

# Wikibase Suite 2023 Survey

Research led by Jon Amar WMDE with expert consultation from Jan Dittrich  
Synthesis and presentation by Keren Ramirez

December 2023



**This is a summary of the 2023 survey  
for Wikibase Suite.**

**00**

# **Executive Summary**

# Executive Summary

## Research Questions

We wanted to understand the use of Wikibase Suite in the community, including **relevant extensions, tools and services**; as well as the size, licenses, and access policies our users have. We explored installation methods as well as **who our users are and how we can support identified needs**. This research was conceptualized as being high level and exploratory, indicating opportunities for follow-on research into more specific investments we could make to enable the community's success.

## Participants

The Suite community survey, conducted during the first quarter of 2023, **collected between 32 and 39 responses per questions**.

The participants came from our existing contact points within the community: partnerships, the Wikibase Stakeholder Group and User Group, the community Telegram channel, mailing lists, and word of mouth.

**01**

# **The human story of community members**

**There's a huge gender gap in  
the Suite community. People  
identifying as men are  
massively over-represented**

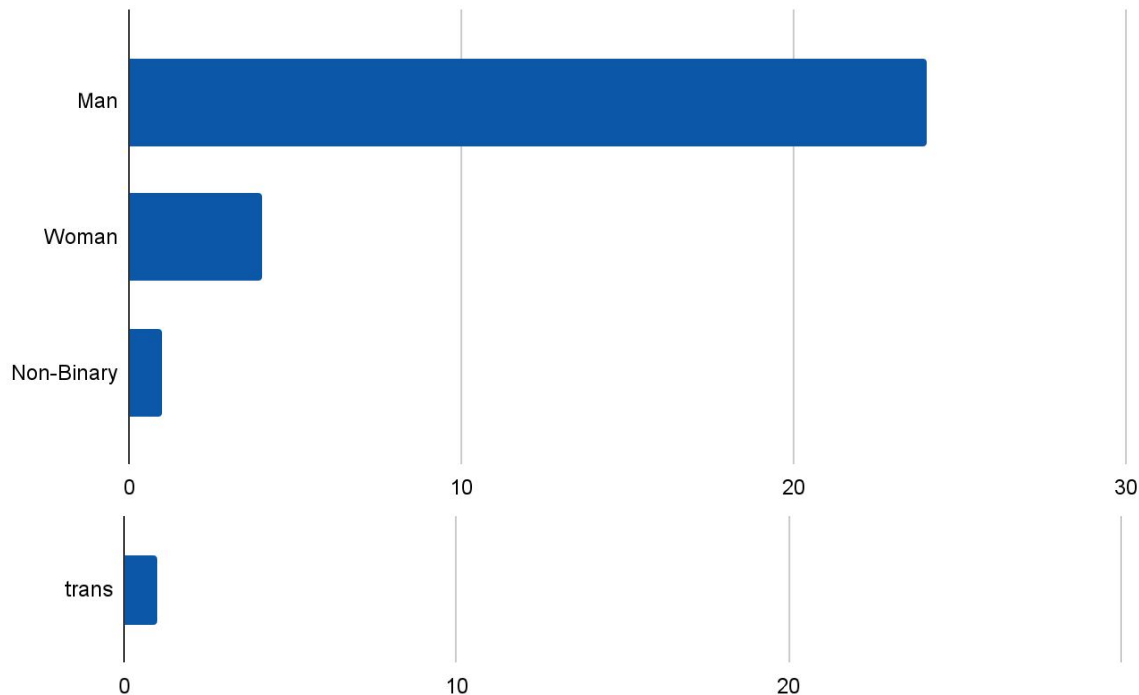
# Gender

## A big gap

The survey findings highlight a significant **gender disparity within the community, with a notable over-representation of men**—six times more people identify as men than women. This gender gap signals a considerable distance from achieving our inclusivity goals, considering also that only one participant identifies themselves as non-binary and only one as trans.

## What is your gender?

multiple answers possible



# Geographical location

## A concentration

The geographical distribution of respondents reveals a concentration with **84% based in the European Union or the United States.**





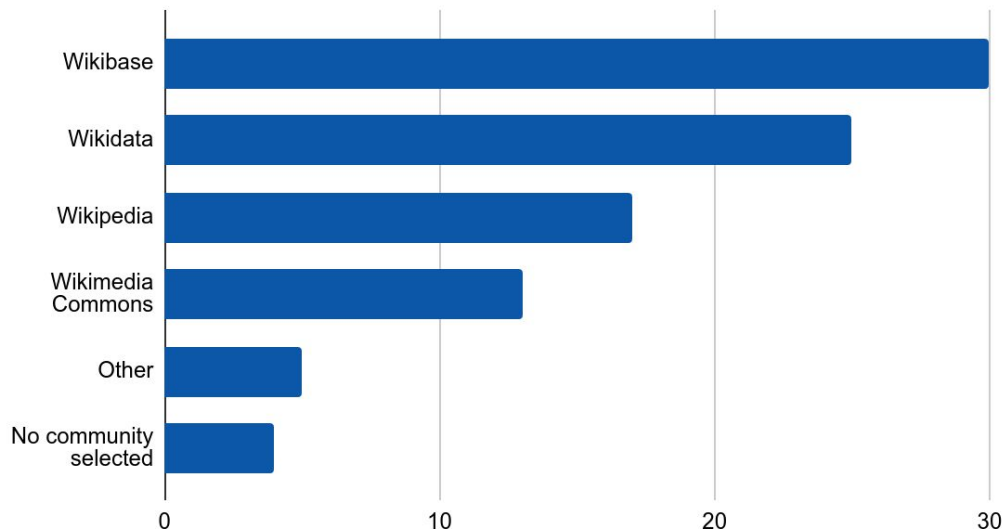
# Community involvement

## Communities

A majority are part of at least one community in the wiki-verse, though most identify as being a part of the Wikibase community itself.

Four people skipped the question entirely.

Which of the following communities do you identify as being a part of?

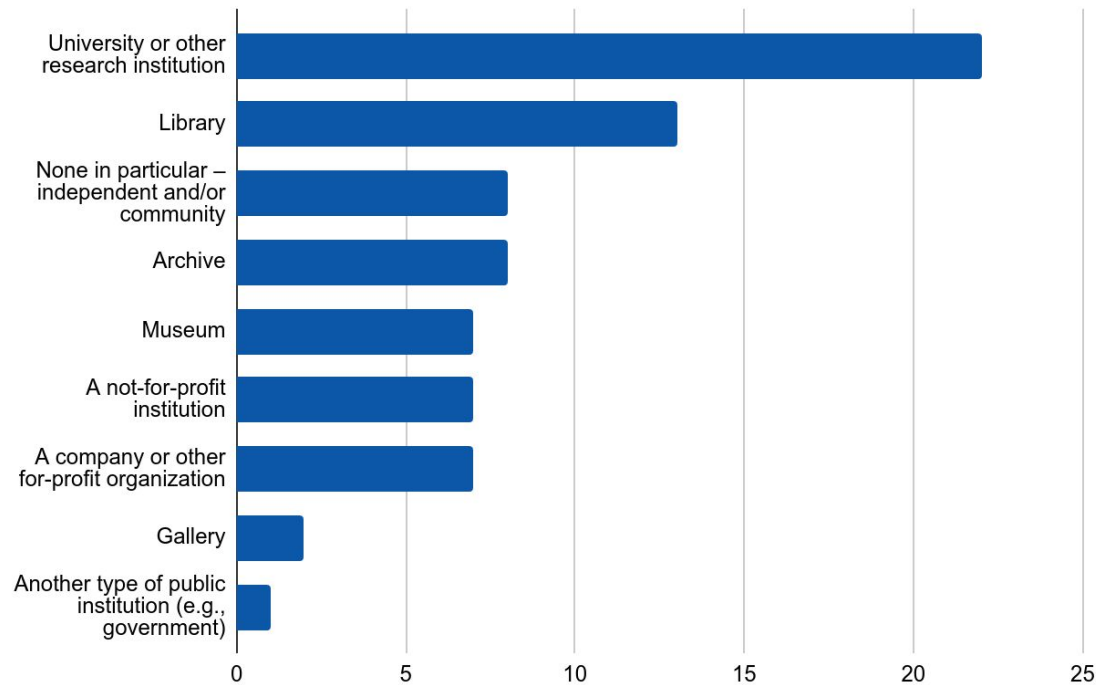


# Organizations using Wikibase

- GLAM institutions, frequently lumped together, see a wide spread in usage numbers. Libraries alone, for example, make up nearly as much as the rest combined (13 vs 17).
- Not-for-profit and public institutions make up the vast majority of respondents whereas for-profit represents only 9% of overall use cases. This has important implications for funding cycles, the longevity of projects, and the availability of resources to individual projects.

In what sort of organization do you primarily use Wikibase?

multiple answers possible

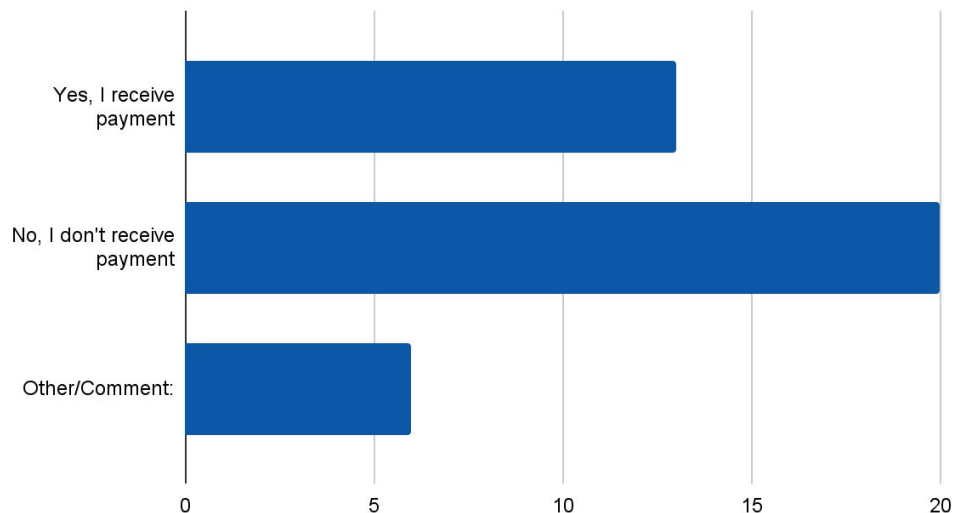


# Resources are a challenge

Organizations are constrained on resources across two key dimensions:

- Small teams: 75% of instances have 1-5 people working on it total
- Low ratio of paid workers: 61% of respondents don't get paid for their WB work (despite 75% of respondents being advanced to expert users)

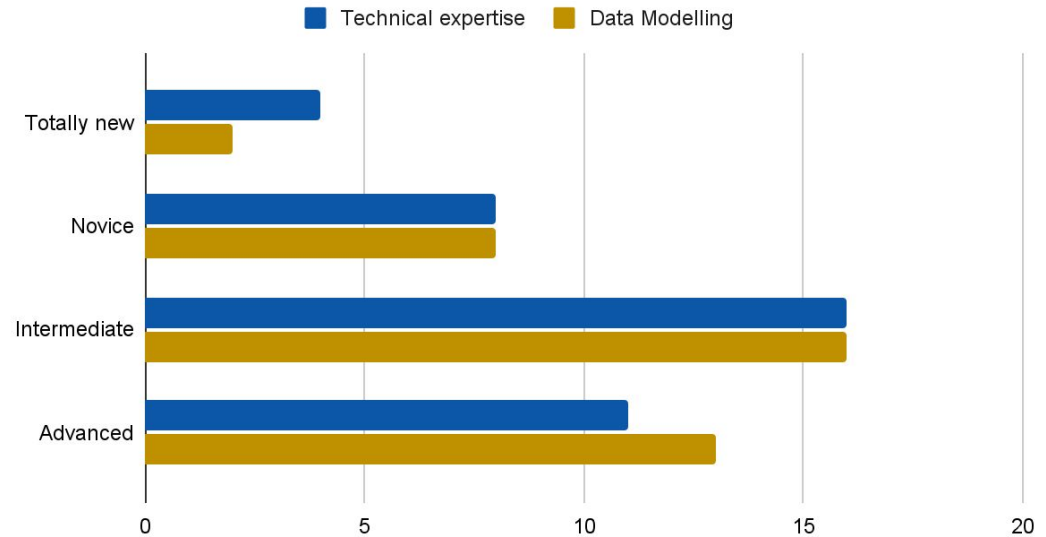
Do you receive payment for your work with Wikibase?



# Community involvement and expertise

- Despite the large percentage of unpaid work, the Suite community nonetheless trends towards experts.
- Despite the meaningful separation in skill sets required to support the Wikibase technical infrastructure vs modelling data, we see a strong trend towards people learning both skills together. This likely reflects the value and necessity of generalists in small teams.

Wikibase technical expertise and Data Modelling expertise



Out of total 32 participants.

# 02

# The basics of using Wikibase Suite

# Installation and maintenance

Among the installation/update methods, Docker containers are the most frequently *tried* (28 in total) method while Cloud is most frequently *preferred*.

## Manual install methods

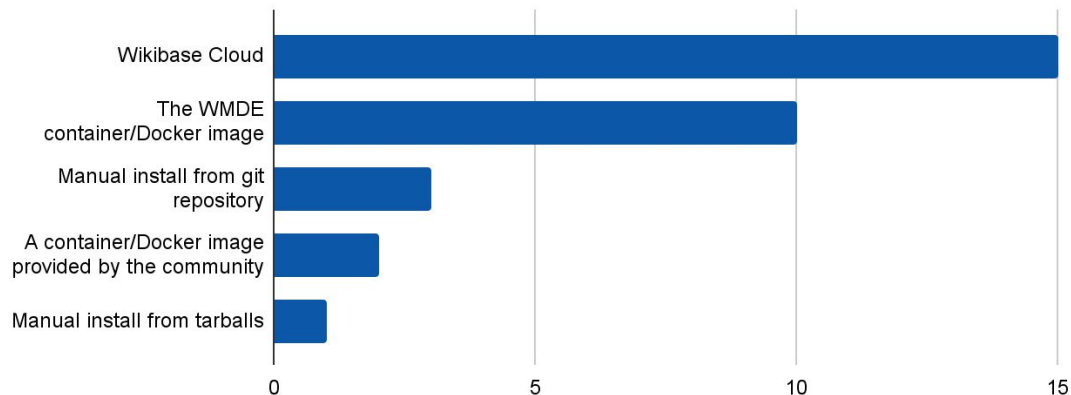
METHOD	TRIED	PREFERRED
WMDE Docker	19	10
Docker provided by the community	9	3
Manual install from git repository	6	2
Manual install from Tarballs	4	1

# Installation methods users preferred

When asking about their preferred method to install Wikibase, we got confirmation that we are betting on the right track:

- Wikibase Cloud wins with 50% (15 respondents). Confirming that the easiest install is no install.
- Docker got 3 times as many as manual installs (from git and tarballs) (12 for Docker vs 4 for manual).
- The WMDE docker got 5 times as many as community Docker images (10 vs 2).
- Higher technical expertise levels were seen in those using self-install methods, with 4/5 self describing as intermediate or above.

Which are your preferred method(s) of installing and updating Wikibase?



**03**

# **Wikibase Suite in practice**

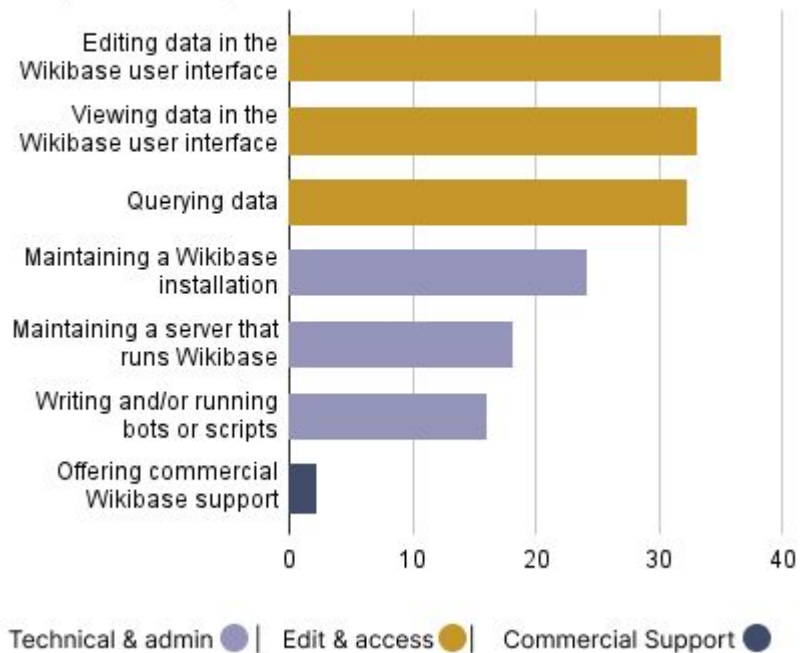


# Ways Wikibase is used

Editing & accessing data accounts for 63% of use-cases. Whereas technical & administrative tasks account for only 36%.

## In what ways do you use Wikibase?

multiple answers possible

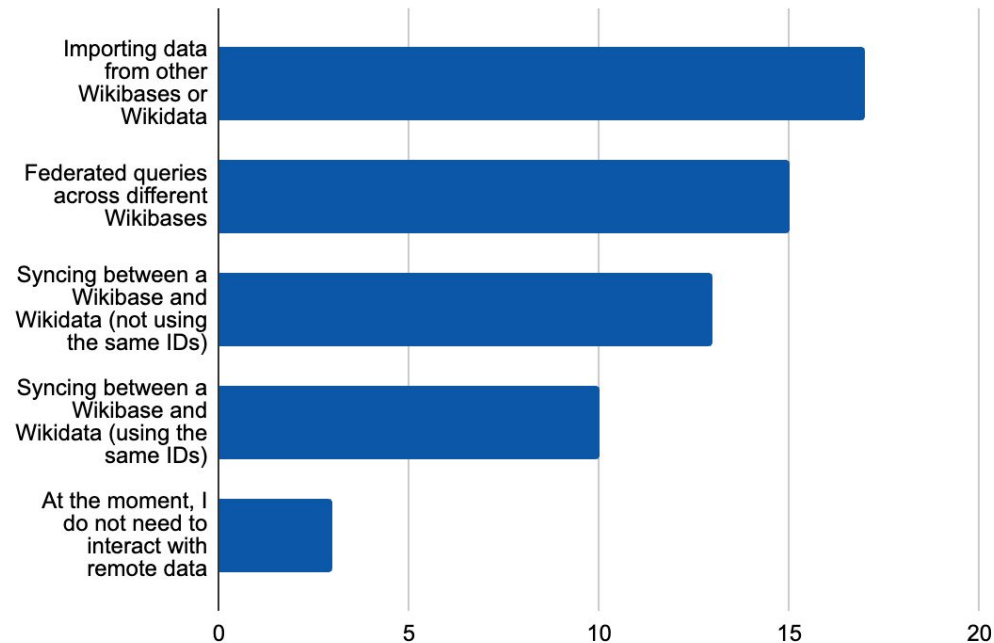


# The importance of remote data

- Fewer than 1 in 10 respondents had no need for interacting with remote data. This is a strong nod towards the persistent appeal and value of Wikibase as Linked Open Data software (as opposed to, for example, solely for semantic structuring of data or as collaborative software for community databases).
- The two use-cases involving syncing data with Wikidata make up a combined 23 selections.

What are the most important ways in which you need your Wikibase to interact with remote data, if any?

Multiple answers possible

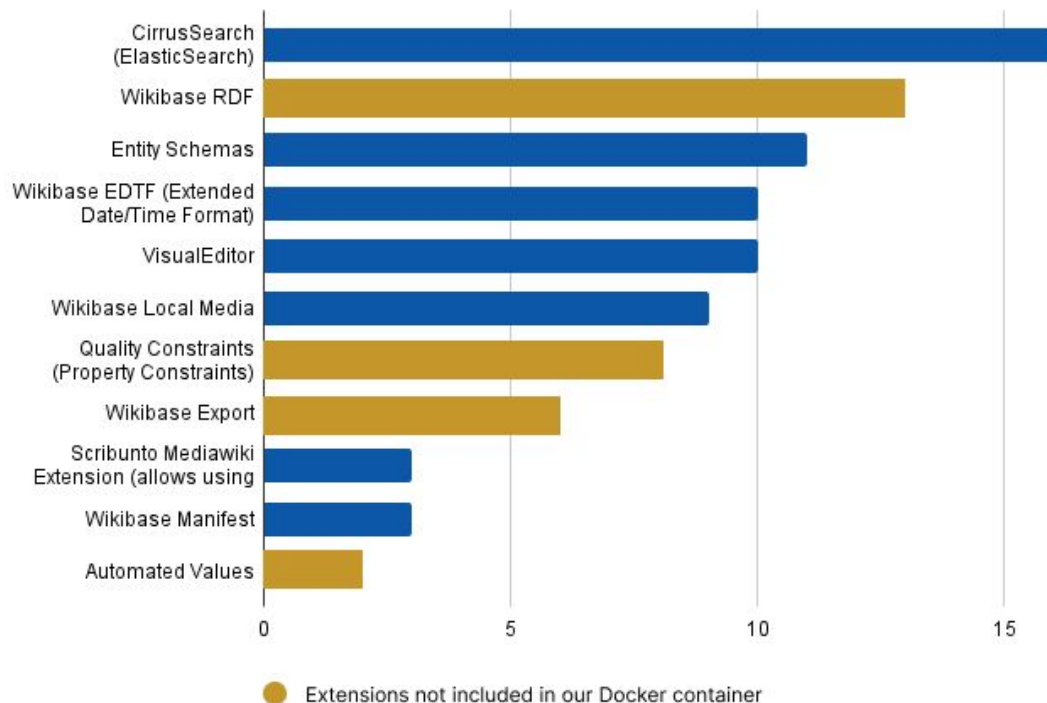


# Important Extensions in Wikibase Use

- By far the most important extension missing from the Suite Docker container is Wikibase RDF at 41% of respondents.
- For longer tail extension uses cases which are not included in our Docker distribution the opportunity here is rather to improve the ease of including additional extensions through easier configuration and upgrade pathways.
- In terms of tools which contribute to overall data quality it's noteworthy that Entity Schemas rank significantly higher than Quality Constraints

## Which extensions are important in your use of Wikibase?

Multiple answers

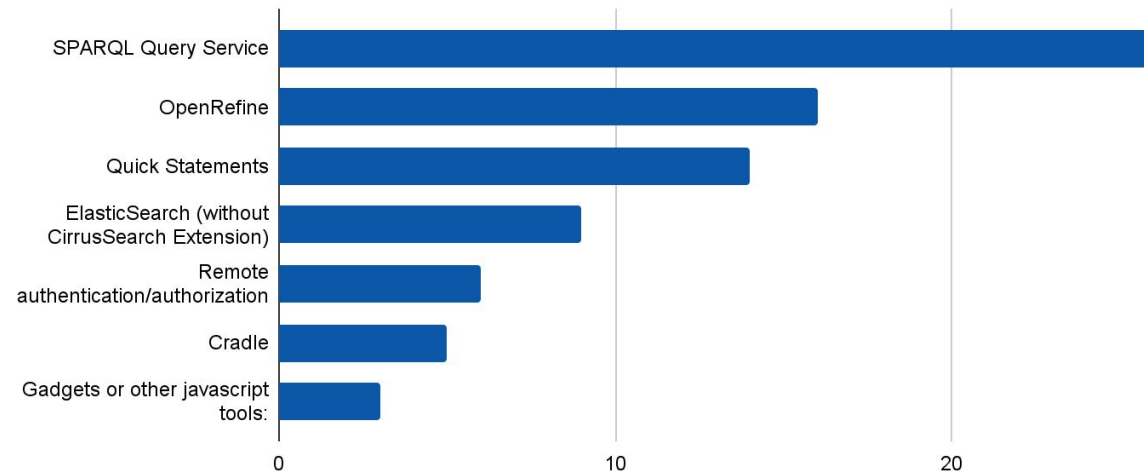


# Which tools or services are important in your use of Wikibase?

One interesting note is that Gadgets, an essential toolbox for the Wikidata community, are relatively unimportant within the Wikibase community.

## Which tools or services are important in your use of Wikibase?

Multiple answers possible



# 04

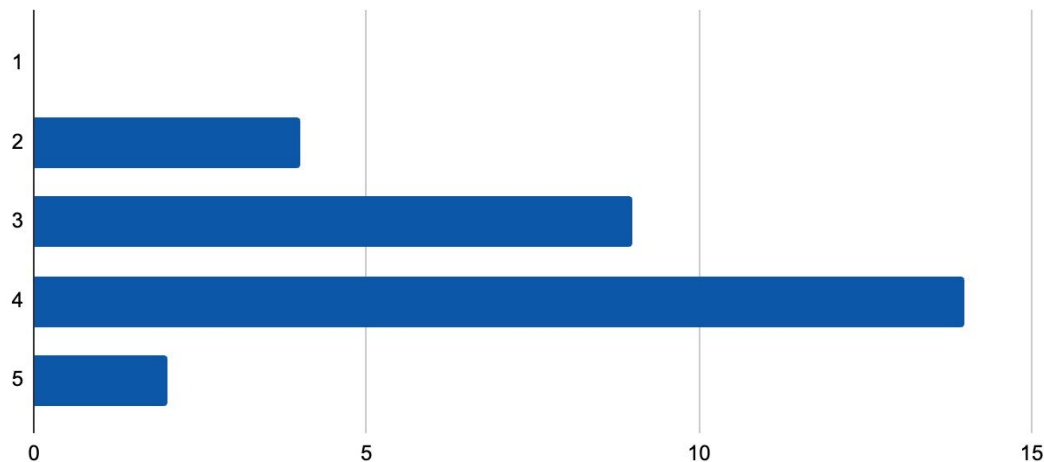
# Challenges and obstacles of using Wikibase Suite

# Meeting objectives

This chart looks good but there's certainly a lot of headroom for improvement: our average score is 3.5 with a median point of 4.

Do you feel like you're able to currently meet your objectives using Wikibase?

Where 1 = "Not at all" and 5="completely"

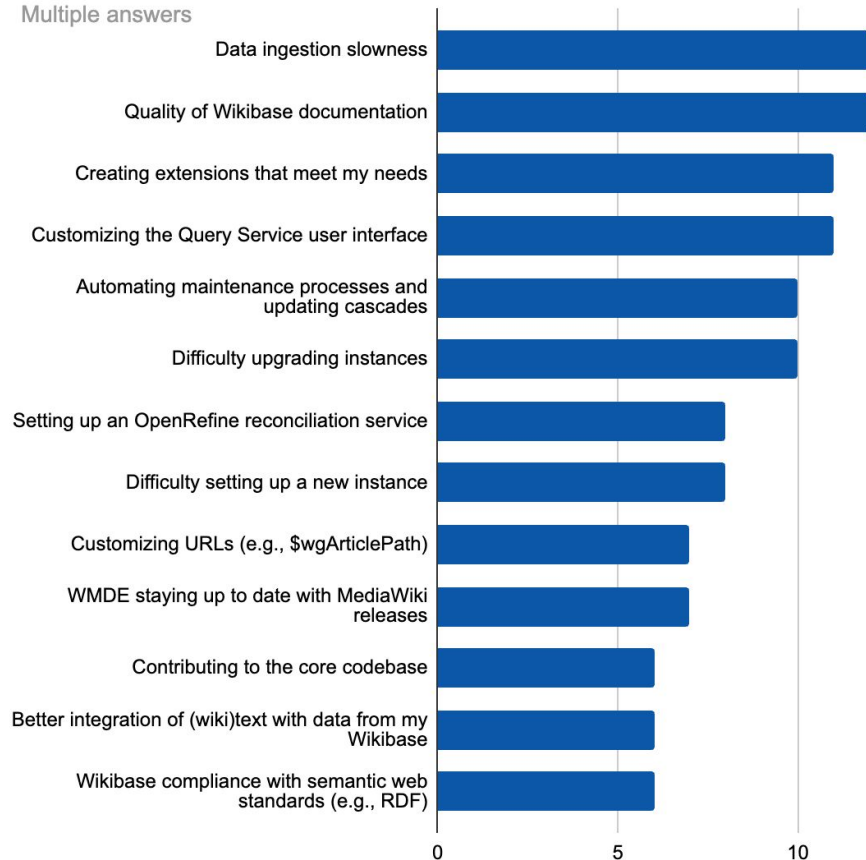


# Main Challenges

- Data ingestion speed and documentation are the biggest challenges participants face.
- Although most participants interact with remote data in some way, their use-cases vary widely.
- Tied for 3rd place is, “Creating extensions that meet my needs”. This is a multiplier. If we get this right we could empower the community to solve their problems in ways which are useful for others in turn.

What are the biggest obstacles to using Wikibase the way you want?

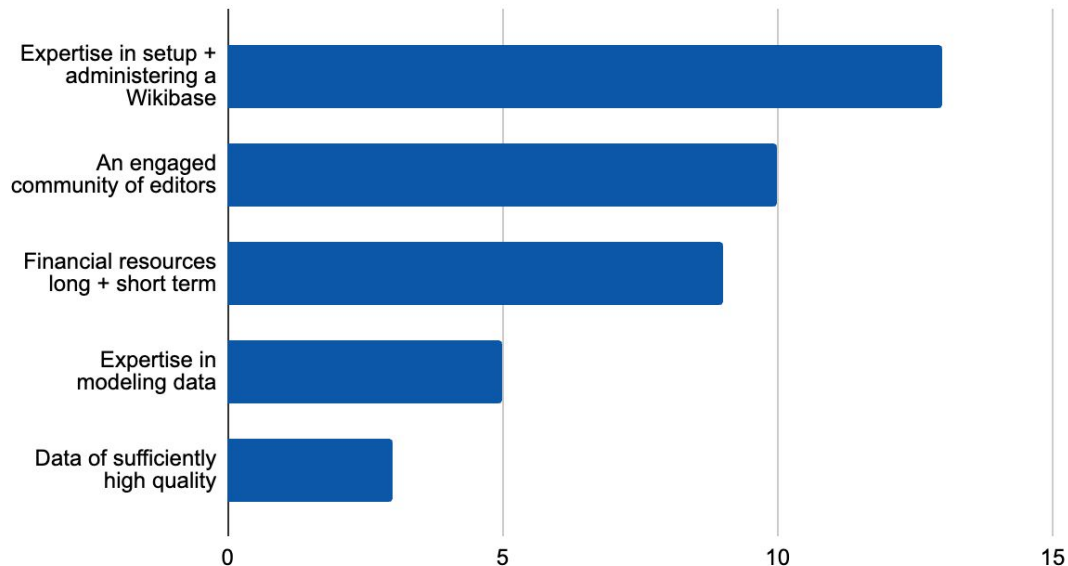
Multiple answers



# Lack of resources

- The challenges associated with setting up and administering a Wikibase highlight the value of investing in lowering the bar to entry in the “first mile” of the user’s experience
- Seeing “an engaged community of editors” prioritised so highly underscores the importance of Wikibases’ social features to our users.

What resources are you lacking in your work with Wikibase?





# Our users' instances

## 1. Location

Most of the instances are in the same location as their participants with 48% of respondents. Only 28% are in a different location and 25% commented that they are not a base with location specific or not particular connection with their own location.

## 2. Topics

There are a wide variety of topics in the respondents' knowledge bases. From art, to biodiversity, culture, government data, music, enslavement & historical data, science & engineering, and more.

## 3. Access and License

Exclusively open licensed: 64% (14)  
Hybrid or closed licenses: 36% (8)

Even some users that mentioned that even when everyone can access and only authenticated users can edit, they are experimenting with fully closed instances.

**A big thank you to the  
community members that  
helped completing this survey.**

Wikibase Suite Team



**WIKIMEDIA**  
DEUTSCHLAND