Decolonising Wikidata

WHY DOES KNOWLEDGE JUSTICE MATTER FOR STRUCTURED DATA?

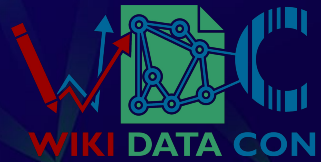
Anasuya Sengupta
Co-founder and co-director,
Whose Knowledge?

Whose Knowledge?

WIKI DATA CON
2021 - a sustainable future for Wikidata

Parabéns, Wikidata e WikiMovimento Brasil!

Knowledge Infrastructures



2021 - a sustainable
future for Wikidata

Who is online?

- **60%** of the world is now online
- **75%** from the Global South (Asia and the Pacific Islands, Africa, Latin America and the Caribbean, the Middle East)
- **48%** of women are online; men are 21% more likely to be online (on avg.)

Who builds our network infrastructures?

Gender & Race...

- **5%** of leadership positions in the technology sector are held by women
- USA: women hold only **25%** of computing roles; of that 25%, only 5% are Asian women, 3% are Black, and 1% are Hispanic
- Only **6%** of technical workers at Apple are Black (even fewer at Facebook and Google)

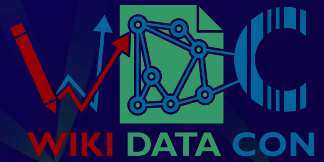
Who builds our network infrastructures?

Geography & Gender…

- Vast **majority** of documents published by the IETF originate in North America and Europe
- At the last in-person IETF meeting, only **6** attendees from Africa (out of 1,108)
- At the 2019 ICANN meeting, only **27%** of participants were women

The internet does not look or sound like many of us...

Knowledge Injustices



2021 - a sustainable
future for Wikidata

Whose Knowledge?

- **~130** million books in 480 languages
- There are **7000+** languages in the world
- **~7%** of the world's knowledge via language is in books

Whose Knowledge?

- Most online content (websites): English (59.3%) Russian (8.3%). Most users: English (25.9%), Chinese (19.4%).
- Most scholarly (incl. digital) publications in English
- 80% of scientific journals, 90% of social science journals on Scopus and JSTOR in English

Whose Knowledge in the UK?

- **140** Black professors (0.7% from 21,000; 18,000 white, 1360 Asian, 2000 other)
- **68%** of all professors are white men (>14,000), **0.1%** are Black women (**25**)
- **1.2%** of 19,868 UKRI studentships for Black or Black Mixed students (Black students make up 6.6% of further ed students)
- **75** Black governors on University boards (out of 3600)

Whose Knowledge?

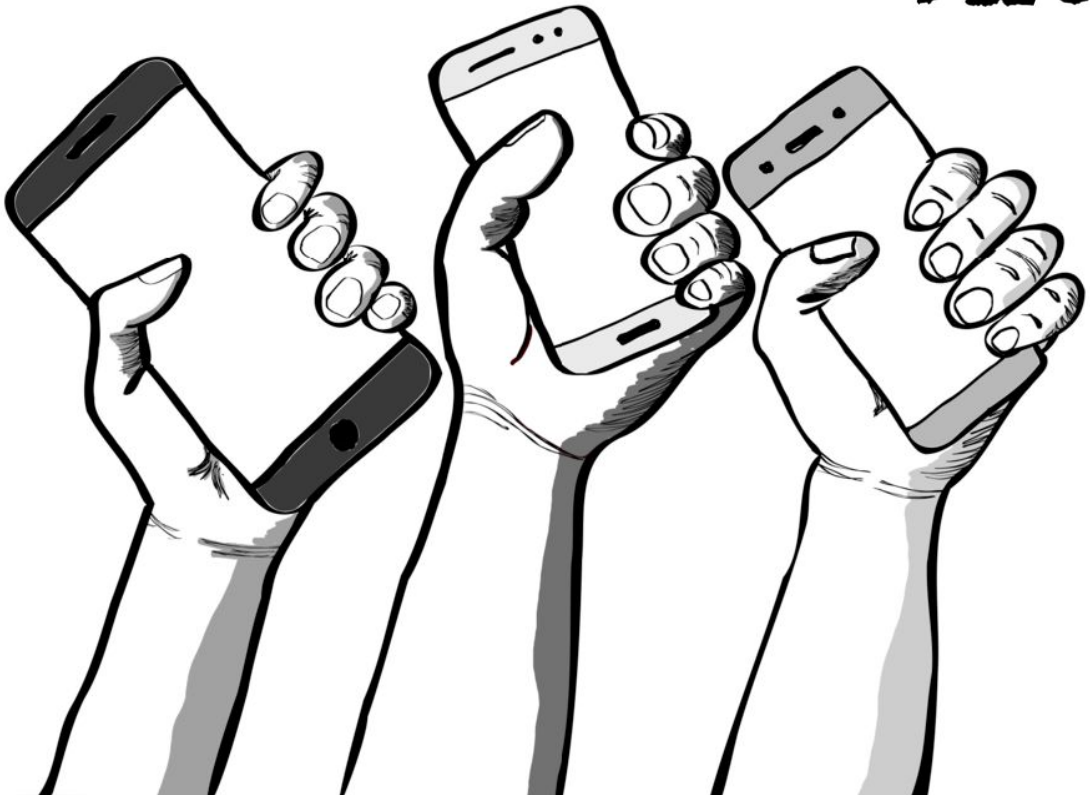
PUBLISHED AND ONLINE KNOWLEDGE,

SCHOLARSHIP AND SCHOLARS...

“RELIABLE SOURCES”

DO NOT READ, LOOK OR SOUND LIKE THE MAJORITY OF THE WORLD

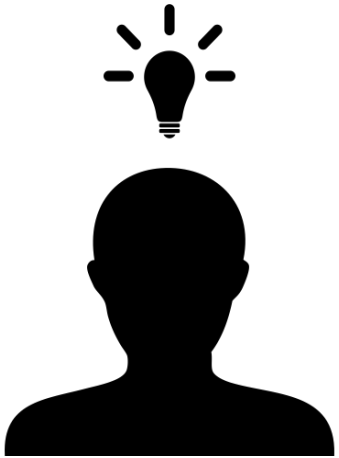
DEcolonize theINTERNET

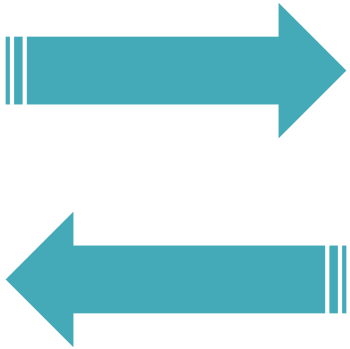


By Alamcsd at wts wikivoyage (Own work), Public domain, via Wikimedia Commons

Embodied (tacit) knowledge

Disembodied (formal) knowledge



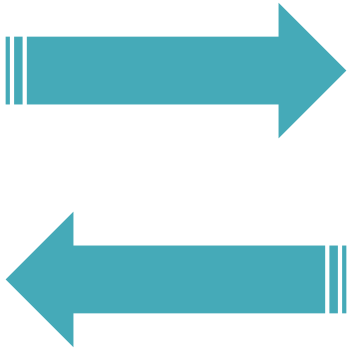


People's
authority

Institutional
authority

Images from Noun Project by Wilson Joseph and Pravin Unagar

Epistemic Injustice



Testimonial

Hermeneutical

Black Lives Matter logo PD, via Wikimedia Commons

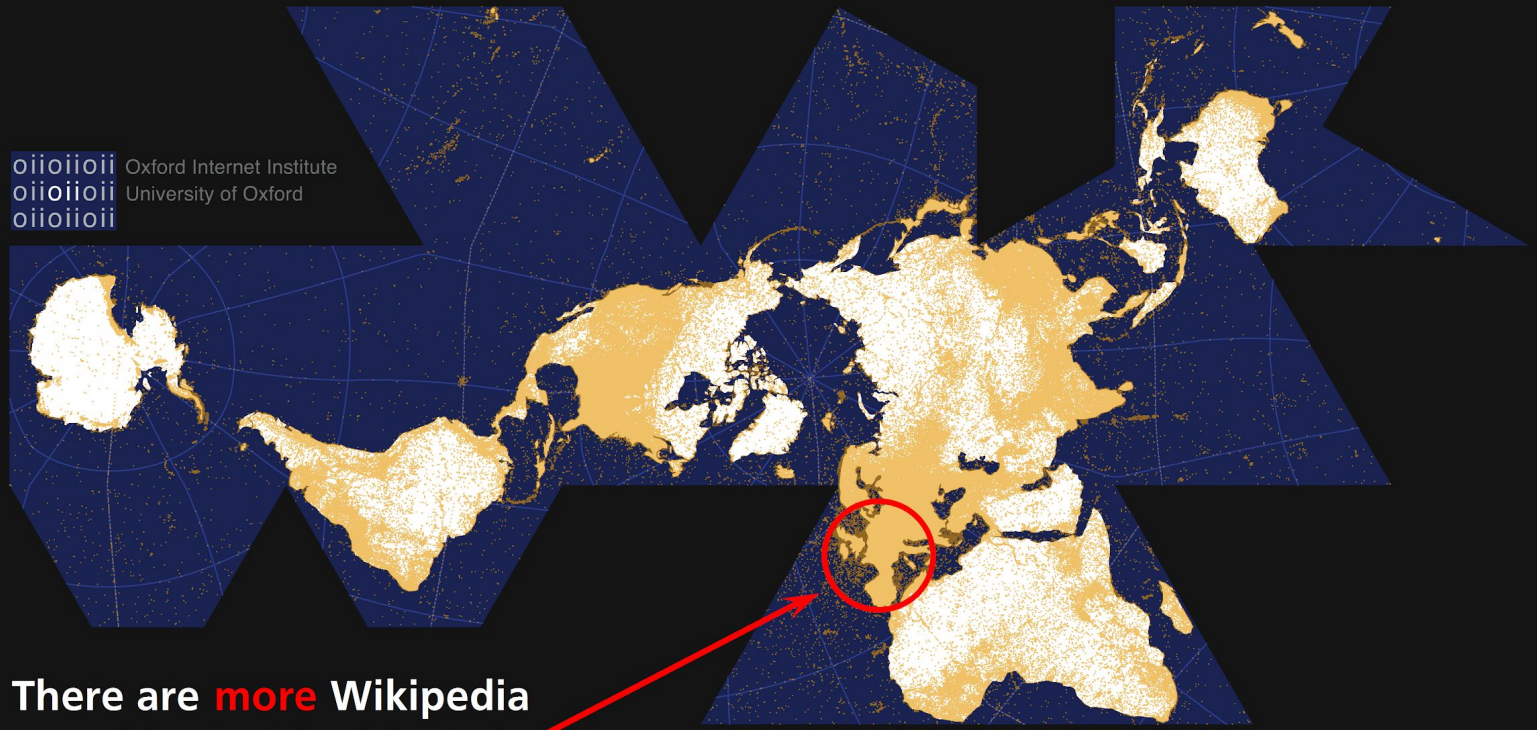
By Marc Nozell CC BY 2.0, via Wikimedia Commons

Using Wikip/media as proxy...

- Only 10% of Wikipedia editors are women or non-binary
- Only around 20% of public knowledge is produced on or by people from the Global South

The Geographically Uneven Coverage of Wikipedia

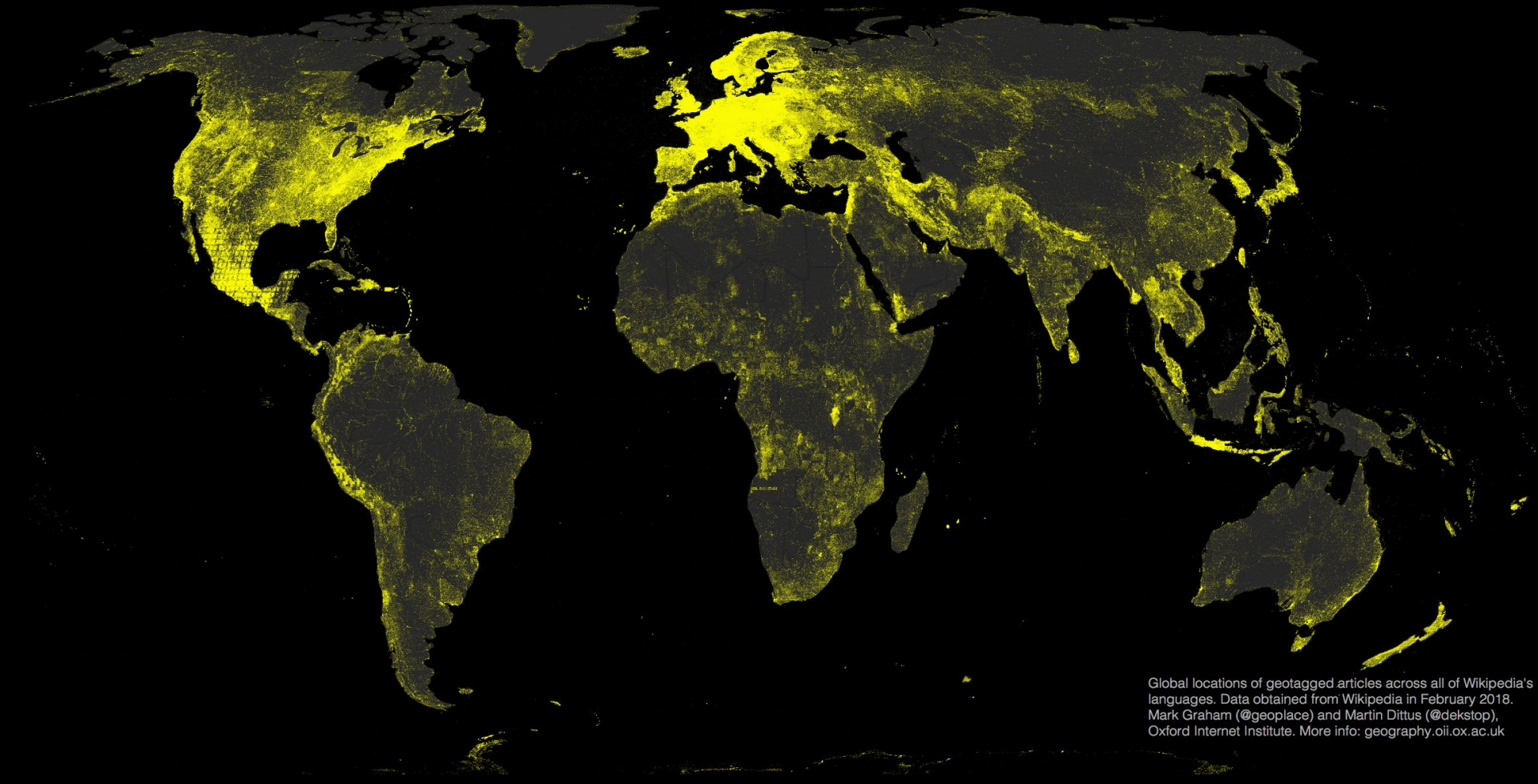
While it is an invaluable resource of knowledge for numerous users, Wikipedia's articles have a strong bias in their geographic distribution. Below map is based on 3,336,473 geotagged articles in November 2012 data dumps of 44 language versions. Each article is represented by an orange dot.



This map is part of the Information Geographies project at http://geography.oii.ox.ac.uk.

CC-BY-NC
Ralph Straumann, Mark Graham
Data sources: Wikipedia, Natural Earth

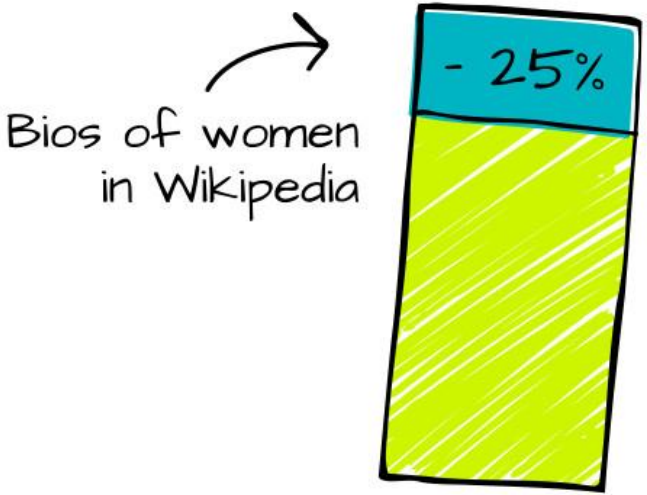
Geotagged Articles in Wikipedia

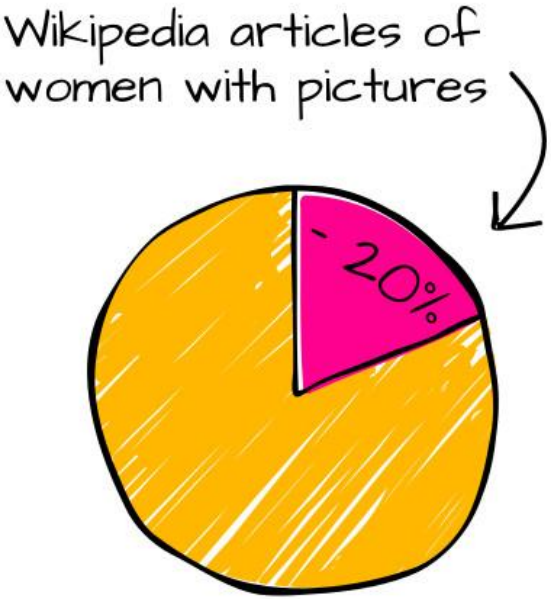


Global locations of geotagged articles across all of Wikipedia's languages. Data obtained from Wikipedia in February 2018. Mark Graham (@geoplace) and Martin Dittus (@dekstop), Oxford Internet Institute. More info: geography.oii.ox.ac.uk

Visualising the gender gap

#VISIBLE WIKI WOMEN



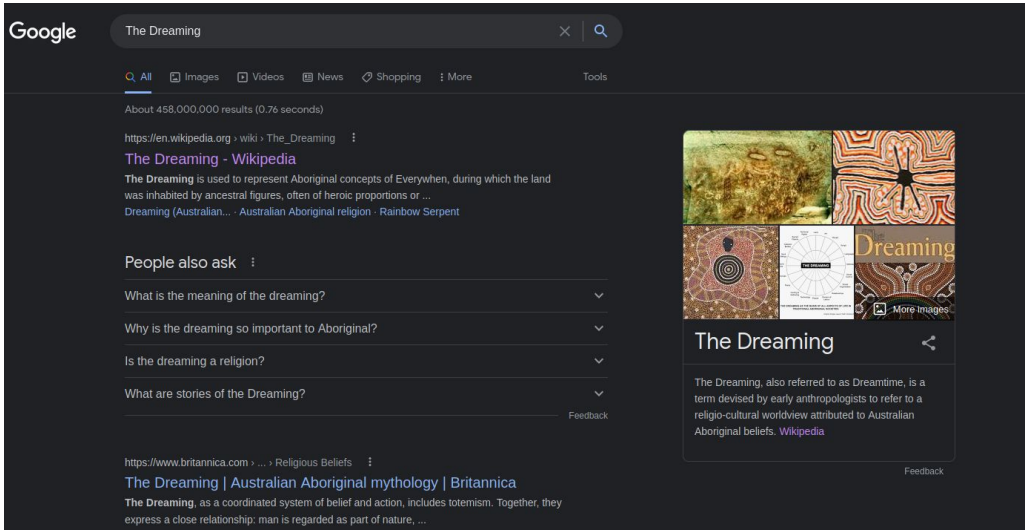


Why does Wikip/media+Wikidata matter so much?

Biggest online free and open information databases

Used by content providers (like Google)

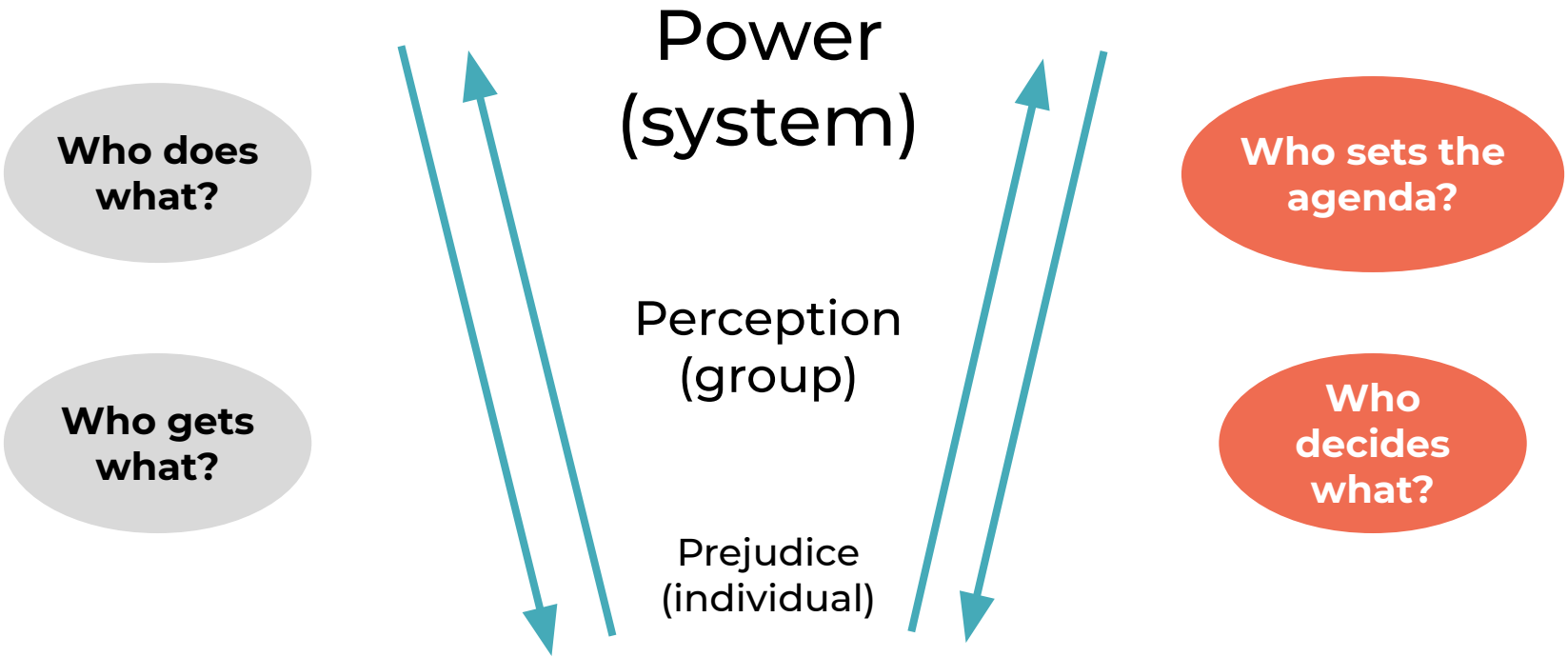
Inequities in Wikip/media spread across the internet



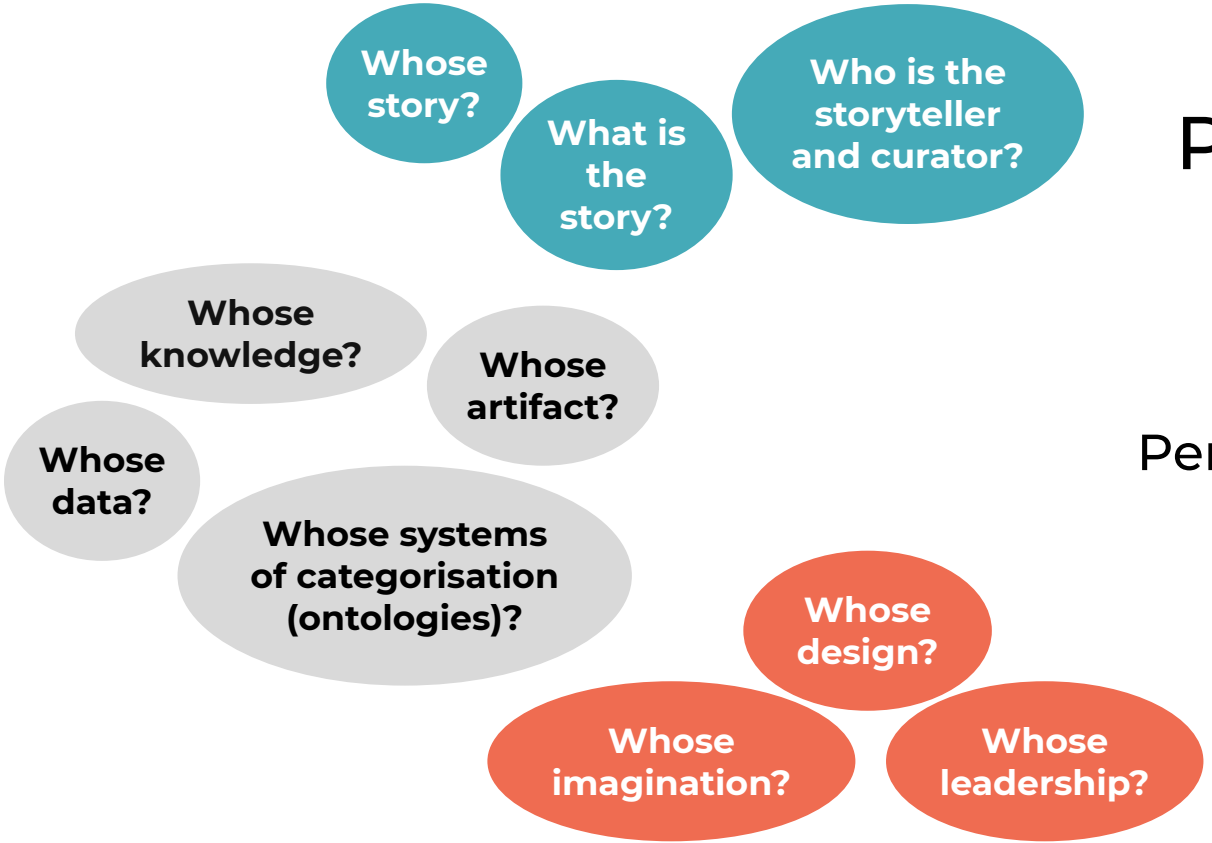
Whose Knowledge?
Design Principles and Protocols

We are rooted in a POWER analysis

Power: what does it mean?



Power and privilege: what does it mean?



Power+privilege (system)

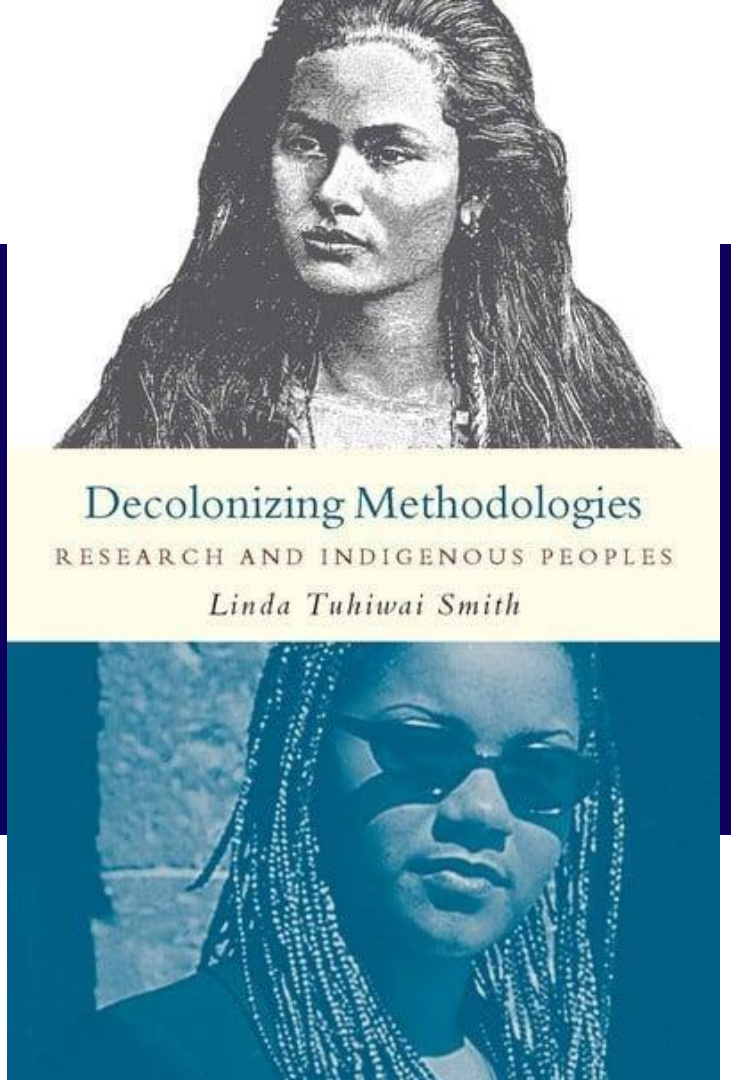
Perception+preconception (group)

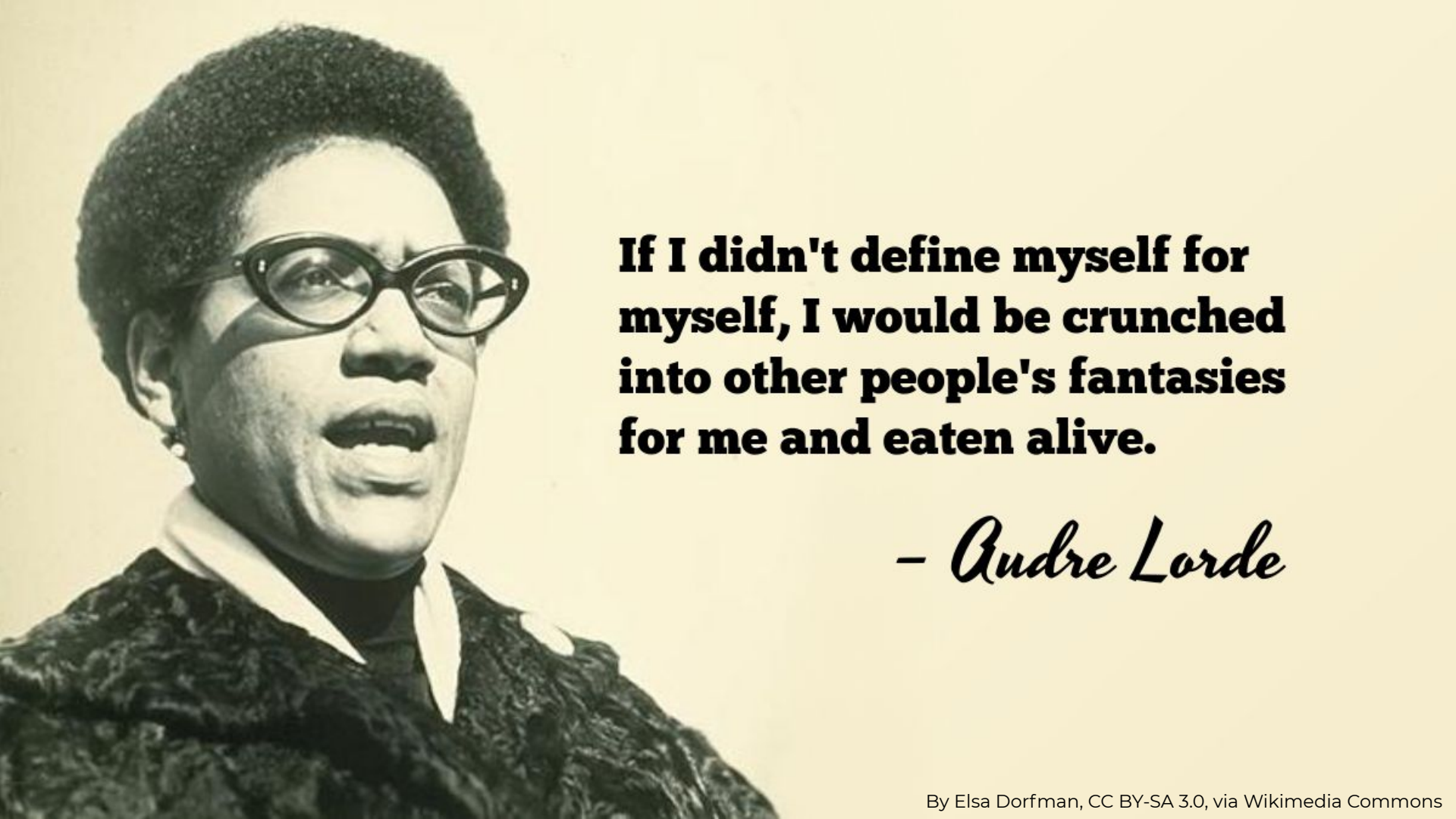
Prejudice (individual)

We centre the leadership+design inspired by the imaginations of the **marginalised majority** of the world

By Tinaral, CC BY-SA 4.0

“ TO RESIST IS TO RETRENCH IN THE MARGINS, RETRIEVE WHAT WE WERE AND REMAKE OURSELVES.





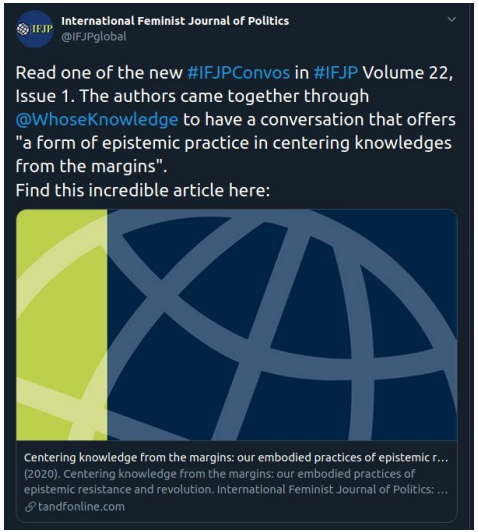
By Elsa Dorfman, CC BY-SA 3.0, via Wikimedia Commons

How we centre and support...

Dalit Bahujan

Queer Bosnian

Native American



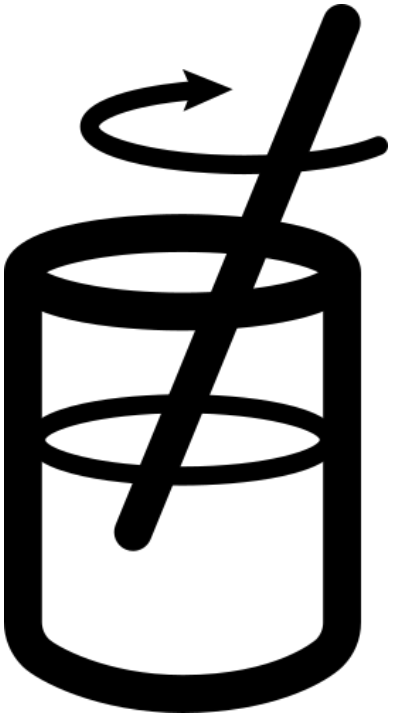
Centering knowledge from the margins: our embodied practices of epistemic resistance and revolution

“ **OUR KNOWLEDGES ARE URGENT... THEY ARE PLURAL. OUR KNOWLEDGES ARE TRANSFORMATIVE. THEY ARE HOPE.**

We recognise the work is of **reparations** and **justice**, not only representation and “solutionism”

By David Dixon, CC BY-SA 2.0 via geograph.org.uk

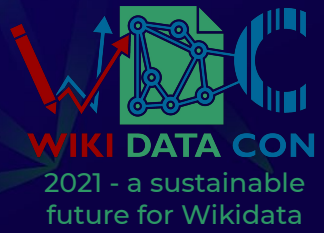
=>This is **NOT** “add women and stir, add Global South and stir…”



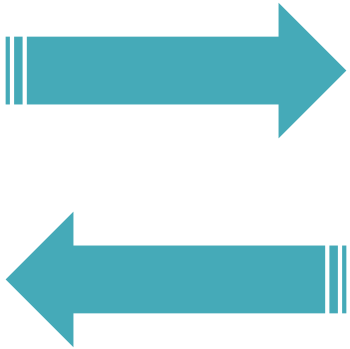
stir by Slowinski Paweł from the Noun Project

We know that scale is a jigsaw puzzle not a factory

Decolonising Wikidata: some principles and possibilities



Epistemic (In)Justice



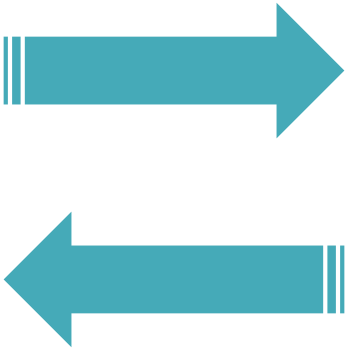
Testimonial

Hermeneutical

Australian Embassy Jakarta, CC BY 2.0 via Wikimedia Commons

By Esther 1721 CC0, via Wikimedia Commons

Knowledge Representation vs. Justice



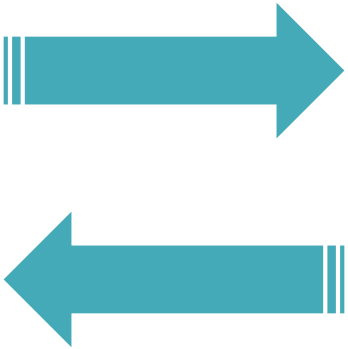
Gaps

Reframes

By NineFiveSeven, CC BY-SA 4.0 via Wikimedia Commons

By Esther 1721 CC0, via Wikimedia Commons

Structured data: frames and reframes



Deductive

Inductive

By NineFiveSeven, CC BY-SA 4.0 via Wikimedia Commons

By Esther 1721 CC0, via Wikimedia Commons

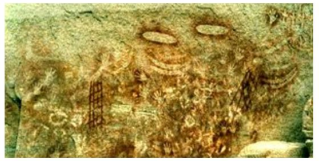
Structured data: frames and reframes

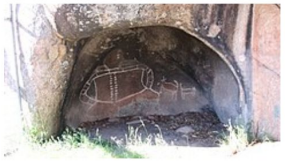
aboriginal art

Search

Images | Audio | Video | Other Media | Categories and Pages

License | File Type | Image size | Community Assessments | Sort by: Relevance | 250 results





Media in category "Dreamtime"

The following 8 files are in this category, out of 8 total.

Bibbulmun Track Markings.jpg
2,540 × 3,387; 4.04 MB

Brooklyn Museum - Painting of Dream Figures on Bark.jpg
473 × 768; 63 KB

Ku-ring-gai Chase - petroglyph.jpg
682 × 1,024; 329 KB

Ku-r… petr…
1,01…

…t.jpg
…2.54

Tjilbruke sculpture.jpg
2,608 × 4,672; 3.6 MB

Uluru Cave.JPG
2,592 × 1,936; 2.35 MB

Wagyl.jpg
3,264 × 2,448; 3.79 MB

…thology of Australian Aboriginals (−) (±) (↓) (↑) | (+)
…ses of Wikidata Infobox | Uses of Wikidata Infobox with no image

WIKIDATA

Item | Discussion

Dreamtime (Q155881)

sacred era in Australian Aboriginal mythology

Dreaming | The Dreaming | Alcheringa | Alchera | Altjira | Mura-mura | Tjukurpa

In more languages

Configure

Language	Label	Description	Also known as
English	Dreamtime	sacred era in Australian Aboriginal mythology	Dreaming The Dreaming Alcheringa Alchera Altjira Mura-mura Tjukurpa
British English	Dreamtime	sacred era in Australian Aboriginal mythology	Dreaming The Dreaming Alcheringa Alchera Altjira Mura-mura Tjukurpa
French	Temps du rêve	thème et mythologie central de la culture des aborigènes d'Australie	Temps du reve Alchera Tjukurpa Mura-mura Intraterrestre
German	Traumzeit	Zentraler Begriff der Mythologie australischer Aborigines und ihrer ethnischen Religionen	Alcheringa Dreamtime Dreaming

Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Concept URI
Cite this page

Structured data: frames and reframes

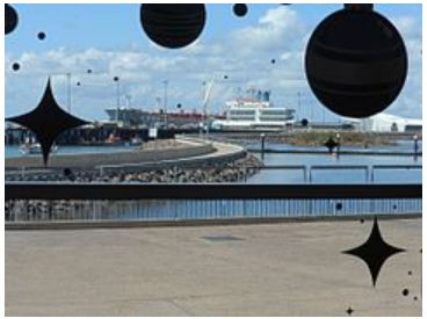
australian reality

Search

Images | Audio | Video | Other Media | Categories and Pages

License | File Type | Image size | Community Assessments | Sort by: Relevance

4 results



No more results found

Terreiros and Ontologies

Possible next steps for Wikidata...

- **Reframe**: Break the silences, have the conversations
- **Invite**: Create an advisory group for multiple epistemologies
- **Test**: Support a few community-led pilots and imaginings
- **Practice**: Begin adopting and adapting...

...Impossible in theory, made possible in practice?

WIKI DATA CON
2021 - a sustainable future for Wikidata

Obrigada, شکریہ, thank you!

GET IN TOUCH WITH US:

Anasuya Sengupta
@anasuyashh

Whose Knowledge?

@whoseknowledge
whoseknowledge.org

Whose Knowledge?

WIKI DATA CON
2021 - a sustainable future for Wikidata

References (beyond Q42)

Internet Stats

- World Internet Stats
- ITU study on gender
- PWC UK: gender gap in tech
- NCWIT: gender and race in tech
- IETF stats, IETF meeting stats
- ICANN stats

Knowledge and languages

- Google: published books
- Languages in the world
- Languages on the internet
- Languages of scholarly publications in science and social science

UK academia

- Race in UK academia
- Staying Power report on race and gender
- Leading Routes: Broken Pipeline report

Frames of reference

- Achal Prabhala, Priya Sen, Zen Marie, People are Knowledge
- Michael Polanyi, The Tacit Dimension
- Miranda Fricker, Epistemic Injustice
- Linda Tuhiwai Smith, Decolonizing Methodologies
- Audre Lorde, Learning from the 60s

Wikipedia

- Gender gap
- Geographical gap
- Oxford Internet Institute research

Whose Knowledge? resources

- International Feminist Journal of Politics article
- Our Stories, Our Knowledges
- #VisibleWikiWomen 2020 wrap up blog post
- More on #VisibleWikiWomen
- Overview of State of the Internet's Languages
- Decolonizing the Internet report 2018
- Decolonizing the Internet's Languages report 2019
- Article in Nature: Build an Internet for, and from, us all
- More resources here
- More media here