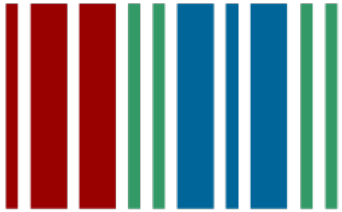


Structured Data for Multimedia

Bringing Wikidata to Commons

Community Discussion at Wikimania 2014

Hosted by the Multimedia and Wikidata teams
8/07/2014

CC-BY-SA-3.0

Agenda

10:00 - **Introductions**

10:10 - **Presentation**

10:30 - **Discussion**

11:30 - **Meeting ends**

Please add your comments on this notepad: **ur1.ca/hx8v6**

* All times above are GMT+1 (UTC)

Goals & Uses

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 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

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

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

Search results

orange

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

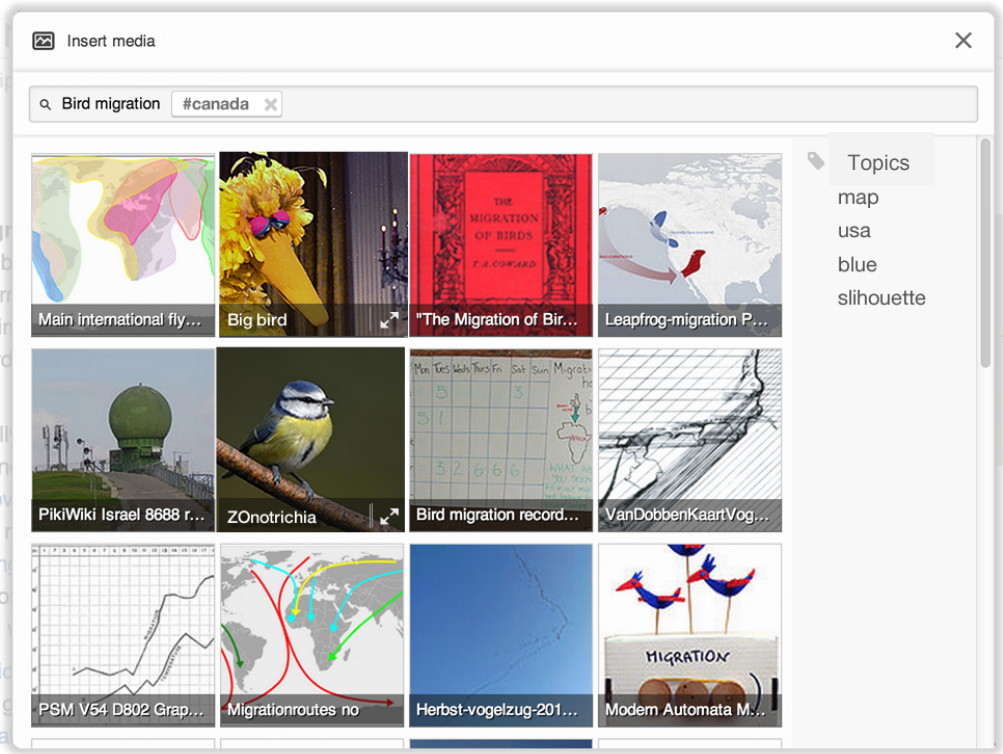
Orange
If you are looking for the **orange** colour, please see: **Orange** (color) and Category:**Orange** . If you are looking for the **orange** tree, please **...**
568 B (75 words) - 18:38, 14 March 2012

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

Search by multiple topics



Learn more with related information

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

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

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	

Edit information from other pages

Cancel | Save page

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 Iriomote 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,

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

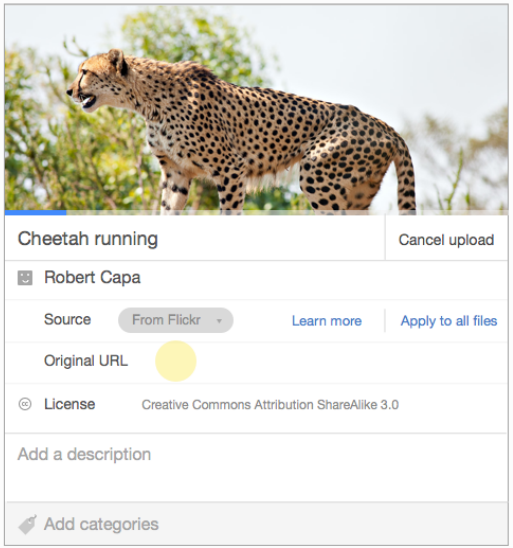
License: Creative Commons Attribution ShareAlike 3.0

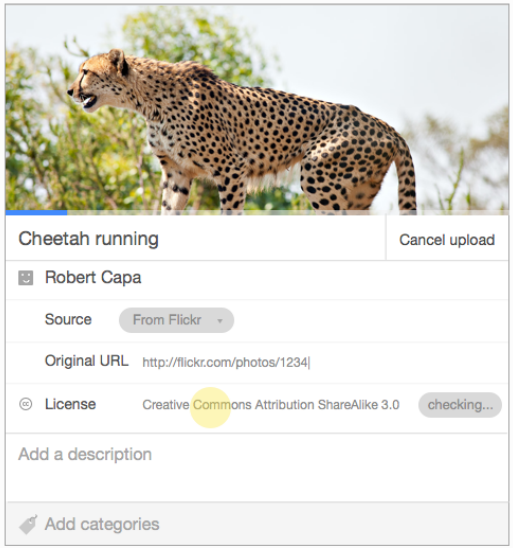
Add a description

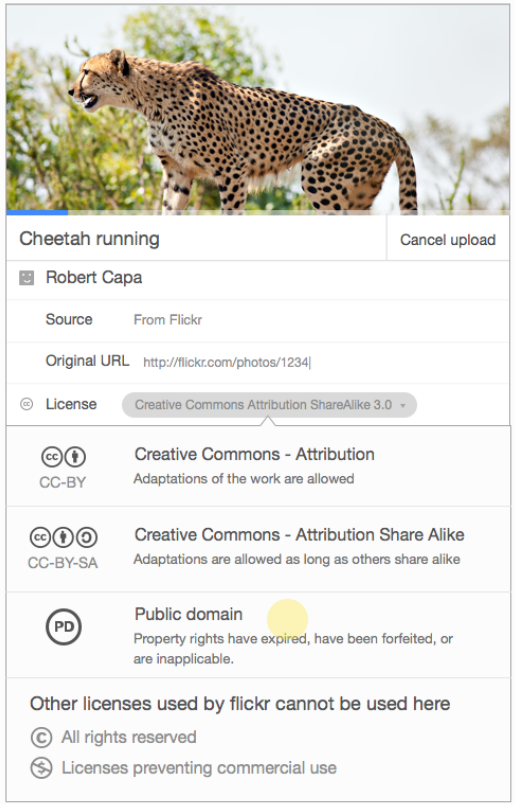
English | Created today | More

Add categories

Get smart suggestions during upload ...







Overview

Example

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

Title (Label): Hamster Berta
Description: My hamster Berta in 2007.
File name: Berta2007.jpg
File type: image/jpeg
Resolution: 800x800
Statements:

- Contributor: http://commons.wikimedia.org/wiki/User:Hugo85
    - Real Name: Hugo Masters
    - Time of contribution: March 2007
    - Role: Photographer
- Rights: http://creativecommons.org/license/CC-BY-SA-3.0-DE
- Terms of use: BY-SA
- Topic: hamster
- Topic: pet
- Type of work: photograph
- Coordinate location: 52’31 N, 13’22 E
- Location: Berlin

See more examples

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.

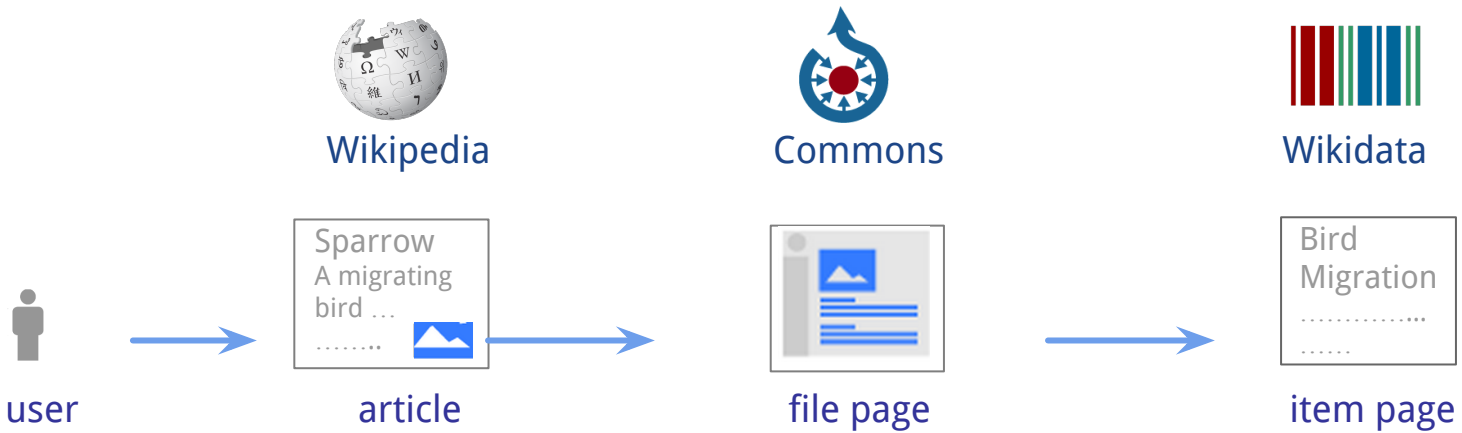
Contributor
- name
- role
- etc.

License
- title
- description
- etc.

Topic
- title
- description
- etc.

Touchpoints

Where will people find structured data?

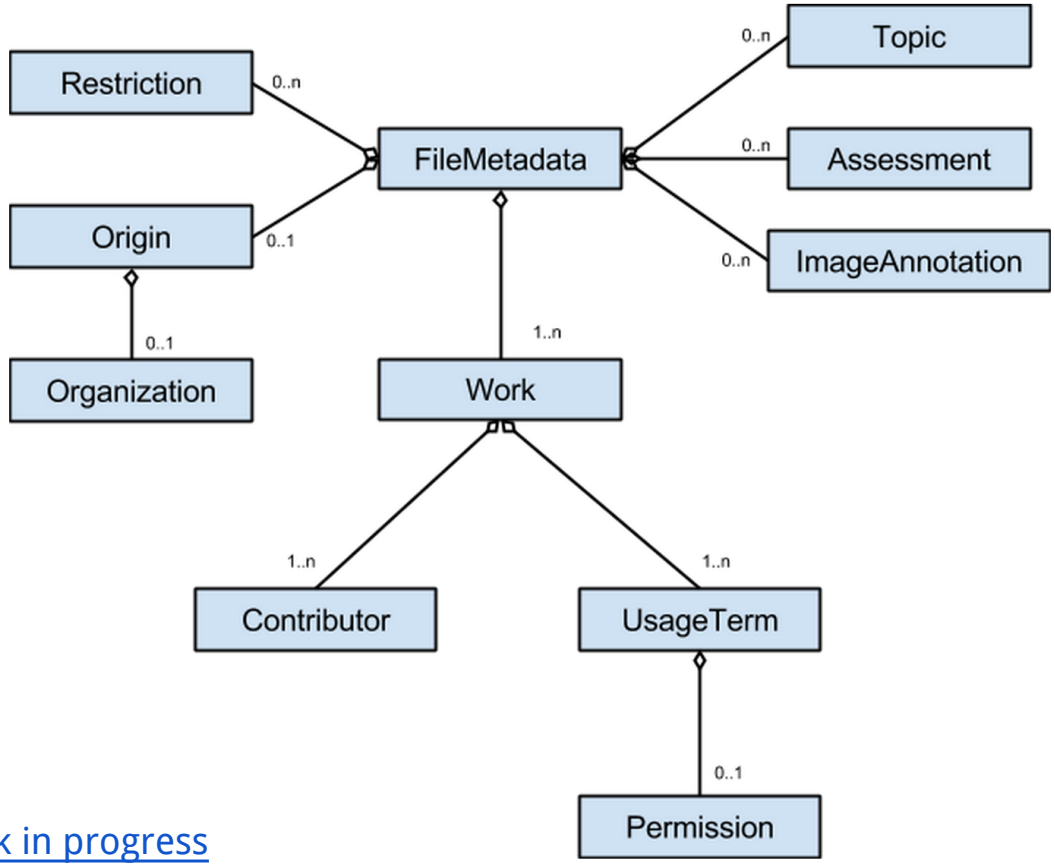


Storage

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]

Next steps

Tasks

- **Plan & discuss** this project with our community
- **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	• planning • design	• specs • metrics	• observe • update
developers	• code review • prototype	• first tests • small features	• info / file pages • editing tools
community	• discuss • guide	• discuss • experiment	• 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

- What are some of your **main issues** with unstructured data?
- Are your current **workflows** addressed by this first proposal?
- How do we define a basic **data structure** for key use cases?
- 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:

mediawiki.org/wiki/Multimedia/Structured_Data

A collaboration between
Wikidata Team (Wikimedia Deutschland)
Multimedia Team (Wikimedia Foundation)
Community (Wikimedia Commons, others)

CC-BY-SA-3.0