

# **Informační prameny a fondy**

Obrazové dokumenty

8. 4. 2022



**Obrazy v knihovnách?**

# Artotéky

- veřejné půjčovny obrazů
- myšlenka z USA (1893)
- často součástí knihoven a muzeí
- knihovny dodávají kontext
- podpora umění a umělců
- MKP, MG Brno

**MORAVSKÁ GALERIE ARTOTÉKA**



# Obrazový dokument (OD)

Dokument, jehož základní charakteristikou je dvojrozměrné obrazové vyjádření. Obrazové dokumenty mohou být statické (podle katalogizačních pravidel nazývané souhrnně grafika - např. fotografie, malba, umělecká reprodukce aj.) nebo pohyblivé (audiovizuální záznamy a kinematografické filmy).

# OD v knihovnách

Paměťové instituce mají fondy OD

- Fond je vždy nějak organizován
- Většinou má specifickou ochranu a pravidla přístupu
- Všechny typy dokumentů nejsou na jednom místě

1. [Náprstkovo muzeum](#)
2. [Krajská vědecká knihovna Liberec](#)

V knihovně máme  
fondy fotografií  
a pohlednic.

**Co s nimi?**



# Sbírký OD

- *jak vznikají?*
- co všechno se s nimi musí udělat?
- jak se s nimi dá pracovat?
- jak k nim pak přistupuje uživatel?

Prints & Photographs Online Catalog

https://www.loc.gov/pictures/

LIBRARY OF CONGRESS

ASK A LIBRARIAN DIGITAL COLLECTIONS LIBRARY CATALOGS

Search Search Loc.gov GO

Library of Congress > Prints & Photographs Reading Room > Prints & Photographs Online Catalog

### PRINTS & PHOTOGRAPHS ONLINE CATALOG (PPOC)


## Prints & Photographs Online Catalog

Print Subscribe Share/Save


Search All GO Advanced | Help

#### Featured Collections


Civil War



Johnston (Frances Benjamin) Collection















Highsmith (Carol M.) Archive



Lijenquist Family Collection of Civil War Photographs

#### All Collections | View All with Description

	<a href="#">Abdul Hamid II Collection</a>		<a href="#">African American Photographs Assembled for 1900 Paris Exposition</a>
	<a href="#">Ansel Adams's Photographs of Japanese-American Internment at Manzanar</a>		<a href="#">Architecture, Design &amp; Engineering Drawings</a>
	<a href="#">Bain Collection</a>		<a href="#">Baseball Cards</a>
	<a href="#">Brady-Handy Collection</a>		<a href="#">Brumfield Collection</a>
	<a href="#">Cabinet of American Illustration</a>		<a href="#">Carnegie Survey of the Architecture of the South</a>
	<a href="#">Carpenter Collection</a>		<a href="#">Cartoon Drawings</a>

BLOG: PICTURE THIS



Objevte inspirující evropské kul X +

https://www.europeana.eu/cs

HOME SBÍRKY PŘIHLÁŠENÍ/PŘIPOJENÍ

BLOG

### The ways of counterculture

British, Dutch and German counterculture movements of the 1960s were much more than just being against the mainstream

BLOG

### Europe in space

Images and videos featuring Europe's space flights, European astronauts and cosmonauts

### Chinese heritage

Explore stories and objects from China's rich history in European collections

GALERIE

### Sporting portraits

In the early 20th century, people's passion for sport led to them having formal portrait photographs taken

GALERIE

### Cherry blossoms

Beloved in Japan and found across Europe, cherry trees welcome spring with bright and beautiful blossoms

BLOG

### Long+short European places

Take a virtual tour to linguistic length-based locations

BLOG

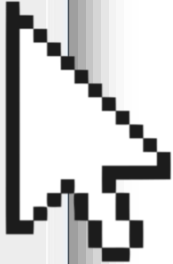
### Style sites in Europe

Exploring places across Europe that gave names to items of clothing, textiles and more

BLOG

### Codices Fennici

Explore medieval and 16th century manuscripts written or used in Finland



# Sbírky OD

- *degradace nosiče*
- digitalizace / *born-digital*
- **formáty**, úpravy,...
- tvorba **metadat**
- výběr platformy
- zpřístupnění



# Formáty obrazu

- klasické formáty a jejich vlastnosti (*JPG, PNG, GIF...*)
- my se setkáváme i s [dalšími formáty](#)
- RAW – proprietární (*NRAW...*)
- DNG – digitální negativ, otevřený (2004)
- TIFF – tagged image, ISO norma, trvalost (80. léta)

1. [Ukázka formátů na příkladu LoC sportovních kartiček](#)





# Metadatový popis

- nutné přidat **metadata** – *ručně nebo automatizovaně*
- jaké mají obrázky metadata?
- *interní/externí*
- *EXIF*

## **EXIF**

<http://exif.regex.info/exif.cgi>

*exchangeable image file format*

specifikace metadat pro uložení do obrazu



Camera:	Nikon D3100
Lens:	<a href="#">AF-S DX Zoom-Nikkor 18-55mm f/3.5-5.6G ED II</a> Shot at 26 mm
Exposure:	Auto exposure, Not Defined, 1/125 sec, f/8, ISO 280
Flash:	Auto, Did not fire
Focus:	AF-A, at <u>2.7m</u> , with a depth of field of about <u>5.4m</u> , (from about <u>1.0m</u> before the focus point to about <u>4.4m</u> after) AF Area Mode: Single Area
Date:	<b>October 17, 2019</b> 10:58:56AM (timezone not specified) (5 months, 9 days, 2 hours, 24 minutes, 17 seconds ago, assuming image timezone of 1 hour ahead of GMT)
Time Zone Offset:	+01:00
Location:	Latitude/longitude: <b>49° 5' 49.2" North, 14° 46' 2.2" East</b> ( 49.096992, 14.767277 )  Map via embedded coordinates at: <a href="#">Google</a> , <a href="#">Yahoo</a> , <a href="#">WikiMapia</a> , <a href="#">OpenStreetMap</a> , <a href="#">Bing</a> (also see the Google Maps pane below)  Timezone guess from earthtools.org: <a href="#">1 hour ahead of GMT</a>
File:	<b>4,608 × 3,072 JPEG (14.2 megapixels)</b> 6,893,234 bytes (6.6 megabytes)
Color Encoding:	<b>WARNING:</b> Color space tagged as sRGB, without an embedded color profile. <b>Windows and Mac browsers and apps treat the colors randomly.</b> <hr/> Images for the web are most widely viewable when in the sRGB color space and with an embedded color profile. See my <a href="#">Introduction to Digital-Image Color Spaces</a> for more information.





# **Ruční** tvorba metadat

- ruční tvorba různých metadat
- popisné a věcné údaje
- *i pro vyhledávání...*
  
- značně náročná činnost
- *kdo to provádí?*



# Ruční tvorba metadat

- značně náročná činnost
- *kdo to provádí?*
- archivy, muzea, knihovny, galerie
- komerční subjekty, které potřebují přesnost
- většinou kombinují ruční a automatizované přístupy
- [Příklad komerční fotobanky GettyImages](#)



PEOPLE

dog



Creative Images

NUMBER OF PEOPLE

- No people
- One person
- Two people
- Group of peo...

AGE

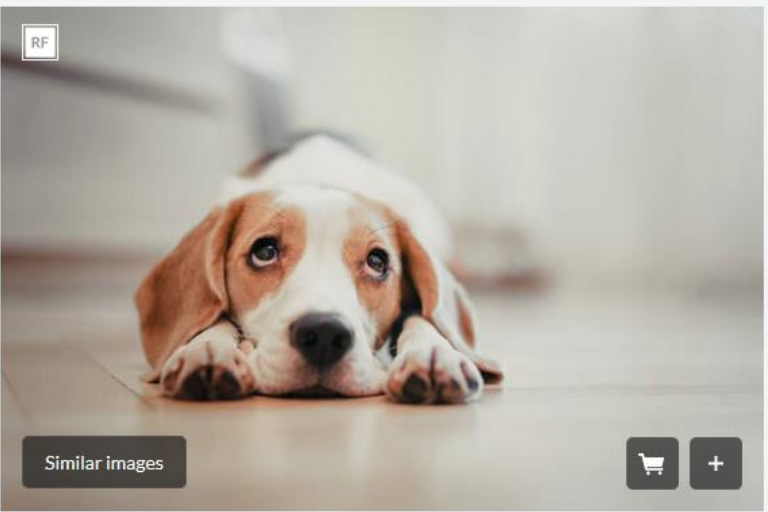
- Baby
- Child
- Teenager
- Young adult
- Adult
- Adults only
- Mature adult
- Senior adult

PEOPLE COMPOSITION

- Head shot
- Waist up
- Three quarte...
- Full length
- Looking at ca...
- Candid

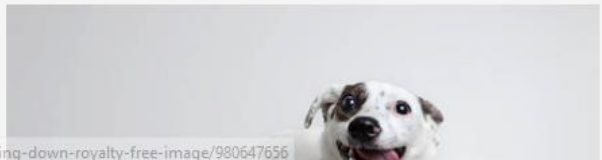
ETHNICITY

- East Asian
- Southeast As...
- South Asian
- Black
- Hispanic/Lat...
- Caucasian



RF

Similar images

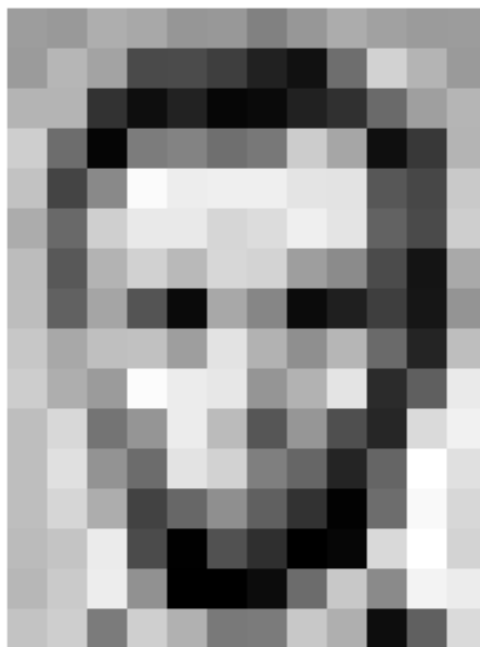


# **Automatická tvorba metadat**

- různé úrovně složitosti
- *od extrakce příznaků (např. barvy)*
- *až po automatické rozpoznávání obrazu*
- *počítačové vidění, neuronové sítě*

# Automatické rozpoznávání obrazu

- jak naučit počítač rozpoznat, co je na obrázku?
- počítačové vidění



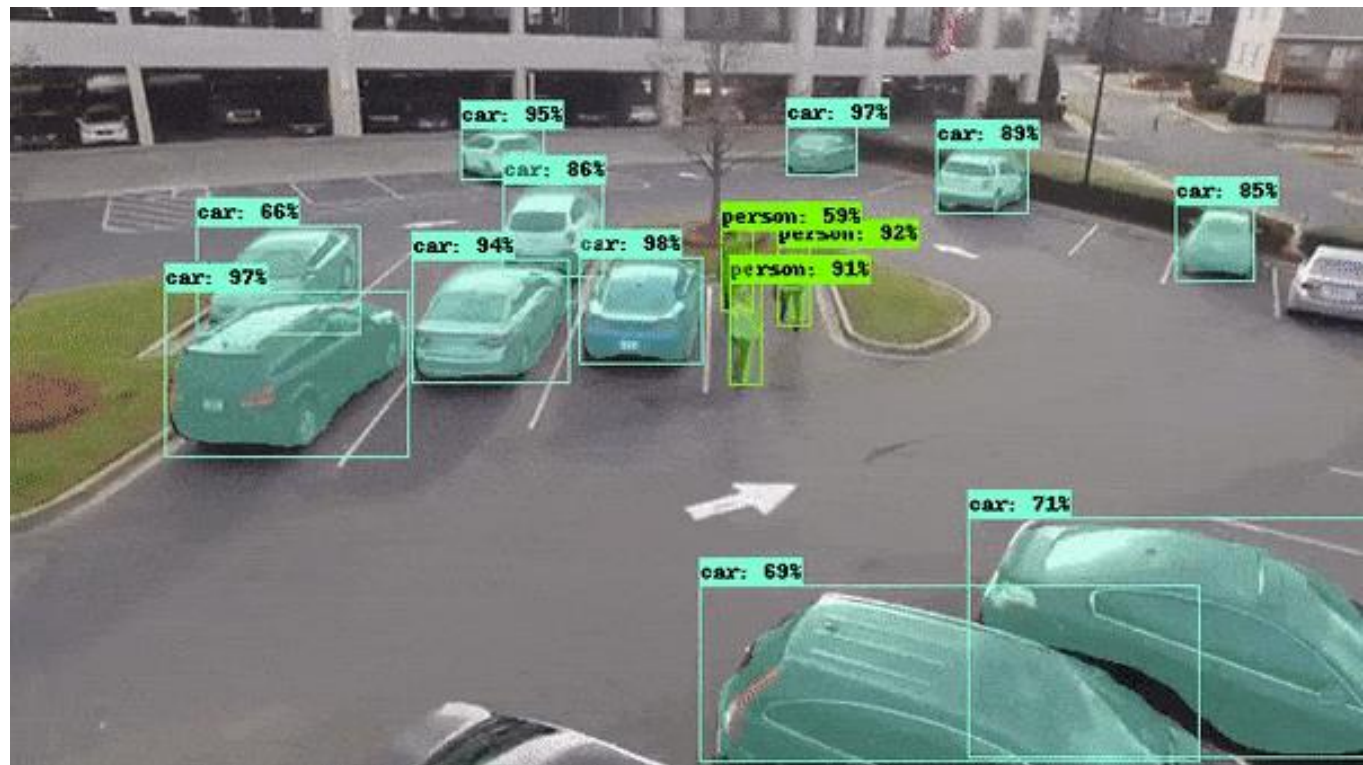
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	80	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

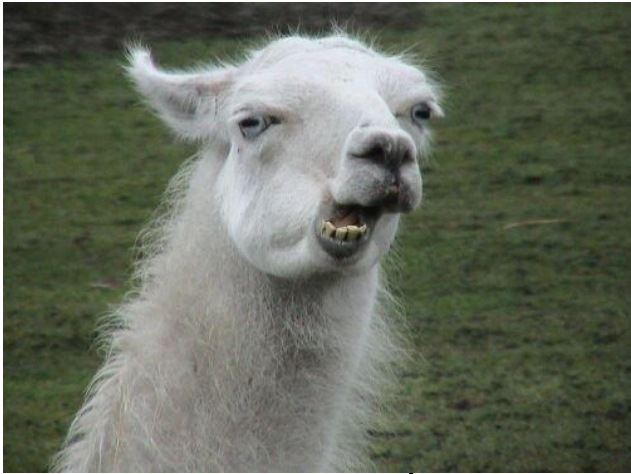
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	80	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



# Automatické rozpoznávání obrazu

- neuronové sítě
- *vstup* – obrázky
- *výstup* – klasifikace, s určitou pravděpodobností
- všude, od taggování na FB po autonomní auta
- *supervised learning*
- *unsupervised learning*





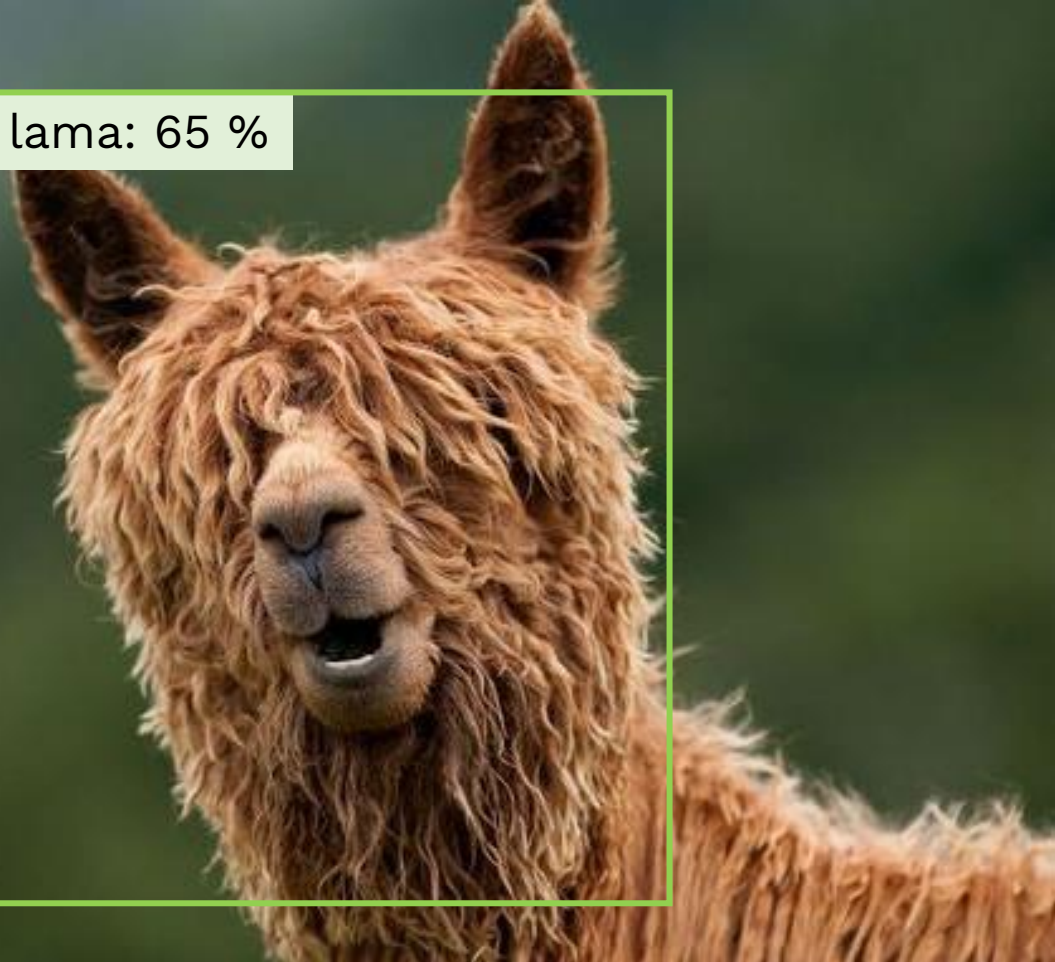
všechno je **lama**



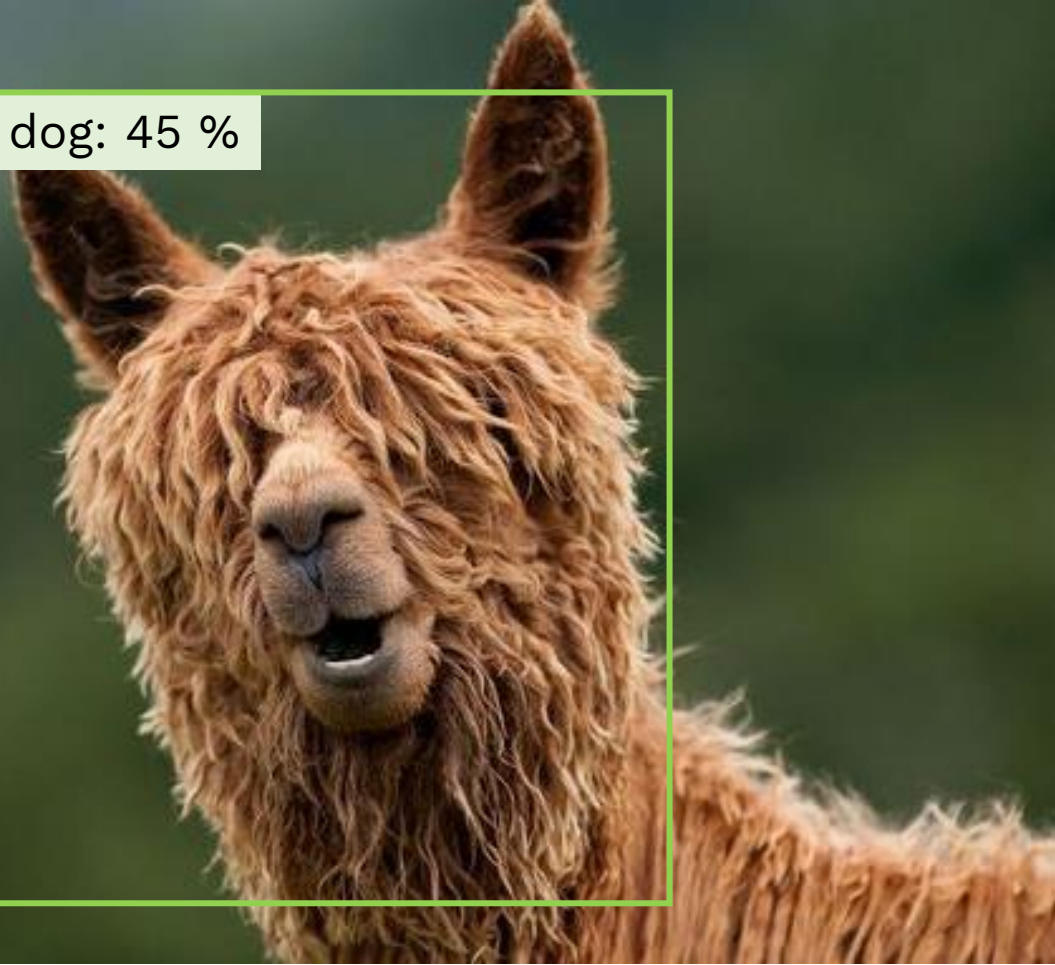


co je tohle?

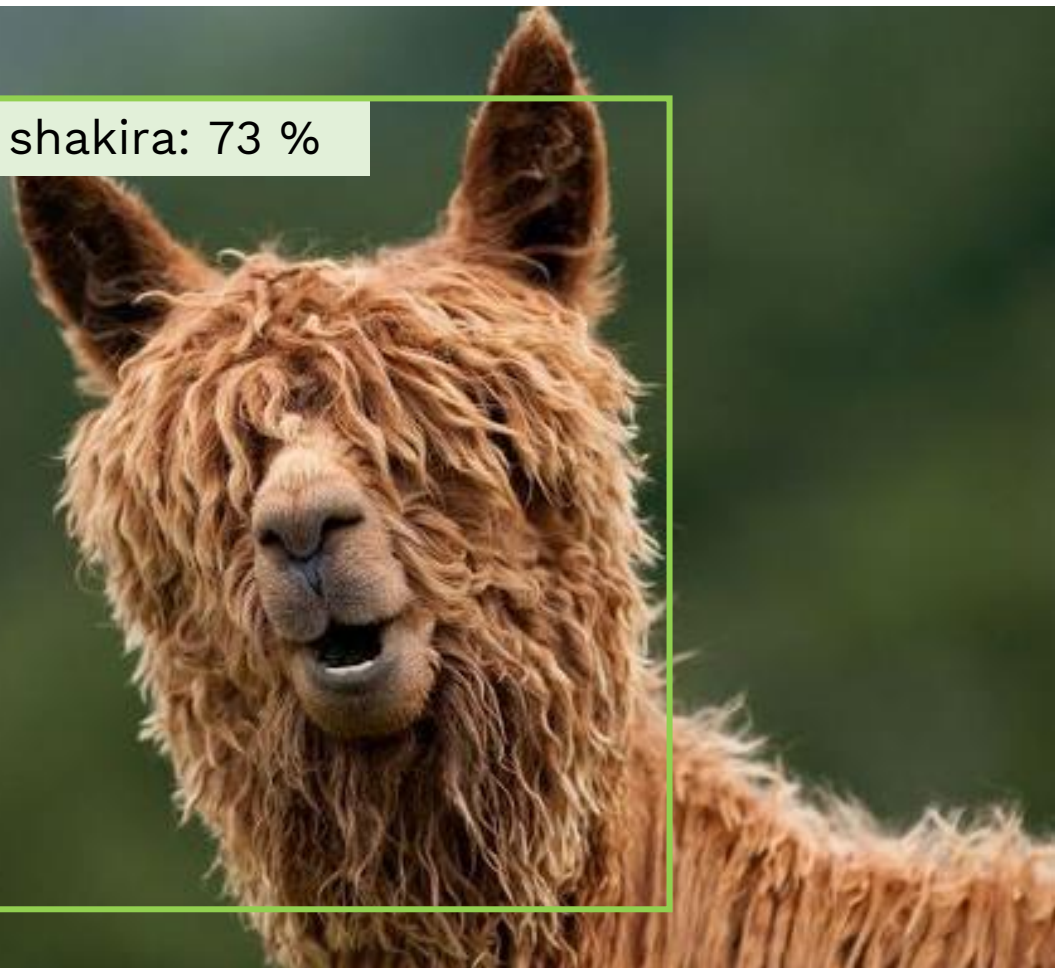
lama: 65 %



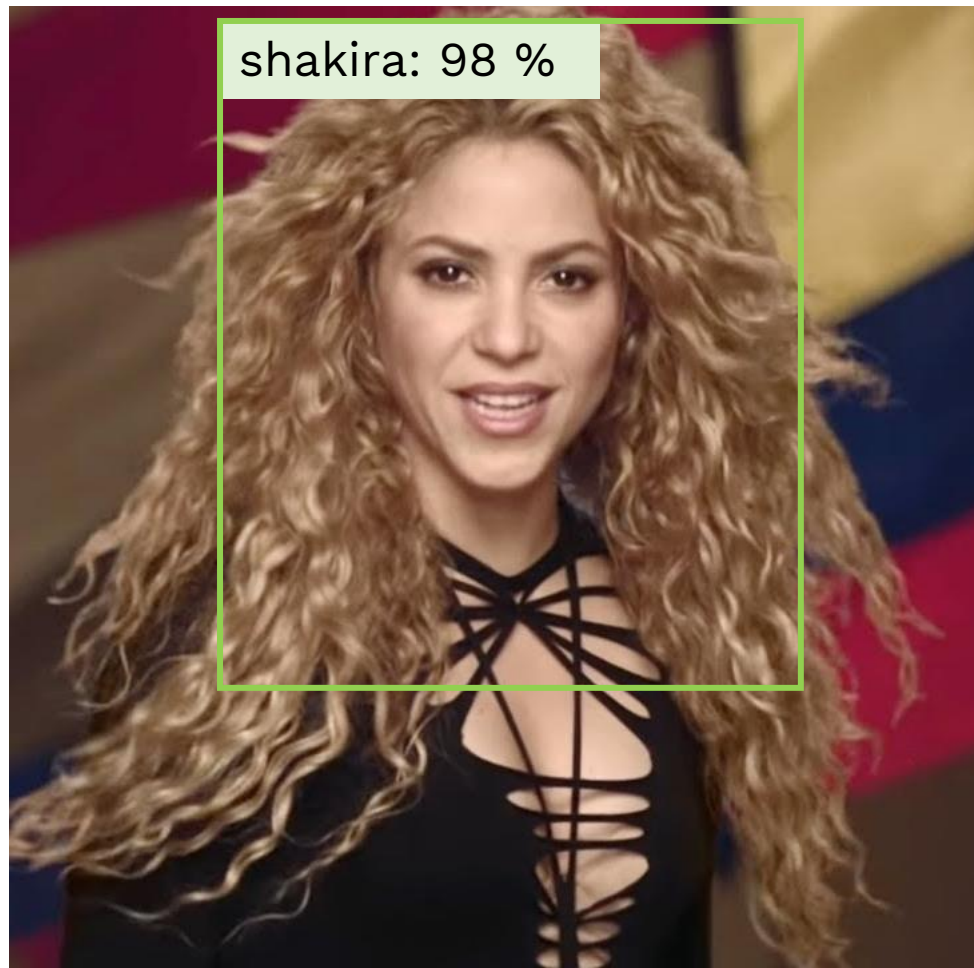
dog: 45 %

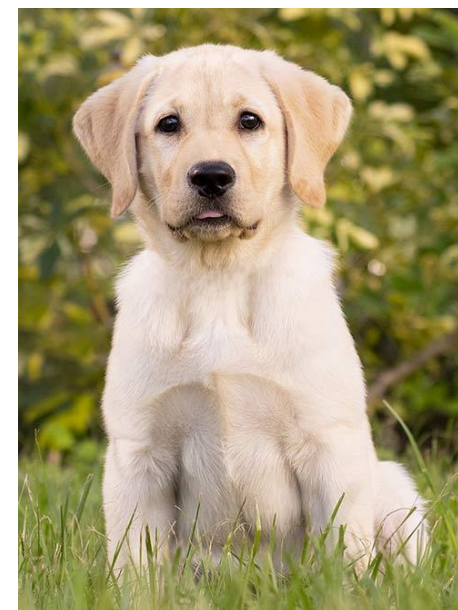
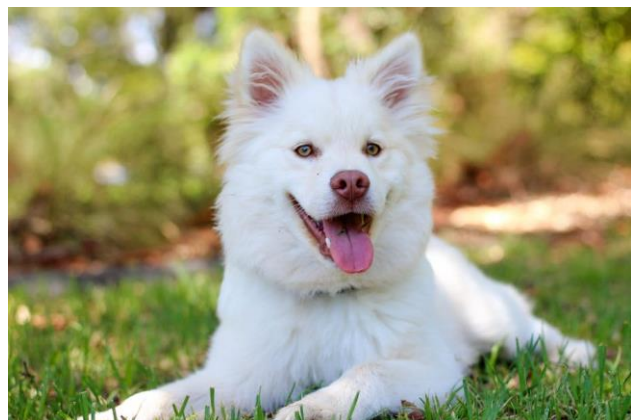


shakira: 73 %



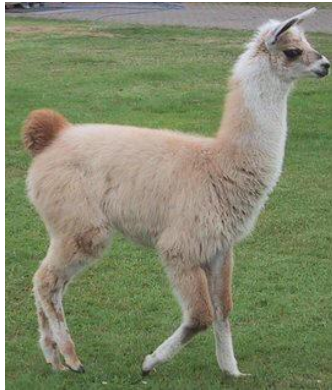
shakira: 98 %



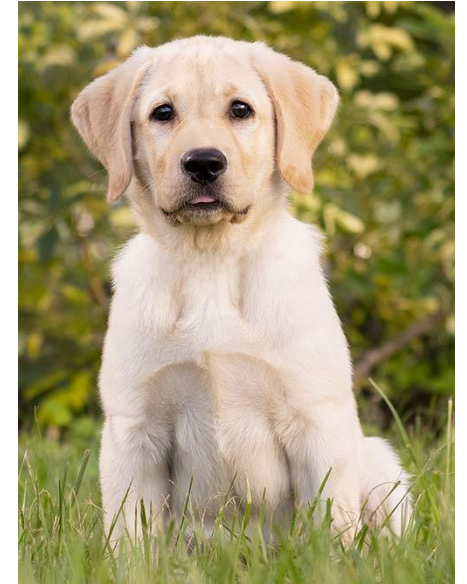


na, tady máš **obrázky...**

group\_A



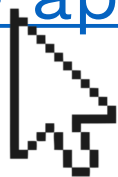
group\_B

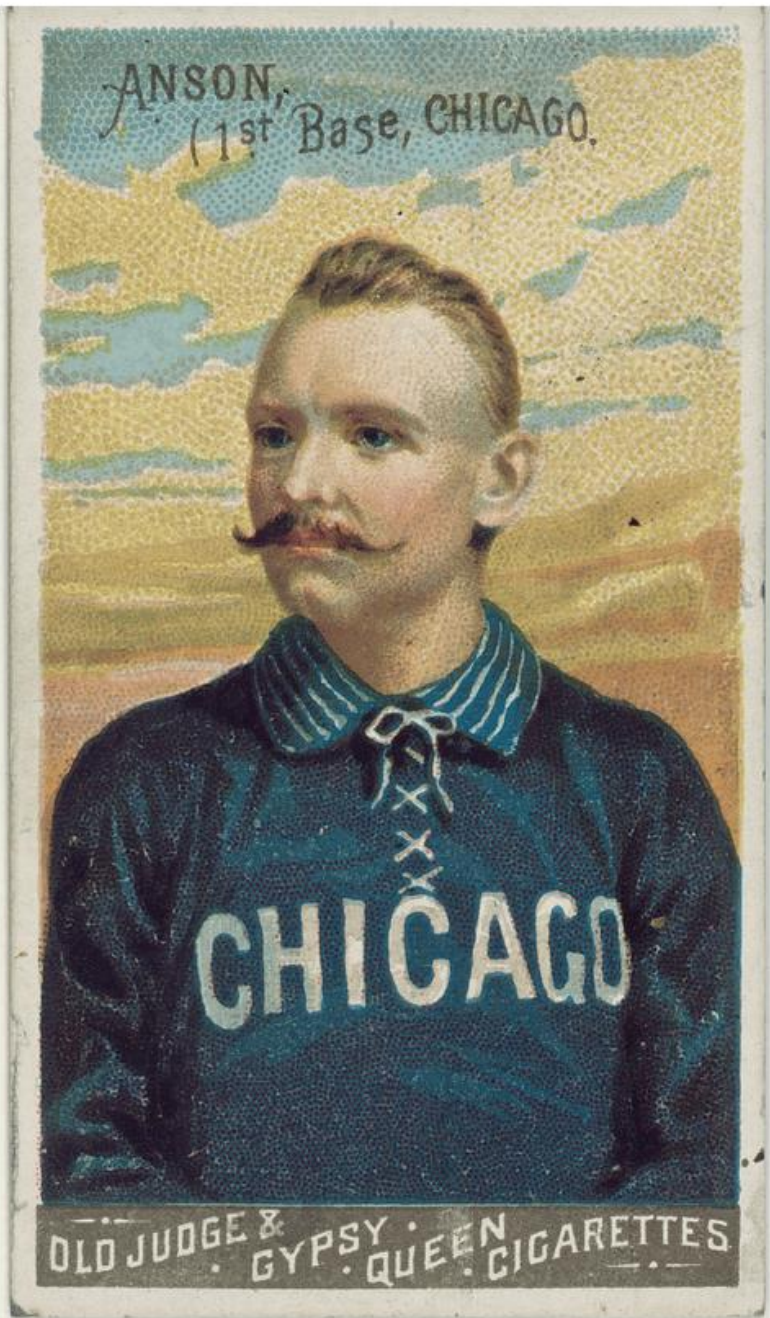


# Automatické rozpoznávání obrazu

*Několik ukázek aplikace mimo knihovny:*

- PowerPoint využívá automatický popis pro ALT texty
- [Ukázka komerčního řešení pro automatické rozpoznávání](#)
- [Ukázka praktické aplikace \(překladač\)](#)





PREDICTED CONCEPT	PROBABILITY
people	0.997
famous	0.993
popularity	0.982
vintage	0.982
retro	0.969
antique	0.966
old	0.961
portrait	0.950
person	0.946
man	0.925
talent	0.924
pop	0.923
<b>mustache</b>	<b>0.922</b>
stamp	0.912
post	0.910



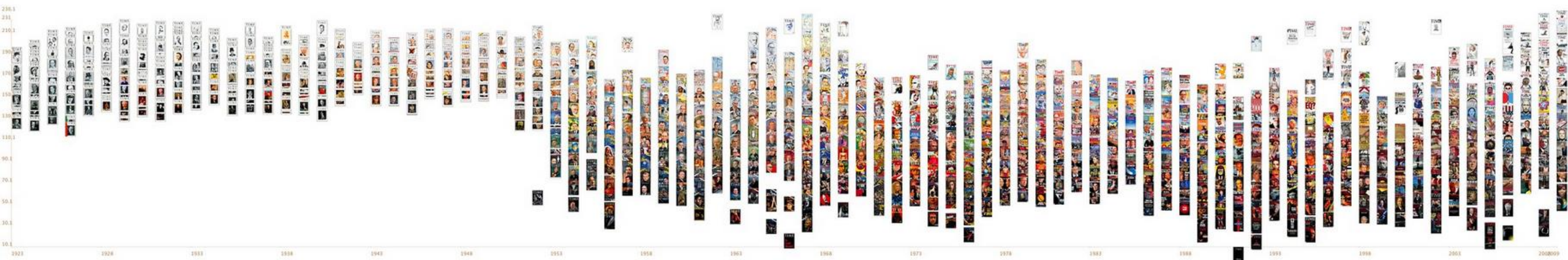
# Jak se s nimi dál pracuje?

Tvorba kolekcí a kurátorství

- *Library of Congress Pictures*
- *Europeana*

Význam pro digital humanities (DH).

[Ukázka SW ImagePlot využívaného v oblasti DH](#)





**Možnosti přístupu uživatelů k OD**

# Přístup uživatelů k fondům OD

- kolekce obrazů rostou neskutečným tempem
- nutnost v nich vyhledávat roste s nimi
- vyhledávání pomocí textových metadat je jasné
- hledáme ale i různé jiné způsoby vyhledávání v OD
- velkou oblastí je tzv. *content-based* vyhledávání
- podrobněji v předmětu *Vyhledávání informací*

# Content-based vyhledávání

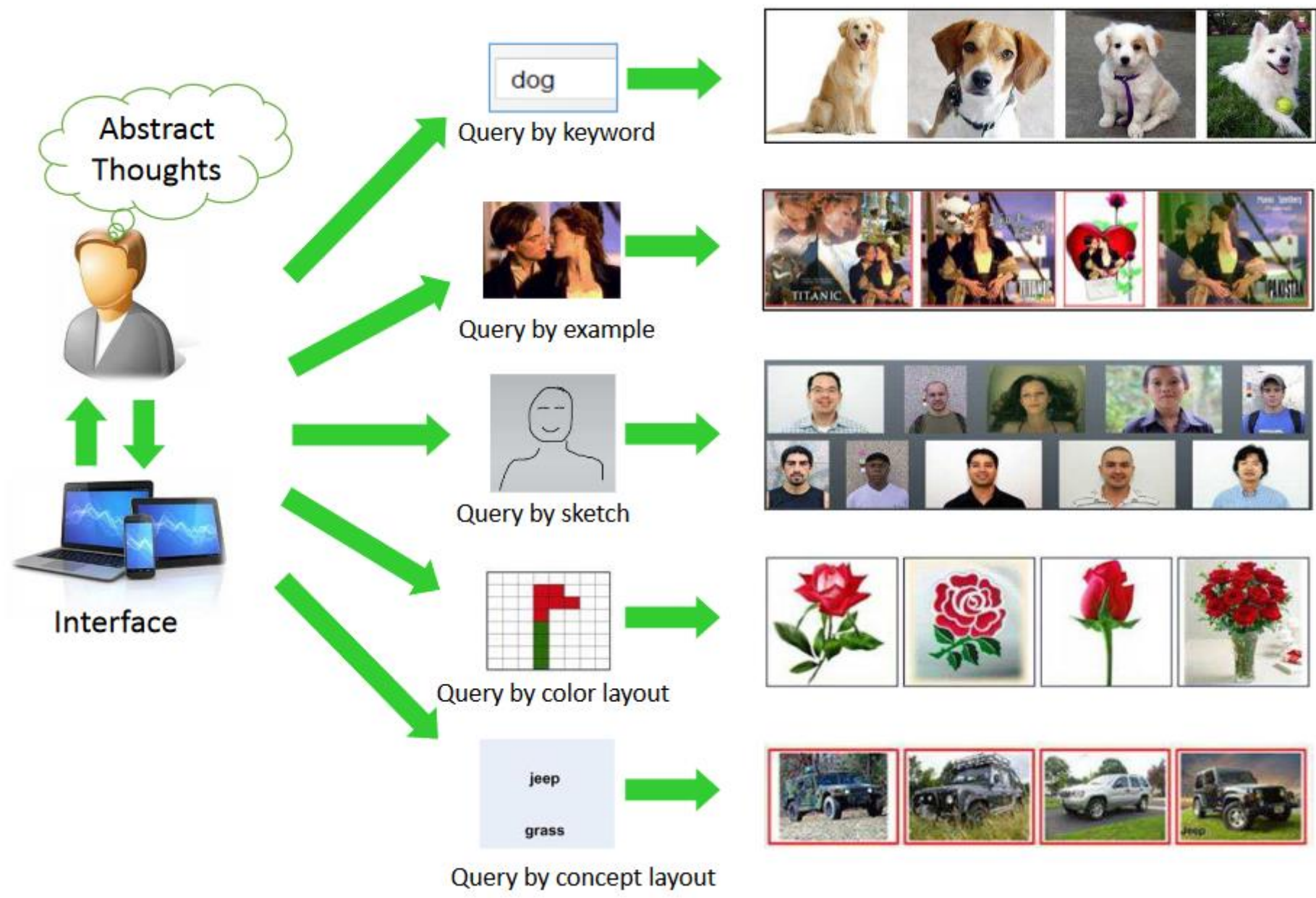
- metadata mimo textové charakteristiky
- prvním krokem je identifikace těchto charakteristik
- poté hledání způsobů, jak je zachytit/popsat

Co ještě může  
**charakterizovat**  
obrázek kromě jeho  
textového popisu?



# Content-based vyhledávání

- metadata mimo textové charakteristiky
- prvním krokem je identifikace těchto charakteristik
- poté hledání způsobů, jak je zachytit/popsat
- co může charakterizovat obrázek kromě textového popisu?
- např. barva, textura, tvar...



# Vyhledávání barvou

- implementováno v různých systémech
- využívá se často u digitalizovaných sbírek
- používá např. Europeana

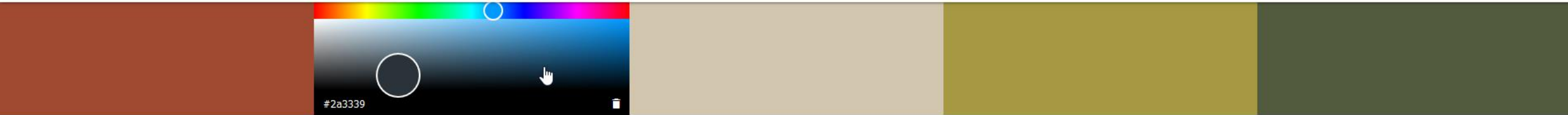


[Ukázka extrakce barvy z uměleckých děl k vyhledávání](#)

[Vyhledávací systém TinEye MultiColor](#)

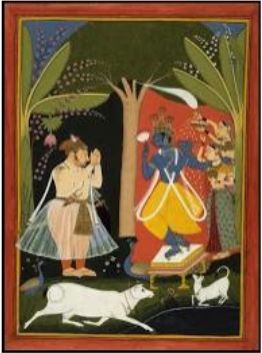






FROM AN IMAGE

SURPRISE ME



# Multicolr: Search by color

A TinEye Lab powered by MulticolorEngine

We used MulticolorEngine to extract the colors from 20 million Creative Commons images on Flickr to make the images searchable by color. Enjoy! Multicolr is addictive and very likely the best color search engine in the world\*!



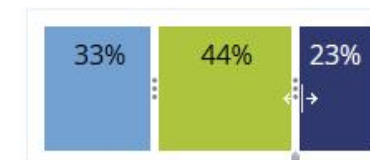
## Step 1

Select up to 5 colors



## Step 2

Slide dividers to adjust color composition



## Step 3

Add tags to refine your results

# Vyhledávání příkladem

- také jako reverzní vyhledávání
- *předkládám systému místo textu obrázek jako dotaz*
- umí to i Google Image Search
- první byl ale nástroj TinEye – reverzní vyhledávání
- [Aplikace Lens od Google](#)

Filtr rozvržení

Vyhledávání kompozice

W

C x

cheese

Zahrnout místo pro text

Vyhledávejte a klíčová slova se zobrazí v poli výše, připravena k přesunutí na místo.

Proč je tato stránka částečně v angličtině? [?](#)



# Vyhledávání nákresem

- [Projekt QuickDraw od Google](#)



# Specifické způsoby procházení

- *image collection exploration*
  - celá oblast, která hledá alternativní způsoby procházení obrazových fondů a kolekcí
1. [Archiv OD z dovolené jako 3D model](#)
  2. [OD jako galerijní výstava](#)
  3. [Zahrada v interaktivní podobě](#)
  4. [Přístup ke sbírkám skrze selfie](#)

# Případovka | Archivy NYPL

1. [Vstupní web do kolekcí NYPL](#)
2. [Obrazové dokumenty ve veřejné doméně](#)
3. [Vizualizace dokumentů ve veřejné doméně](#)
4. [Pátá Avenue tehdy a dnes](#)
5. [Speciální péče o stereografii](#)
6. [Stereogramimator](#)
7. [Mansion Maniac nad volně dostupnými plány](#)



**Tak zase (přes příští) pátek!**