

Verifiability on Wikimedia Results Report

**Sarah Gelb-Wiegand
Spiegel Institut Mannheim
May 2021**



Study Background



- The non-profit **Wikimedia Foundation (WMF)** provides the **essential infrastructure** for free knowledge.
- **Wikipedia**, the free online encyclopedia, is **created** and **edited by volunteers** around the world, as well as many other vital community projects.
- Wikipedia's **reputation for reliability** rests on its **foundation of external sources** which give readers a **clear understanding** of **where information** is **coming from**, and where they can **learn more**. To best support Wikipedia communities in building a trustworthy and reliable encyclopedia, editors need to be equipped with the **skills** to do **effective research**.
- This research project should **help to understand** how the community presently goes about **researching a subject**, what **pain points** they're facing, and **how different communities approach** the concept of '**verifiability**' on Wikipedia.

Project Setup







Method:

- N=17 remote **video-interviews**
- Duration: **90 minutes**



Projects/ languages:

-  **English:** very large project, country of origin
-  **German:** large, established project
-  **Korean:** growing project
-  **Basque:** small, developing project



Sample: N=17

- **English: n=3** Wikipedia community experts from WMF
- **German: n=5** community members
- **Korean: n=5** community members
- **Basque: n=4** community members

Preliminary Notes



Results in general:

- Results are displayed in **descending** order from **top to bottom** and **left to right**
- To underline an **importance** and/or **frequency** of a certain finding or mention, those findings might also be **emphasized** with a bigger **font size** and written in **bold** letters

Projects/Languages:

- The **results** represented in this report are **aggregated** over all **projects/languages**
- Results and quotes are only **marked** as particularly relevant for a respective **project or language**, if seen as **specific** for this respective **region** and considered as **not applicable** for **other projects** in the sample
- Project/language findings are **marked** with the respective **regional flags**:



English



German



Korean



Basque

Wikipedia Member Profiles



Backgrounds:

- **Nearly all Wikipedians in the sample have an academic background**, such as French studies, history of knowledge and science, old languages, lecturer for humanities, computer sciences
 - Therefore, most have a high to very high research literacy
- Some worked with information/ in libraries before
- Non-academic background



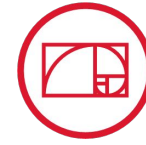
Wikipedia experience levels:

- **Vast majority: High to very high (10+ years)**
- Few: Medium



First touchpoints with Wikipedia:

- **Using Wikipedia for personal or work topics**
- **Research for school/ college projects**
- 🇰🇷: First fixing typos and doing translations
- 🇺🇸/🇩🇪: Doing small edits



Work patterns:

- Mainly **two patterns of work distribution** can be observed, that distributes equally (might also vary within one person):
 - Condensed work: Several consecutive days/ `Spurts`
 - Evenly spread work: Several hours a day/ week
 - Few: Scattered here and there

`I think I work in spurts. If I work on a topic of my interest such as animation or arts, I work in spurts of heavy activity, I edit like 10 articles in a short time, or in a day. Then I take a rest.`

`It depends entirely on what I'm doing and what my schedule is like. Wikipedia is my most flexible day-to-day business for me. It also depends on the task I am on`

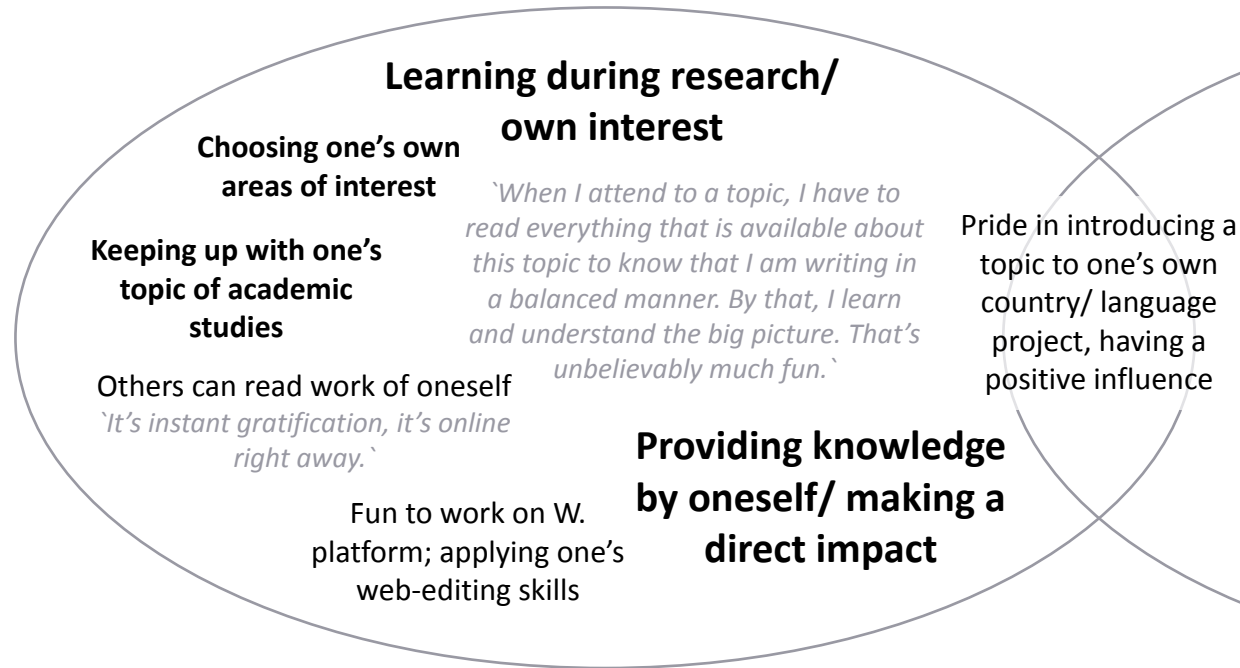
- Also, **two main patterns of project load** appear, equally distributed:
 - Many topics in parallel (e.g., one big and other smaller, related topics)
 - Only one topic at a time
- For some, Wikipedia work is their biggest hobby

`I spend my weekends taking pictures when I am not working in articles`

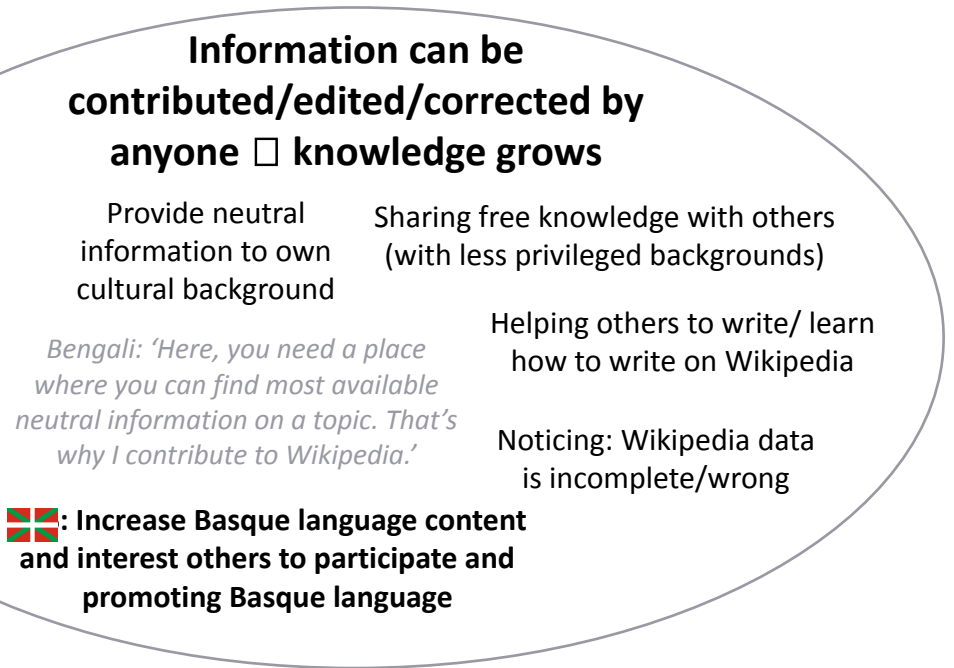


Motives For Working On Wikipedia

Personal Reasons

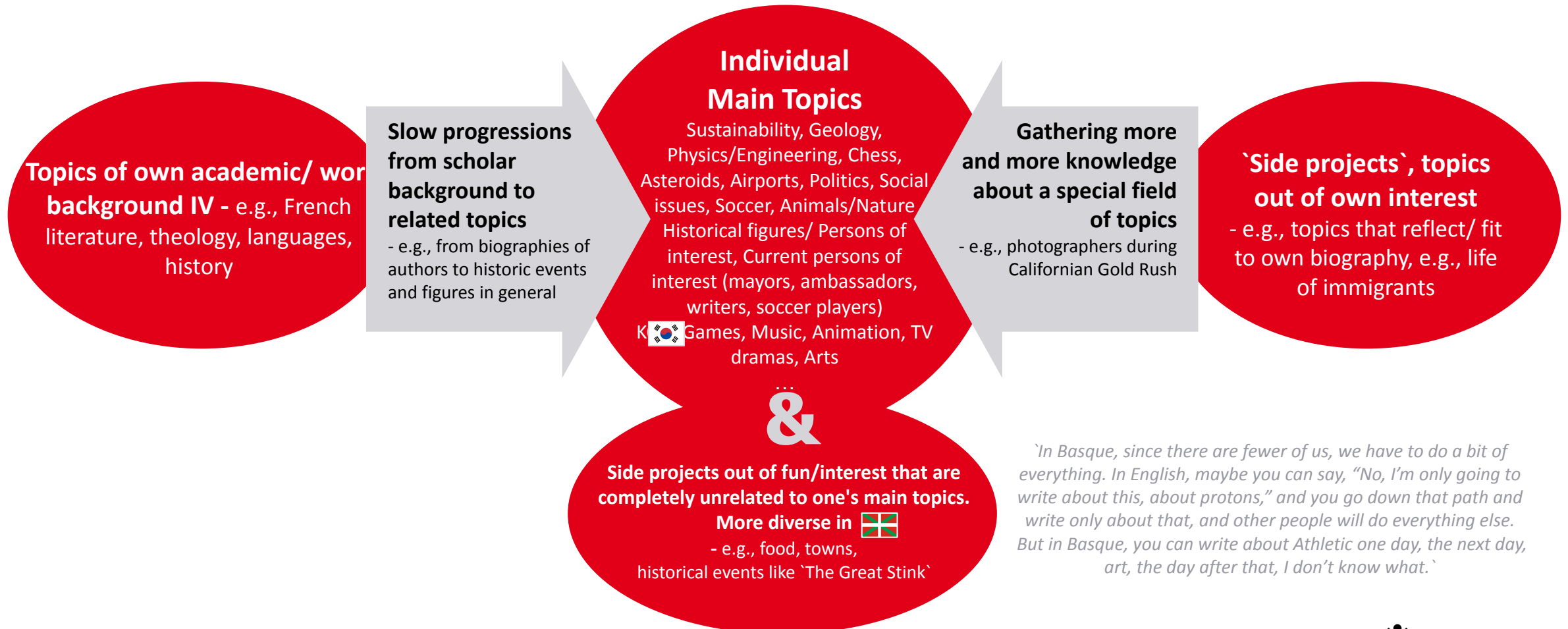


Altruistic Reasons







Motives to contribute on Wikipedia are as diverse, as the authors themselves. Motives can roughly be distinguished by personal reasons, making people enjoy their contributions for their own sake, and altruistic reasons, where motivation roots in the benefit for others.

Topics Of Authors In The Sample



Languages

<p>Mainly writing in:</p>	<ul style="list-style-type: none"> English  Hebrew <i>'Hebrew Wikipedia needs me more.'</i> Bengali 	<ul style="list-style-type: none"> German  	<ul style="list-style-type: none"> Basque  Translating from English, Spanish, French, German/Russian (also automatic translation) 	<ul style="list-style-type: none"> Korean  Translating from English, Japanese, Chinese (few); (also automatic translation)
<p>Seldom working in:</p>	<ul style="list-style-type: none"> Spanish English – only fixing 	<ul style="list-style-type: none"> English – correcting mistakes Swedish – writing and adding Dutch – adding Spanish – writing and deleting 	<ul style="list-style-type: none"> Spanish – editing English – editing simple things or contributing to discussions; documenting in Wikipedia Commons/Wikidata Catalan, Galician, Portuguese, Italian, Occitan – editing with translation tools 	<ul style="list-style-type: none"> English (editing existing articles/ fixing small issues) Japanese
<p>Capable of foreign languages:</p>	<ul style="list-style-type: none"> Hebrew Bengali Latin, Greek 	<ul style="list-style-type: none"> All: English Some: French, Latin, Greek Individually: Swedish, Dutch, Spanish 	<ul style="list-style-type: none"> All: English Read: other Romance languages (e.g., Italian, French, Catalan, Galician, Portuguese) 	<ul style="list-style-type: none"> All: English, Japanese (some only read) Individually: Chinese
		<p><i>'I don't care whether I read German or English technical literature. This is also very practical for Wikipedia Library. Because if you concentrate on German, then you don't get much out of it.'</i></p>	<p><i>'Yes, English is almost mandatory. You can work on Wikipedia without it, but to delve deeper and so on, you need English.'</i></p>	<p><i>'Yes, English is very helpful. Most of my textbooks in university are in English, and if you can't read English, it's really hard to study. Same for Wikipedia editing.'</i></p>

English is regarded as very important in all non-English language projects in the sample. However, editors do only seldom write or edit in English, as they refrain from working in a language that had a great many contributors and is not their own native language. A few only though only hesitate to work in foreign languages, as they are not familiar with conventions (e.g., citation requirements) of the individual projects. But for research, especially in academic sources (the preferred kind of source), English skills are indispensable. No respondent had an issue with research in English, as nearly all had an academic background and were used to working in English anyway.



Main/Most Frequent Activities On Wikipedia

Frequency



Writing articles



Research for writing



Adding pictures



Contributing to existing articles

- Adding references and making changes/ additions/corrections
- : "I sometimes just can't stand by and witnessing the misery of bad articles."

Though most respondents mention, that there are Wikipedians who only do research, take photos, make corrections, etc., nearly all in the given sample saw writing/ starting articles as their main task on Wikipedia



Administrator Role

- On Wikipedia (most) or Wikimedia commons (one)
- e.g., deleting articles that are obviously not valid, such as marriage proposals or pictures with wrong licenses
- Mediating conflicts in discussion about articles



Research for adding references

- When too few references are there
- When content seems to be controversial/ wrong
- When old articles do not yet have references (



Keeping track of own articles

- Using the watch list to get notification about changes
- Less in : Articles are not closely followed, establishing an article base has higher priority. Also, criticism/discussion culture is less established



Advisor for other Wikipedia projects

- Small/ developing projects

: "I check it with the automatic translator to try to see which language it's better in. Sometimes some texts are better in German, they're very meticulous. Sometimes in Russian too, but that can be hard to tell."



Translating articles into own language

- / : Adapting content/ POV according to own cultural background (e.g., aspects of race, gender and culture)
- partly: Reviewing sources when translating
- Taking sections from different languages for one topic to translate an article into Basque

All Wikipedians in the sample have in common: Writing articles and researching for writing are their main activities. All who write do the related research as well.

Nearly everybody will add pictures and also contribute to existing articles in one way or another. Many also work as an administrator in their project

Corrections And Individual Activities On Wikipedia



Corrections/ Patrols:

- **Fixing dead URLs/Links**
- **Correcting content (when apparent)/ language corrections**
- Looking for articles that are marked as missing references and research for sources
- Fixing articles when source information has changed (information rot) – encountered during translation work
- Fixing articles that contain authors’ opinions – biased content (esp. religious and political topics)
- Review: Checking references for reliability
- Tracing vandalism, blocking SPAM
- Updating/ fixing old articles from times when standards were lower
- Correcting articles that only contain popular/ internet sources

Corrections in others’ articles are frequently done, but systematic `patrolling` does rarely happen. A major issue in corrections is fixing dead links to online sources. Opportunities for corrections are rather encountered in passing. Besides commonly shared activities, every Wikipedian has some hobbyhorses of activities they pursue.

: `Problematic articles appear in our online room called ‘Sarangbang’ and in our association member chatting room. As Korean Wikipedia is quite large, there is a monitoring person who is in charge of patrolling such issues, so I don’t need to patrol other people’s edits.`



Individual activities:

Steering:

- **Discussion of overall Wikipedia issues**
- **Driving projects within Wikipedia**
- Director of local Wikimedia Association
- Secretary general of local Wikimedia Association
- Member of Wikimedia association
- Coordinating User Groups
- Spokesperson of Association of Wikipedians
- Presswork for German Wikipedia
- Set guidelines for other editors
- / Supporting offline events

Particular detail tasks:

- **Contributing to Wikimedia Commons** (e.g., digitalization of pictures of old books)
- Adding Links between articles using ‘Findlink’
- Consulting in own field of expertise (mostly own academic background)
- Identify currently popular topics for new articles
- Creating short articles with a few references for others to start working
- Technical work (templates, software, helping others with technical issues)
- Corrections/additions in Wikidata
- Contribute to gender equality initiatives

Training/ teaching:

- **Teaching how to write articles in trainings/ introductions to Wikipedia**
- **Mentoring new editors/ students, e.g., teaching the need for references**
- : **Introducing Wikipedia to students in own lectures at university, e.g., in programming**
- Contributing videos of courses (Latin, Wiki Data)
- Maintaining blogs with trainings (Wikidata)
- Review/ helping others to improve their contents and skills (e.g., tagging sections for improvements)

Personal tasks:

- Attending contests in Wikipedia

Need/ Reasons For Research

Often (in all projects) one activity leads to another, as interesting topics/sources/pictures are found.



Commonly shared reasons for research:



Start an article and structuring content



Adding pictures to articles (nearly all)



Adding information to existing articles



Encountering an article with too few references instead of marking it as `missing source` or `check for deletion`



Looking for references to back facts (that are known by own expertise/knowledge)



Looking for additional references when own article is marked with `missing source` or `check for deletion` (esp. early articles, when beginning with Wikipedia work)



For translation articles (English -> Hebrew; English/Japanese -> Korean) (official policy of Hebrew Wikipedia) -> does source really (still) fit information in article?

Frequency

Individual reasons for research:

- Research on photos for upload on Wikipedia Commons, digitalization of pictures of old books (check for copyrights)
- Adding information on Wikidata
- Checking articles that might have controversial content to see if all facts/ points of view are covered
- Check reliability of sources and replace unreliable sources
- Include information on newest development of article topic
- 🇰🇷: Back information that author already knows by sources
- See, if article topic is noteworthy enough



'If there is no reference, I'll stand the article one or two references. This way, the article at least does not get deleted.'

Researchers agree, that the need for sources and therefore research has grown. In the early times of Wikipedia, much less importance was placed on sources and verifiability.








'You can see the earliest version (of my article) was a summary of the book and the plot summary. I knew nothing about sources, at the time, you could get away with that.'



Research Sources I/II


- **Google/ : Naver News (or Naver news archive)/Daum News**
 - To get a broad overview and obvious sources (starting point)
 - **Most powerful research engine**
 - To look for new sources for otherwise well-known topics
 - : Google for overseas data
 - Occasionally: Google News

- **Academic sources/ scientific online databases:**




- **Often accessed via Wikipedia Library**
 - **Wikipedia Library less often reported to be used in / **
- Science direct
- JStore : *‘Since I graduated from university now, it’s not easy to have access to academic database behind paywalls. If I want to use those data, I visit the National Assembly Library.’*
- BioOne
- Academia.edu
- University libraries
- Self-paid
- : Korean National Assembly Library
- : RISS (research Information Sharing System)
- : Online meta-catalogues; Karlsruher Virtueller Katalog (KVK)
- : Sci-Hub (shadow-library)

: *‘There is not much information in Korean Wikipedia Library. English Wikipedia Library has some information, but I don’t need English information. If I need law information, I can go to the Korean Law Information Center, or look for certain websites.’*



- **Physical books:**

- **Search in reference library**
 - Only if not available as online source
 - : Maybe also interlibrary loan is necessary
- **Self-bought books for topics that are hard to find in other ways or need to be used frequently (e.g., specialized encyclopedia):**
 - Geography in developing projects/ countries
 - Antiquarian books (not found in libraries anymore)
 - Cultural sites
 - Archeological sites
 - (Local) history
 - Theology


- **Sources from same topic articles in other languages**

- : Frequently sources from English and Japanese, sometimes Chinese project
- Esp. : piecing together from different languages
- : Seldom, scruples about quality
- Bengali: copying from larger projects

- **Reliable news websites**

- : frequently overseas publishers (BBC)
- : Yonhap News (regarded neutral)




- **Sources from related topic articles in same project**



: *‘Yes, of course. If I find something tricky in an English article, I use other (Korean) Wikipedia articles to find references.’*

‘The issue is the English Wikipedia. It is not as good in quality. Many German Wikipedians consider it’s a giant scrapheap. It is only raggle-taggle information without coherence. I’d not look anything up there. It’s a catastrophe!’



Research Sources II/II

- **Google Scholar**
- **Google books:**
 - For ‘quick and dirty’ snippets, but not for thorough articles
 - Read summary to know information is there, not read actual article
- **Other Wikipedians or experts in the respective field**
 - Give recommendations which sources to use or where to look
 - : Asking for definitions or how to use a specific term
 - : Wikipedia Community site where other Wikipedians with access to certain sources (e.g. university libraries) can be asked to look for publications of interest
- **Specialized databases**
 - Astronomy
 - Ornithology
- : Government websites
- Sources that are known by own expertise, e.g.
 - Inter-governmental panel on climate change
 - Sites on specific topics:
 - Climate-ADAPT
 - Subject-based websites with layout like Wikipedia
 - Research for information, but no citation of that source, verifying information by other sources
 - NGO reports

- Newspaper archives for research on towns/ local history
- Person that is written about – but information needs to be verified
- Citations of 2nd order: Source that cited a source than cannot be directly accessed; mediocre reliable source
- Encyclopedias, compendiums, specialized books
- Wikidata
- Video searches/ YouTube videos
 - : “top notch source”, but cumbersome to use
 - : Not valid

‘I actually use it (Videos) as a second resort, if not the last resort. For my process, it is inefficient. You can’t search.’

Other Wikipedia articles as a quotable source:

- All agree, that another Wikipedia article cannot be used as a reference itself

Many Wikipedians - especially those who stay within the same topic range and/or write in their own academic field – establish something like an own personal ‘library’ for themselves, consisting of fundamental works, reference works, articles that were used as references, links and self-bought physical books.

Online Vs. Offline Sources



Online


‘I don’t have any reservations. If it is the decisive information on the topic, I use it.’
‘Sometimes, It looks like there’s a very good book on a topic, because it’s very often cited. I’ll try to see if it’s in the library, but first I try to find it online. Technically, it’s easier’


- **Mostly preferred and used due to convenience for the author**
- Can be more up-to-date, than the printed version of a source
- Believed to be less reliable (more changeable)
- Links in articles (e.g., to digitalized content):
 - Seen as a good addition for readers, but not top priority
 - 🇰🇷: Positive, fosters verifiability
 - Problematic, due to information rot*
- Pure online sources:
 - More difficult to cite
 - Information rot is an issue*

‘For example google docs, it constantly changes links. You end up nowhere, it is not reliable.’
‘It’s frightening how many websites change their URLs. Well, there are then the perma links, that can then still work. But that is an incredible maintenance effort. That’s why the access only to websites is problematic.’



Offline

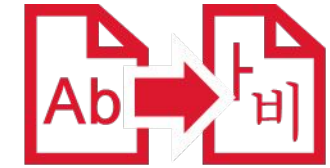
: *‘If I have a physical source, I’d prefer that, but it’s quicker and easier to see on the internet. You don’t always have all the books you need.’*

- **For sources than cannot be accessed online, like old books or when no access to a digital version is available**
 - A digital version of a book or magazine can also be considered an ‘offline’ source ()
 - Believed to be more reliable (less changeable)
- 🇰🇷: ‘It’s a matter of accessibility. Whether I can access the information or not, is the most important thing. If the data doesn’t exist on the internet, that’s acceptable. But as for physical research, for example if I can’t borrow or find books for free at the library, I would have to make an application to copy or borrow the book to eventually get that source.’*

The topic of online vs. offline sources is not discussed with respect to accessibility for readers (thus in terms of verifiability by the reader), but with respect to accessibility for writers and credibility of the source. Academic books and magazines may be considered an ‘offline source’ that can also be accessed online. Neither is there a differentiation regarding credibility, as both sources must withstand the same demands on verifiability and credibility.

***Even renown and widespread online sources (e.g. nation-wide online news sites) are subject to information rot: the link information referring to a specific article becomes invalid or has changed. Some projects use bots that search for such dead links and try to regain access to such sources (e.g. through internet archives). But in general, information rot of online sources poses big difficulties to verifiability.**

Replacing Of Foreign-Language Sources



Replacing foreign-language sources?

- **No (majority)**
 - ‘Common’ languages mostly accepted (English, French; 🇰🇷: Japanese and Chinese)
 - Quality trumps language
 - Esp. scientific papers and research findings are only available in foreign languages, commonly accepted by the research community
 - References are chosen fitting to the language of the article topic
 - E.g., some information can only be accessed in other languages, like famous foreign authors
 - Translation websites work well, if someone wants to look up the original source
 - 🇩🇪: Some community members do not want foreign-language sources as they are not comprehensible
- **Yes (only occasionally)**
 - Adding local references, if more up-to-date
 - If equally valid, sources in project language are used
 - 🇬🇧: Replacing references, if easily possible and a suitable source is at hand

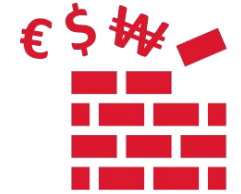
‘It’s okay to use English, Chinese or Japanese citations but when using citations from other language sources unfamiliar to us, like from Mongolian language, Russian or Greek, only few people in Korea speak those languages and it would be hard to verify such articles.’ 🇰🇷

🇬🇧: ‘With a lot of the references I include, I translate them, but I include the original. With a lot of topics, there are no references in Basque, so I use whatever I read, in English or in Spanish, I use that.’

Bengali: ‘To be honest, most translators use the existing sources believing that they have been checked by that community. They normally do not cross check the source most of the time.’

🇬🇧: ‘I try to look up topics to see if there are any references in Basque. In Elhuyar or in Berría, or an event, something contemporary. Instead of including a BBC news story, I’ll put in the one in Berría. That’s something I try to do. But if there’s a book, or if there’s a scientific journal, it’s not usually in Basque.’

Paywalls



- **All Wikipedians in sample agree, that it is acceptable to use paywall articles, as they are accessible by everybody in principle**
 - Some Wikipedians are willing to pay for articles and already do so, e.g., charged newspapers or databases
 - Some state, that hardly anybody will try to access the cited sources anyway
 - 🇩🇪: Using Sci-Hub to bypass paywalls
 - **Advantages** of paywall articles:
 - Are most probably verified by the publisher of the article, thus reliable information
 - Might have better information and Wikipedia articles therefore get better
 - Using these sources allows others to get the important information that is behind the paywall
 - **Disadvantages** of paywall articles:
 - Cannot be accessed by everybody

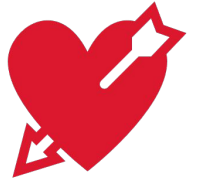
- **Though, Most prefer sources without paywall**
 - **Most: Paywall articles won't be used, if not freely accessible by one's own possibilities (through Wikimedia Library, university accesses, connections who can provide the article)**
 - **No willingness to pay on their own**
 - No willingness to pay for scientific information/ should be free to the public
 - Bengali/ 🇰🇷: Paywalled newspaper articles will become free anyway after a certain time

'I can use them. For example, to read a paper, you need to pay for some amount, and a paid source doesn't mean that it is not accessible, so I can use those references.'

'I have also bought myself access to databases that were not available in the Wikipedia Library. It's 100\$ per year.'









'We can create open content that synthesizes what is written behind the paywall. Most academic articles and their content do not matter for the public.'


'I do not like the extortionate rates that academic paywalls charge even though I can afford them, ideologically it upsets me. If can't prove it otherwise and if I encounter a paywall, I give up on that citation.'




General Pain Points

Pain points are as individual, as are Wikipedians (all single mentions):

- Finding sources, esp. when not available online *‘Either you have to buy books or pay news organisations to access information that you need or some research project.’*
- Getting an overview about a topic and how to approach it
- Writing an overview of an article *‘When I write a biography, then it takes the most time to think about what marked that person. What was the golden thread in his life?’*
- Verification process of conflicting sources
- Paraphrasing source content into own wording
- Distinguishing good from bad information/ sources
- Coming up with the right keyword for searches
- Page numbers cannot be cited when books are read in e-reader format (only chapters)
- : Encountering poorly organized data in Wikidata (also equivalences between languages, complicates translations)
- : Finding the first few sources (esp. when topics are not covered in other languages)
- : Need to find a source for every basic fact, esp. public places, landmarks
-  only: Citations in Citoid:
 - Dynamic documents, hard to be cited, only URL available, no details like author, date of publication etc.
 - Citing external links like podcasts in : Verifying sources when translating *‘The interface doesn’t give you a workflow for external links like podcasts.’*
- : When `opening` articles for others to write, no mechanism exists, that those can be marked as a starting point or `draft` for others to take over. These articles are treated as an incomplete published article/faulty and are always threatened by deletion by other users
- : Writing about topics, that are not scientifically covered. Only officially published sources are valid
- : Translating proper nouns into Korean

: *‘Why can’t Wikipedia be used as the first place where something is written down? We don’t have a reference for everything. We do it with the photos, don’t we?’*

: *‘There is a regulation in Wiki to cite from renowned books, news articles or research data from popular journals. You cannot just cite any information from online blogs. This is the biggest challenge. It’s relatively easy in science, but for topics like games or the internet etc., people rarely work on those topics or write journals or thesis. I can rarely find articles related to games, and most game-related information are on online blogs, but I can’t rely on blog information. So, for some topics, I can’t do editing on Wiki.’*

Tools Used




Writing

- Most authors directly write in the Wikipedia interface
- One respondent mentioned to write in a text editor for programming



References:

- References are typically added right into the unpublished version of the article
-  only (occasionally): Citoid



Some use other methods/ places to track their sources. All methods have in common, that they are mostly simple lists with only a few comments, if at all:

- Simple notepad file to gather links
- Links in a Google doc
- Piles of books sorted by topic, marked with Post-Its
- EverNote
- No tool, all in memory
- Paper notebook
- Downloaded PDFs in a local folder

`It's shockingly inconsistent. I think the two standard things are, I create an article draft or I download the PDFs.`



Evaluation Of Sources

- Authors are generally **trying** to get the **most credible source**
'I avoid yellow journalism media or articles having those tones. I can't exactly explain how I define yellow journalism. I just learn from my experiences..'
- **Credibility** of a source is often judged by **own experience** and **knowledge** which source is credible in one's **specific field**
- Trying to **find out**, if a source is **reliable** or somehow **influenced** by an interest group or stakeholder can be very **cumbersome**, but is **mostly done thoroughly**
- The **less information** can be **found** about a topic, the **less particular** researchers can be about the **source** they cite
As for this person, there were only news articles, and no other sources like books, so I only used sources on the internet.
- When **translating**, the work of the original authors is mostly **trusted**, sources are only **checked**, if **appearing** very **unusual**
'It took me an hour to find the stuff because I needed to separate and see if these were serious and trustworthy.'

Most trustworthy sources

Scholarly sources (officially published as book or magazine)

- Verified by peer-review or publishers
 - Written by experts in field
 - Cite sources themselves
 - No citation rot
 - Governmental sources (🇰🇷, 🇩🇪)
 - Neutral/abstract
- 'Korean news on overseas topics are low in quality so I try not to use them. I prefer research data and government sources.'*

Topics covered by academic fields

Mostly Credible sources

- Big and/ or reputable newspapers/ media; wide circulation, established
 - NGO reports
 - News agencies
- 'You can only assume some kind of peer-review, when the circulation is wide enough.'*
- ### Secondary or tertiary sources
- E.g., books; first-hand data already verified by third party

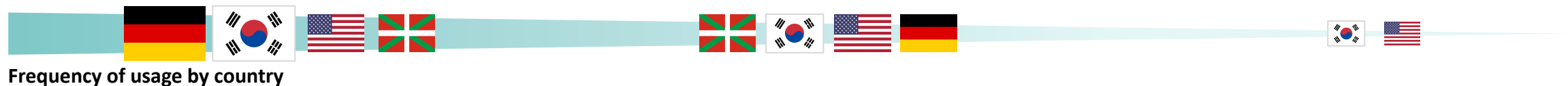
Events and narrative facts
Persons of interest

Least trustworthy sources

Non-backed sources/ doubtable sources

- Blogs
- Small newspapers
- Primary sources (like self-statements)
- Predatory journals
- Governmental sources with doubtable interests (e.g. Bengali)







Non-academic topics, e.g., food
Persons of interest



References




Content without reference:


- **It is the common opinion, that content without references is not valid and contradicts the Wikipedian principle of verifiability**
 - Those articles are marked as `references missing` and then (but seldom) deleted (by others), if no references are added
 - Some add a few references by themselves to prevent the articles from deletion
 - : Still existing in old articles when referencing was not demanded – slow renovation of those articles
 -  (occasionally): Better a short article (e.g., a definition) on a notable topic with no reference, than no article. Also, new editors must not be discouraged of editing
 - : Differentiating aspect between Wikipedia and [Namuwiki](#), where authors may write out of their own opinions and beliefs
 - : Not possible to find reference for every notable topic
- **It is seen as `OK` by many to publish articles with only a few references and add more later but can cause conflicts. Preferably, an article should be published with references from the beginning or references should be added as soon as possible**
- , : Simple definitions/ obvious facts should not need a reference

‘Sometimes it’s absurd to have a reference for example: `This or that is a neighborhood in this town.`’






‘If I can’t find that information anywhere, it won’t make it into the article.’

 *‘The editor would have edited the article with good intentions, wanting to be helpful, but if I just delete that article, it would be just kicking out the editor from Wiki editing.’*

‘If the editor can add reference in a short time after writing the content, it would be okay, it’s not a big deal having no reference for a short time, and it’s reasonable to allow that for editors.’

 *‘In the US, every sh** is has a reference. To me, this looks like a beginner’s work. We have different science traditions.’*

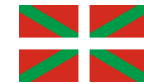
References in different language projects:

- More references in the  comp. to  (more frequently in text) – : References at end of paragraph are sufficient
- : Believe, that German references need to be more reliable/ verifiable, than in English Wikipedia
 - More non-academic online sources in English => More citation rot/ broken links
-  esp.: Dispute within project, about how many sources ought to be added as references
 - Some: One good source for one fact is enough; others (not in the present sample) in contrast: Demand 3-4 sources instead of just one
- Articles in other languages might sometimes be biased by political views of country, e.g., article on [Liaoncourt Rocks](#) where sovereignty is questioned, is called `Dokdo` (solitary island) in Korean, but `Takehima` (bamboo island) in Japanese

Project Maturity



‘There are around 530,000 articles in Korean Wikipedia, and there are around 6.28 million articles in English Wikipedia., it’s almost 12 to 13-fold. So I concentrate on Korean articles’



Mature

- Enhancing a solid foundation of articles
- More detailed topics/ marginal topics are added □ mechanisms to exclude non-notable topics
- Trying to keep up-to-date with and adding articles on current affairs
- Re-working old articles from the beginning of the project
- Guiding/ counselling smaller or starting language projects
- Authors mostly staying within their fields of expertise

Establishing

- Still concentrating on building article base on basic topics
 - Articles are often created and elaborated, but not so often edited by others
- Translating articles from bigger projects into own language
- Might accept articles with no references (yet) to build said article base
- Articles are often created and elaborated by one person, but not so often edited by others
- Not (yet) dealing with ‘quality assurance’ topics like information rot
- Believe to have less strict rules, than English Wikipedia project

Starting

- Still concentrating on building article base on basic topics
- Translating articles from bigger projects into own language, often piecing together parts of articles/ sources from various languages
- Authors often working on various topics, also such that are far away from their own field of expertise
- Discussions often focus on translation issues

 *‘I am not an expert in many topics. For me, the most difficult thing is generally figuring out what the initial outline should look like.’*

Prioritization Of Sources



Conflicting sources:


- The **predominant approach** do deal with conflicting sources is to **describe the contradiction/** conflict and substantiate **with sources for each point of view**
 - Sometimes, a discussion in the footnote is added about which source is presumably the more reliable one
- **Occasionally**, the **decision** is made by the author, which source to use over the other. In doing so, authors try to **judge by themselves**, which source and information is the **most credible or preferable one**. Sometimes, it is reported to be an obvious choice. Criteria might e.g., be repetition of one fact in multiple sources
- Few report, that conflicting sources are hardly encountered. This of course depends on the field of writing
- Few might go to considerable length of following back to original sources (e.g. birth certificates) to validate one of the facts or resolve contradiction

‘Many editors misunderstand the concept of neutrality. As you can see here, the right concept of neutrality is to cautiously and critically analyze various reliable sources.’

‘I will write ‘There is debate regarding this point’. And I add the references and that’s it. I’m not going to get into judging one or the other.’



Multiple sources:

- Often, one or two sources are selected. Sources are prioritized by:
 - **Reputation/circulation, judged at own discretion or by official Wikipedia guidelines**
 - **Novelty**
 - Well-accessible for author and reader
 - Neutrality (not affiliated to an interest group) Knowledge needs to be built, which source is trustworthy
 - : (Also) Use Basque source, if equally valid and reputable
- Few might be tracking down the original source of the information when several sources are based on it
- Some will **cite multiple sources to emphasize credibility** of information (also to prevent a new/short article from deletion by demonstrating notability of topic)

: *‘There is a reliable data list, which is good to refer to.’*


Collaboration




Solitary Work/ Individual Contributions

- **Most frequent form of collaboration: Anonymous collaboration**

Collaboration on articles is slow and long-term, often with no contact between individuals

- **Collaboration is not coordinated or planned**
- Articles of others are reviewed
- Coordination is done via comments and nominations (for deletion) after publishing
- Experts in respective topics and with access to relevant sources contribute on their own accord
- Lists of subject matter experts (in Hebrew and  Wikipedia)
 - ☐ user groups can be tagged

‘Contrary to the belief that Wikipedia is collaborative, and it happens in groups, actually a lot of high-quality content gets written by individuals.’







 *‘It’s strange that you’ll always find people, no matter how odd the topic. You find people who know about it.’*



Mutual Open Exchange

‘There is a lot of this long-term slow collaboration. People leave questions, problems, shortcomings etc. on the page. I am a big believer in leaving these little love notes to the future.’

Ways of seeking help/ support by community:

- **Exchange happens, when problems or irregularities occur on a general or personal level (e.g., approach to sources, : tone of writing (polite or honorific), how to deal with topics, religious and governmental topics, BAS: corrwording in Basque...)**
- **Most frequent form of exchange: Discussion Sites on Wikipedia**
- Asking for information in specific topic portals in Wikipedia
- Asking for technical/programming help in Wikipedia Workshop (e.g., programming information box)
- Some use communication platforms outside Wikipedia like Twitter, Facebook, Twitch, Instagram, Telegram, WhatsApp...(few in given sample)
 - : All users in sample communicate via Discord or Kakaotalk voice chat
- : Discussions are not always seen as a good thing, as *‘feelings are hurt during the process.’*
- : Posting article on social media and getting hints on further sources from others
- : Only project to have a real collaboration in writing articles together (in given sample), coordination is done publicly, via Wikipedia talk pages

 *‘Yes, sometimes we work together. To expand something perhaps, I’ll work on one section, and I know someone else will do another.’*

Collaboration



‘This is the logical consequence of the Christianity project. Something can come from the Catholic colleague. You ask ‘how is this topic seen in the Protestant church. Do you know?’. I come in and add to that. There is a certain exchange. We support each other.’

Networks/ Closed Exchange

Some German Wikipedians are organized in topic-related communities:

- Collaboration in a group of themes (closest form of collaboration)
 - Working on articles with related topics, exchanging pictures and sources
- Discussion of topics and relevance/ validity/ necessity of sources -> strong quality control
- Communication via Wikipedia topic site, telephone, e-mail or text messages

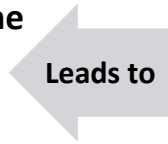
Other, than helping each other out occasionally with tips and clues for references, no work is shared in the research process, not even if editors work on a common project like an edit-a-thon. No real collaboration in writing or researching happens, rather a network-approach with experts and discussions when encountering issues and needing help.



One-to-one exchange:

Some Wikipedians have partners/ personal contacts who contribute certain steps in the editing process:

- Esp. : Providing sources/ articles/ literature due to access to specialized databases
- : Promoters of the community are very closely linked together and in constant exchange
- Screen for notability, before creating an article
- Proof-reading/ linguistic improvements
- Draw pictures for articles in Photoshop



Offline meetings

Some projects do offline meetings:






- : ‘Crackerbarrel’ meetings of Wikipedians; **not work-focused meeting**
 - Connections are formed in those meetings and relationships/ friendships form
- : **Working together during ‘edit-a-thons’; work-focused meetings**
 - Partly: Everybody works on an own text there (as success of funded event is partly measured by number of articles)
 - Partly: Collaboration finding references and editing a specific topic together

: *“We know each other, we are like a tight-knit community, that’s not visible to the outside. We have real-life meetings where you talk about Wikimedia all evening long. The English-speaking Wikimedia is very much different to the German. They don’t have that.”*

Images – Sources



How images are attained:

- **In most cases, researchers engage in a direct search for pictures. Sources are:**
 - **Most search in Wikimedia Commons**
 - : Some believe, that Wikipedia Library and Wikimedia Commons cannot be used in Korea due to licensing issues
 - **Google** – looking for license-free pictures
 - **Government-affiliated accounts, embassies, ministries, national associations** (such as NASA)
 - **Flickr:**
 - Ask user for permission
 - Museums - authors directly ask museums for usage permission in Wikipedia context
 - But: The majority refrain from using images from museums and galleries, as the copyright status is not always clear or would have to be enquired
 - : [Pxsphere](#)
 - : [Europeana](#), e.g., for old maps
- Sometimes, pictures are found incidentally when doing other work/ research
- **Other Sources are:**
 - **Taking pictures on one's own:**
 - **Some researchers intentionally go out to take pictures of their desired motifs (esp. KOR)**
 - Occasionally, suitable image motifs are found by chance when passing by (e.g., buildings, animals)
 - Buying pictures on Ebay (e.g., postcards) and scanning them on their own
 - Draw pictures in Photoshop (but content is not reliable)
 - : [Acofran](#)
 - : Organizing photo contests on Wikipedia to enlarge image base

‘Adding photos would be perfect but it’s hard to search and edit photos. For example, there’s a program used by engineering students for drawing graphs in maths and science, it’s easy to add images in those sectors. But as for ordinary images like the moon, you have to go out and take a photo of it which is really hard. Photoshop can be used but that’s pretty hard for me. So, I just add equations and graphs.’

‘Most flicker users do not know that they can share the pictures in the Wikimedia commons website. I motivate them.’

‘When I look at Korean Wiki articles, I can see that there are not many images, compared to English Wiki articles. So when I go for a business trip or go traveling, I always think about using photos taken there for Wiki articles. For example, as for the ‘Hell Chosun (Hell Korea)’ article.’

Images – Verification



Verification of image content:

- **Mostly, there is not direct effort made to verify the picture content. A rather indirect verification takes place through the appraised credibility of the picture source,** e.g., Wikimedia Commons, government-affiliated accounts, embassies
- Sometimes, pictures content is simply judged by own knowledge about the article topic, e.g., no doubt for known historical building
- Esp. in 🇰🇷, researchers state, that authors do not pay much attention to the verification of images
- 🇩🇪: If necessary, researchers are checking alleged picture content with other material that can verify the image (e.g., position plans of archeological digs and ground plans)
- Some are comparing picture content with other pictures that claim to show the same content

Verification of image copyright status *‘Licensing is essential. We can’t upload anything that’s not free.’*

- **More often, than the content, the copyright status is verified. Typically, not the picture itself is checked for copyright status, but the picture source itself is chosen as providing pictures with suitable licenses**
- Occasionally: google picture search is done and a picture is verified, if not search hit comes up
- Occasionally: Verification of day of taking
- 🇰🇷: Some believe, that Wikipedia Library and Wikimedia Commons cannot be used in Korea due to licensing issues

Pictures of other Wikipedians’ articles are usually not checked systematically for validity of content. The ‘good faith’ rule applies.

Pictures are in general seldom verified. Historic pictures might be undertaken a more thorough check, but current pictures are judged by common sense.

The situation is very different for copyright status. All Wikipedians in the sample are very careful with licensing issues and regulations.

Some complain, that standards for how to illustrate articles and verify images are missing (e.g. to prevent the usage of deep fakes).

🇰🇷: *‘Acrofan website is the only source that I can use images freely without copyright regulations, so I mostly use images from Acrofan.’*

🇰🇷: *No, because using those sources (museums) would be copyright infringement. Other countries work on Wikipedia Library and Wikipedia Gallery and many sources can be used from those websites, but in Korea, sources cannot be freely used.*

Images – Purpose, Demands, Captions And Abstract Topics



Usage of images:

*‘Pictured are very necessary, humans are very visual.
Images really help you understand an article’*

- **Illustrative use of image**
- **No use of image-content as a direct source for information** *‘It just looks better and more convenient with a picture.’*
- To lighten long passages of text, make it more ‘digestible’
- Digitalization of pictures of old books for Wikimedia Commons
- Depicting of history topics: Old artefacts like coins from respective eras, no paintings are used (as they are already an interpretation of the historic event)

Pictures for abstract topics:

- **Mostly, editors refrain from adding pictures to abstract topics**
- **Seen controversial, whether and how to do illustrate abstract topics**
- Done occasionally, e.g.: war ☐ people fighting, democracy ☐ people voting
- 🇰🇷: Copied from English Wikipedia article with same topic

Decisive criteria between pictures/ demands on pictures:

- **Highest resolution/ good quality**
(but also: **Mostly not an issue anymore**)
- **Looks good ‘stands out’ / by intuition/ layout**
- Source: License free/ published under correct license
- Detailed/ high information value/ relevance
- Motive must be clearly visible and discernible
- Must add value to the article
- No personal information, e.g. license plates visible

Picture caption:

- **Picture captions are important, but the scope of information is subject to discussion**
- **All agree: An explanation of picture content is necessary**
 - Some: Relation to article content should also be established
 - Too extensive descriptions get criticized by other editors
- Year of picture taking
- For art: Artist, museum, year
- 🇰🇷: Date of taking and photographer

‘All images are managed in Wikimedia Commons, they store all images on a server, so if a Korean article is uploaded, that is stored on this global server. If only an image is uploaded without any explanations, users from other countries won’t understand about the image.’

‘I want to make the picture content comprehensible for the reader’

Image Associations



Research going well:

- Structured work packages that are all in relation to each other. Nothing left unclear, the spaces just need to be filled
- Research means exploring things and having insights
- Knowledge that is built: a lot of space still to cover
- Having a flow
- Needs to focus his aim, the central focus point is visible when research goes well
- Accuracy
- Moving straight and fast without problems
- Feels good
- A lot of data piles up
- Things are well-linked, all is interconnected and well-organized



Research going poorly:

- A lot of failing before the end result is valid. A lot got destroyed in the process that cannot be repaired
- Spent a lot of time, but going nowhere
- Found a few sources by the skin of one's teeth, but nothing is really suitable. It's okay, but not really good
- A lot of literature, very similar to each other, nothing sticks out, no orientation what is the central point
- Chaos
- Data not being linked together
- Too much material
- Being all alone with a lot of unclear data
- Feeling lost
- A lot of things can go wrong like when building a bridge



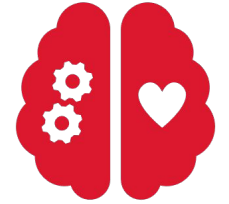
Important Skills For Researchers



Important Skills in the eyes of researchers:

- **Background Knowledge of topic**
- **Diligence/ attention to detail** – thoroughly processing all information
- **Patience/ perseverance**
- Scientific work skills
- Learning to `let go` of an article that one wrote for others to change/ amend
- Ability to compromise
- Being critical and questioning everything
- Neutrality: Not interpreting what one reads
- Being able to leave things out that are not verifiable
- Quick reading
- Coming up with good keywords for searches
- Knowledge about credibility/neutrality of sources, e.g., publishers
- Quick-wittedness – understanding quickly though being unfamiliar with a topic
- Knowledge about citations of articles and how to read them
- Filtering information well
- Know how to verify information
- Good writing skills in own language
- Being organized
- Curiosity

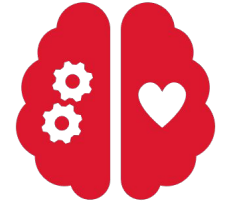
`What you need is to enjoy knowledge and to work well on Wikipedia with attention to detail. Do it properly, because you're publishing, after all, and a lot of people will read it.`



Important skills for being a good Wikipedia researcher are manifold.

Most frequently mentioned is the importance of background knowledge of the topics that one wants to write about. This knowledge helps in finding and evaluation sources and structuring articles.

Skills



How research skills were acquired:

‘Well, there’s no resource. I just have to learn the skills in the course of my work. I can learn about editing skills, about extensions, about the Wikipedia guidelines etc. by constantly editing on Wikipedia projects.’

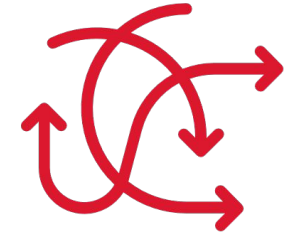
- **Develop skills by researching, ‘learning by doing’**
 - **Learning about guidelines and rules in ones Wikipedia project (e.g., using the ‘Help’ sections on Wikipedia)**
 - Start with one topic and get from source to source
 - Learning to find the right search keywords for searches
 - Develop knowledge about the topics to better grasp concepts
 - Learning what the right sources are for a certain group of themes
 - Learning how and where to find reliable sources (e.g., from an official Wikipedia [list](#))
 - Learning not to be too specific (e.g., case studies) and give too many examples
 - Learning that sources might contradict each other
 - Learning topic-specific peculiarities/ nomenclatures
 - Talking to other Wikipedians
 - 🇰🇷: Learning by reading others’ articles, e.g., structure
- **Research skills developed during own academic career or already at school**
- Research skills developed during personal ‘research projects’, like writing a book

🇰🇷: *‘Wiki editors vote for the best articles, and after reading the best articles, I can benchmark them, I that’s how I learned about writing and editing skills.’*

‘I learned about ‘good writing, logical writing’ in university.’


Two main ways of acquiring the necessary skills for doing research are mentioned: Learning on the task itself and benefitting from the experience that was gathered through one’s school/ academic education. In line with this, it is mentioned, that researchers/ editors without an academic background need more support and time to build good skills.

Mistakes And Initial Difficulties



Mistakes

- **Writing articles without sources (also reported to be a frequent phenomenon in early Wikipedia articles of all projects)**
- **Interpretation of texts by authors/ not writing neutrally**
- Deleting words while fixing typos
- Uploading pictures with the wrong size
- Relying in google books (only cooperates with certain publishers)

: *`I wasn't cautious about references and keeping neutrality. That's when I learned about the Wikipedia requirements and criteria regarding references. I learned that subjective descriptions should be ruled out and neutrality is required in Wikipedia.`*




Difficulties

- Knowing where to look for good sources
- Not knowing about neutrality and objectivity
- Staying neutral


Most mistakes and difficulties mentioned are about the need and problems of staying neutral while writing. This can hardly be taught by manuals and help sites but is either learned from own scholarly experience or by reading existing articles and thus experiencing the tonality of the work.



Advise for beginners


- **Start with a topic that is familiar to oneself**
- **Resort to Wikidata, Wikimedia Commons and Wikipedia Library**
- Read other people's articles
- Read help sites of Wikipedia
- : Be nice and humble, established Wikipedians can be rude

`Start with a topic you like, a topic you're familiar with. Or write about your town. Something you can do easily.`

: *`You can learn article structures, tones, description styles etc. by reading other articles. Just reading the manual won't help in writing good articles.`*

Recommendations/ Wishes For Oneself



: `I think it's urgent to increase the amount of freely accessible sources, maybe by handing out subscription rights of paid sources, or turn paid sources into free sources upon discussion.`

: `As for the source of information, I think the Foundation should pay more interest in Korean Wikipedia and East Asian countries. The Foundation is in U.S., so they seem to only focus on Britain and the States, making investments there. I feel that Korean Wikipedia has been left out.`

Expand Wikimedia Library database base/ access to more/smaller publishers


esp.  

Also invest more funding of sources in smaller projects, e.g., for building an image database

Make investments to make finding images easier, esp. current affairs


 Improve configuration of Wikipedia translation tool

UI improvements of Wikipedia sites

 Provide a way to be able to use non-referenced information, as not everything is written down in small languages, e.g., adding `note by the editor xy`

KOR: Improve old-fashioned Wikipedia discussion sites

Make Wikidata data sets usable as templates to create new/similar datasets for related topics, e.g., persons

 More research on Wikipedia user behavior (results often used in debated and discussions)

Work on standards for using images on Wikipedia

Implement a meta-search function in Wikimedia Library to speed up database search (like available in [KVK](#))



Overall, the interviewed Wikimedia users are very satisfied with their situation. Many wishes concentrate on improving the situation for others (less experienced Wikipedians, smaller/less developed projects – see next slide). The wishes of authors for their own research are on the availability of sources and usability of tools provided by the foundation.




Support For Others/ Smaller Projects



For beginners:

- Help with:
 - Judging quality of sources
 - Expected citation practices for a specific topic/ content
 - Guide especially non-academics to approach Wikipedia
 - Inform people about sources like Wikipedia library
 - Improve onboarding materials, esp. for English Wikipedia
 -  *'This is perversely the mature and the high bar Wikipedia. It is almost the most difficult to get integrated into.'*
 - : Provide more information on Korean Wikimedia Foundation to attract more new researchers

For smaller language projects:

- More funding of single books/sources for authors; funding of (borrowable) photo equipment
- Improve usage on mobile devices for younger communities that do not use actual PCs and currently use alternative platforms to communicate (WhatsApp, Facebook), contradicting Wiki principles
- Foster literacy skills of emerging communities
 -  *'I have travelled extensively and in emerging communities. (...) They could contribute in English even if they do not have quite native English. (...) What they lack is information literacy skills, they have next to none of that in their K12 education. When you tell them hey, please only rely on reliable sources, they get the idea of what you want, but they don't know how.'*

For both:

- Clarify/explain notability and verifiability standards
 - Revise page on standards and make it coherent and more comprehensible, e.g using alternative ways of transporting knowledge like videos
 - 'This is a generation that wants everything explained in 5-minute videos.'* *'You post something, you are told that it is not notable and sent to a page on notability, which is 20 pages, and you get that you did something wrong, but I don't know what needs fixing or how. Is there a tutorial?'*
 - Centralized steering of such activities by the foundation, as it is a big task/responsibility
- Provide funding For US and poorer language projects: More funding of single books/sources for authors; funding of (borrowable) photo equipment

Observations English




- Some Wikipedians working in the English project were from a different cultural background and non-native English speakers, though at least occasionally working in English
- As the English project is very large with many members, the researchers prefer to contribute to the project of their cultural heritage
- Particularities of the non-English projects that Wikipedians worked in:
 - **Hebrew:**
 - All sources need to be checked when translating from another language (project policy)
 - `Expert groups` exist that can be tagged in an article that needs help
 - **Bengali:**
 - Information can be influenced by the government, esp. on news websites lands on Wikipedia and misinformation need to be balances from outside the country (strong reason for this Wikipedian to work in Bengali)

Observations Germany



- The German project as a very **strong scholarly approach**, all German Wikipedians in the sample were **academics** and working in an academic manner and along **academic standards** was of **greatest importance** to them
- Some (especially very long members) engage in a **strong community** that also organizes offline get-togethers
- **Communication outside** of Wikipedia happens **occasionally**
- Often **organized in thematic groups** with other Wikipedians who work in the **same** or **closely related field**
- German Wikipedians often **compare** their project to the **English project**, if asked for customs in other projects:
 - German project believed to have **less strict duty** to add **references** (fewer, less dense)
 - German project believed to have **stricter rules/higher quality standards** in terms of acceptable **sources**
- **Different custom of citations**, than other language projects, esp. English: less citations, quoting longer text passages

There is a section for weblinks and one for sources. If everything is already in the weblink section, like the person's whole biography, then I don't have to verify every single sentence anymore. It's all there in the weblinks.'

: *The Germans are very... They are very inflexible in a lot of things. They use a different system for references in German. They're much more likely to put a bibliography at the end and quote very long sections. In English they tend to be much shorter; the Germans do that much more, take whole paragraphs. The custom may make academic sense, but it's different.'*

Observations Korea



- Korean Wikipedia seems to be a **very young project**, even the **director** of Korean Wikimedia is still a **university student**
- **Articles** are either **barely existing** or **fully elaborated** when published
- Korean editors often report to **learn** about editing by **looking at articles of others**
- **Some** Korean Wikipedians **focus** more on **translating** and little on writing own articles
 - **English articles** are possibly **corrected/supplemented** during translation with **Korean POVs**, esp. when dealing with cultural/political topics
- **All** Korean respondents have **participated** in **edit-a-thons, offline meetings** where one or more topics are **edited together**. Some editors only join to work on topics independently. Otherwise, all work **individually without cooperation** though having gathered in the same location
 - They **learn the editing skills** at the event and **gain more interest** in Wikipedia
- **Communication outside** of Wikipedia happens **frequently**, often via **Kakaotalk** or **Discord**
- **Which language** to use can be a big topic with **heated debates** in Korean society and therefore in Wikipedia as well. E.g., using honorific or polite way of describing things *‘I’m for using words that are commonly used by Korean people. Some people engaged in emotional conflicts for 10 years for that issue.’*
- The Korean community believe to **have less strict rules, than English Wikipedia** project
- Some Korean editors reported to **refrain from discussions** with others (supposedly to avoid impoliteness)
- [Namuwiki](#) is a **competitive Wiki Platform** in Korea with **less neutrality**, allowing **personal interpretation** of content by authors

‘There are two types in Korean articles, they are either in just one sentence, or they are perfectly in full content. For the latter type, nothing needs to be done. If it’s in one sentence, you need to start from scratch, it takes so much effort.’

Observations Basque Region I/II



- More **diverse topics** are approached by every **individual**, as the **community size** does **not allow** to wait for an **expert** on every topic
- As situated in the **European language space**, authors **trust themselves** to draw upon various **European languages** when **translating** and **looking for sources**. Since the Basque language area is very small, editors will **research in many languages**, even if they are **not capable of understanding** them
- Basque respondents were the **only ones** to report using the **automated Wikipedia translator**
- **Discussions** in Basque Wikipedia are often focused on **translation/ wording** matters in the Basque language
- As it is a very **small project**, **questions** to the community might **remain unanswered**
- **Observations** of **other language projects** by Basque editors:
 - **English:** believed to have the best quality articles and references
 - **French:** often staying within own bibliography
 - **German:** being very strict and having different citation principles (less references, larger paragraphs of quotations)
- Basque editors sometime **struggle** to find **sources** for very **basic facts**, like **buildings** or **landscape features** and **quarrel** with the need to **cite a source** for well-known facts or facts that no source besides e.g., narrations exists. Though it is an issue in the Basque project, authors **refrain** from **using non-backed** or **little credible sources** and keep to Wikipedia guidelines of verifiability

`If I have to trust someone, I usually look at what there is in English. That's perhaps the most reliable.`

`If you go somewhere, you see a landscape with a new building and you can say: «After March 2020, this valley was filled with works, and they built a bridge over that river». Where is there information about that? In your photo you can see they are building a bridge, right? For example. And that is a reference itself. Because you have taken a photo and it has a date, you can see they are building a bridge. (...) In Basque, the information source the person who told me about it. Then, not everything is published, there are things that are not published. They are there, waiting for someone to collect them.`

Observations Basque Region II/II

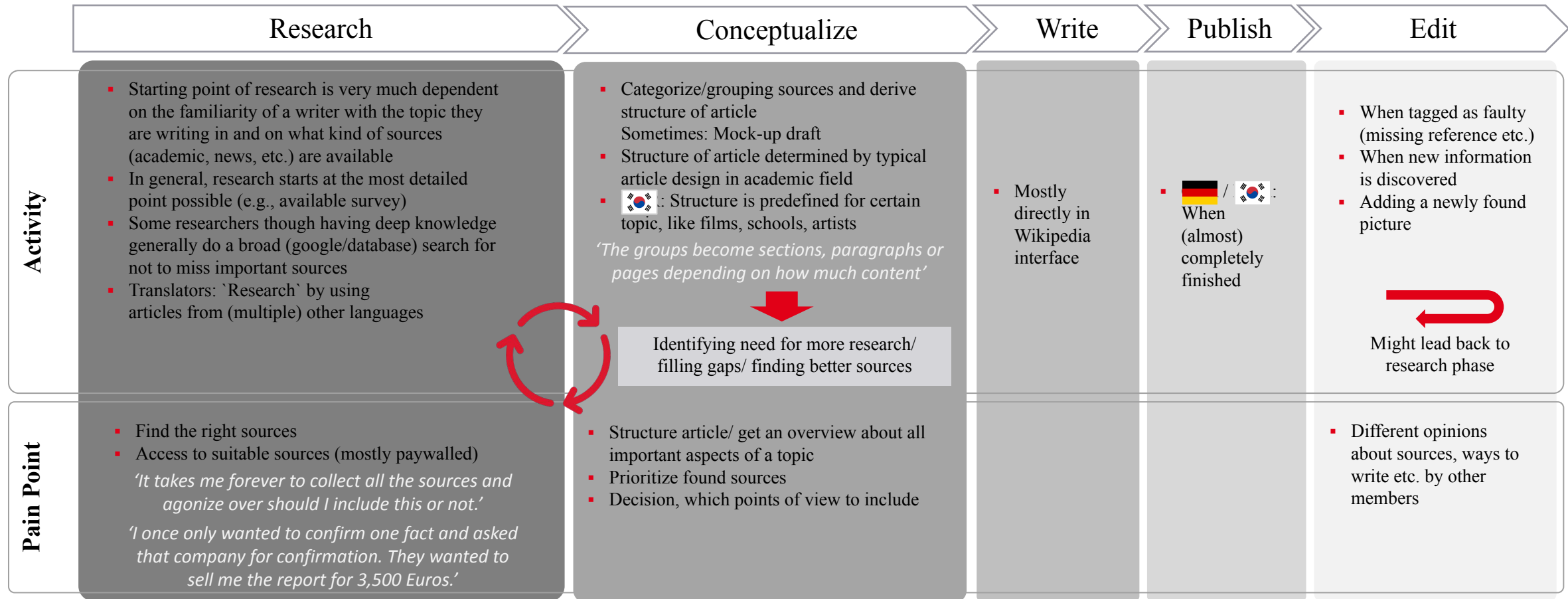


- As the Basque region is **not a sovereign country** itself, the Basque **langue** is used to **amplify the cultural impact** of the region and get a **stronger standing** within the Spanish peninsula.
- The **`Basque government`*** uses Wikipedia as a **promotion tool** for their **language and culture**:

`The Basque Government has a commission for promoting Basque language. It gathers important people from the Basque world and, that committee, some 5 years ago, highlighted that Wikipedia could be an interesting tool for strengthening the Basque language. The Basque Government, when hearing that, made an investment, and defined a 3-year project to develop it. And then, it gave grants, and not a few, of about 100,000 euros, with the following aim: Initially, to straighten up some 1,000 articles so high school students or from 12 to 16 could use them for their schoolwork. Then, they chose 1,000 articles that could be used as a base. And we worked on them a lot for 3 years, to improve them. And it was great improvement.`

*Expression used by a respondent, it officially called the Basque Autonomous Community

Workflow Map – General Workflow Of Wikipedia Researchers



Personas

The Scholar

- **Main Activities**
 - Writing articles
 - Engage in topic-related discussions and providing expert knowledge
- **Research Literacy:**
 - High, acquired by own academic history
 - Very much effort is put into researching and verifying sources and images
- **Domain Knowledge:**
 - Very high in own academic field
 - High in additionally acquired fields, high motivation to thoroughly dig into each topic, great attention to detail
- **Wikipedia experience:**
 - High, 10+ years
- **Other characteristics:**
 - Neutrality is very important: two contradictory (credible) sourced will be included equally and depicted as disagreement in the field
 - Won't accept statements of persons (e.g., own birthplace), unless information can't be found elsewhere; no usage of self-statements

The Teacher

- **Main Activities**
 - Teaching and guiding others in the World of Wikipedia
 - Enable other/small projects to expand their skills
 - Marks/flags contributions of others instead of deleting them to allow others to learn and gain experiences
 - Uses almost exclusively online sources
 - Starting articles for others to work on
 - **Research Literacy:**
 - High
 - Started with editing and expanded their scope of activities for others to benefit
 - **Domain Knowledge:**
 - Very high
 - **Wikipedia experience:**
 - High, 10+ years
 - **Other characteristics:**
 - `Meta`-view on Wikipedia and the meaning and purpose of data and knowledge transfer
 - Teaches and guides by actively helping and communicating with others
- 'For a new editor there are a lot of other things to do, learning to write encyclopedically, neutrally. All you help them with are the sources, they you get them to do the editing.'*

The Pragmatist

- **Main Activities**
 - Writing articles
- **Research Literacy:**
 - Good
- **Domain Knowledge:**
 - Medium/high
 - Not necessarily an expert in the field of writing/ might adopt new topics, that are not in own area of expertise
 - Needs to identify good sources for topics, as he/she is not an expert in the field
- **Wikipedia experience:**
 - Medium/high
- **Other characteristics:**
 - Rather casual demands on sources, e.g., occasionally accepts videos, self-statements, lectures about a topic that is based on a paper as citable sources
 - Might choose for themselves which source is more reliable and include it
 - Any guidance in finding good sources is highly valued

Personas

The community worker

▪ Main Activities

- Writing articles
- Building and engaging in topic-focused networks
- Engaging in topic-related discussions and driving projects within the own topic network

▪ Research Literacy:

- High to very high

▪ Domain Knowledge:

- Very high

▪ Wikipedia experience:



- Medium to high

▪ Other characteristics:

- Working together with others in a topic-centered program
- Close exchange about topics, references, etc.
- Well connected in their field of expertise
- Ambitious with regards to own work (might take part in contests on Wikipedia)
- Draws a lot of motivation out of the communal aspect of Wikipedia and the feeling of “common achievements”

The translator (;)

▪ Main Activities

- Translating articles to quickly build knowledge base of own Wikipedia project
- : eclectically piecing together bits from various languages to form an own article;  mainly translating one foreign-language article at a time

▪ Research Literacy:

- Medium to high
- Rather little to medium effort is put into researching
- Content of pictures are assumed to be correct judging by the credibility of the source or judged by own discretion

▪ Domain Knowledge:

- Low/medium
- Topics are chosen out of interest, translating articles means learning about the topics

▪ Wikipedia experience:

- Medium/ high

▪ Other characteristics:

- Working solitary
- Motivation to push own language project by expanding size

The Improver/ Supervisor

▪ Main Activities

- Adding content to existing articles
- Caring about policies/formalities and actively mentioning them to other members
- Contributing technical solutions like programmed bots/ translation tools etc.
- Working on templates and guidelines to introduce structures and systemization

▪ Research Literacy:

- Very high

▪ Domain Knowledge:

- High to very high

▪ Wikipedia experience:

- Very high, started with writing/translating articles and moved on to higher-level topics as time went by

▪ Other characteristics:

- Rather working in the background, provision of solutions and guidance without actively advising others