

Structured Data for Commons

Project Overview

Prepared by
the [Multimedia](#) and [Wikidata](#) teams,
with the [Wikimedia Community](#)

8/21/2014

CC-BY-SA-3.0

Goals

Why are we doing this?

Goals

We aim to:

- **use structured data** for all media files on our sites
- make it easier for **users** to **read / write file info**
- enable **developers** to **build better tools**

This will be our multimedia team's **main infrastructure project** this year.

Users

This impacts **all users**:

- **readers**
- **contributors**
- **editors**
- **curators**
- **developers**

Notes:

- initial focus on **Commons**, then **Wikipedia** and other sites.
- during development, experienced users and developers most involved.

Benefits

How structured data can help our movement:

- offer a **better user experience**
- **find relevant content**
- drive **editing and re-use**
- provide **new ways to contribute**
- improve our **infrastructure**

Many different uses



Search



Viewing



Editing



Uploading



Curating

User stories

As a user, I want to:

- **Search for files** by multiple criteria
- **Learn more** with **related information**
- View file information **in my language**
- **Edit file information** more easily
- **Edit file info** from other pages
- **Add more topics** and data types
- **Upload / transfer files from any wiki**
- **Get smart suggestions**

[See more user stories](#)



WIKIMEDIA
COMMONS

[Main page](#)
[Welcome](#)
[Community portal](#)
[Village pump](#)
[Help center](#)

Language select

English

Select

Participate

[Upload file](#)
[Recent changes](#)
[Latest files](#)
[Random file](#)
[Contact us](#)

Tools

[Special pages](#)
[Printable version](#)

Better search: Help me find what I'm looking for

Search results



Search

Results 1 - 20 of 58,418 for orange

[Content pages](#) [Multimedia](#) [Translations](#) [Everything](#) [Advanced](#)

[Help](#) • [Search categories](#) • [Show other tools](#)

There is a page named "**Orange**" on this wiki

Orange

If you are looking for the **orange** colour, please see: **Orange** (color) and

Category:**Orange** . If you are looking for the **orange** tree, please ...

568 B (75 words) - 18:38, 14 March 2012

Category:**Orange**

Category:Types of red Category:Types of yellow Category:Colors by name

Category:Single colors.

335 members (12 subcategories, 322 files) - 16:06, 3 October 2011



File:California map showing **Orange** County.png

Map of **Orange** county in California. From en:migration relicense. Category:California county locator maps (yellow)Category:Maps of **Orange** ...

(200 × 231 (22 KB)) - 06:15, 29 May 2012



File:California county map (**Orange** County highlighted).svg



WIKIMEDIA
COMMONS

Main page
Welcome
Community portal
Village pump
Help center

Participate
Upload file
Recent changes
Latest files
Random file
Contact us

Tools
What links here
Related changes
Special pages
Printable version
Permanent link
Page information
Nominate for deletion

In Wikipedia
العربية
Català
Deutsch
English
Español
Eesti
فارسی
Français

Learn more with related information

From Wikimedia Commons, the free media repository

Download Use this file Use this file Email a link Information



Size of this preview: 419 × 600 pixels. Other resolutions: 168 × 240 pixels | 335 × 480 pixels | 537 × 768 pixels | 715 × 1,024 pixels.

Add a note

[Original file](#) (1,861 × 2,663 pixels, file size: 391 KB, MIME type: image/jpeg); (request rotation); ZoomViewer: [flash/no flash](#)

Expand view

Artist	Johannes Vermeer (1632–1675)	
	Alternative names	Johannes van der Meer, Jan Vermeer, Jan Vermeer van Delft, Johannes Reyniersz. Vermeer
	Description	Dutch painter and art dealer
	Date of birth/death	31 October 1632 (baptised) 15 December 1675 (buried)
	Location of birth/death	Delft Delft
	Work period	1653–1675
	Work location	Delft
	Authority control	VIAF: 51961439 · LCCN: n79006488 · GND: 118626590 · BnF: cb11940023b · ULAN: 500032927 · ISNI: 000000010901268X · WorldCat

Title *Het meisje met de parel* (*Girl with a Pearl Earring*)

Paintings of "[The Girl With The Pearl Earring](#)" and "[Jan Vermeer van Delft](#)" are both by Johannes Vermeer and available as public domain works on Wikimedia Commons.

Edit information from other pages

round black spots measuring from 2 to 3 cm (0.79 to 1.18 in) across, affording it some **camouflage** while hunting. There are no spots on its white underside, but the tail has spots, which merge to form four to six dark rings at the end. The tail usually ends in a bushy white tuft. The cheetah has a small **head** with high-set eyes. Black "tear marks" running from the corner of its eyes down the sides of the nose to its mouth keep sunlight out of its eyes and aid in hunting and seeing long distances. Its thin and fragile body make it well-suited to short bursts of high speed, but not to long-distance running.

Agility, rather than raw speed, accounts for much of the cheetah's ability to catch prey. Cheetahs can accelerate four times as fast as a human (thanks to greater muscle power) and can slow down by 14 **kilometers** per hour in one stride. They can hunt successfully in dense forests. ^[16]

The adult cheetah weighs from 21 to 72 kg (46 to 159 **lb**). Its total head-and-measure 60 to 84 cm (24 to 33 in) in length. ^{[17][18][19][20]} Cheetahs are 66 larger than females and have slightly bigger heads, but there is not a great apart by appearance alone. Compared to a similarly sized **leopard**, the che averages about 90 cm (35 in) tall) and so it appears more streamlined.

Some cheetahs have a rare fur pattern **mutation** of larger, blotchy, merged constitute a separate subspecies but are in fact African cheetahs; their unu "king cheetah" has only been seen in the wild a handful of times, but it has



Comparative illustration of a leopard (left) and cheetah (right)

The cheetah's paws have semi-retr the **flat-headed cat** and the **Irriomote** structure of the cheetah's claws is t and fur present in other varieties, a always visible. The dewclaw is muc

Adaptations that enable the cheetah increased oxygen intake, and an enlarged heart and lungs that work together to circulate oxygen efficiently. During a typical chase, its respiratory rate increases from 60 to 150 breaths per minute. ^[11] While running, in addition to having good traction due to its semi-retractable claws, the cheetah uses its tail as a rudder-like means of steering ^[citation needed] to allow it to make sharp turns,

1 with




Image details missing
The image will not be visible for readers until the details are provided

✎ 🗑️ 🗑️ 🖼️ +

Cheetah running Delete image

👤 Cronopio

Source This is my own work


© License Creative Commons Attribution ShareAlike 3.0

Add a description 🚩

English + 📅 Created today 🌐 More ▾

🏷️ Add categories

Get smart suggestions during upload ...



Cheetah running Cancel upload

📷 Robert Capa


Source From Flickr [Learn more](#) [Apply to all files](#)

Original URL ●

© License Creative Commons Attribution ShareAlike 3.0

Add a description

🏷️ Add categories



Cheetah running Cancel upload

📷 Robert Capa


Source From Flickr

Original URL http://flickr.com/photos/1234|

© License Creative Commons Attribution ShareAlike 3.0 checking...

Add a description

🏷️ Add categories




Cheetah running Cancel upload


📷 Robert Capa


Source From Flickr

Original URL http://flickr.com/photos/1234|


© License Creative Commons Attribution ShareAlike 3.0


 **Creative Commons - Attribution**
CC-BY Adaptations of the work are allowed

 **Creative Commons - Attribution Share Alike**
CC-BY-SA Adaptations are allowed as long as others share alike

 **Public domain**
Property rights have expired, have been forfeited, or are inapplicable.

Other licenses used by flickr cannot be used here

 All rights reserved

 Licenses preventing commercial use

Product

Example

Here's an example of how structured data can be used for a typical file:

Title (Label): White Russian Hamster

Description: The Djungarian hamster, also known as the Siberian hamster, Siberian dwarf hamster or Russian winter white dwarf hamster, is one of three species of hamster in the genus Phodopus.

File name: Hamster ruso blanco.JPG

File type: image/jpeg

Resolution: 4,320 × 3,240 pixels

Statements:

- Contributor: [Nanny99](#)
 - Contribution date: 17 May 2013, 04:10:09
 - Role: **Photographer**
- License: creativecommons.org/licenses/by-sa/3.0
- Terms of use: CC BY-SA 3.0
- Topic: **hamster**
- Topic: **pet**
- Type of work: **photograph**



[See more examples](#)

"Hamster ruso blanco" by [Nanny99](#). Licensed under [CC BY-SA 3.0](#) via Wikimedia Commons.

Key data

In a **first phase**, we would like to focus on these key data for each file:

- **title**
- **description**
- **work**
- **contributor**
- **role**
- **organization**
- **license**
- **category**
- **topics**

[See multimedia data list](#)

Key concepts

Here's one way to visualize some of the structured data clusters:

File

- title
- description
- metadata

Links to:

- work
- contributor
- license
- topics

Work

- title
- description
- etc.

Contributor

- name
- role
- etc.

License

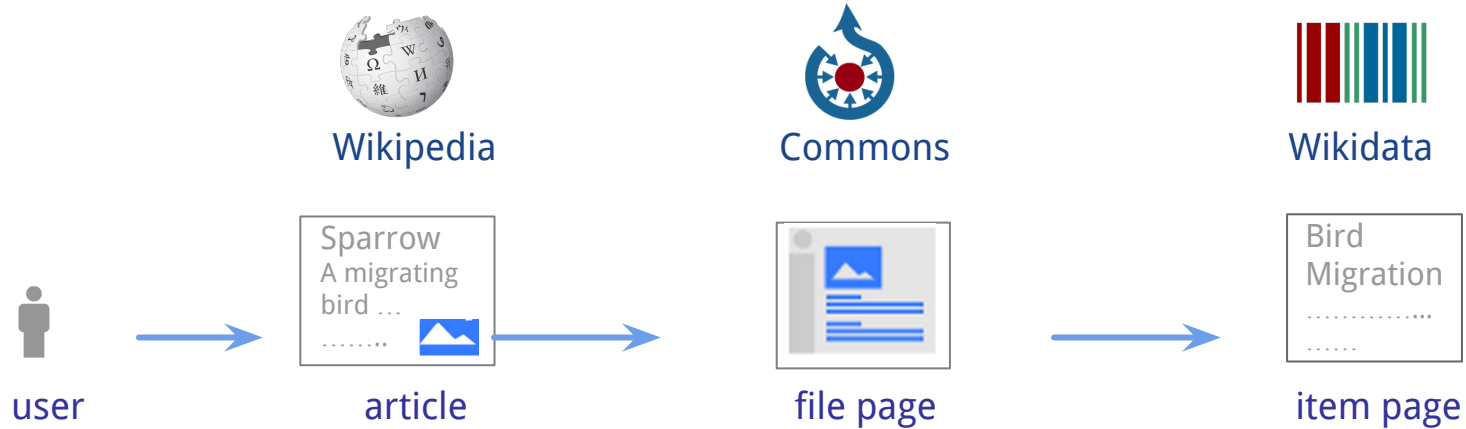
- title
- description
- etc.

Topic

- title
- description
- etc.

Touchpoints

Where will people find structured data?

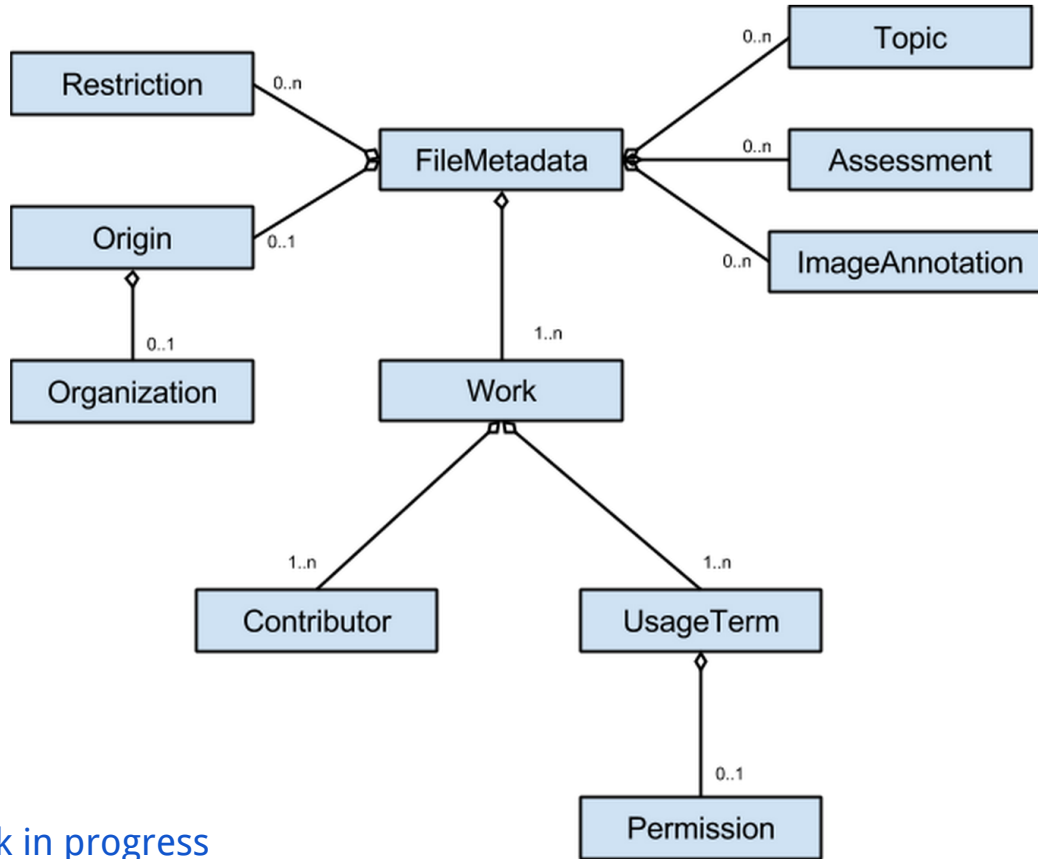


Storage

Where might data be stored?

Wikibase on Commons	Wikidata.org	Commons MediaWiki
title	work	file name
description	contributor	format
work link	license	resolution
contributor link	category	size
license link	topic	date
category link		geo-tags
topic link		version
file link		EXIF metadata

API Class Diagram



[See API work in progress](#)

Editing metadata - Before

The screenshot shows the Wikimedia Commons interface for editing the metadata of a file named 'Berta2007.jpg'. The browser address bar shows the URL: `https://commons.wikimedia.org/w/index.php?title=File:Berta2007.jpg&action=edit`. The page title is 'Editing File:Berta2007.jpg'. The main editing area contains a code editor with the following metadata:

```
=={{int:filedesc}}==
{{Information
|description={{en|1=My hamster Berta in 2007.}}
|source={{own}}
|date=2007-03
|author=[[User:Hugo85|Hugo Masters]]
}}

{{Location|52|31|N|13|22|E}}

=={{int:license-header}}==
{{cc-by-sa-3.0-de}}

[[Category:Uploaded with UploadWizard]]
[[Category:Molla Nasreddin]]
[[Category:1909 paintings in Azerbaijan]]
```

Below the code editor is an 'Edit summary:' field and two checkboxes: 'This is a minor edit' and 'Watch this page'.

The left sidebar contains the Wikimedia Commons logo and navigation links:

- Main page
- Welcome
- Community portal
- Village pump
- Help center
- Participate
 - Upload file
 - Recent changes
 - Latest files
 - Random file
 - Contact us
- Tools
 - What links here
 - Related changes
 - Special pages
 - Page information
 - CMD metadata
 - Nominate for deletion

The top navigation bar includes language options (English, Tgr (WFM)), user preferences (Talk, Preferences, Beta, Watchlist, Contributions, Uploads, Log out), and a search box.

Editing metadata - After

The screenshot shows a web browser window with multiple tabs. The active tab is 'Editing File:"Molla N...'. The address bar shows the URL 'https://commons.wikimedia.org/wiki/MediaInfo:Berta2007.jpg'. The page content includes a sidebar with navigation links, a main header with the Wikidata logo, and a main content area with a table of statements.

Item [Discussion](#) [Read](#) [View history](#) [More](#)

Hamster Berta ^(Q6573)

[\[edit\]](#)

My hamster Berta in 2007. [\[edit\]](#)

Statements

contributor	http://commons.wikimedia.org/wiki/User:Hugo85 [edit]
	role photographer
	[add]
time of creation	March 2007 [edit]
	[add]
topic	hamster [edit]
	photograph
	[add]
license	CC-BY-SA-3.0-DE [edit]
	[add reference]
	[add]

commons-wiki...html [Show all downloads...](#)

Next steps

Tasks

- **Plan & discuss** this project with our communities
- **Design** the data structure and user interface
- **Develop** the code and tools needed for this project
- **Migrate** unstructured data to the new format
- **Test, observe** and adjust plan as needed

Roadmap 2014-15

	Stage 1	Stage 2	Stage 3
product	<ul style="list-style-type: none">• planning• design	<ul style="list-style-type: none">• specs• metrics	<ul style="list-style-type: none">• learn• update
developers	<ul style="list-style-type: none">• code review• research	<ul style="list-style-type: none">• prototypes• small features	<ul style="list-style-type: none">• info / file pages• editing tools
community	<ul style="list-style-type: none">• discuss• guide	<ul style="list-style-type: none">• discuss• experiment	<ul style="list-style-type: none">• migrate data• edit data

These preliminary goals are under review and subject to change.

Components

Here are some of the first modules we might work on:

- **High-level API**
- **Media Info Page**
- **File Page**

- **Search**
- **Upload Wizard**
- **Media Viewer**

Questions

- Which **workflows** do we develop first?
- How do we define the **data structure**?
- How can we **measure** and **validate** each feature?
- How do we gradually **roll out** these features?
- How do we coordinate the **data migration**?

What now?

What will happen in coming months:

- **more discussions**
- **planning, design, specs**
- **developer bootcamps**
- **first prototypes & tests**
- **learn, adjust, try again**

- **be patient: this will take time ...**

Learn more:

commons.wikimedia.org/Commons:Structured_data

A collaboration between:

- **Wikidata Team** (Wikimedia Deutschland)
- **Multimedia Team** (Wikimedia Foundation)
- **Wikimedia Community**