WIKIMEDIA FOUNDATION

Cross-Wiki Uploads on English Wikipedia

April 2024

Skip to the Arabic Wiki report.

Bethany Gerdemann, Eman Yahia
For the Structured Content Team

WIKIMEDIA FOUNDATION

01 About

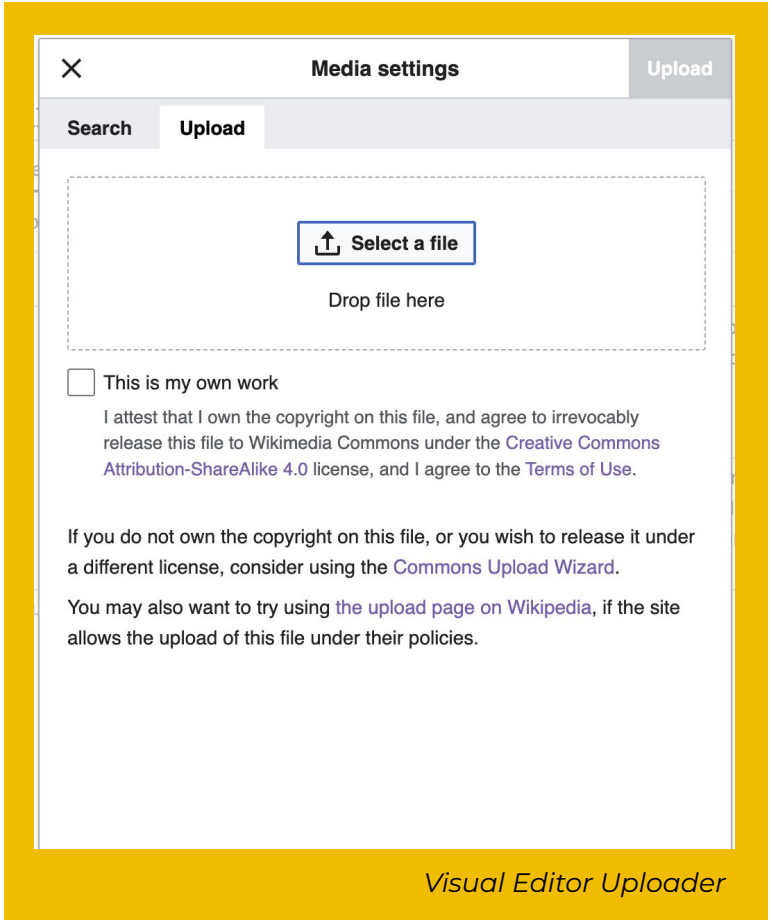
Context

Interviews with Wikimedia Commons Administrators in 2023 revealed the perception that a disproportionate number of “own work” copyright violations are transferred to Commons through cross-wiki uploads*.

The Structured Content team hypothesized that that current visual editor UI encourages users to upload copyrighted material because the only option for authorship presented in the UI is “own work”. Uploaders proceed by selecting “own work” when the work is not their own.

Because the available data to validate the phenomenon is incomplete, this project proposes a qualitative investigation of the experience of users who have recently uploaded their first image to Wikipedia, focusing on how they interpret the UI and their understanding of Wikimedia’s image policy, including what constitutes “own work”.

These findings can be applied towards UI changes and other solutions aimed at reducing improper files on Commons, thereby reducing the workload for moderators.

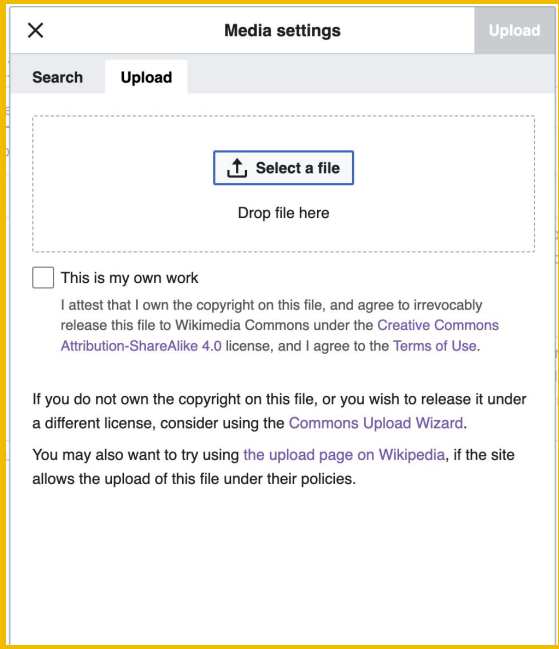


Visual Editor Uploader

*Cross-wiki uploads are files uploaded via Wikipedia, often to illustrate an article, and then automatically transferred to Wikimedia Commons.

Research Questions

- ➔ What are the mindsets of users who upload images to Wikipedia?
- ➔ What are their goals and intentions when uploading images?
- ➔ What is their experience whilst uploading images?
- ➔ How do users interpret and interact with Visual Editor upload UI?
- ➔ How do users interpret “own work” and its requirements?
- ➔ What is the awareness, understanding, and sentiment regarding copyright?
- ➔ What is the awareness, understanding, and sentiment regarding the image policy?
- ➔ What do users need in order to comply with the image policy?



Visual Editor Uploader

Method

We conducted 12 semi-structured interviews with users of en.wikipedia.org.

- 11 participants have uploaded images to Wikipedia in the past 60 days.
- 1 pilot participant who is a new editor, but has not uploaded an image.

We spoke with one Wikimedia Commons Admin who provided context and advice.

We conducted one follow up interview with one of the participants.

Additionally, we conducted 5 interviews with users of Arabic Wikipedia who have recently uploaded images. The results from those interviews are presented in a separate report.

WIKIMEDIA FOUNDATION

02 Participants

Participants

Gender
10 male
2 female

Location
4 United States
3 Britain
2 India
2 Nigeria
1 Uganda

Profession
10 working professionals

- IT professional
- Telecommunications marketer
- Communications strategist
- Lawyer
- Trade industry leader
- Digital project manager
- Engineer
- Software engineer
- Legal Marketer
- Biology researcher

1 PhD student

- Classics and philosophy

1 undergraduate student (pilot)

- Electrical engineering

Age range
19-66 years

User Experience Level
10 new uploaders
1 experienced uploader
1 new editor, never uploaded

Mindsets, Beliefs, and Values

Mindsets

- All participants have good intentions.
- Consider Wikipedia to be a trustworthy source of information.
- Confused about Wikipedia's scope.
- Uninformed about policies on Wikimedia.
- Operate under many false assumptions, often based on social media platforms.

Beliefs about rules on WP

- All participants respect rules.
- All understand that WP needs rules in order to maintain its quality and integrity.
- Participants want to follow rules, and do so to the extent that they understand them.
- Participants intend to upload properly, according to what they believe are the rules and requirements.
- Participants who have violated policy are unaware that they did so, or were unaware at the time.

Beliefs about copyright

- All participants respect the concept of copyright.
- Many are creators and other professionals who have direct experience with copyright concepts.
- All recognize the need to respect ownership and authorship of intellectual property.
- All understand that WP needs to respect intellectual property.
- No one understands the technicalities of copyright, especially when it comes to their image.

WIKIMEDIA FOUNDATION

03 Key Findings

Wrongly selecting “own work” is a pervasive problem on the platform, and occurs more frequently than we are aware.

12 participants total

- ➔ 1 uploaded their own work.
- ➔ 1 uploaded someone else’s work correctly.
- ➔ 1 (pilot) participant hasn’t uploaded any images.
- ➔ **9 uploaded others’ work as “own work”.**

This pattern suggests that new users:

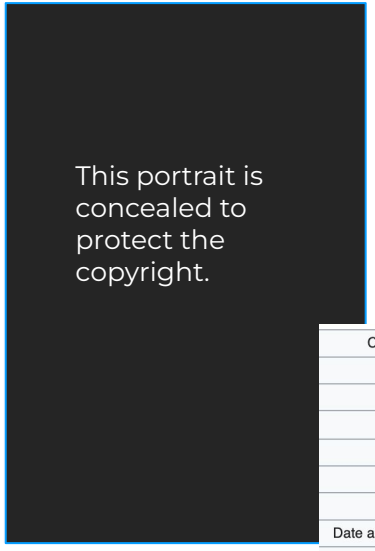
- don’t understand **how to properly upload “not own work” images;**
- don’t understand **when they *should not* upload an image.** They don’t consider *whether* they should upload an image, they are focused on ***how* to get it done**.
- **overlook the other uploading methods** as they instinctively check the only box on the screen in order to move through the uploading process;
- or they are **deterred from selecting “not own work”** because of the multi-step process of the Upload Wizard and the confusing selections and requirements needed to complete it.

Of the 11 participants, two user patterns emerged:

The Self-promoter and the Marketer.

The Self-Promoter

(5 of 11 participants)



Camera manufacturer	Canon
Camera model	Canon EOS 60D
Author	Arun_Anna_Shiva
Copyright holder	Shiva_arun_anna
Exposure time	1/125 sec (0.008)
F-number	f/10
ISO speed rating	100
Date and time of data generation	17:09, 2 February 2014
Lens focal length	50 mm

Description	**English:** Rj Kranthi
Date	2 February 2014
Source	Own work
Author	Username

Goals:

- Writes an article about self.
- Adds image of self to article.
- Some are using the platform for a professional or personal profile online (academic researcher).

Beliefs:

- Some believe they are notable (local radio jockey).
- One met notability requirement.

Upload process:

Uploads image of **self** using **VE** upload, as "**own work**".
Photo was taken by someone else.

Outcome:

- Article is likely to be deleted or rejected (but viewable in their sandbox).
- Image is **deleted** or **not yet moderated.**
  - Personal image violation (F10: Speedy).
  - Clear Copyvio (F1)-when © is in metadata.
  - "Own work" CR violation.

Self-promoters violate the personal image policy, and often copyright too.

> *It's my image. It was taken with my phone*. -P12E

- Self-promoters are unaware of the "notability requirement", and that uploading an image of themself violates the personal image policy.
- They comfortably use the VE uploader believing that their image is *own work*, believing the image is *theirs* because it is of them and often taken with their phone.
- Self-promoters are unaware that the copyright is with the person who took the photo, thus uploading as "own work" violates the copyright policy.
- All 5 self-promoters in our study uploaded someone else's work as their own.
- The file either goes unmoderated or is deleted by moderators, which leads to confusion and discouragement.
- Rather than addressing the copyright violation related to "own work" for self-promoters, it may be more effective to address the personal image violation.

The Marketer

(5 of 11 participants)

This image depicts the exterior of a hotel. It is concealed to protect the copyright.

Camera manufacturer	Canon
Camera model	Canon EOS 5D Mark IV
Author	Ray Main
Exposure time	1/200 sec (0.005)
F-number	f/11
ISO speed rating	100
Date and time of data generation	17:35, 1 June 2023
Lens focal length	16 mm
Online copyright statement	www.raymain.co.uk
Copyright holder	© Ray Main 2023. All rights reserved. Contact ray@raymain.co.uk - 00 44 7831 200 901
Image title	Maybourne Hotel Group01/06/2023
Short title	Claridge's State Rooms
Show extended details	

Description	**English:** The exterior of Claridge's
Date	1 June 2023
Source	Own work
Author	Username

Goals:

- Often employee editing on behalf of employer or someone representing another party.
- Creating an article for their company or adding small edits on employer's article.
- Replaces existing image with one that is better, uploads a company logo, or adds the first image to the article.

Beliefs:

- They are on work assignment to upload the image or they have permission to upload the image.

Upload process:

Many marketers navigate to the **Upload Wizard** because they know that the image **isn't "own work"**, but are deterred by the license step and ultimately return to VE to upload as "**own work**".

Outcome:

- Images with copyright metadata are likely to be **deleted** or **not yet moderated**.
  - Clear Copyvio (F1)-when **©** is in metadata.
  - "Own work" CR violation
- Logos are likely to be **deleted** or **not yet moderated**.
  - Logos are usually considered copyrighted, and therefore risk being deleted as a Copyvio.

There is no pathway for the Marketer to upload their company's images.

> *I got permission from the company's communications department, and they sent me the file directly.* -P5E

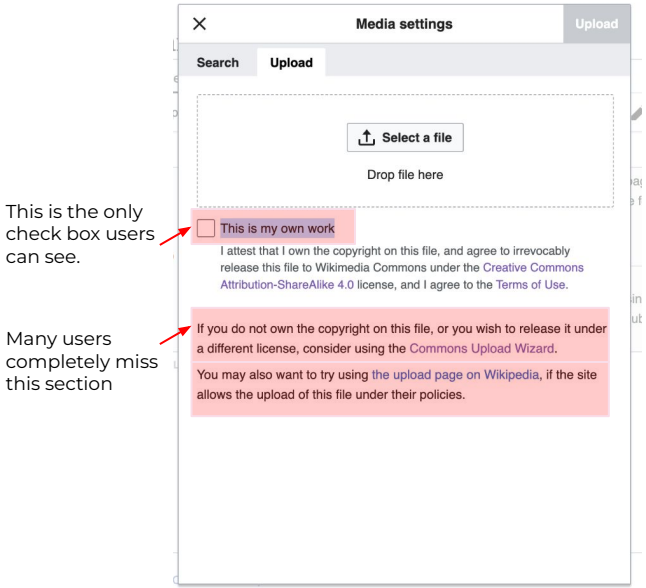
What marketers do:

- Marketers do not know the copyright status of the files, but they believe their company owns them.
- These images can be logos or professional images that the company uses across the internet.
- Marketers struggle to find the right author / license selections, and don't realize that, in some cases, they should not upload the file.
- This leads to frustration and discouragement. They upload as "own work" for lack of better options.
- When the file is deleted by moderators, participants experience more confusion about why and how to resolve the issue.

What they should do:

- The image policy demands that the © owner proves ownership status through the VRT process, which is not built in to the upload process, nor is it explained during the process, and in itself is cumbersome.

Upload Wizard is too complicated and the Visual Editor is too simple.



- Users come with the **goal of uploading an image as part of a task** they intend doing on Wikipedia (e.g. writing or editing an article.)
- They expect the process takes no more than a few seconds but **it turns out to be complicated**.
- They use **the simplest and easiest method** to get their job done.
- **With good intentions**, they click on “my own work”, as the design suggests, to proceed with the uploading process.
- Many don’t even see the other upload options, and **only see the checkbox**.
- **Some users don’t spend lots of time** thinking about what the phrase “my own work” means. These users **overlook the other options.**
- **Other users attempt to use the Upload Wizard**, but are discouraged by the long, complicated process, and return to the visual editor uploader to complete the task.

Participants do not know and have not seen the image policy.

1	**Has seen the image policy**
5	**Think that the licenses step or other prompts constitute the policy**
6	**Are not sure or never encountered the policy**

- This contributes to many violations that could be prevented by presenting the policy before upload.
- Most participants confound copyright, licensing, the image policy, the instructions for uploading, and the terms of use, demonstrating their confusion about what each of these are and how they should interact with them.
- Most participants assume that all of the image requirements are built into and mediated by the upload process.
- Many assume the an automated detection system will filter out ineligible images.
- All participants expect to see the image policy, or key points concisely summarized, during the upload process. In the absence of that, they assume they are meeting all requirements.

> *There was definitely a lack of clarity that distinguish the instructions from the rules. I was fending for myself, trying to work it out.* -P8E

Because they do not know the image policy, participants unknowingly violate it.

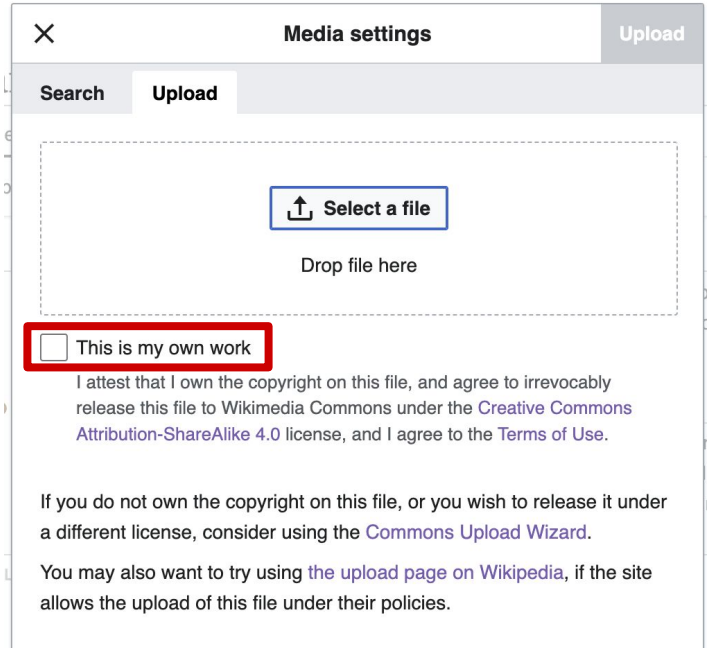
- Because they don't know the policy, they intuit the rules based on other platforms and what they observe on Wikipedia, which leads to false assumptions.
- Participant's assumptions about the rules are mostly inaccurate, and show there is little understanding of how copyright is applied on the platform.
- No one is aware that personal images are not allowed on Wikipedia, yet it is one of the most frequent violations.

What do you think are the rules for uploading images to Wikipedia?

8	Not lewd, offensive, or controversial
5	Related to the article topic
4	Not copyrighted
4	Must attribute author
3	Must be a good image (well composed, good image quality, HD)
2	File description: naming convention, description, keywords, etc.
2	Must provide source of the file
2	No doctored or misleading images
1	Must have permission to upload
1	Useful for Wikipedia
1	For public use
1	Must own rights to the work
1	Own work only
1	Anyone featured in the photo needs to give consent
0	No selfies or personal images

Policy in green

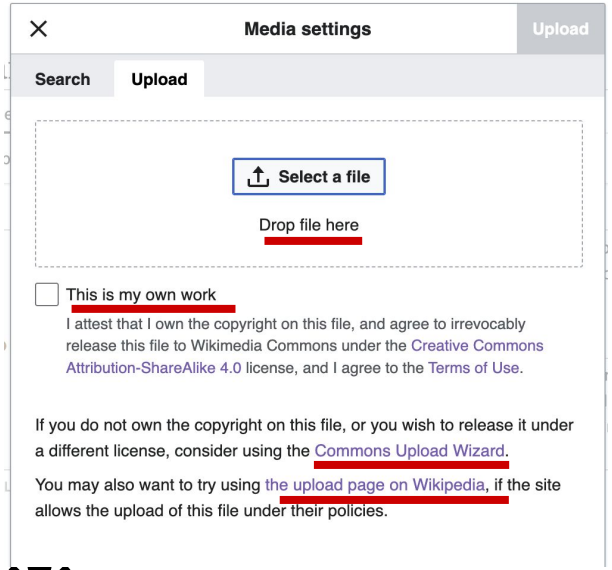
Participants think “own work” means different things.



6	I have rights/authority over the use of the image
4	I created the image myself only
1	I'm involved in the creation of this work, therefore it is mine
1	Image is "free to use"

> *It’s kind of hard because it means I created the work myself. I have ownership. Or in this context, my company’s work? In this case I assumed that my own work referred to my own company’s work.* -P8E

Participants misinterpret the text on the Visual Editor Uploader, and do not understand when to use the 3 upload options.



0	Correctly interpreted all statements
4	Correctly interpreted “own work”
2	Correctly interpreted “I attest that…”
4	Correctly interpreted when to use Commons Upload Wizard
0	Understood what is the upload page on WP and when to use it.

> *Why is this an option? Why are there multiple options to upload media on this page? Why can’t I just do it here? Why do I have to upload it somewhere else? I thought I was on the upload page on WP. So, where am I?* -P12E

Everyone understands and respects the concept of copyright, but participants are mistaken about how applies to images on Wikipedia, especially their uploads.

> *I work at an Intellectual Property law firm. We do a lot of copyright cases. It's important.* -P11E

> *I've dealt with a lot of idea theft, so I take copyright very seriously.* -P5E

- All participants showed understanding of the basic concept of copyright, and recognize its importance for protecting creators.
- Despite this, participants are largely mistaken about who owns the copyright.
- The UI assumes that users know that the author of an image is the CR holder by default. Understanding that is necessary to proceed with the upload forms correctly. But that is not explained anywhere and users aren't aware.
- Many believe that files that are used around the internet are okay to upload to Wikipedia (e.g. images available on Google image search, used by their company for marketing on other sites, uploaded to social media, etc.)
- Many expect that the upload tool to detect and block copyrighted files. Many assume that their file is okay to upload if it passes through the upload process.

Participants do not understand what they are agreeing to when they upload. They have a have a false impression of what will happen to their image.

> *I don't need to know what CC A-SA 4.0 is. I trust Wikipedia to do the right thing.* -P7E

> *If Wikipedia is using my image, it's only authorized for particular things. When uploading image, I'm giving the copyright to Wikipedia for use on that particular page, not for general use.* -P10E

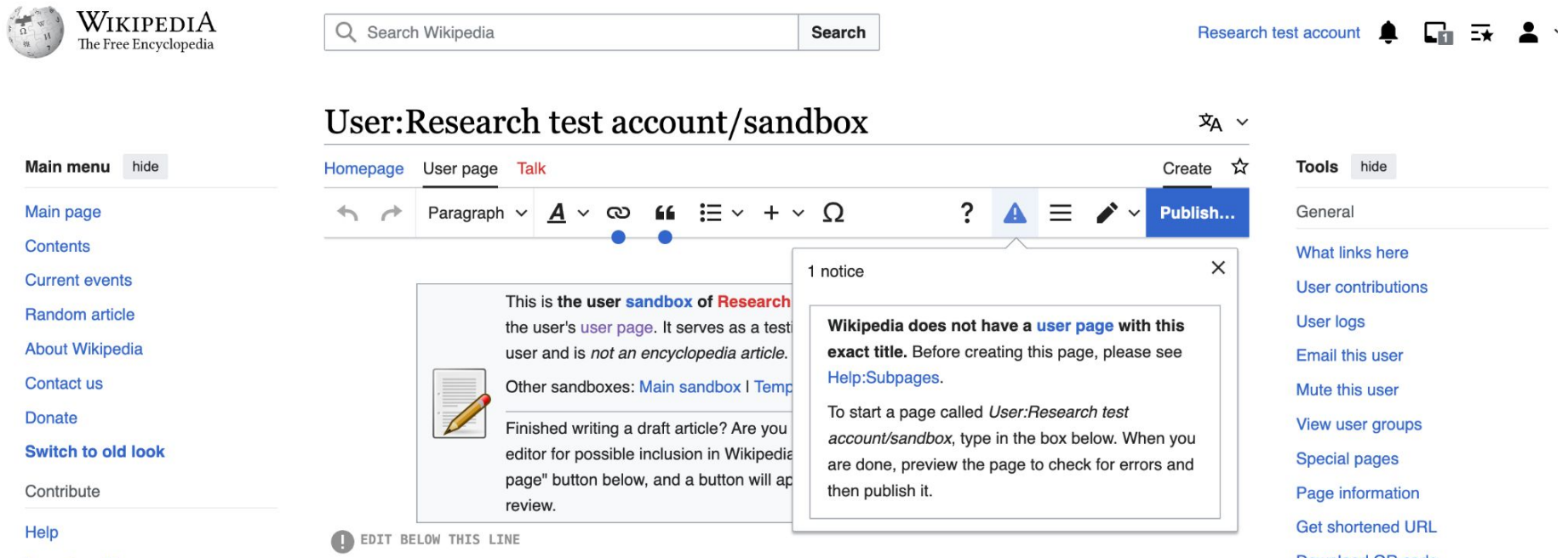
- Participants are mistaken about:
  - their authority for uploading the image.
  - what will happen to the copyright holder's ownership of the image.
  - what can be done with the image once uploaded, and by whom.
- Participants don't want to read a lot of text to achieve their goal of uploading.
- They assume that learning about all aspects of the agreement will be time consuming and confusing.
- The stakes are low enough to blindly agree.
- Several said they *trust* that Wikipedia will do right by them, so they can comfortably speed past the fine print.

WIKIMEDIA FOUNDATION

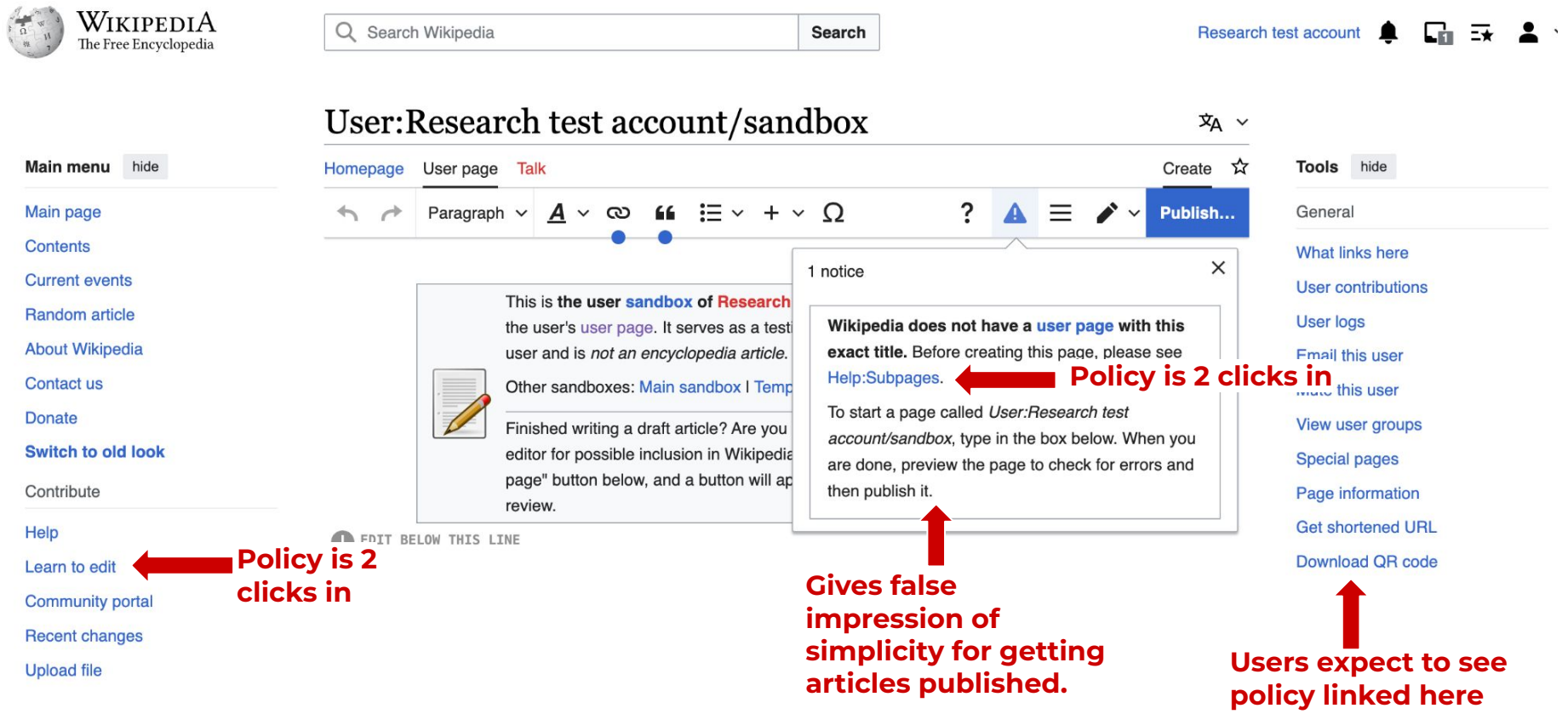
03 Visual Editor

Media Upload

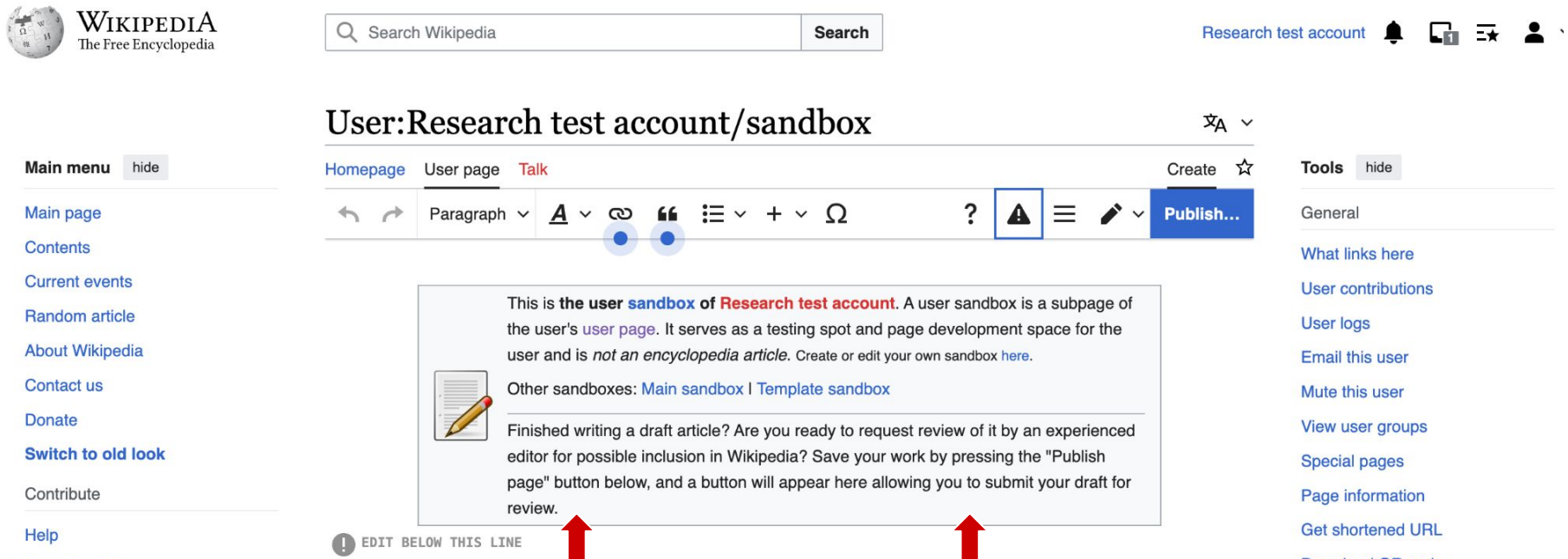
No easily discoverable policy or guidelines.



No easily discoverable policy or guidelines



No easily discoverable policy or guidelines.



Users spend time drafting articles that are rejected, leading to discouragement and confusion. Still no mention of policy here.

Users expect to see policy and instructions, written in simple language, prioritized during the editing process.

Users expect to see policy in Tools list

Half of participants were mistaken about what the media search will return.

More awareness of its function could reduce redundant images on Commons and lead to better categorization of images during upload.

What do you think this search is for?

6 **search for media that only I uploaded**

7 **search media other users uploaded**

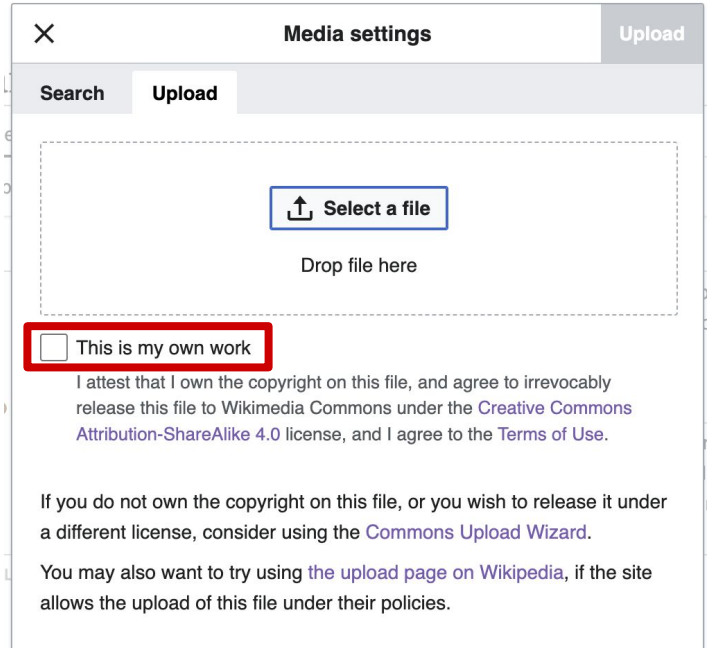
2 **search for images from the internet**

2 participants have used the search bar, including 1 experienced editor. Neither found relevant images.

> *I believe that's to search for images and videos from Google and Bing.* -P12E

> *That to search images that I've previously uploaded before on this particular WP article. If it's a long article, then it's easier to search for my images than scroll through to find them.* -P1E

Participants interpret “own work” differently, leading to wrongly attributed uploads.



What does “own work” mean here?

4	I created the image myself only
6	I have rights/authority over the use of the image
1	I'm involved in the creation of this work, therefore it is mine
1	Image is "free to use"

> *I created the art or media, I took the photo, I am the owner, or I have the copyright of the photo.* -P3E

Few participants could articulate what they are agreeing to when they upload an image.

This confusion contributes to misunderstandings about who has the authority to upload and convert the copyright to a free license, and what happens to the image after uploading.

What does checking the box signify?

- Agreeing I took the picture myself.
- Granting WP to use my copyrighted image.
- Putting the image into public domain.
- I own the file and can do what I want with it, including upload it to WP.
- Ceding the rights to WP to use it.
- Giving WP rights to use my file on WP.
- I own the copyright, I have overship over it.
- I have the ability to use this photo with full rights over the copyright.
- The work is taken by me, I own the rights, and the photo only contains me, not anyone else, from whom I would also need consent to upload
- **It's my own work, and I'm the sole owner, and I'm accepting the terms and conditions linked here.**
- **I give license for WP to use the image, under this specific license, and it'll be attributed to me.**
- **You're legally affirming that its your own work, and placing it under creative commons license, where anyone can use it.**

Bold are the most accurate interpretations.

Participants don’t know what to do with files that are not their own creation.

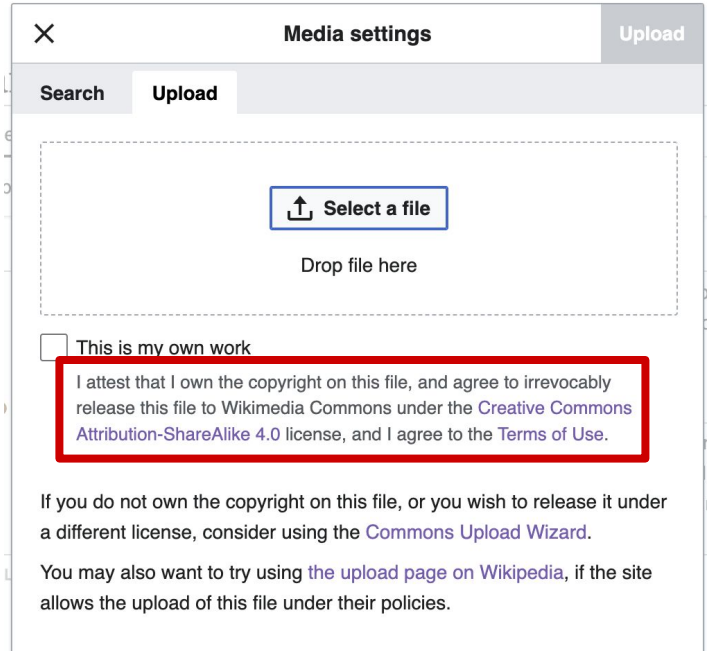
Describe any confusion you may have about “own work” here:

3 **say they have no confusion, but their responses to other questions suggest otherwise.**

9 **don't know what to do when the image isn't exclusively "own work", with confusion about who *owns* the file and has authority over it if created by someone else.**

> *I might be a bit confused. It only says “own work”, but if the image is free of use, or someone gives me permission, then I would assume it would be okay to upload it. But here, this only says “own work”; I would be confused if I were trying to upload a free of use image. I wouldn’t know what to do.* -P1E

The text is ineffective at conveying to users what happens to the image as a result of uploading to Wikipedia.

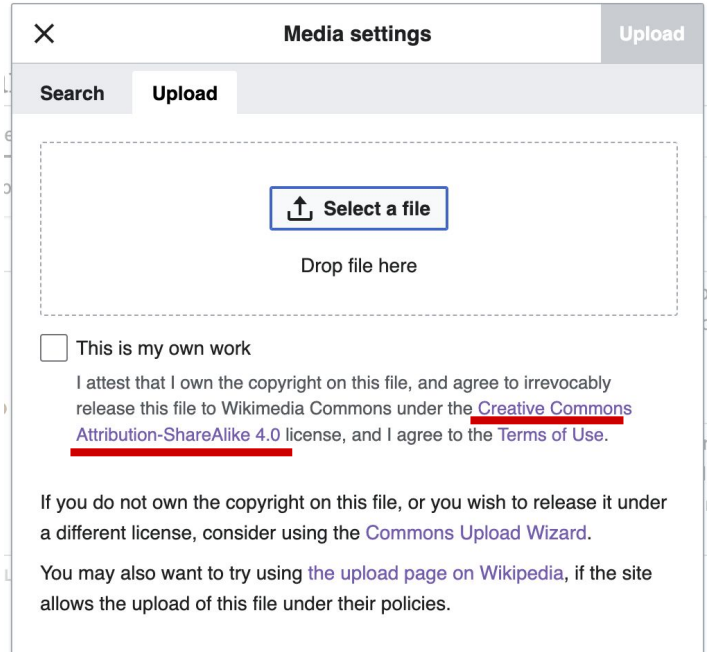


What do you think this means? Why do you think it's there?

2	**understood that they are converting the copyright of the image to a CC license for anyone on the internet to use.**
5	**stated they are agreeing to share the file with WP for use on WP only.**
5	**expressed that they aren't sure what it means, but this situation is low stakes enough for them to accept whatever conditions to move forward.**

> *I'm specifying that for my image, I own the copyright, and I have permission to upload it, and agreeing to share this data with Wikipedia.* -P12E

Participants do not feel the need to read about the Creative Commons license, due to misunderstanding who has authority over the image and what will happen to it as a result of uploading.

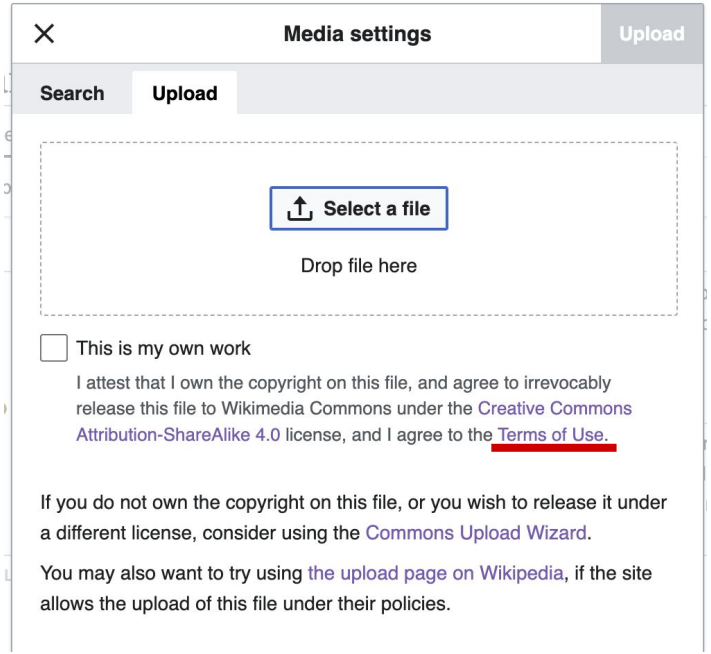


Do you feel the need to read and understand the default license in order to upload your image?

4	**I trust Wikipedia, so no need to read it.**
4	**This situation is low stakes for me, so no need to read it.**
3	**Yes (includes those who read it while using Upload Wizard).**

> *WP is the most trusted site. I don't need to go into detail. WP doesn't violate people's privacy.* -P10E

Participants do not read the Terms of Use, and wrongly assume what it contains.

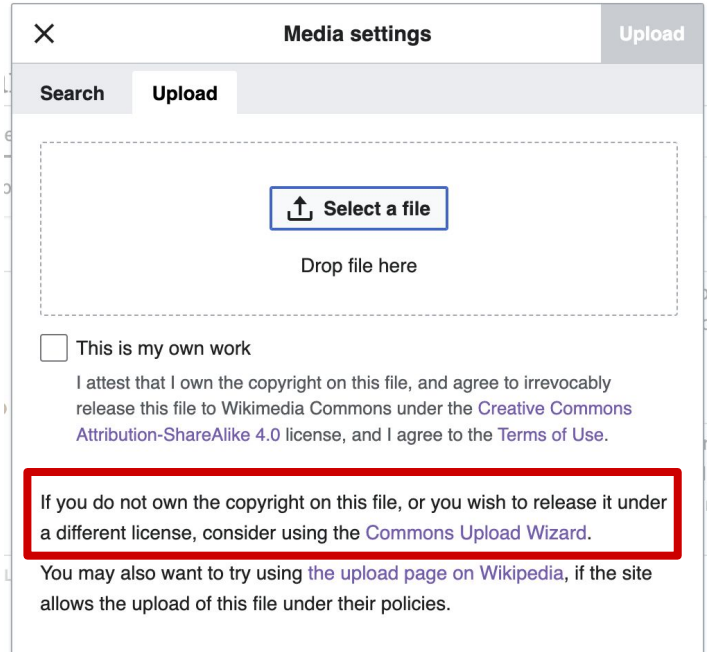


- **No participants read the Terms of Use.**
- **There is ubiquitous confusion about the differences between Terms of Use, licenses/copyright, the image or editing policy, and the instructions for uploading images.**

> *I would think that the Terms of Use contains what's allowed and not allowed, which to me is like a code of conduct, which is where I would expect to see what you want and don't want on WP.* -P9E

> *The confusion is still- and I've not read the Terms of Use, I need to read that- but does it mean that, if I am to upload any image, do I need to get one copyright license? Is that mandatory?* -P2E

The text is not effective at conveying when to use the Upload Wizard.



What do you think this means? Why do you think it's there?

2 if it's not your own work.

2 to determine if you can use someone else's work.

1 if you don't own the copyright, but the image is free to use.

1 for images found on the internet, or sources other than self.

4 to upload others' copyrighted work.

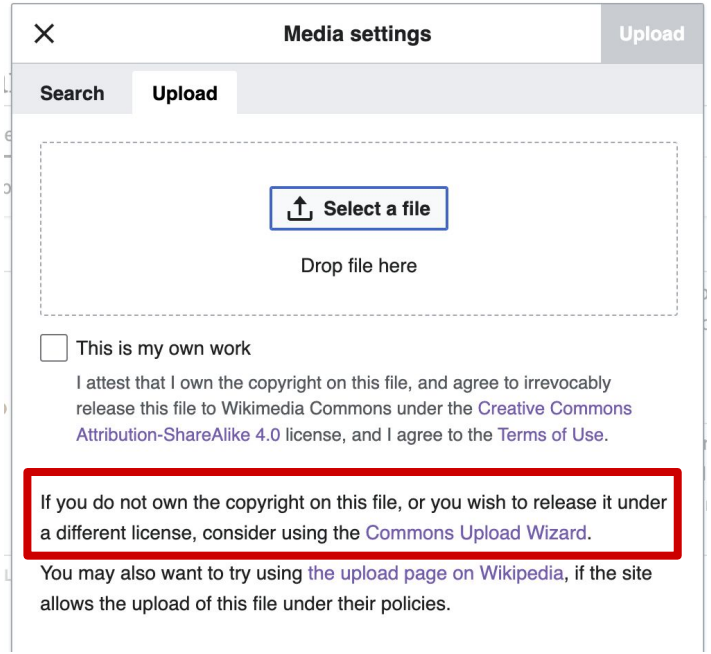
1 to be presented with different *copyright* options.

4 to be presented with different *license* options.

2 if you are experiencing a blocker/aren't successful at uploading on the VE uploader.

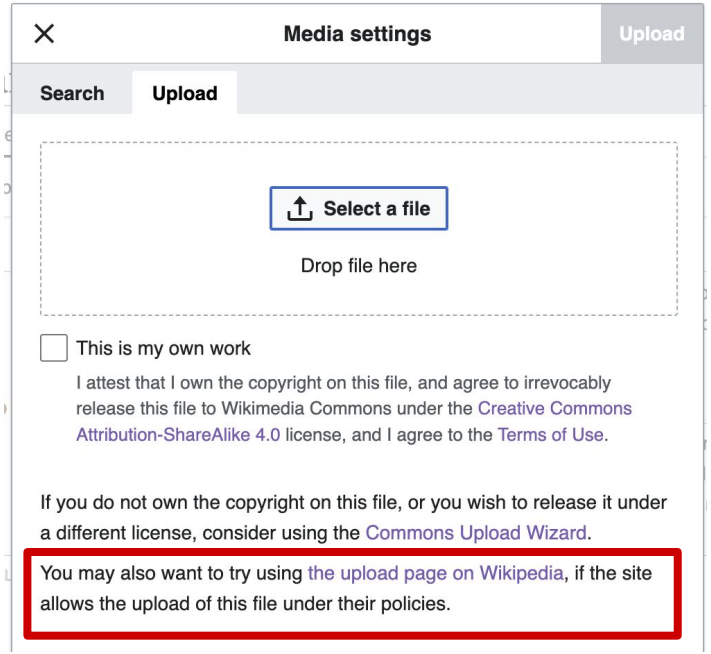
Colors used to group similar responses.

Participants expressed a wide variety of confusions and inaccuracies about this statement.



- Participants don't understand what "uploading under a different license" means. This confusion is rooted in not understanding the basics about copyright and licenses, and not understanding the copyright status of their file.
- The wording of this statement lead some participants to wonder if that they need to obtain a copyright for the file in order to upload it.
- Participants regularly confound Creative Commons with Commons Upload Wizard (because the word "Commons" suggests it is linked).
- Many think the Commons UW link takes them to Creative Commons to choose a license.
- No one understands why it says "consider", rather than being more directive.
- Many believe the Upload Wizard will perform checks to determine the file's eligibility.
- Many believe that the Upload Wizard is where to go if you are unsuccessful at uploading using the VE uploader.

The “upload page on Wikipedia” statement is the most confusing of all.



- **All participants misinterpreted this statement, and all expressed confusion about what it means.**
- **Participants are unaware that there are different Wikipedias with different policies, making that statement ineffective.**
- **Participants do not know when to use the upload page on Wikipedia versus the other two upload options presented on this screen.**

I thought I was on the upload page on WP? So, where am I? -P11E

Why does it say “you may also” instead of stating when to use it? -P4E

“If the site allows.” I don't know which site they are referring to here. -P3E

Why are there three places to upload media to Wikipedia? -P1E

If you fail to upload using the previous two, then try the third.-P6E

New users are not aware that their images have been nominated or deleted. No one expects to be notified via TalkPage message.

As of April 2024, 2 participants images have been deleted. Both were unaware their images were nominated and deleted.

1 participant is unaware that their image was nominated for deletion.

How do you expect to be notified if your image is rejected?

7	**Machine detection in the upload tool should tell me instantly that my image doesn't comply.**
6	**Users should receive an email**
3	**A notification on the home screen or top-right.**
0	**Talk page message**

*** Participants expressed desire for an explanation, solutions, and the opportunity to contest the decision.**

> *I didn't realize it had been deleted until my boss, the founder of the law firm, asked me what happened to the photo. Then I had to look at the article history, where I saw some cryptic comment by the person who deleted it, saying something about "OTRC." It was a nightmare to figure out how to resolve this. It took weeks.* -P11E

WIKIMEDIA FOUNDATION

04 Visual Editor Recommendations

Recommendations

1. Use machine detection to filter and reject personal and copyrighted images.

2. Unify the upload process across all platforms so that users are not tempted to use the easiest option.

There should not be one simple process and another complex process. Given that we have evidence that most cross-wiki uploads may not be the uploader's own work, the VE UI should have all of the same steps and options as the UW.

Recommendations

3. The UI should have a complete set of questions and options to lead users through the entire process.

More details are explained in section #5 of this report.

- A series of questions and response choices should lead users to take the correct action.
- Users need to see all of the options at once to engage with the right one.
- Users should not have to complete the process outside of the tool (i.e. VRT).
- Give options to upload other's work, specifically addressing the Marketer use case by giving pathways to upload images as a representative.
    - Build the VRT process into the upload process, or design a new process to handle these cases, and make it abundantly clear which files require it.
- Currently the Upload Wizard requires users to know which CC license the file is released under. Build a workflow that assumes users uploading "not own work" don't know the CC license and guide the to the correct action to take.

Recommendations

4. Fine-tune the wording.

- The wording in the VE uploader is ineffective and causes confusion. Use clear, directive language to explain when a user should take each action presented on the screen. Eliminate ambiguous language, such as “consider” or “you may also want to.”
- The licences names (e.g. CC A-SA 4.0 Intl) are not necessary and cause confusion. Use simple, everyday language or visuals to guide users to the correct selections. See examples in section 5 of this report.
- Clarify “own work”. Users interpret “this is my own work” differently, so consider stating it as “I created this work entirely myself” or similar.
- Currently “not own work” on the UW does not address some of the common cases, such as personal photos taken by someone else or uploaders representing another entity. Consider adding more options to clarify different scenarios within the “own work” category.
- Use tooltips to provide expanded information without sending users to other pages to learn more.

Recommendations

5. Educate and inform during and outside of the upload process.

- A concisely written image policy with accompanying videos should be strategically placed around the platform, especially during the editing and uploading processes.
    - Provide complete translations.
- Users need to encounter the image policy during the upload process, just as they need to encounter the editing policy during the editing process..
- The policy should be conveyed in simple, concise terms accompanied by visual media. Users want the option of watching a video.
    - YouTube has a policy video that some participants referenced as being helpful for understanding policy.
- Minimize the links that send users away to other pages or make them open multiple tabs. Users assume these external links contain long, dense texts, and will not engage with the links.
    - Use hover on tooltips instead. See section 5 #6 of this report for details.

Use deletion nominations and deletions to educate users, rather than punish. Consider a short quiz where users interactively learn about the policy and rules through situational questions.

Recommendations

6. Modernize and update the image policy.

- Modernize and update the image policy, using images, videos, and concise summaries, and leading with the most common issues.

- For example:
  - Selfies and pictures of people: *Pictures of people, including selfies, need to meet the notability requirements. Any selfies or images of ordinary people will be rejected per the image policy. Learn more [link to policy].*
  - Images created by someone other than yourself: *The image creator is the copyright owner. The photographer is the copyright holder, regardless of whose phone or camera was used.*
  - Logos: *Logos are copyright by default. The copyright holder is the logo creator or the entity that commissioned the logo. To upload logos, you must…*
  - Pictures of buildings, monuments, and other architectural works: …

- This should be done by a communications specialist, with eye-catching images and graphic communication methods that have been proven effective.

- Include video explanations, perhaps adopting the approach of the “A Wiki Minute” videos produced by the WMF.

Survey Admins for short list of the most common problems, and present the most common problems first in the revamped image policy.

Recommendations

7. Update tools for moderators and processes for alerting users.

- Participants whose images have been all reported feeling confused and lost by the moderation of their image. They did not know why their image was deleted, why some images are deleted while other similar images are not, how the policy applies to their image, and more. Part of the problem is a lack of consistency in the messaging when an image flagged or deleted.
- Build a template that all moderators will use, that clearly informs the user and states which policy is violated. This template should also include resources for receiving help and contesting the decision.
- Participants should be alerted in all possible places when their image is flagged or deleted, including but not limited to email and talkpage. Ideally ineligible images would be caught during the upload process.
- Currently, many users feel attacked by moderators for making mistakes. Moderators should be trained to understand that the goal behind deletion is not punishing or censoring users, but educating them. By implementing a mandatory template message system, users will receive notifications that are both educational and neutral in tone.

Participant recommendations for helping them comply with the image policy.

> *It's like filling out a government form, which makes you fatigued by the time you're done. The majority of the other parts of WP is a creative process: you enjoy writing, you're in the creative zone, doing your best work- then you have to figure out all of these check boxes. I'm not thinking in terms of copyright.* -P3E

- Make the policy accessible by putting it in multiple places.
- Build the requirements into the upload process, including copyright release (VRP).
- Explain what is/isn't allowed in few words and bullets, and also link to full policy.
- Summarize key points of the CC A-SA 4.0 license.
- Use images, infographics, animated visuals, and/or a short video to explain.
- Provide a guide, accessible during the editing and upload process, to findings public domain and other free images.
- Present instructions in terms of image types (i.e. logo, self-portrait), not a list of different licenses.
- Implement machine detection during the upload process to accept or deny images.
- Explain what to do in special situations and where to go for help.
- Force users to complete a tutorial before proceeding with the upload.

WIKIMEDIA FOUNDATION

05 Upload Wizard

Upload Wizard Insights

Although this project does not evaluate Upload Wizard, the insights and recommendations presented in this section are based on the issues and confusions surfaced in the user interviews.

6 of 11 participants, and mostly the *Marketers*, attempted to use the Upload Wizard during their upload process. Out of discouragement, these users returned to the VE uploader to upload as “own work”. Only one participant correctly uploaded someone else’s work as “not own work,” but they also reported feeling negatively about the experience.

Self-promoters do not use the UW because they truly believe their image is their own work. Users who truly feel that “own work” applies to their image use the VE uploader and consider the process quick and easy.

The difficulty in using the Upload Wizard is determining the right author and license combination for images that are not their own. This is because their image is not eligible to upload under the rules of copyright and they are unaware of that, or because they must go through the VRT process to make it eligible. Neither solution is apparent to users in the current UI.

Regardless of upload tool, participants have intensely negative experiences when their image is deleted by moderators. Most don’t understand why their image was deleted, and the information given about the deletion is insufficient to understand and rectify the situation. Participants cannot easily find help, explanations, or solutions.

Recommendations

1. Present a revamped and updated image policy in the beginning of the Upload Wizard, as detailed Recommendation #5 in section 4 of this report.

- Consider making new users complete a short quiz before uploading an image. This idea was suggested by both Admins in the Commons Admin Interviews research and some participants in this project.
- Consider defining “new users” as users who have uploaded fewer than X images.
- Determine X by surveying the moderator community.

Recommendations

2. Use machine detection, wherever possible, to identify image categories and route users to the appropriate action.

3. Focus on curtailing personal image uploads, a low-hanging fruit and an extremely common violation, in the upload process.

Nearly 50% of participants uploaded personal images and no one was aware that such images violate policy. Selfies and personal images are easy for users to understand and it is relatively simple to communicate the policy. Build a step into the upload process that either detects or asks if the image is a personal image. Use the opportunity to informs users that images of people are not allowed, unless they meet the notability requirement. Inform them that their personal image will be deleted if they don't abandon the upload.

All participants who uploaded personal images also incorrectly claimed them as own work. Tackling personal images directly will help to reduce copyright violations.

Recommendations

4. Present common scenarios that are easy for users to apply to their image.

Example of a common image scenario:	*This is a logo*
Example of a common ownership scenario:	*This work is owned by an entity I represent, such as a company or employer.*

5. Insert the policy into the recommended action, especially when uploading is discouraged.

This is a logo➜	***Logos are copyrighted by default and Wikimedia does not host copyrighted media.*** *To upload this logo, the copyright holder must convert the copyright to a free license. Proceed to the next step to release the copyright and upload the logo. ➜*	*next screen contains a form to a streamlined VRT process.*

Recommendations

6. Use hover over tooltips to present more information to avoid sending users away from the screen, in addition to a link to the revamped policy page.

This is a picture of a person, including self-portraits. ➔	Do the person or people featured in this image meet the notability requirement? [**tooltip:** more about this policy; **link:** to section on the policy page] ➔	DO NOT UPLOAD. Only images of people who are deemed notable as defined by the policy can be uploaded to WP/WMC. These platforms do not host images of common people, even on your profile page. ➔	Dead end or close window. When possible, tag these images to flag to moderators in case the user backtracks to make a different selection.

Recommendations

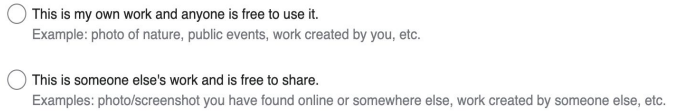
7. Decouple author from licensing in the first step. Address each part separately with expanded options for each.

- Present multiple options for authorship presented in a decision tree format.
- Better specify “own work”
- Add more options to the “someone else’s work” selection that reflect common scenarios.

Ideas

I created this work entirely myself.	• Uploading to WMC makes this image available for anyone to use. Proceed ➔
This is someone else’s work.	• I represent the copyright owner. ➔ • I don’t know who the copyright owner is. ➔ • I found this on the internet.➔
This work is a mix of my own and other’s work, including...	• This contains images of other’s work, such as book or album covers, architecture works like buildings or monuments, part of someone else’s photo, ... ➔

Current UI



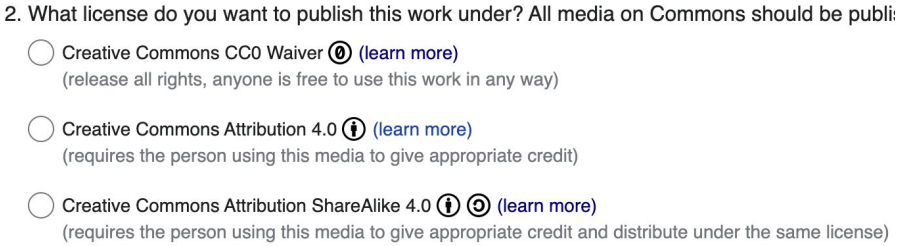
Recommendations

8. Use everyday language to explain the key points of the licenses. No need to prioritize the names of the licenses.

- In doing so, inform users of what is happening and who is authorized to take such actions.
- Consider including some other common scenarios in the selection choices that inform users about the policies and purposes of the platform, such as “I only want this image used on WP” and “I want to maintain the copyright on this image”, both of which would lead to a “do not upload!” message with an explanation.

Ideas	
I’m releasing the copyright, which restricts others from using this work, in favor of a free license, which allows anyone to use it for any purpose. **Only copyright owners can take this action.** →	• Anyone can use this image for any purpose, as long as they attribute the author. ***[tooltip: key points of CC A 4.0]*** • Anyone can use this image for any purpose, as long as they do not copyright the work and attribute the author. ***[tooltip: key points of CC A-SA 4.0]*** • Anyone can use this image for any purpose, as long as they do not copyright the work and attribute the author. ***[tooltip: key points of CC 0 Waiver]***
I only want this image available on Wikipedia and its sister platforms.→	• All images uploaded to WP/WMC have free licenses, meaning anyone can use them for any purpose. Do not upload this image if you wish to restrict the use of the image.

Current UI



Recommendations

9. Surface the VRT process in the upload process.

- Consider revamping the VRT process, either to explain and streamline submitting proof of copyright and permission to release it, or overhaul the process and requirements altogether.
    - This will require buy-in and participation from the moderator community.
- This should be presented for all of the “someone else’s work” options.

I represent the copyright owner. ➔	WMC does not host copyrighted images. Many images that appear on other websites and in image searches are copyrighted and cannot be uploaded to WMC. To upload a copyrighted image, the copyright owner must **convert their copyright to a free license**. **[tooltip: releasing the copyright implies that the copyright holder no longer has exclusive rights to this work.]** Proceed to the next step to submit permission to release the copyright. ➔	*next screen contains a form to a streamlined VRT process.*

WIKIMEDIA FOUNDATION

Cross-Wiki Uploads on Arabic Wikipedia

May 2024

Bethany Gerdemann, Eman Yahia

For the Structured Content Team

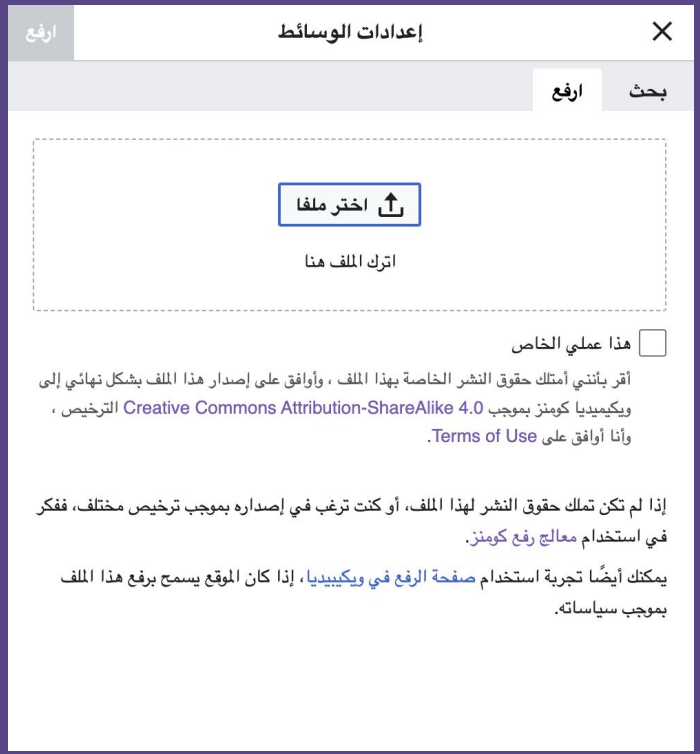
Arabic User Interviews

This is a report on a subset of research conducted as part of the Cross-Wiki Uploads project.

To understand how experiences differ across language communities, we conducted 5 interviews with Arabic Wikipedia users in addition to the 12 interviews with English Wikipedia.

Interviews for the ar.wiki segment were conducted in Arabic, with users who had uploaded images to ar.wikipedia.org in the past 6 months. Additionally, we interviewed one moderator for Arabic wiki.

This report highlights the similarities and differences between the two user segment, and brings recommendations that are specific to the Arabic user experience.



Visual Editor Uploader on ar.wikipedia.org

Participants

Gender
5 Male

Location
2 Egypt
1 Yeman
1 KSA
1 Palestine

Profession
- PhD holder and governmental employee
- Marketing manager
- Training manager in educational institution
- Web developer
- Engineer

Age range
19-66 years

User Experience Level
4 new uploaders
1 experienced uploaders

Arabic participants exhibited most of the same patterns as English participants.

Overall, the results of the Arabic interviews were very consistent with the results of the English interviews.

“This is my own work”

- 4 of 5 Arabic participants uploaded other’s work as “own work”.

Confusion about the moderation process

- Both sets of participants feel perplexed about the image review process. Some elements such as dimension, size, and type are automatically checked by the system, while human moderators handle other aspects like copyrights and policies. This leaves participants unsure about how Wikipedia assesses images.
- For those whose images have been deleted, none were aware of the nomination, and were therefore taken by surprise when the image was deleted. All participants feel discouraged from contributing as a result of their deleted contributions.

> *The word “my own work”. I feel this is related to copyrights. I here admit that this is my own work. I made it by myself or, in my case, my institution made it. This is ours. I here testify that this logo or these photos are ours and I give people the freedom to download it or publish it somewhere else but they won’t edit, steal, or give the credits to themselves.* -**P3A**

> *Okay, I got into this image uploader and I uploaded an image of a piece of WPC wood. That’s the industry I work in, then I clicked on this is my own work. I thought it means this is what I work in or this is my business. But, now I feel it may mean that I’m the one who owns the image. Honestly, I was on rush and was doing many tasks at once back then and that’s why I didn’t pay attention to any details here.* -**P2A**

Understanding of copyright

- Though they respect the concept of copyright, they understand its rules differently**.**
    - Some think that copyrighted images are only the ones labelled with copyright information or attached warnings.
    - Others believe that citing the place or the person they get the image from is “saving its copyrights” as long as they don’t claim themselves as the image owners.
    - Others think that getting an oral permission from their employer, friend, or whoever owns the image is enough to use it freely.
- Like English participants, none of the Arabic participants understood the technicalities of copyright, especially when it comes to their image.

Understanding of policies and procedures on Wikipedia

- Arabic participants are similarly confused and uninformed about the rules for uploading images and editing Wikipedia.
- Though they respect the copyright concept and agree that WP needs rules in order to maintain its quality and integrity, they have no tolerance for reading text-heavy technical content to upload images.
- All Arabic participants were unaware of the volunteer response team (VRT) process. The one instance where a participant uploaded their own work, the participant’s image was deleted, most likely because it required the VRT process to verify ownership. This was not explained and the participant, to this day, remains uninformed about the correct process.

Mindsets, beliefs, and values

- Similar to English participants, Arabic participants view Wikipedia as a trusted and highly respected source of information.
- When they contribute, they do so in good faith, with the intention of uploading images properly, according to what they believe are the rules and requirements, and they are unaware when they violate Wikipedia policy.
- With good intentions, they contribute to Wikipedia, but they are also under-informed about its scope, purpose, and policies.
- They operate under many false assumptions, often based on other platforms, such as social media and YouTube.

> *"I respect copyrights a lot! This is so important to me! I'm a freelancer photographer too and think it's unfair if someone takes a photo I worked so hard on and refers to it as his own work, but if it was shared publicly on social media and the person cites my work or refer to it under my name, that is not a problem."* -**P1A**

> *When I uploaded the first image, I got blocked because the system rejected the image dimensions, then I went to the Photoshop and adjusted this! I thought that's it the system accepted my image and it will be published. But later got surprised that someone deleted my image and I have no idea why!* -**PA4**

Pain Points

- Participants feel that uploading is needlessly complicated. It requires following opening multiple tabs and moving back and forth between the VE and Upload Wizard tabs.
- They feel frustrated that such a simple task, which can easily be done on any other social media platforms, is so complicated and overwhelming on Wikipedia.
- They believe that Wikipedia should take on the responsibility of simplifying the process of learning about its policies by providing a streamlined experience, visuals, videos, a YouTube channel, etc. They feel that they are left alone to figure out how to learn about how Wikipedia works.

> *Usually, in other platforms like YouTube and Facebook, when you upload an image that violates one of their rule like "inciting violence", they warn you first that they will be deleting it. This doesn't happen on Wikipedia, we pay high price to learn what is allowed or not allowed. A moderator can delete my hard work out of the blue without any consideration to all the effort exerted in creating such an article. This is so frustrating and it discourages me from editing any content on Wikipedia. I haven't even saved a copy of my article for myself.* **-P1A**

> *I'm a good reader by the way. But, I don't have time to read these long pages. . . Do you see how we are moving here and there between screens just to upload an image. I do this in 2 seconds on other platforms. I understand that Wikipedia has rules and I'm willing to obey them but you need to make things easy for us, too! I never read policies anywhere else and I doubt if anyone does! I just need to upload an image and that's it.* **-P2A**

There are a few key differences between Arabic and English users experiences.

- In contrast to the English language user patterns, which were evenly divided between Marketers and Self-Promoters, we saw a third pattern in Arabic participants: Storytellers.
  - **2 of 5 participants were Storytellers**
  - **3 of 5 participants Marketers**
  - **0 participants were self-promoters**

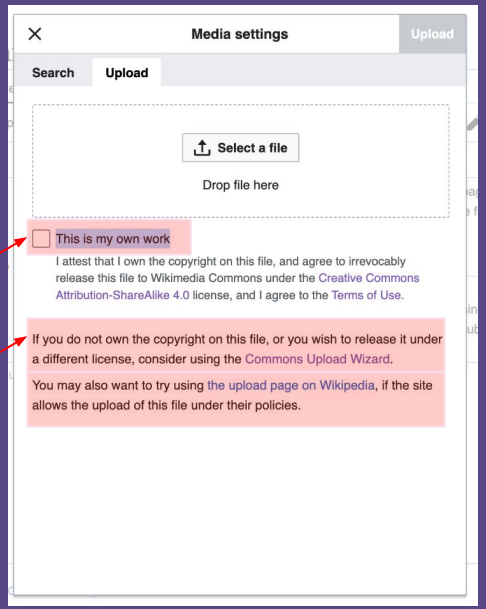
Wikimedia Commons and the Upload Wizard

- In contrast to English wiki users, none of the Arabic users tried or even saw the Upload Wizard. They all went with the first impulse of testifying the image is their own work to proceed, without paying any attention to any of the sentences and links below.

> *Well, I've never seen this instruction photo before. It wasn't there when I uploaded my image. You know, even this progress bar at the top that shows the steps wasn't there, either. I believe that's not how I uploaded my image.* **-P3A describing seeing the Upload Wizard for the first time in the interview.**

This is the only check box users can see.

Many users completely miss this section



Incomplete Arabic language translation and documentation

- There is less documentation available in Arabic regarding policies, explanations for deletions, and Creative Commons license information.
- Some of the upload UI and Creative Commons pages are not fully translated to Arabic, which confuses and frustrates users.

> *Can you see how some sentences are written in English? I hate this. What does this mean? Too difficult words. Attributions.. Waiver.. It's hard for me to understand this.. It's really confusing.. I just want to upload an image.. Why it is so difficult here? It should only take a few seconds.* -**P3A**

Partially translated Commons Upload Wizard. Arabic participants were shown the UW to assess if they had seen it before.

Partially translated Creative Commons license page.

Moderation in Arabic

- There are few Arabic speaking administrators, which leads to poor interactions and less thorough explanations for deletions.
- These moderators are all volunteers and have varying levels of tolerance when it comes to educating users about the correct and incorrect policies. This ultimately leads to inconsistencies in their responses and explanations for deletions, which causes confusions amongst users.
- Due to the high volume of images uploaded daily to Commons and the lengthy review process for image deletion, some images remain unreviewed. This is usually misinterpreted by users as a double standard or inconsistency. In the worst case, this is misinterpreted as censorship, especially among users who feel they are part of marginalized group (e.g., a Palestinian user telling stories about Palestinian martyrs).

> *We get thousands of images, audios, and videos everyday on Wiki Commons and as I showed you it takes a long time to review, moderate, nominate, or delete one image. And all moderators are volunteers, that's why many images are moderated so late and sometimes never get moderated at all.* -**Arabic Wiki Moderator**

> *Actually, all the information about the image licenses I was just explaining to you I learnt on my own the hard way through trial and error. No one taught me any of it. I made all these mistakes myself. I uploaded selfies, images from the internet, I did all of it. It took me years to learn these stuff and I can tell you that most of the Arabic users are unaware of these rules and unfortunately there isn't enough efforts exerted to educate Arabic Wiki users about the rules.* -**Arabic Wiki Moderator**

The Storyteller

(2 of 5 Arabic Participants)

This portrait is concealed to protect the copyright.

File:Mahmoud-titi محمود طيطي.jpg: Revision history

View logs for this page (view abuse log)

There is no edit history for this page.

This page does not exist. The deletion, protection, and move log for the page are provided below for reference.

- 12:06, 28 March 2024 Yann (talk | contribs) deleted page File:Mahmoud-titi محمود طيطي.jpg *(Copyright violation; see Commons:Licensing (F1): https://w fbid=502772903232322&set=pcb.502773053232307)* (global usage; delinker log)

Description	Mahmoud-titi محمود طيطي.jpg.
Date	28 March, 2024
Source	Own work
Author	Username

Goals:

- The uploader is often a relative or an acquaintance of someone who had an exceptional life experience or who is a person with extraordinary knowledge and expertise in their field (e.g., religious scholar).
- Uploaders create an article for those they believe have stories that need to be told, often a marginalized figure.
- They upload an image for the story hero to give it more credibility and increase reader engagement.

Upload process:

Storytellers navigate to the Visual Editor in the process of creating their articles. They are confused by what “my own work” means, but feel okay to upload their images as “own work” because they either took the image by themselves or got oral permission to publish it from the person in the image.

Outcome:

Images with are likely to be **deleted for either** Clear Copyvio (F1) when © is in metadata, or “Own work” copyright violation if the moderator doubts “own work”.

The Storyteller is driven by the passion for sharing an important story.

> *“Before I create this article and upload the image, I asked for this person’s permission. He gave me an oral one. He even shared with me some photos to choose the best of them for the article.”* -**P5A**

What storytellers do:

- These images are used to illustrate the article about “the story hero”.
- Storytellers are driven by their mission of creating an article about this hero, so they don’t think too much about license selections, and don’t realize that, in some cases, they should not upload the file.
- Storytellers do not know the copyright status of the files, but they believe they either own it or their relative or acquaintance owns it.
- They believe they are authorized to upload because they have a personal connection to the subject of the photo.
- When the file is deleted by moderators, participants experience frustration, discouragement, and confusion. In some cases, they request the image be restored after deletion.

What they should do:

- In the cases where the image is not own work, the images should not be uploaded.
- One participant uploaded their own work, but their image was deleted. In this case, it had appeared on social media first. When an image is on another site, the copyright holder must also go through a VRT process to explain copyright status. This information is not explained to users.

The Marketer

(3 of 5 Arabic participants)

تعليق
Uploaded while editing "Arab Center For Arbitration" on ar.wikipedia.org

الوصف	**العربية:** شعار المركز العربي للتحكيم
التاريخ	١٨ أبريل ٢٠٢٣
المصدر	عمل شخصي
المؤلف	username

Goals:

- Often employee editing on behalf of employer or someone representing another party.
- Creating an article for their company and uploading a company logo, or adding the first image to the article.
- Contributing to ar.wikipedia by adding their company information and profile out of their trust, appreciation, and respect to Wikipedia.

Beliefs:

- They are on work assignment to upload the image or they have permission to upload the image.

Upload process:

All marketers used the Visual Editor. Since the image is their company's and they have permissions, they considered the file their "own work".

Outcome:

- Images with copyright metadata are likely to be **deleted** or **not yet moderated**.
  - Clear Copyvio (F1)-when © is in metadata.
  - "Own work" CR violation
- Logos are likely to be **deleted** or **not yet moderated**.
  - Logos are usually considered copyrighted, and therefore risk being deleted as a Copyvio.

There is no pathway for the Marketer to upload their company's images.

This logo is ours. I mean it's my employer's. I'm the one responsible for doing marketing and social media content. All I wanted was to create a page for the institution as we have other social media pages like on Facebook. I didn't think the process would be that complicated. By the way, I saw similar pages for other institutions on Wikipedia and that encouraged me. I wanted to do this because we know how Wikipedia is trusted and respected from people. I didn't know that this is against the rules. I just wished that when someone search for our institution name on Google, the first thing he gets is our Wikipedia page -**P3A**

What marketers do:

- Marketers do not know the copyright status of the files, but they believe their company owns them.
- These images can be logos or professional images that the company uses across the internet.
- Marketers don't put too much attention to find the right author / license selections and don't realize when they should not upload the file.
- When images are deleted, it leads to confusion, frustration and discouragement.
- Participants don't fully understand what went wrong, how they violated policy, how to resolve the issue, or what they should have done differently.

What they should do:

- The image policy demands that the © owner proves ownership status through the VRT process. This process is unknown to users, it is not built in to the upload process, nor is it explained anywhere. Users are unaware that this is the proper way to upload images that they have not created.

Arabic Wiki Specific Recommendations

The following recommendations focus on the Arabic user experience, however all of the recommendations that were given for the English user segment also apply to Arabic users. For a more comprehensive set of recommendations, please see the English segment report in the first section of this document.

1. Onboarding is the key to a good user experience.

- Improve the WP onboarding strategies for new users and create accessible, visually appealing, and concise explanations of policies and processes. This will help new users feel confident about contributing in a productive manner, rather than left to figure things out through trial and error.

2. Recruit and train more Arabic moderators.

- When some of a users images are moderated but others are not, it is interpreted as arbitrary enforcement. In some cases, this leads users feeling censored or singled out. It is not apparent to users that there may be a long backlog of images that require Arabic moderators, which is why images are inconsistently moderated.

3. The VRT process needs to be more prominent, or redesigned.

- The VRT process is not easily discoverable and certainly not understood, yet all Arabic participants should have used it to properly upload their images. Either the process for justifying such images needs to change, or the VRT process needs to be more prominent in the upload tools.

Arabic Wiki Specific Recommendations

4. Complete translations.

- Complete Arabic translations for interfaces, policy pages, licenses pages, reasons for deletion, and template messages that moderators use to explain why images have been deleted.

5. Show contextual tips/rules relevant to users' images.

- Utilize machine detection that can provide proactive help that aims to familiarize users with the rules of the image upload process. Aim to show contextual information relevant to the user's images rather than displaying all policies in lengthy, text-heavy pages at once.

6. Improve the designs of image uploaders and merge VE and Wizard.

- Remove the need to move back and forth between tabs and help users choose the right option without overwhelming users with copyright technicalities.

Arabic Wiki Specific Recommendations

7. Templates for moderators to ensure consistent messaging.

- Develop standard message templates with all necessary information to understand the problem and how to rectify it. This will eliminate confusion by ensuring uploaders are given complete information, promote learning of the policies, and reduce feelings of being singled out or punished.

8. Warn before deletion, and improve the warning system.

- A deletion nomination alert should be sent in advance of the deletion. This should allow users to take some precautions (e.g., save their content somewhere else), understand the rule they violated and modify their image when possible.
- In some cases, users are warned via talk page messages, but most users are unaware of their talkpage.

9. Deletion is a chance to learn, not to punish.

- Notify users of the deletion and use it as a good chance to allow users to learn about Wikipedia's policy.
    - This could include a short quiz where users interactively learn about the policy and rules through situational questions.
    - YouTube Policy Training could be an inspiration for us.

Thank you!


Thank you to Eman Yahia, Sneha Patel, Alexandra Ugolnikova, Luca Martinelli, Daisy Chen, Mike Raish, Connie Chen, and User:Kadı for your work and assistance on this project!