

Wikidata: Status and Future Plans

Oct 2023

Alan Ang
Partner Manager (Wikidata)



WIKIMEDIA
DEUTSCHLAND

Session outline...

- About Wikimedia Deutschland
- Wikidata Strategy and Roadmap
- Wikidata Updates
- Wikibase for Research



...and a few words about the presenter



[Alan Ang](#), Partner Manager (Wikidata)

- responsible for connecting with global institutions and organizations, and onboarding them to Wikidata and Wikibase.
- point of contact and information exchange for Partners, WMDE's software department, and Wiki Communities.
- love working with socially-minded people and organizations, especially on issues relating to opening, diversifying, decolonizing and democratizing knowledge.

Wikimedia Deutschland



We are a **non-profit** organization based in **Berlin**. The **German chapter** of the global **Wikimedia Movement**.

We work to provide more **people** with more **access** to more **knowledge**.

Software Development

Wikimedia Deutschland aims to give **everyone free access to more knowledge** by delivering innovative and useful open-source software.

We develop and maintain **Wikibase** and **Wikidata**, that are run on the servers of the Wikimedia Foundation.

To ensure we build the right products, **users must play a central role** in our product development processes. We adopt a **community-centered development approach** where our community of users informs product development, and us building transparency around our work.



Our Linked Open Data vision

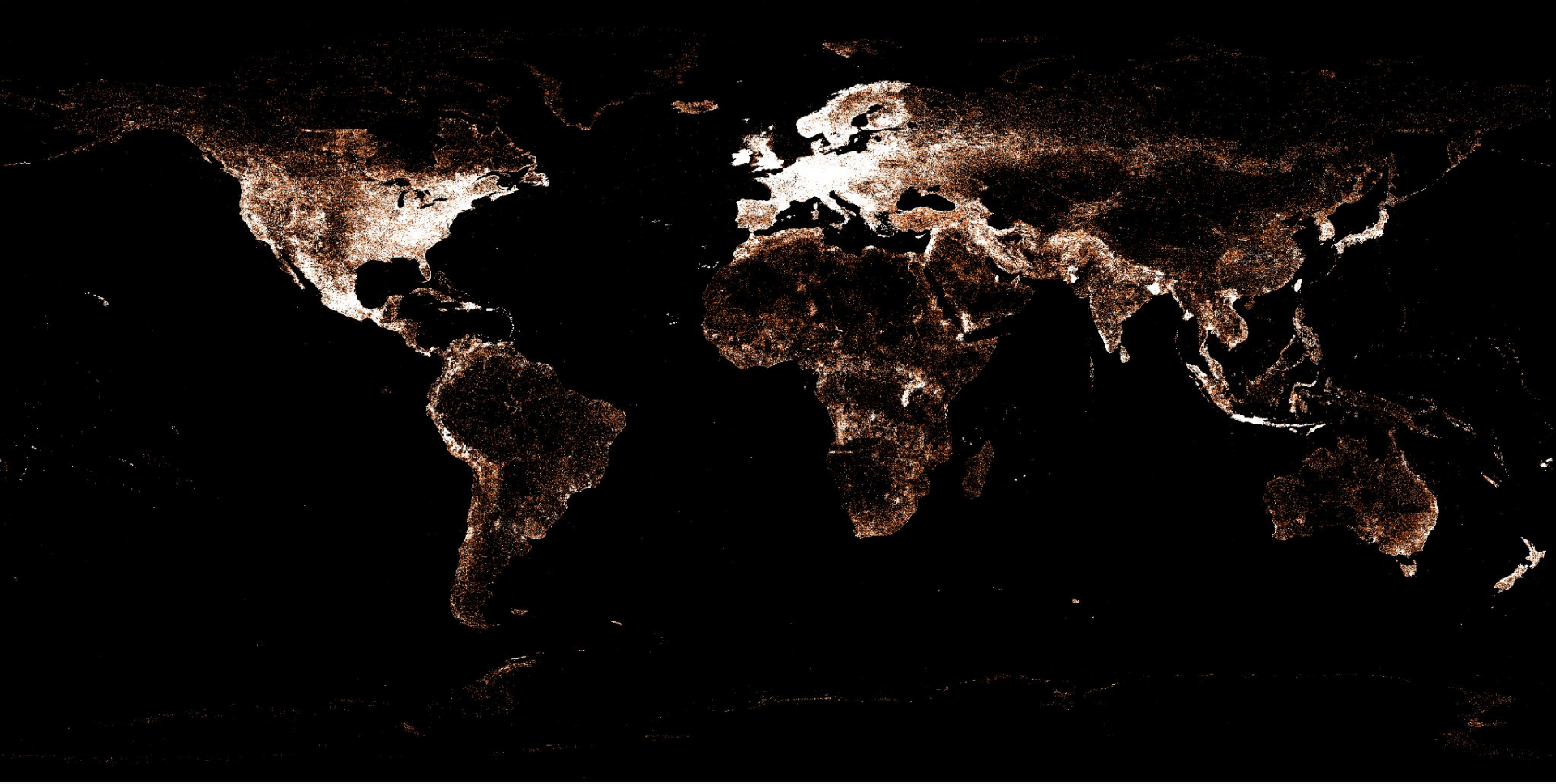
“We work towards a future in which people share the power to collect and organize the data that shapes humanity's understanding of the world.

*Diverse communities around the world participate in Wikidata and in a network of specialized Wikibases, co-creating an **open and free** global knowledge graph in a thriving **Linked Open Data** web.*

This is the basis upon which people, companies and institutions of all sizes can generate new insights, build new apps and services, and change the world for the better.”



Source: [Wikimedia Linked Open Data Strategy 2021](#)



100+ million data Items

Wikidata Strategy & Roadmap



WIKIMEDIA
DEUTSCHLAND

Wikidata Strategy & Roadmap

1

Empower the community to increase data quality

We must ensure that our socio-technical system helps editors increase the quality of Wikidata's existing data and contribute new high-quality data.

- [Feedback loops with data re-users](#)
- [Entity Schema build-out and integration \(development\)](#)
- [Wikidata Data Modelling Days 2023](#)

2

Facilitate equity in decision-making

We want to ensure that diverse perspectives come into play as fundamental decisions are made for and about Wikidata. By doing so, we can better support knowledge often underrepresented in knowledge graphs.

- [WikidataCon 2023](#) with Wikimedia Taiwan

3

Increase re-use for increased impact

We want to make sure that anyone can use the data in Wikidata to make the world a better place. While everyone can re-use our data, we give priority to organizations and projects that align with our values and have a high impact.

- [First version of REST API released](#). Continue to build out first version based on feedback.
- [Build out with focus on operations on Properties](#).

Wikidata Strategy & Roadmap

4

Strengthen underrepresented languages

More people need access to knowledge and technology presented in their own language, and content in that language should be accessible to all. Language data is a fundamental building block in reaching that goal.

- [Lexicographical Data UI improvements](#)
- [Wikidata Software Collaboration with Wikimedia Indonesia and Igbo Wikimedians User Group](#)

5

Enable Wikimedia Projects to share their Workload

Wikimedia Projects should be able to rely on Wikidata to share their workload across language and project family boundaries.

- [Hiring and onboard initial team members](#)
- [Getting closer to the community through a series of conversations to understand where we can help to improve their workflow.](#)

Wikidata Updates



WIKIMEDIA
DEUTSCHLAND

Q3

EntitySchemas:

- We have made EntitySchemas more legible by displaying them by the label instead of their ID (e.g. 'Human' instead of 'E10')
- Also to increase legibility we have added a language fallback for EntitySchemas

WiKit to Codex Migration:

- We have planned and begun our migration from WiKit to Codex to align with Wikimedia front end standards

Mismatch Finder:

- We began working on expanding the types of mismatches that can be uploaded

Update to the Item UI:

- We have begun alignment on assumptions and priorities for Item UI improvements

Language fallback (MUL):

- Make improved fallback with default values for Labels ready for community testing

Wikidata Analytics:

- Deep dive into the size of WDQS sub-graphs to support the workaround Blazegraph
- Wikidata REST API analysis and reporting first steps



Q4

WikidataCon and Data Modeling Days:

- Bring the community together for Wikidata's 11th birthday online and in Taiwan
- Host an online event for the Wikidata community to discuss data modelling

EntitySchemas:

- Update the EntitySchema Termbox to reflect the same mobile termbox UI of Items and Properties

WikiKit to Codex migration:

- Continue aligning with Wikimedia front end standards by migrating Mismatch Finder to Codex

Update to the Item UI:

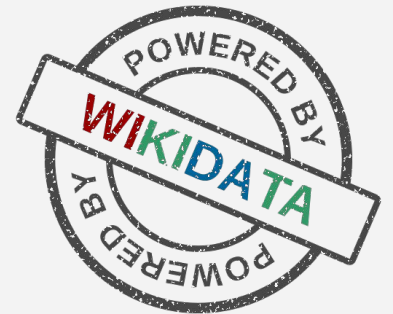
- Continue the discovery phase preparations by detecting risks and defining mitigation strategies. Start aligning with other LOD teams to review goals and assumptions

Language fallback (MUL):

- Community testing before proper release

Wikidata Analytics:

- Learning and rolling out Airflow data pipelines for reporting and regular data aggregation
- Preparations for analytics support around the Item UI



Wikibase Use cases and collaborations

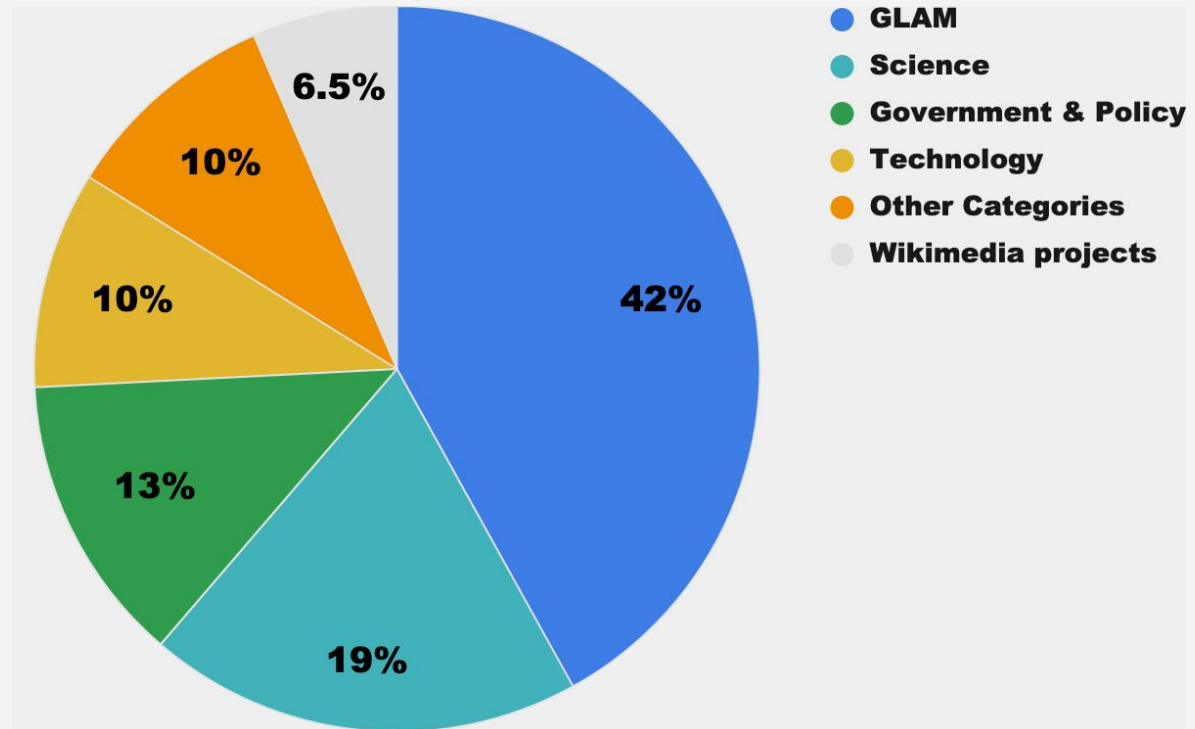


WIKIMEDIA
DEUTSCHLAND

Wikibases at a glance

The active Wikibases instances that we currently track span a number of disciplines.

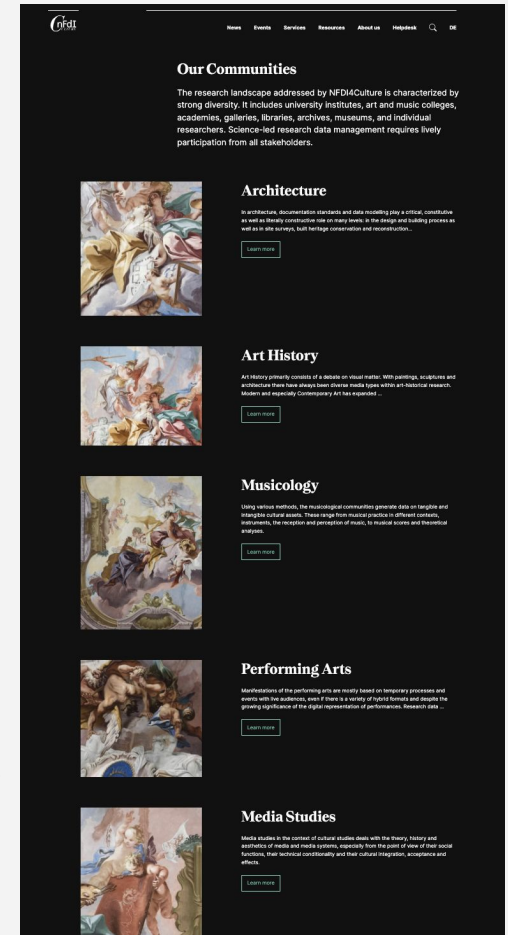
This pie chart represents the % of Wikibase instances in each category, not the % of data they contribute to the decentralized knowledge graph.



*GLAM stands for galleries, libraries, archives, and museums

Wikibase4Research and NFDI4Culture: Multimodal Cultural Heritage

- NFDI4Culture is a German consortium for Research Data on Material and Immaterial Cultural Heritage
- Open Science Lab at TIB, Hannover is a participant in the consortium leading the development of Wikibase as a software service for research data
- The main development has been the orchestration of Wikibase within a toolchain that includes a media viewer for 2D, 3D, and AV media and annotation – Semantic Kompakkt



The screenshot shows the 'Our Communities' page on the NFDI4Culture website. The page features a dark background with a grid of five community cards. Each card includes a small image, a title, a brief description, and a 'Learn more' button. The communities listed are Architecture, Art History, Musicology, Performing Arts, and Media Studies. The website header includes the NFDI4Culture logo and navigation links for Home, Events, Services, Resources, About us, and Helpdesk.

Our Communities

The research landscape addressed by NFDI4Culture is characterized by strong diversity. It includes university institutes, art and music colleges, academies, galleries, libraries, archives, museums, and individual researchers. Science-led research data management requires lively participation from all stakeholders.

Architecture

In architecture, documentation standards and data recording play a critical, constitutive as well as timely constructive role on many levels. In the design and building process as well as in the survey, built heritage conservation and reconstruction.

[Learn more](#)

Art History

All history presents complex 4D objects or virtual models: With paintings, sculptures and architecture there have always been diverse media types with art-historical research. Modern and especially Contemporary Art has expanded...

[Learn more](#)

Musicology

Using various methods, the musicological communities generate data on tangible and intangible cultural assets. These range from musical practice to cultural context, instruments, the reception and perception of music, to musical scores and theoretical analyses.

[Learn more](#)

Performing Arts

Manifestations of the performing arts are mostly based on temporary processes and events with live audiences, even if there is a variety of hybrid formats and despite the growing significance of their digital representation of performance. Research area.

[Learn more](#)

Media Studies

Media studies in the context of cultural studies deals with the history, theory and aesthetics of media and media systems, especially from the point of view of their social functions, their technical conditionality and their cultural integration, acceptance and effects.

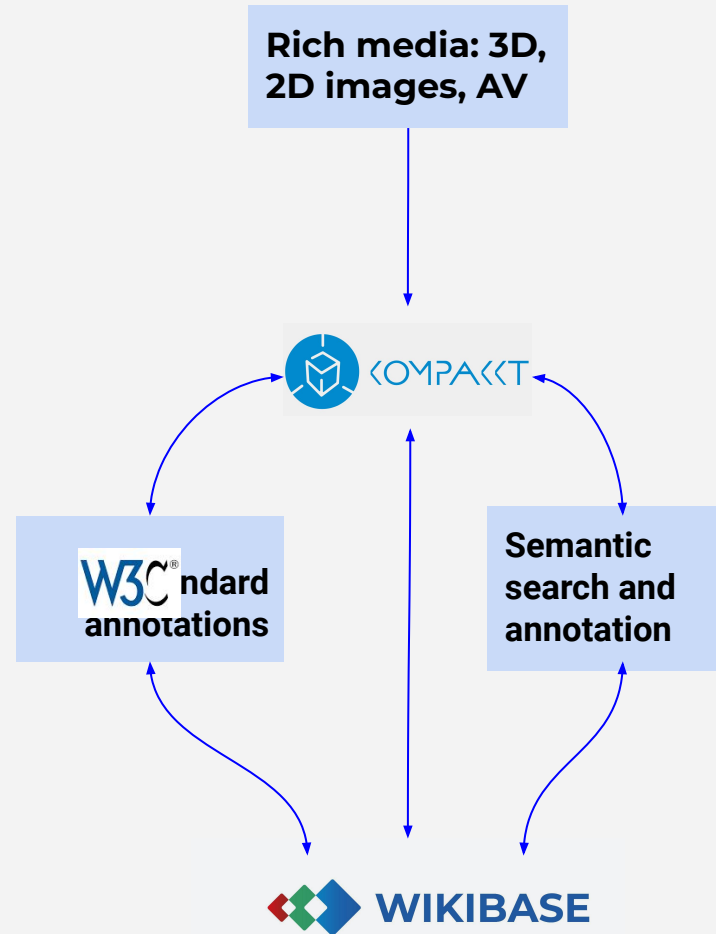
[Learn more](#)

Why use Wikibase?

According to NFDI4Culture

- Open-source software with a large international community
- Allows to LOD-ify existing research data to take advantage of semantic capabilities in ongoing work with natural language text annotations, as well as media files
- Lower learning curve compared to other existing KG tools
- Not a single tool, but an “umbrella” of services

See video [here](#)



Thank you!



WIKIMEDIA
DEUTSCHLAND



Reach out to us!
alan.ang@wikimedia.de
swe_partnerships@wikimedia.de



WIKIMEDIA
DEUTSCHLAND