

Wikimedia Commons tutorial

by Jarek Tuszynski
(user:Jarekt)

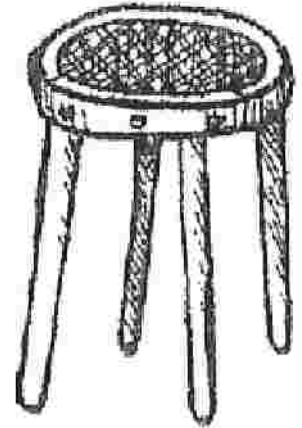
Commons in wiki universe

Commons is different than other projects

- Different focus: On Wikipedia everything revolves around articles, on Wikisource around books and on Wikimedia Commons everything revolves around files, most of which are images
- Different rules:
 - No [Neutral Point of View](#)
 - We keep files with extreme opposite points of view and let the other project decide which one to use
 - No [No Original Research](#)
 - We do “original research” all the time: identifying and categorizing animal species, locations, people, etc.
 - Ok to have usernames named after institutions
 - User:BrooklynMuseum

Commons

- Successful image has to have multiple attributes:
 - Graphics: good quality, high resolution, [in scope](#) (educational)
 - License: free, verifiable, complete, clear, concise
 - Categories: covering many aspects (orthogonal), detailed
 - Metadata: as full as possible, clear, concise, internationalized (i18n), organized in infoboxes



Graphics

- Scope:
 - [3.1 Must be a media file](#)
 - [3.2 Must be an allowable free file format](#)
 - [3.3 Must be freely licensed or public domain](#)
 - [3.4 Must be realistically useful for an educational purpose](#)
 - [3.4.1 File in use in another Wikimedia project](#)
- Quality:
 - Resolution as high as possible (even if image will be displayed as an icon in the article)
 - Better to have fewer high quality images than a lot of low quality ones

Images from previous partnerships with GLAM institutions



Licensing

- Free
 - Usually Creative Commons or Public Domain
 - Public domain due to: age, publication, official government materials, created by US government employee, etc.
 - No Fair use materials
- Verifiable
 - Provide: date of author's death, publication date or sources if license relies on it
- Complete
 - Image must be free in country of origin and in the US
 - Many authors in PD in their own country are still copyrighted in the US due to ULRA
 - All creators must be covered
 - We might have a scan of a famous photograph of a sculpture, which is a copy of earlier sculpture, so we might need to have license or 2 sculptors, one photographer and explain why person digitizing does not need a license
- Clear and concise

Licensing examples (2D)

{{Walters Art Museum license/2D}}

This is a faithful photographic reproduction of an original two-dimensional work of art. The work of art itself is in the [public domain](#) for the following reason:



This work is in the [public domain](#) in the United States, and those countries with a copyright term of life of the author plus 100 years or less.

This file has been identified as being free of known restrictions under copyright law, including all related and neighboring rights.

This digital reproduction has been released under the following licenses:



This work has been released into the [public domain](#) by its author, [Walters Art Museum](#). This applies worldwide.

In some countries this may not be legally possible; if so:

Walters Art Museum grants anyone the right to use this work for any purpose, without any conditions, unless such conditions are required by law.

In many jurisdictions, including the United States, faithful reproductions of two-dimensional public domain works of art are not copyrightable (see [Commons:Reuse of PD-Art photographs](#)). In these jurisdictions, this work is actually in the public domain and the requirements of the above license are not compulsory.

This is a faithful photographic reproduction of an original two-dimensional work of art. The work of art itself is in the [public domain](#) for the following reason:



This work is in the [public domain](#) in the United States because it was published before January 1, 1923.

The author died in 1926, so this work is also in the public domain in countries and areas where the copyright term is the author's life plus 30 years or less.

[Češky](#) | [Danok](#) | [English](#) | [Français](#) | [Italiano](#) | [日本語](#) | [Македонски](#) | [සැසඳුම](#) | [Nederlands](#) | [Română](#) | [Slovenščina](#) | [Svenska](#) | [中文 \(简体\)](#) | [中文 \(繁體\)](#) | [العربية](#)

The official position taken by the Wikimedia Foundation is that "faithful reproductions of two-dimensional public domain works of art are public domain, and that claims to the contrary represent an assault on the very concept of a public domain". For details, see [Commons:When to use the PD-Art tag](#).

This photographic reproduction is therefore also considered to be in the public domain.

Please be aware that depending on local laws, re-use of this content may be prohibited or restricted in your jurisdiction. See [Commons:Reuse of PD-Art photographs](#).

{{PD-Art | PD-old-100-1923}}

Licensing examples (3D)

Object:




This work is in the **public domain** in the United States, and those countries with a copyright term of life of the author plus 100 years or less.

This file has been identified as being free of known restrictions under copyright law, including all related and neighboring rights.

Photograph and description:

This file is licensed under the [Creative Commons Attribution-Share Alike 3.0 Unported](#) license.

Attribution: Walters Art Museum




You are free:

- **to share** – to copy, distribute and transmit the work
- **to remix** – to adapt the work

Under the following conditions:

- **attribution** – You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **share alike** – If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.



Permission is granted to copy, distribute and/or modify this document under the terms of the [GNU Free Documentation License](#), Version 1.2 or any later version published by the [Free Software Foundation](#); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled [GNU Free Documentation License](#).

Categorization

- Categories allow groupings of similar images
- An image could be categorized by multiple criteria:
 - Creator(s)
 - Subject: non-obvious features of the context (identification of animals, vehicles, weapons, geographical features)
 - Time: depicted time, time of creation
 - Location: depicted location, origin
 - Medium used: etching / painting / rare photographic process / rare type, etc.
 - Institution
 - Source: images from ... manuscript
- Many categories combine or intersects one or more basic criteria
 - [Islamic manuscripts in the Walters Art Museum](#)

Category intersections

[Category:Painted portraits of men of France](#) which seems like intersection of 4 basic categories:

1. [Category:Paintings](#) (1)
2. [Category:Portraits](#) (2)
3. [Category:Men](#) (3)
4. [Category:France](#) (4)

Logically it should be a subcategory of 4 intersections of 3 of above categories and 4 intersections of 2 of above categories:

1. [Category:Painted portraits of men](#) and [Category:Portrait paintings of males](#) (1-2-3) sub of
 1. [Category:Portrait paintings](#) (1-2)
 2. [Category:Portraits of men](#) (2-3)
 3. [Category:Paintings of men](#) (1-3)
2. [Category:Paintings of men of France](#) (1-3-4) sub of
 1. [Category:Men of France](#) (3-4)
 2. [Category:Paintings of people from France](#) (1-4)
 3. [Category:Paintings of men](#) (1-3)
3. [Category:Portraits of men of France](#) (2-3-4) sub of
 1. [Category:Portraits of people from France](#) (2-4)
 2. [Category:Men of France](#) (3-4)
 3. [Category:Portraits of men](#) (2-3)
4. [Category:Painted portraits of people from France](#) (1-2-4) sub of
 1. [Category:Portraits of people from France](#) (2-4)
 2. [Category:Portrait paintings](#) (1-2)
 3. [Category:Paintings of people from France](#) (1-4)

Serious issues with current system:

- All categories in English while many Wikipedia users do not speak English
- Intersecting categories produces a lot of extra categories and extremely specialized ones like [Category:Painted portraits of women from Spain in national costume](#)
- Wikidata properties might be the solution

Metadata: Roles of File Infobox Templates

- File description
 - In case of large collections of images from GLAM institutions, already compiled metadata should not be lost
- Information allowing verification of the license, like source, author, etc.
- Internationalization
 - Wikimedia Commons is shared among many projects using many languages. Users might not know English.
 - metadata should ideally be presented in the language of the user
 - links should point into most relevant Wikipedia
- Repeated metadata should be stored in as few places as possible to ease maintenance
- Machine readability

Three tiers of infobox templates

1. Most general: **Information** template (used by 88% of files 11.6M)
 - Fields: author, description, date, source, permission and other-versions
 - often (2.8M) accompanied by **Location** template for geo-location
2. Specialized infobox templates for specific types of images:
 - **Artwork**: for 2D and 3D museum artworks (360k)
 - **Photograph**: for museum photograph collections (82k)
 - **Book**: for images of scanned books (500k)
 - Other templates for: images of airplanes, busses, stamps, etc.
3. Customized templates for mass-uploads of images from a single source, are created based on one of above templates
 - Created for images from Bunesearchiv, Deutsche Fotothek, Walters Art Museum, National Archives and Records Administration, many other
 - Used in case none of the four standard templates is a good match to the provided metadata

Example of Artwork template

```
{{Artwork
|artist          = {{Creator:Cristofano dell'
  Altissimo}}
|title           = {{en|Portrait of
  Christopher Columbus}}
|description     =
|date            = 1556
|medium          = {{technique|Oil|on=
  canvas}}
|dimensions      =
  {{size|unit=cm|height=59|width=42}}
|institution     = {{Institution:Uffizi}}
|location        =
|references      =
|object history  =
|credit line     =
|inscriptions    =
|notes           =
|accession number =
|source          = {{WGA link|ID=244|pic-
  url=http://www.wga.hu/art/a/altissim/columbus
  .jpg|info-
  url=http://www.wga.hu/html/a/altissim/columbu
  s.html}}
|permission      = {{PD-art|PD-old-100}}
|other_versions  =
}}
```

I18n of {{Artwork}} fields

Artwork template field name	Recommended i18n template
author	{{Creator}}
institution	{{Institution}}
date	Dates should be in YYYY-MM-DD format or {{Other_date}} template should be used for phrases like around..., before..., between..., etc.
medium	{{Technique}} can be used for i18n of phrases like “adjective noun and adjective noun on adjective noun”. We have a vocabulary of 500 nouns and 60 adjectives translated into about 30 languages
dimensions	{{Size}} template is used for i18n of artwork dimension. It can translate width, height, depth, diameter, etc. and display numbers in metric or SI units depending on language
object history	{{ProvenanceEvent}} template can i18n phrases like “date: acquired by name, from name”

Creator and Institution templates

- Metadata about creators (painters, sculptors, writers, etc.) is stored in [Creator](#) templates (15.6k)
- Metadata about GLAM (Gallery, Library, Archive & Museum) institutions is stored inside [Institution](#) templates (1.8k)
- Both sets of templates occupy their own specialized namespaces
- Both sets produce in file namespace an collapsed template with mostly a name visible, which when clicked expands into informative infobox
- Most templates from both sets introduce links to Wikidata, Wikipedia and/or Wikisource pages, and links to Authority control databases.

Creator templates

```
{{Creator
| Name = {{LangSwitch
| de = [[[:de:Stanislaw Ignacy Witkiewicz
|Stanislaw Ignacy Witkiewicz]]
| en = [[[:en:Stanislaw Ignacy Witkiewicz
|Stanislaw Ignacy Witkiewicz]]
| ru = [[[:ru:?????????, ??????????
????????|????????? ?????????? ??????????]]
}}
| Alternative names =
  {{{name|pseudonym|Witkacy}}}
| Nationality = PL
| Gender = male
| Occupation = painter / photographer
/ writer
| Birthdate = 1885-02-24
| Birthloc = Warszawa
| Deathdate = 1939-09-18
| Deathloc = [[[:pl:Jeziory
(Ukraina)|Jeziory]]
| Workperiod =
| Workloc =
| Image = Witkacy Autoportret
1924.jpg
| Sortkey = Witkiewicz, Stanislaw
Ignacy
| Homecat = Stanislaw Ignacy
Witkiewicz
| Linkback = Creator:Stanislaw
Ignacy Witkiewicz
| Wikisource = s:pl:Autor:Stanislaw
Ignacy Witkiewicz
| Option = {{{1|}}}
| Authority = {{Authority
control|PND=118807706|VIAF=88039095|LCCN=
n/79/86775|ULAN=|TSURL=Stanislaw_Ignacy_W
itkiewicz|bare=1}}
```


Institution templates

```
{{Institution
| name          = {{Walters Art Museum}}
| native name  = Walters Art Museum
| location     = Baltimore, Maryland
| latitude     = 39.296077
| longitude    = -76.615305
| established  = 1934
| homecat     = Walters Art Museum
| website      = [http://thewalters.org/
  thewalters.org]
| linkback    = Institution:Walters Art Museum
| image       = Walters-museum-building_1.jpg
| authority   = {{Authority
  control|LCCN=no/00/96900|VIAF=153235247|ULAN=50
  0279110|bare=1}}
}}
```

Batch uploads of artwork images

- Preparation
 - Create needed templates and categories
- Metadata preparation
 - Match artwork artists with Commons creator templates and categories. Create new ones if needed
 - Match institutions with Commons Institution templates and categories. Create new ones if needed
 - I18n of other {{Artwork}} fields
- Possible approaches:
 - i18n before upload using spreadsheet (DB)
 - I18n after upload using multiple passes with a bot
- Both approaches take a lot of time with ~10% of odd cases taking 90% of time

Customization of existing infobox templates

- **Information, Artwork** and **Photograph** templates were designed to allow modifications by child templates
 - new fields can be added by using other_fields field and **Information field** template
 - unneeded fields can be left intentionally blank
- Modifications should be kept to minimum for consistency and common experience
- New fields should be internationalized

Example of [Walters Art Museum artwork](#)
template

Example of NARA-image-full template