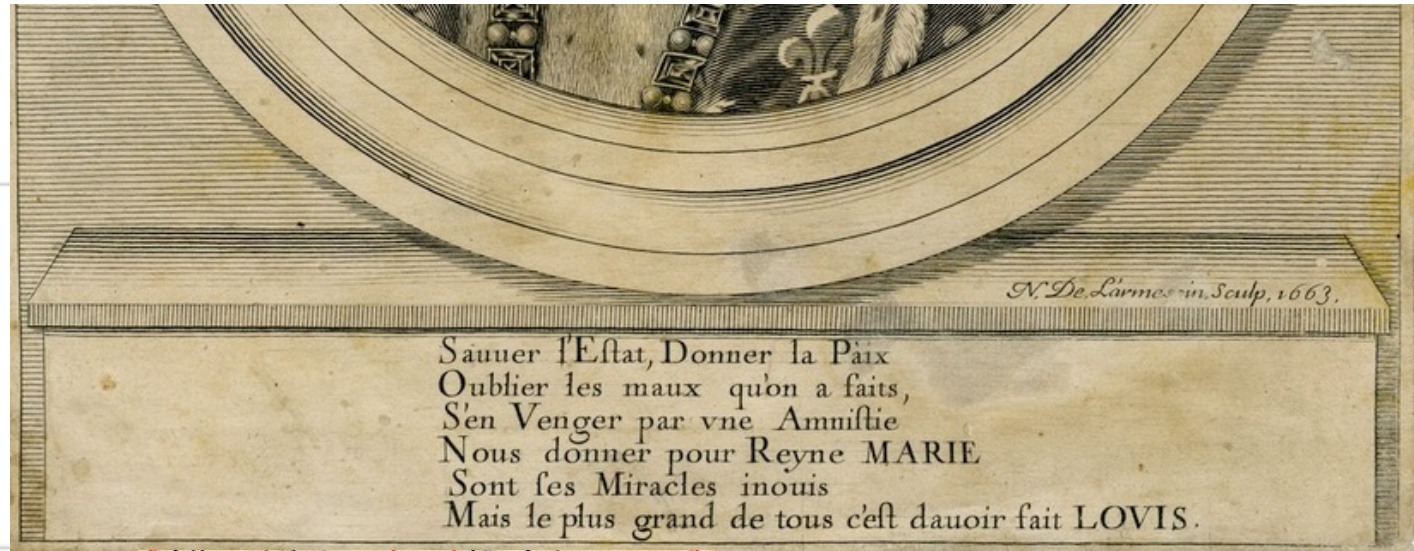
Example data at https://github.com/HeardLibrary/linked-data/blob/master/image_analysis/google_cloud_vision/text_detection.csv

	A	B	C	D	E	F	G	H	I	J	K
1	accession number	locale	description	rel_left_x	rel_right_x	rel_top_y	rel_bottom_y	abs_left_x	abs_right_x	abs_top_y	abs_bottom_y
107	1979.1225P	en	Fight World Famine enroll in The Boys Working Reserve U.S. EMPLOYMENT SERVICE APPLY THER COMPANY, BALTO New You DEPARTMENT OF LABOR	0.026026026026026	0.988988988988989	0.752893124574541	0.981620149761743	26	988	1106	1442
108	1979.1225P		Fight	0.113113113113113	0.347347347347347	0.769230769230769	0.82641252552757	113	347	1130	1214
109	1979.1225P		World	0.359359359359359	0.600600600600601	0.764465622872703	0.820966643975494	359	600	1123	1206
110	1979.1225P		Famine	0.615615615615616	0.897897897897898	0.759019741320626	0.815520762423417	615	897	1115	1198
111	1979.1225P		enroll	0.43043043043043	0.54954954954955	0.82573179033356	0.842750170183799	430	549	1213	1238
112	1979.1225P		in	0.562562562562563	0.601601601601602	0.821647379169503	0.84138869979578	562	601	1207	1236
113	1979.1225P		The	0.118118118118118	0.229229229229229	0.85500340367597	0.89788972089857	118	229	1256	1319
114	1979.1225P		Boys	0.249249249249249	0.408408408408408	0.853641933287951	0.895847515316542	249	408	1254	1316
115	1979.1225P		Working	0.431431431431431	0.675675675675676	0.851599727705922	0.893124574540504	431	675	1251	1312
116	1979.1225P		Reserve	0.68968968968969	0.897897897897898	0.848196051735875	0.889720898570456	689	897	1246	1307
117	1979.1225P		U.S.	0.026026026026026	0.067067067067067	0.938733832539142	0.950306330837304	26	67	1379	1396
118	1979.1225P		EMPLOYMENT	0.073073073073073	0.18018018018018	0.938053097345133	0.948944860449285	73	180	1378	1394
119	1979.1225P		SERVICE	0.185185185185185	0.248248248248248	0.936691626957114	0.947583390061266	185	248	1376	1392
120	1979.1225P		APPLY	0.098098098098098	0.164164164164164	0.968005445881552	0.980258679373724	98	164	1422	1440
121	1979.1225P		THER	0.841841841841842	0.882882882882883	0.923757658270933	0.92716133424098	841	882	1357	1362
122	1979.1225P		COMPANY	0.885885885885886	0.912912912912913	0.923757658270933	0.92716133424098	885	912	1357	1362
123	1979.1225P		,	0.912912912912913	0.914914914914915	0.923076923076923	0.926480599046971	912	914	1356	1361
124	1979.1225P		BALTO	0.916916916916917	0.941941941941942	0.923076923076923	0.926480599046971	916	941	1356	1361
125	1979.1225P		New	0.957957957957958	0.970970970970971	0.923076923076923	0.926480599046971	957	970	1356	1361
126	1979.1225P		You	0.973973973973974	0.983983983983984	0.923076923076923	0.926480599046971	973	983	1356	1361
127	1979.1225P		DEPARTMENT	0.801801801801802	0.904904904904905	0.930565010211028	0.940095302927161	801	904	1367	1381
128	1979.1225P		OF	0.912912912912913	0.928928928928929	0.929884275017018	0.940095302927161	912	928	1366	1381
129	1979.1225P		LABOR	0.937937937937938	0.988988988988989	0.929884275017018	0.939414567733152	937	988	1366	1380

Results

30
SCOO
N De Larmes in. Sculp, 1663.
Sauuer Eftat, Donner la Paix
Oublier les maux qu'on a faits,
S'en Venger par vne Amniftie
Nous donner pour Reyne MARIE
Sont fes Miracles inouis
Mais le plus grand de tous c'eft dauoir fait LOVIS.

1970.019 fr



بار اور کوردو کورکار ملا خاطر کیا کہ کسانیزه
بو وصار وفوا حوالی یک مردی را دید و خودیده کرمانیخره
باران او در یور خوا راسه
دلبر یکتا ندوی صدا شنوا و کیم
جان دیروب دلبره منت ومالوم جاره
وجود کارای مداران باصفا
ظہر کی کہ ملک ایک مسافر زفات خطوط افقری
اختر دارد کہ
وارد ربطونکہ مفاجور
جویر ایوب اور این مادہ در نیایم وقتی
وفرالامن طنطا وندرس ایلان کوری

1968.009 fa



an. Jy loco pafcueibime coffocauit.an.
Delicta iumentutis. Pfalmus.
Dte die leuauī anima mea:de?
meus in te cofido non erubefcam
eqz irrideat me inimici mei:etenim
Bniuerfiqui fuftinent te no cofundant.
onfundatur omnes iniqua agètes:
fuperuacue. ias tuas die demōftra
michi:afemittas tuas edoce me.
irige me in veritate tua & doce me:
q2 tu es deus faluatoz me ate fuftinui
tota die. eminifcere miferationum

1967.003 la



Cost

- First thousand images per month are free.
- I paid \$3.84 to do the remaining 500 or so.

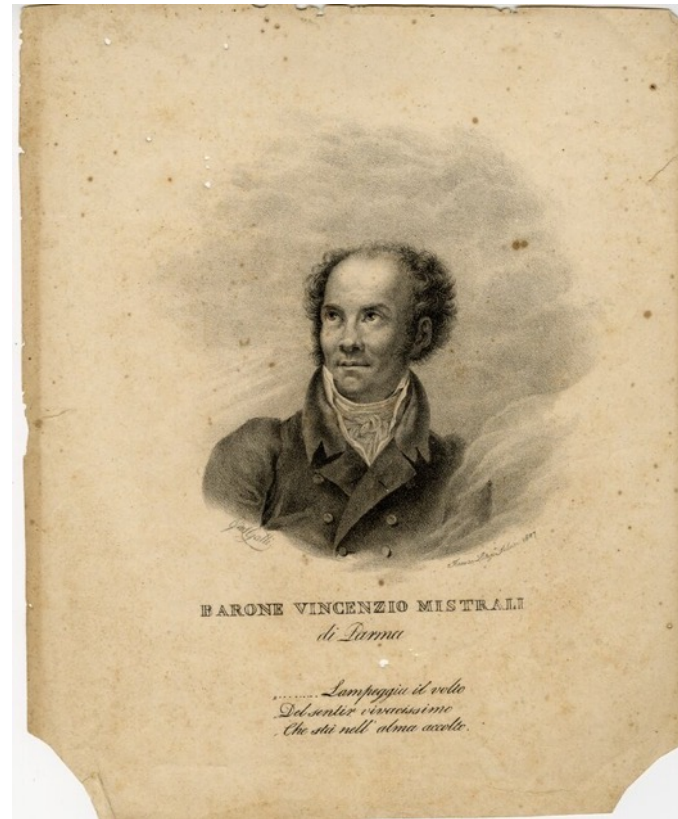
What's next?

- Expose text on Wikidata (as "inscription", P1684 ?)
- Use detected language (along with label language detection) to add non-English titles to Wikidata
- Differentiate between posters and other works with text captions.

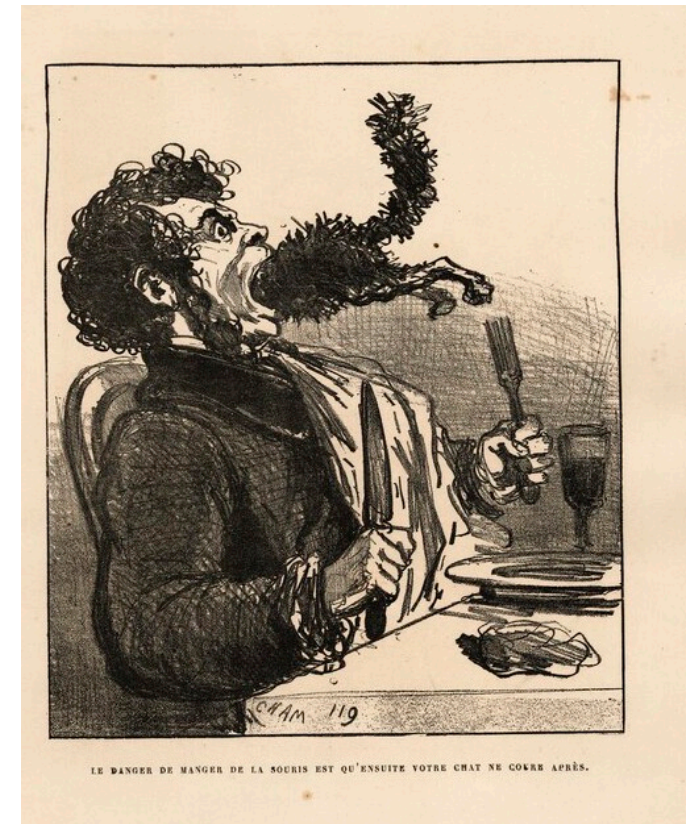
Poster text or caption?



"This Device on Hat or Helmet Means U.S. Marines". 1918. Print by Charles Buckles Falls. Vanderbilt Art Gallery 1979.1218P



"Barone Vincenzo Mistralli di Parma". 1827. Print by Giuseppe Galli. Vanderbilt Art Gallery 1971.025



"Le danger de manger, from Album du Siège". 1871. Print by Honoré Daumier and Cham. Vanderbilt Art Gallery 1999.172t