

# 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!



## Mapeamento dos Terreiros de Salvador

[Apresentação](#) | [Análise](#) | [Terreiros](#) | [Mapa](#) | [Equipe](#) | [Contato](#) | [Link](#)

### APRESENTAÇÃO

Salvador é a terceira cidade mais populosa do Brasil (2.714.119 habitantes) pela estimativa de 2005, depois de São Paulo e Rio de Janeiro. Também é a capital brasileira com o maior contingente negro entre a população. Mais de 70% dos seus habitantes são afrodescendentes e a negritude se expressa através da sua forma de organização geográfica, cultural e social. Tal herança deu origem às religiões de matriz africana e se diferenciam por nações com diferenças rituais e no vocabulário (Angola, Ketu, Jeje, Ijexá). As Secretarias Municipais da Reparação e da Habitação, em parceria com o Centro de Estudos Afro-Orientais da UFBA, realizaram um grande pesquisa sobre as religiões de matrizes africanas na cidade do Salvador. O objetivo foi conhecer os terreiros da nossa cidade: saber quantos são, onde estão localizados, suas condições de documentação, regularização fundiária e infra-estrutura, entre outros aspectos sócio-culturais e demográficos. Esta pesquisa compõe o Programa de Valorização do Patrimônio Afro-Brasileiro, além de ser o ponto de partida para uma série de políticas públicas a serem adotadas nas comunidades. Ações que vão colaborar para a legalização e regularização fundiária desses espaços, diminuir o preconceito sobre as religiões afro-brasileiras e, principalmente, valorizar a cultura afro-brasileira.



# Knowledge Infrastructures

# 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



# 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?

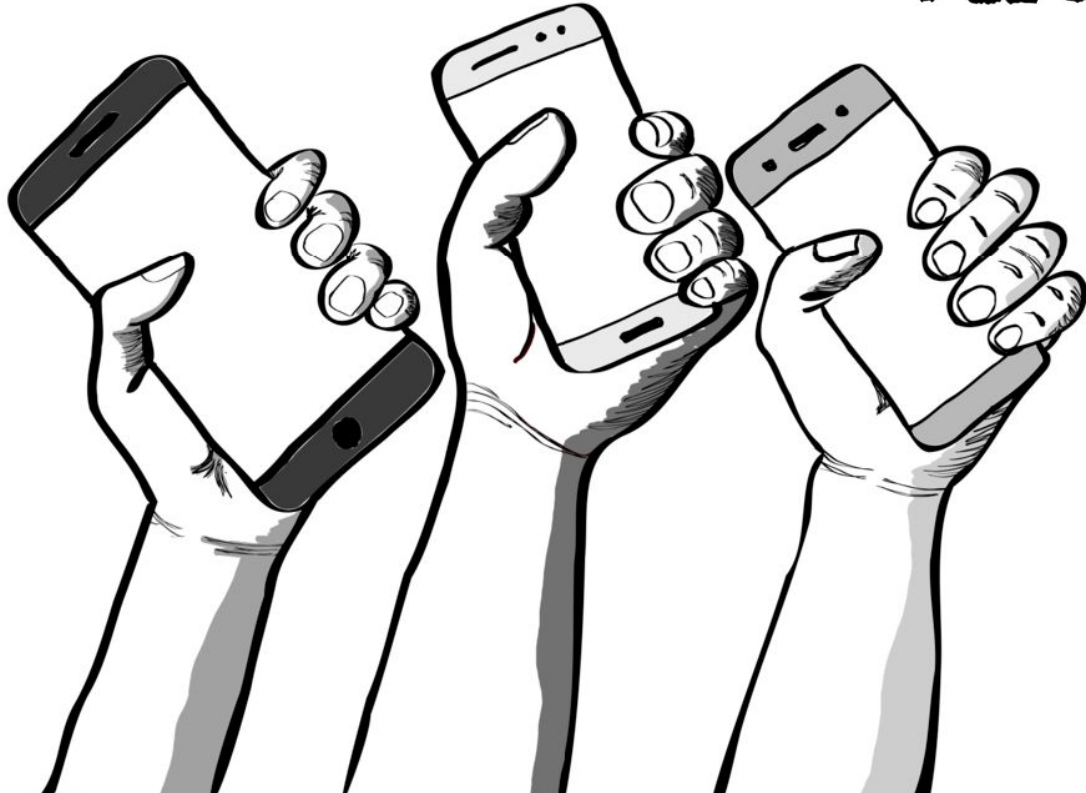
PUBLISHED AND ONLINE KNOWLEDGE,  
SCHOLARSHIP AND SCHOLARS...

**“RELIABLE SOURCES”**

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



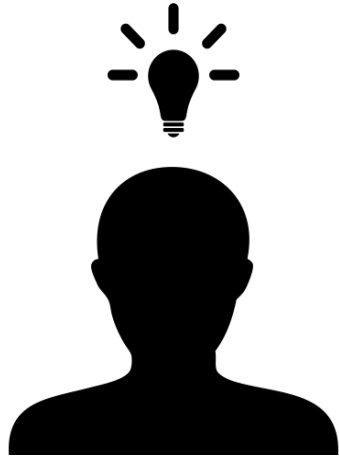
# DEcolonize theINTERNET



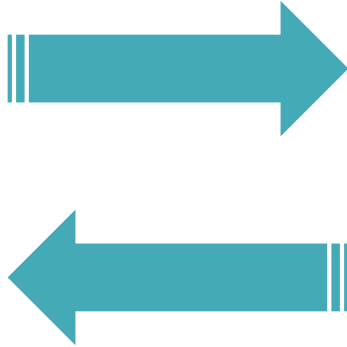


# People are Knowledge(s)

**Embodied  
(tacit) knowledge**



**People's  
authority**



**Disembodied  
(formal) knowledge**

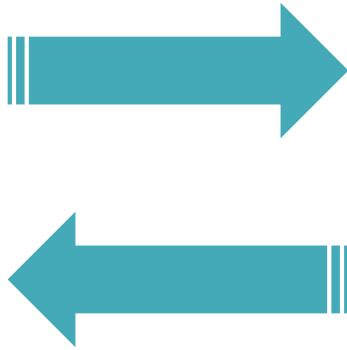


**Institutional  
authority**

# Epistemic Injustice



Testimonial



Hermeneutical



# Using Wikip/media as proxy...

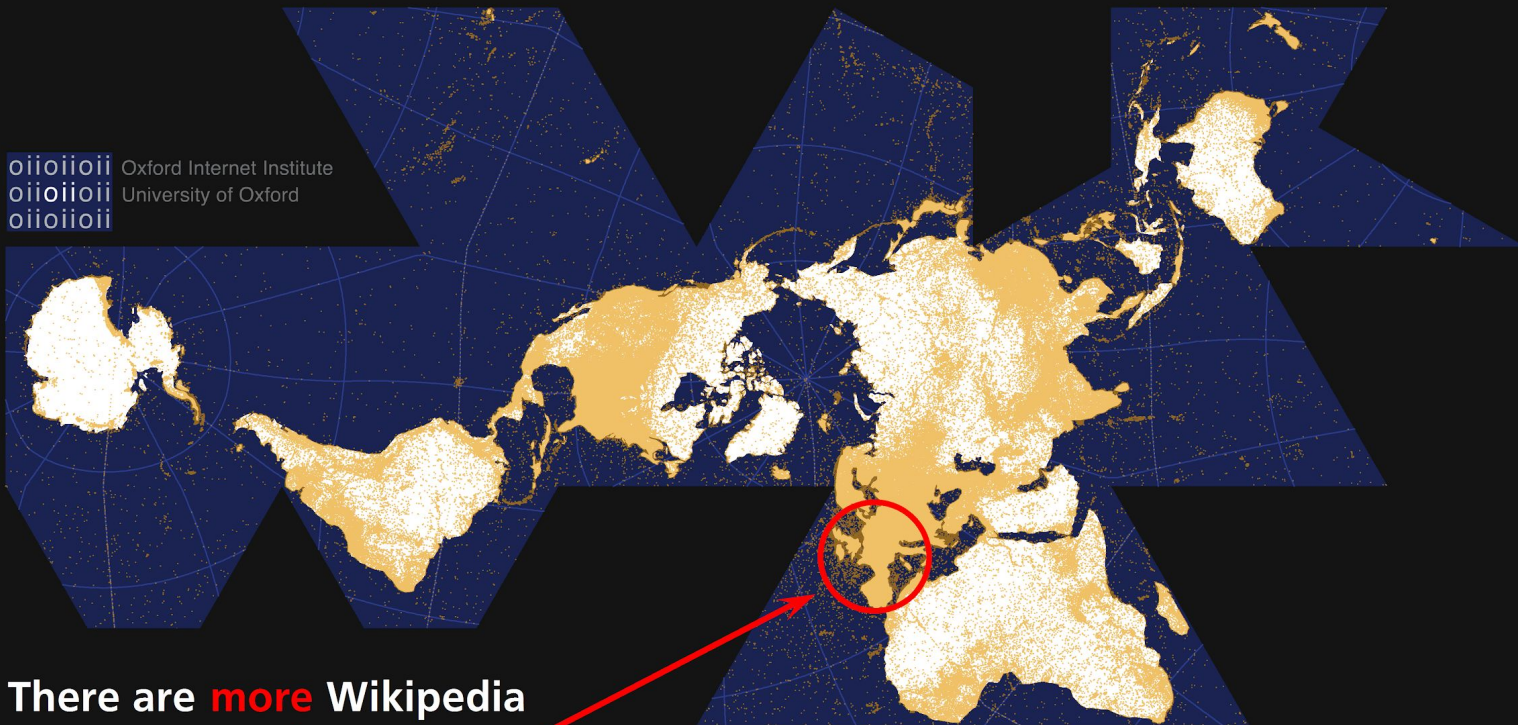
- Only 10% of Wikipedia editors are women or non-binary
- Only around 20% of public knowledge is produced on or by people from the Global South



# 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.

oioioioii Oxford Internet Institute  
oioioioii University of Oxford  
oioioioii

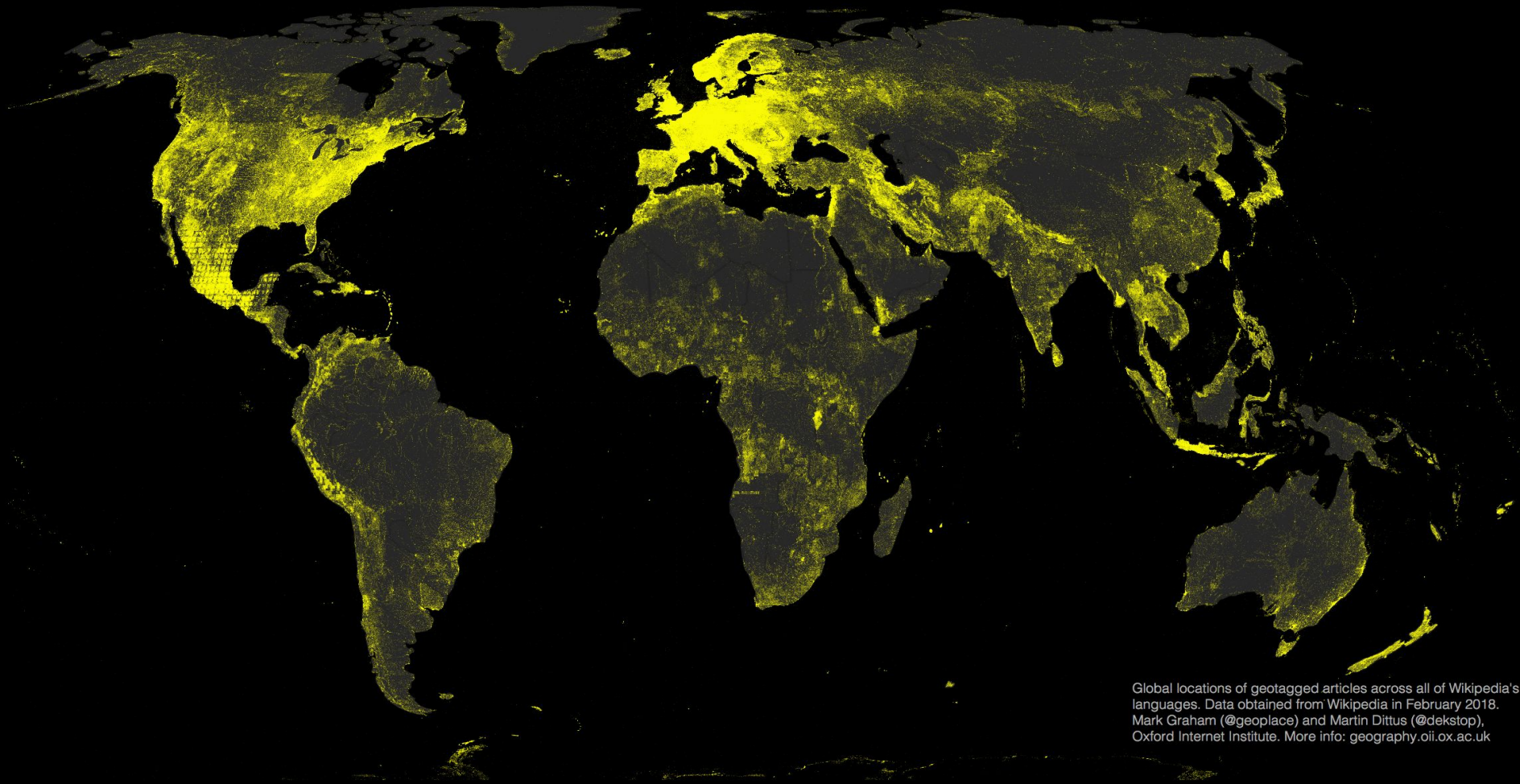


There are **more** Wikipedia articles **inside** this circle **than outside** of it

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](http://geography.oii.ox.ac.uk)

# Visualising the gender gap

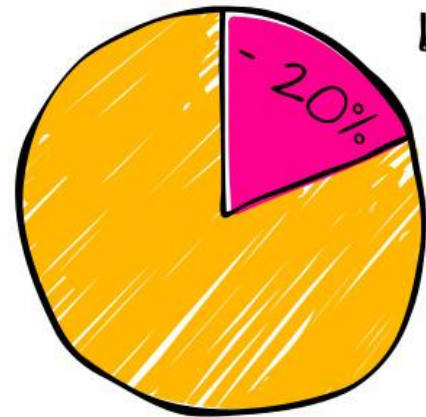


#VISIBLE  
WIKI WOMEN

Bios of women  
in Wikipedia



Wikipedia articles of  
women with pictures

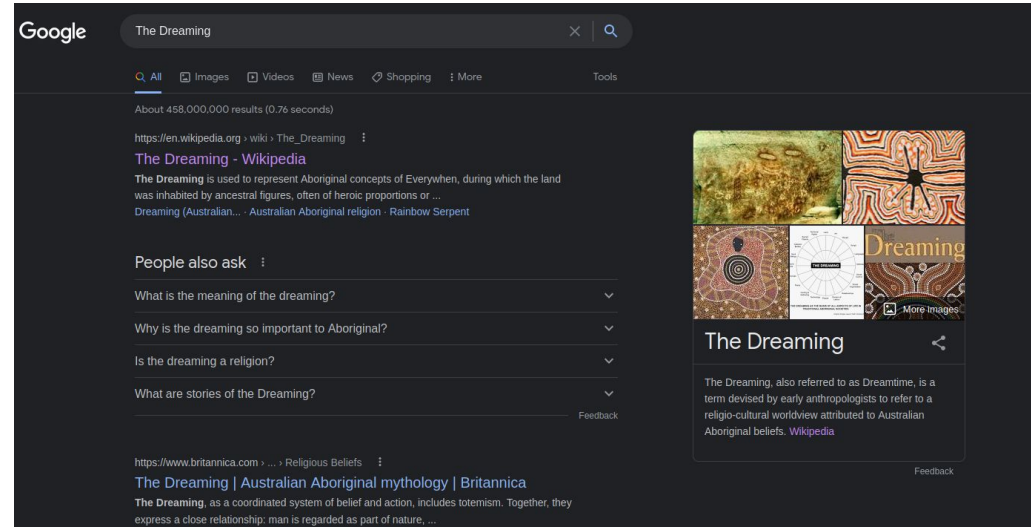


# Why does Wikip/media+Wikidata matter so much?

Biggest online free and open information databases

Used by content providers (like Google)

Inequities in Wikip/media spread across the internet



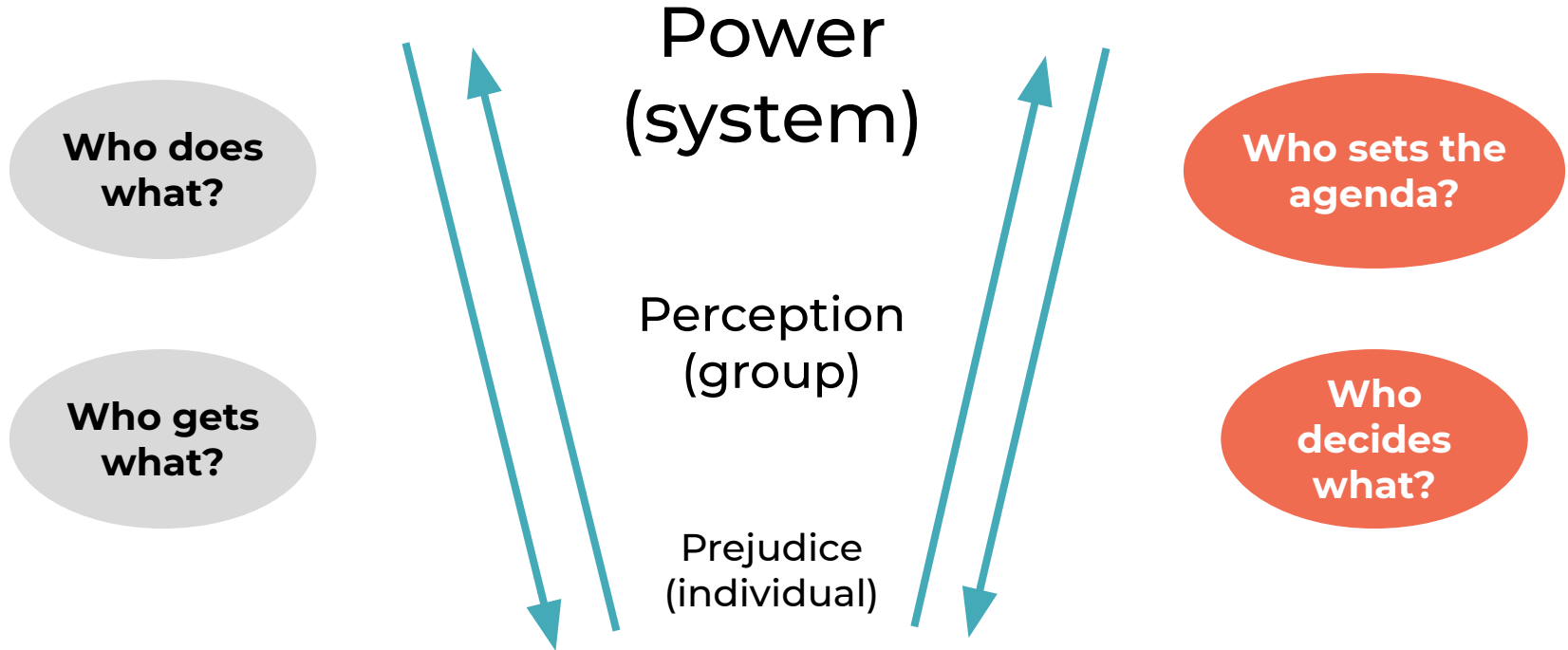
The screenshot shows a Google search for "The Dreaming". The search bar at the top contains the text "The Dreaming" and a magnifying glass icon. Below the search bar, there are navigation tabs for "All", "Images", "Videos", "News", "Shopping", and "More". The search results indicate "About 458,000,000 results (0.76 seconds)". The first result is from Wikipedia, titled "The Dreaming - Wikipedia", with a snippet: "The Dreaming is used to represent Aboriginal concepts of Everywhen, during which the land was inhabited by ancestral figures, often of heroic proportions or ...". Below this is a "People also ask" section with four questions: "What is the meaning of the dreaming?", "Why is the dreaming so important to Aboriginal?", "Is the dreaming a religion?", and "What are stories of the Dreaming?". To the right of the search results is a featured snippet titled "The Dreaming" with a grid of four images: a landscape painting, a colorful geometric pattern, a circular diagram, and a book cover titled "Dreaming". Below the images is a text box: "The Dreaming, also referred to as Dreamtime, is a term devised by early anthropologists to refer to a religio-cultural worldview attributed to Australian Aboriginal beliefs. Wikipedia". At the bottom of the page, there is a result from Britannica: "The Dreaming | Australian Aboriginal mythology | Britannica" with a snippet: "The Dreaming, as a coordinated system of belief and action, includes totemism. Together, they express a close relationship: man is regarded as part of nature, ...".

# 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?

Whose story?

What is the story?

Who is the storyteller and curator?

Power+privilege  
(system)

Whose knowledge?

Whose artifact?

Whose data?

Whose systems of categorisation (ontologies)?

Perception+preconception  
(group)

Whose design?

Whose imagination?

Whose leadership?

Prejudice  
(individual)

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

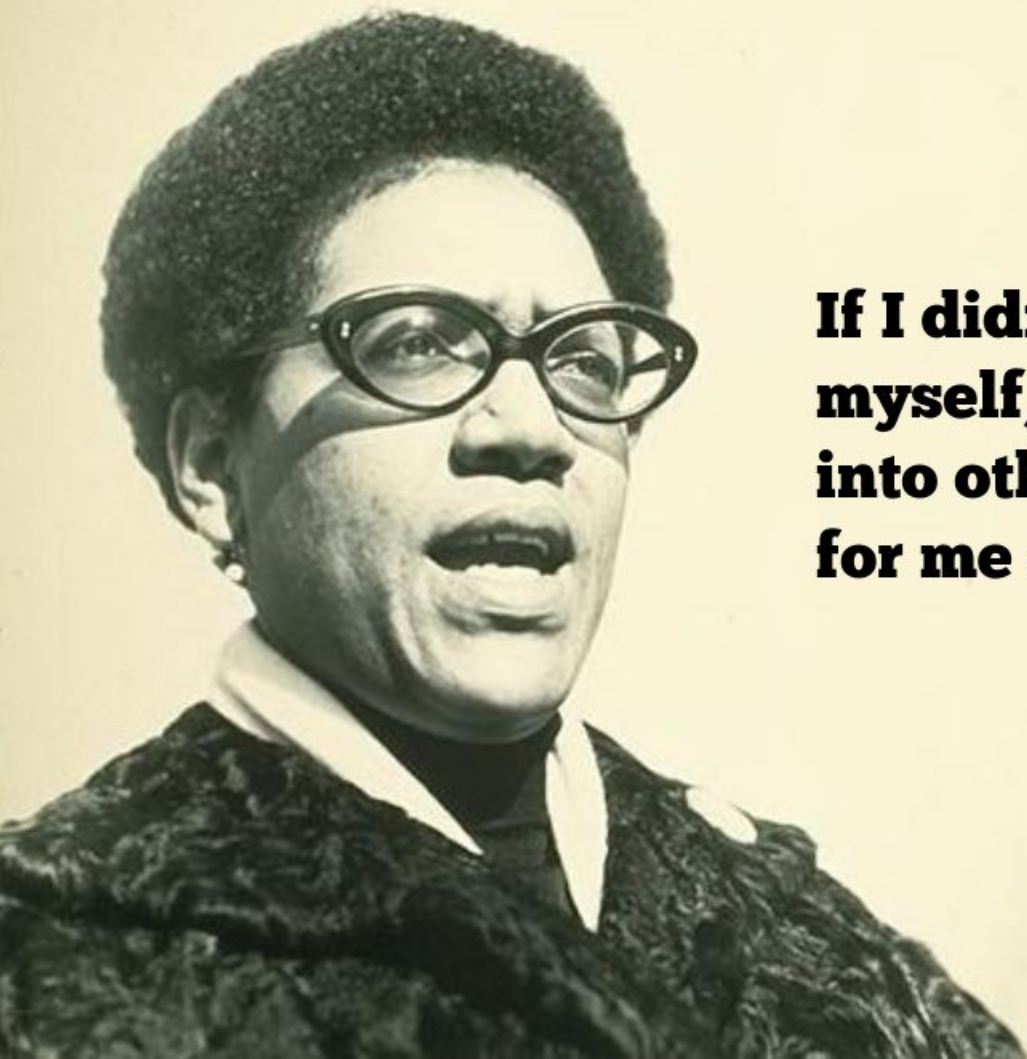


“ TO RESIST IS TO  
RETRENCH IN THE  
MARGINS, RETRIEVE  
WHAT WE WERE AND  
REMAKE OURSELVES.



Decolonizing Methodologies  
RESEARCH AND INDIGENOUS PEOPLES  
*Linda Tuhiwai Smith*





**If I didn't define myself for myself, I would be crunched into other people's fantasies for me and eaten alive.**

*- Audre Lorde*

# How we centre and support...

Dalit Bahujan

Queer Bosnian

Native American



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



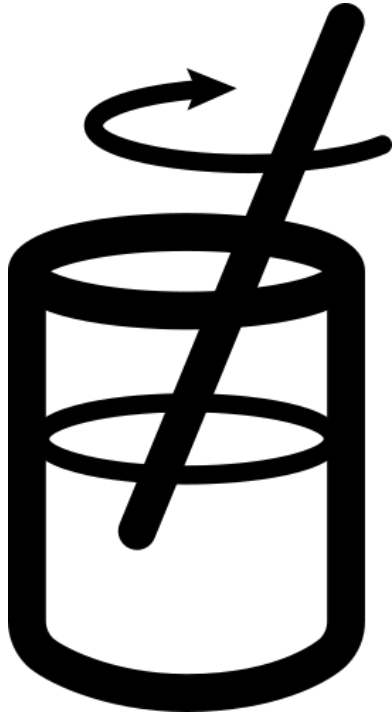
“

**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”



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





We know that scale is a jigsaw puzzle  
not a factory

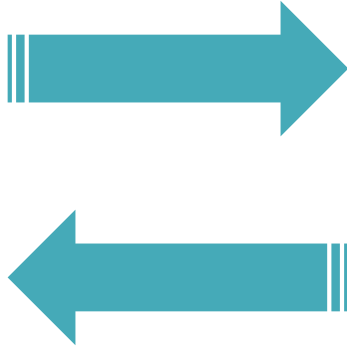


# Decolonising Wikidata: some principles and possibilities

# Epistemic (In)Justice



Testimonial

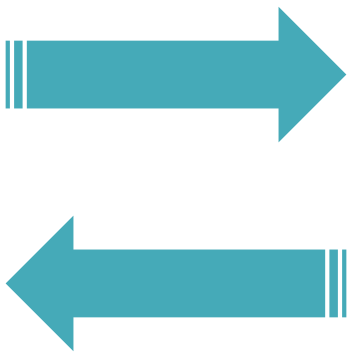


Hermeneutical

# Knowledge Representation vs. Justice



Gaps

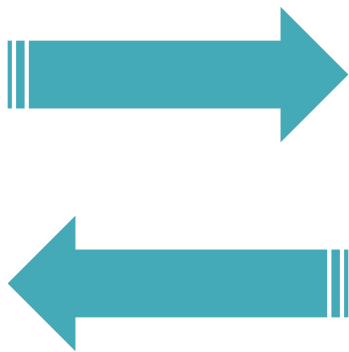


Reframes

# Structured data: frames and reframes



Deductive




Inductive


# Structured data: frames and reframes

## 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-ring-gai Chase - petroglyph.jpg**  
1,012 × 682; 329 KB

Search


Images
Audio
Video
Other Media
Categories and Pages

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



- Main page
- Community portal
- Project chat
- Create a new item
- Recent changes
- Random item
- Query Service
- 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


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 rêve Alchera Tjukurpa Mura-mura Intéroterestre
German	Traumzeit	Zentraler Begriff der Mythologie australischer Aborigines und ihrer ethnischen Religionen	Alcheringa Dreamtime Dreaming



**tj.jpg**  
2.54 MB



**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](#)

# Structured data: frames and reframes

australian reality



Search

**Images**

Audio

Video

Other Media

Categories and Pages

License ▾

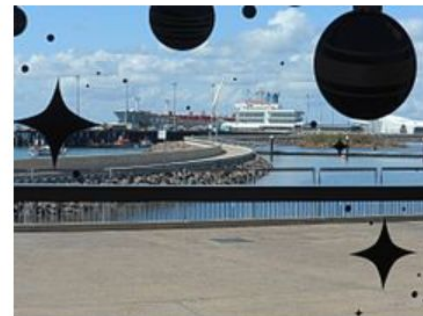
File Type ▾

Image size ▾

Community Assessments ▾

Sort by: Relevance ▾

4 results



No more results found

# Terreiros and Ontologies



## Mapeamento dos Terreiros de Salvador

[Apresentação](#) | [Análise](#) | [Terreiros](#) | [Mapa](#) | [Equipe](#) | [Contato](#) | [Link](#)

### APRESENTAÇÃO

Salvador é a terceira cidade mais populosa do Brasil (2.714.119 habitantes) pela estimativa de 2005, depois de São Paulo e Rio de Janeiro. Também é a capital brasileira com o maior contingente negro entre a população. Mais de 70% dos seus habitantes são afrodescendentes e a negritude se expressa através da sua forma de organização geográfica, cultural e social. Tal herança deu origem às religiões de matriz africana e se diferenciam por nações com diferenças rituais e no vocabulário (Angola, Ketu, Jeje, Ijexá). As Secretarias Municipais da Reparação e da Habitação, em parceria com o Centro de Estudos Afro-Orientais da UFBA, realizaram um grande pesquisa sobre as religiões de matrizes africanas na cidade do Salvador. O objetivo foi conhecer os terreiros da nossa cidade: saber quantos são, onde estão localizados, suas condições de documentação, regularização fundiária e infra-estrutura, entre outros aspectos sócio-culturais e demográficos. Esta pesquisa compõe o Programa de Valorização do Patrimônio Afro-Brasileiro, além de ser o ponto de partida para uma série de políticas públicas a serem adotadas nas comunidades. Ações que vão colaborar para a legalização e regularização fundiária desses espaços, diminuir o preconceito sobre as religiões afro-brasileiras e, principalmente, valorizar a cultura afro-brasileira.





# 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?*

# Obrigada, شكريه, thank you!

GET IN TOUCH WITH US:

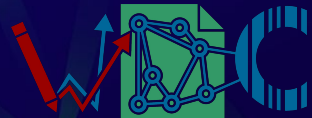
Anasuya Sengupta  
@anasuyashh

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](#)