media matching synthesis: Japanese+Arabic+English

I always start in the other-language versions of Wikipedia to see if there's a picture. If there's a picture, I'll follow it to Commons. For me, it's a key to enter Commons and arrive at the categories.
- Arabic Wikipedia editor

key takeaways

- Experienced editors' biggest media-related challenges are found in the search process, and not in the mechanics of adding an identified image to a Wikipedia article.

- AI-derived image recommendations have the potential to be extremely helpful to media editors in non-English Wikipedias.

pain points   **suggestions**   **search strategies**

summary

New editors often want to add an image to the Wikipedia article they're working on, but don't know how. On the other hand, experienced editors' biggest media-related challenges do not directly stem from the mechanics of adding an image. They instead face challenges in locating media and determining which media will be accepted by other community members.

Non-English editors face additional significant burdens during the media search process that arise from their natural desire to use local languages to search for locally relevant media on Commons. An AI image recommendation tool may ease the burden of searching Commons in local languages and may greatly expand the pool of potentially acceptable media for such editors.

This document summarizes the main areas of difficulty and confusion encountered by experienced editors working with media in three Wikipedias. The individual reports summarizing interviews with media editors in English, Japanese, and Arabic Wikipedias can be found here:

A. English
B. Japanese
C. Arabic

contents

misery loves company: shared pain points

captions and image descriptions are burdensome to compose

Adding appropriate captions and descriptions to images on Wikipedia is a persistent hurdle across languages, and it is most difficult for non-English editors.

- **English** editors view this activity as secondary to the "main" goal of adding an image. They tend to spend less effort on it and gravitate toward the low-effort solution of directly copying the Commons file descriptions, which are often inappropriate for verbatim use on Wikipedia.

- **Arabic and Japanese** editors face the additional burden of frequently lacking easy access even to the Commons file descriptions, which exist most frequently in English. Adding appropriate descriptions to Wikipedia for these editors thus requires either composing an entirely new description or translating an English description, possibly using external tools. This process is additionally burdensome if the editor lacks relevant domain knowledge.

media is hard to find on Commons

Editors of all languages face challenges locating media that has already been uploaded to Commons. Non-English editors—especially those attempting to search Commons using non-Latinate scripts—find themselves translating search terms in order to access Commons media.

What appears is completely different to what I searched. I need to try to use katakana *instead of* kanji.

- *Japanese participant*

- **Experienced English editors** have varying degrees of comfort with navigating Commons, although most participants use Commons categories as their main avenue for arriving at a specific image after performing a search. This widely used strategy allows editors to select appropriate images outside of their areas of specialization, for example, by using human-assessed "Quality images" categories.[1] However, it limits the extent to which most editors can access uncategorized media, which remain out of reach of most editors unless they have developed their own idiosyncratic methods for accessing uncategorized media.

- **Arabic** editors generally believe that Arabic-language search on Commons is ineffective. Searches for media therefore rely on translation of search queries into English or on editors' knowledge of English.

- **Japanese** editors similarly face difficulty when trying to search Commons using Japanese. Editors may try searches using [multiple scripts](#) or combinations of scripts, although these are less effective than searching in English. Arabic participants rely primarily on Google Translate, while Japanese participants additionally use [DeepL](#) and [Excite Translator](#).

---

description of a typical Japanese media search experience during interviews:

They go to Wikimedia Commons and search for the bird's name. They start by changing the [interface] language to Japanese, spending a while searching for it in the language dropdown. Then they search using the top right search bar on the main page. They couldn't find anything based on this search both times they tried. Then they deleted the kanji **and found some birds.**

---

editors are unfamiliar with Commons

Commons' inconsistent support for local languages leads to a situation in which Arabic editors don't understand the relationship between Commons and

---

[1] Experienced editors' navigation strategies may change in response to the recent addition of human evaluation search filters to Commons search, however.

Wikipedia. For many editors, their editing starts and ends on Wikipedia. **Arabic and Japanese** internet users tend to shy away from sites that are even partially in English, and this tendency limits the proportion of editors who are motivated to dive in to working with Commons in order to add media to Wikipedia. One Japanese participant—in spite of being an experienced Wikipedia editor—had never used Commons to search for media before attempting the interview task. For non-English speakers, the navigation of Commons in order to accomplish an editing task can be daunting.

license and copyright rules are unclear

Both Japanese and Arabic editors can encounter a great deal of frustration stemming from lack of clarity around license and copyright policies, and from the differences in such policies between their Wikipedias and Commons.

- **Arabic** editors often lack awareness of the importance and function of license information in general, and may upload media to Commons that doesn't satisfy standards of fair or free use. Incomplete, fabricated, or placeholder license information may be used for expediency. When this media is later removed from Wikipedia articles or from Commons, this removal can be confusing and leave editors with the impression that Commons and media work are best avoided.

- **Japanese** editors, on the other hand, often see media removed from Wikipedia articles because the Japanese Wikipedia community has a more restrictive approach to copyright and license issues than English Wikipedia or Commons itself. If anything, the Japanese community may be hyper-aware of these issues, as minor uncertainties about an image's provenance are often used as an excuse to remove it from Japanese Wikipedia, even if the file satisfies Commons' license requirements.

articles in need of illustration can be hard to find

The **Arabic** community has experimented with mechanisms to alert editors to articles in need of images—including curating categories such as [Saudi Articles Needing Images](#) and including a [placeholder icon requesting illustration](#)—however

a generalized lack of awareness of talk pages limits the usefulness of these features for most editors.

translating diagrams and maps into local languages is very important, but very difficult

The **Arabic** community has made the "Arabization" of media such as maps, diagrams, and subtitled videos a high priority. Many experienced media editors spend a great deal of time on this localizing work. This is specialized work, however, given that artefacts such as maps often include highly stylized, specialized, or otherwise esoteric text. A recent example of this process can be [seen here](#), in which [this diagram](#) was translated [into Arabic](#). A lack of supported tools and general lack of Commons support for Arabic makes this translation work "exhausting," however.

- **Arabic** image translators also note that there is no automated notification system in place to support image translators—if the original English diagram linked above were to be updated with new text, the translator doesn't have an easy way to be notified and update their translation.

verifiability on Commons is not enough for some Wikipedias

Experienced media editors often encounter a mismatch between what counts as "acceptable" media on Commons and what counts as "acceptable" on their Wikipedias. This mismatch generally doesn't stem from local Wikipedias' cultural or social norms—individual Wikipedias have developed accepted practices for managing these—rather Wikipedia admins often either mistrust Commons' licensing information or suspect user-created images of being unreliable.

- **Japanese** participants feel personally responsible for ensuring the verifiability of anything they add to Wikipedia, often going so far as to consult numerous external and print sources. Given the Japanese community's reluctance to accept Commons' licensing information at face value, this additional burden of verifiability also leads to Japanese editors taking more of their own photographs to avoid deletion.

- **Arabic and English** experienced media editors note that admins on their Wikipedias often mistrust user-created diagrams and illustrations, and ensuring their inclusion in Wikipedia often requires engaging in discussions about their accuracy.

media supply < media demand

Arabic and Japanese editors both view their communities as in need of additional media, and they are also aware of the fact that other Wikipedias such as English tend to be more heavily illustrated.

- **Arabic** community members tend to attribute their lack of media to the difficulties associated with integrating media from Commons into Arabic Wikipedia. New and experienced editors alike face linguistic and technical challenges throughout this process.

- **Japanese** community members tend to attribute the lack of media on Japanese Wikipedia to the community's strict approach to copyright and licensing issues, in the understanding that the Japanese context is different than other national contexts such as the U.S.

- **English** community members tend not to be concerned about the amount of media on English Wikipedia. Some believe that, if a good image for an article is on Commons, it has likely already been added to Wikipedia. Non-English Wikipedias tend to have very different relationships with Commons, however.

new editors *want* to do media work, but they don't know *how*

Adding an image to an article is a high priority activity for many new editors, however it remains challenging.

- **Japanese** editors are particularly interested in uploading images to Commons from mobile. Given that Japanese Wikipedia is reportedly strict regarding image licensing, many media editors take and upload their own images. Interview participants believe that, in this context that prioritizes

working with one's own media, improvements to the mobile uploading process may help newer editors become more involved.

- **Arabic** new editors generally want to add images, and community mentors note that new editors' most frequent questions are about how to do so. Given the difficulties that face Arabic editors as they try to access and use Commons, encouraging and training them to take and upload their own images may be an effective way for them to begin learning the media ropes.

more of Commons needs to be localized

The task of localizing Commons templates and interfaces (i.e., translating them into local languages) is time-consuming, specialized work, and remains incomplete for many languages. However, high-quality translations greatly contribute to users' comfort. Skilled translations generally give the impression that this community was created and is populated by like-minded humans.

- **Arabic** participants note that poorly translated interfaces are actually *less useful* than ones presented in English. Furthermore, given the high importance that the Arabic community places on linguistic quality, poorly translated interfaces may also be off-putting for potential Arab media editors.

participation in Commons discussions requires English

Commons is a global project that uses English as a language of communication. In addition to interfaces and templates which have been incompletely localized, a range of core Commons functions occur primarily or exclusively in English, including deletion and nomination discussions.

- **Japanese and Arabic** editors have access to some tools to improve access, such as Google Translate, however the necessity of English is a highly salient barrier for large groups of editors seeking to use Commons for any purpose. The difficulties inherent to following and joining discussions on Commons may prevent large numbers of editors from becoming more involved in media use.

suggestions for image recommendations

The prospect of automatic or AI-derived image recommendations was received positively by interview participants in all three languages.

This tool may be of greatest value in non-English Wikipedias where editors face additional linguistic, technical, and social challenges when searching for and adding media (social challenges are related to individual communities' stances toward licensing).

Make copyright/license information visible or easily accessible in image recommendations

Many Wikipedias have different norms regarding the acceptability of licensing information, and acceptability on Commons often isn't enough for use on local Wikipedias. Making the licensing information on Commons more visible to editors on Wikipedia may help streamline the process of negotiation that occurs when adding material from Commons to Wikipedia.

Include related audio files in image recommendations

Many editors ignore audio files, or only use them when they come across them coincidentally. The inclusion of relevant audio files in image recommendations may help non-image media to be used more broadly.

Include a "snapshot view" of a related article in the same Wikipedia

Many editors already check related articles as part of their media workflow. Checking related articles allows them to get a quick view of image placement, image-to-text ratio, and types of images (i.e., scientific, depicting specific regions vs. global views, etc.) that are accepted in a given genre and in a given Wikipedia. Individual Wikipedias often have different norms regarding image placement and number, and new editors adding media may be unaware of these. Experienced editors working outside of their domain of expertise may also be unaware of these

norms in their new genre. A "snapshot view" of a related article would help editors of all experience levels to quickly align their work with local norms.

Include a "snapshot view" of the same article in other Wikipedias

Many editors in non-English Wikipedias freely draw on other-language articles to locate media. Editors from less-illustrated communities tend to check English first, however participants in these interviews also reported checking Spanish, French, Portuguese, Swedish, and Dutch. Most editors limit themselves to languages they can read, however some editors travel widely, finding media on Wikipedias with which they have no prior relationship. A snapshot view of other-language Wikipedias would help editors to:

- Identify appropriate media that they might miss if searching Commons in a local language,
- Save them the effort of laborious non-English Commons searches, thereby helping them to add more media overall, and
- Get a sense for any norms that may have emerged at the "super-language" level—i.e., a snapshot view of other languages may reveal that most languages prefer photographs in infoboxes about geological phenomena, and diagrams in the article body.

Include support for machine translation in image recommendations

Non-English editors would benefit from machine translation at several points in the media search process. This includes the act of searching Commons (during which editors already lean on external tools), navigating Commons and selecting an appropriate image, and adding caption and description information for the selected image to Wikipedia. Machine translation could be used at several points:

- Suggesting English search queries for use on Commons,
- Translating Commons category names into local languages to aid the search process,
- Translating Commons file descriptions into local languages to aid the search process,
- Translating Commons file descriptions into local languages for use as captions or descriptions, and
- Translating captions or descriptions used on other Wikipedias for a selected image into local languages.

reservations

Participants who expressed reservations about AI-derived image recommendations focused on the potential for AI tools to reinforce the biases of their creators. Most participants who expressed this view did so in the form of a caveat—in other words, they are cautiously optimistic, but would like to have more detail about this tool's implementation.

how did participants search for and add media?

Arabic

Searching: Arabic participants were experienced media editors already active to varying degrees on Arabic Wikipedia and Commons. Participants' first inclination was to directly access Commons categories from the task article, if possible. Some immediately translated the task article title into English in order to search on Commons; others navigated to adjacent Wikipedia articles to access their media and follow the trail back to Commons. One participant ended up with "landslip," instead of "landslide" after translation, and used this term to search on Commons.

Adding: All participants used the text editor, rather than the visual editor. Several were surprised that the visual editor automatically accomplished some of the commands they usually manually type in the textual editor, like "align left."

Japanese

Searching: Japanese participants also started by accessing images in related Wikipedia articles and following that trail back to Commons. Some had never used Commons to search for media before, and some had difficulty identifying an appropriate image on Commons due to lack of familiarity with English. While on Commons, participants tried multiple combinations of scripts and search queries that were regarded as ineffective, and success was generally achieved after participants managed to locate and use the appropriate categories.

Adding: Experienced Japanese editors tend to prefer using the textual editor, rather than the visual editor. Some participants copied formatting—such as the infobox template—from related articles, while others added formatting from memory. As in other languages, the actions through which editors add an image are individualized.

English

Searching: English editors were drawn from a range of backgrounds and include editors who work primarily on Commons, as well as some who work primarily on Wikipedia. Most editors who were comfortable with the interview task were quite confident in using Special:Search on Commons, typically using the Latin species name as the query. Other editors who are primarily uploaders noted that they usually conduct Google searches, since the media they work with often doesn't exist on Commons yet. Once inside Commons, however, participants tended to use categories to navigate their way to the image they ultimately selected.

It's kind of a bad habit of mine that I don't check Commons. Commons is notoriously bad for having horrible search.
- English editor

Adding: English participants preferred using the textual editor, rather than the visual editor. Most described this as an issue of personal comfort.