University of Massachusetts Amherst

February 5, 2013

Mapping Participation Gaps in Wikipedia

Laura Quilter, *University of Massachusetts - Amherst*

Mapping 'Participation Gaps' in Wikipedia

Laura Quilter / Copyright & Information Policy Librarian
University of Massachusetts Amherst, Feb. 5, 2013
397M "New Media Technologies & Social Change"

aka: "Systemic Bias in Wikipedia"
What it looks like, and how to
deal with it

part 1
questioning wikipedia

Social Mediation

Design & Use

- Recombinant change
- Network organization of Society & Technology

Artifacts ⟷ Practices/ Use

Social Arrangements

Social Consequences

- Ubiquity
- Interactivity
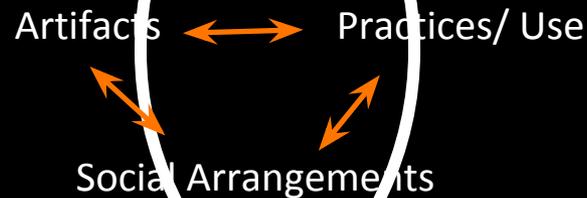
New Media as a Socio-Technological System

(Lievrouw, 2011)

what this talk is NOT about :

- should academics or students use Wikipedia?
  - *any time an encyclopedia would be okay ...*
- is Wikipedia accurate?
  - *studies show ... as much or more than other encyclopedias*
- is plagiarizing Wikipedia okay because it doesn't belong to any one other person?
  - *NO, because student and academic plagiarism is about the plagiarist's violation to the trust invested in them -- not just harms possibly suffered by plagiarized subjects.*

what this talk IS about :

"Systemic bias is the inherent tendency of a process to favor particular outcomes. The term is a neologism that generally refers to human systems..."

"Systemic Bias", Wikipedia, quoted as of Oct. 19, 2012; http://en.wikipedia.org/w/index.php?title=Systemic_bias&oldid=493766084

types of SB in wikipedia

- systemic bias on Wikipedia replicates the systemic biases of culture at large: racism, sexism, heterosexism, ableism, majority religion, etc.
- replicates systemic biases of Internet culture: Anglophone, class privilege, "Libertarian" politics, WM
- amplified / modulated by the medium of the Internet:
  - pseudonymized communications
  - self-selected class of privileged, well-educated, technologically-savvy, folks with time on their hands
- various additional biases: pop culture; tech topics; non- academic; current events; corporate / MSM sources; etc.

what systemic bias looks like

content bias
- dearth of articles
- inadequate articles
- perspective bias

- editor demographics
- editor experiences
- editor participation, contributions, & retention

what it looks like: missing content

User:Dsp13/Biographical Dictionary of Women in Science

From Wikipedia, the free encyclopedia

< User:Dsp13

The Biographical Dictionary of Women in Science is a 2-volume work containing entries on nearly 3,000 scientists. Below is a list of its entries in Volume II, the text of which is accessible online on bookrags.com

Ogilvie, Marilyn; Harvey, Joy, eds. (1999), *The Biographical Dictionary of Women in Science*, II, Routledge, ISBN 978-0415920407

- each red link is a woman scientist without an article
- *Biographical Dictionary of Women in Science*, User: Dsp13

what it looks like : inadequate articles

- shorter & less content
- fewer references
- fewer links in or out
- fewer editors & fewer eyes ==> more bugs, typos, bad writing, uncorrected vandalism
- perspective bias ("criticism" sections; unlinked and uncredited family members of men; non-global perspectives; undue weight)
- ... fewer "Featured Articles", "Good Articles", "Do-you-know" highlights; "Vital Articles".

what it looks like: inadequate content
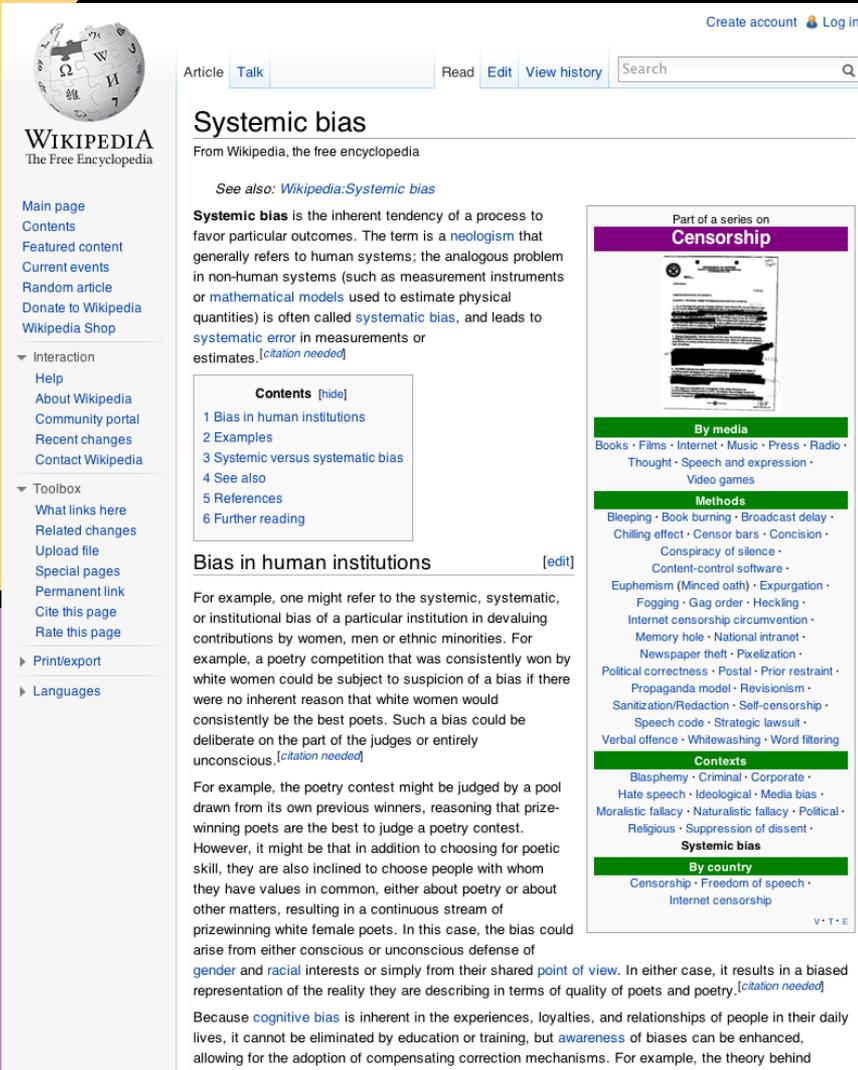
- the article exists!

BUT:

- inadequate article: overall small article relative to topic's importance

- perspective bias: very little discussion of gender / ethnicity

what it looks like : geographies

Mark Graham (2009) "Mapping the Geographies of Wikipedia Content" http://www.
zerogeography.net/2009/11/mapping-geographies-of-wikipedia.html
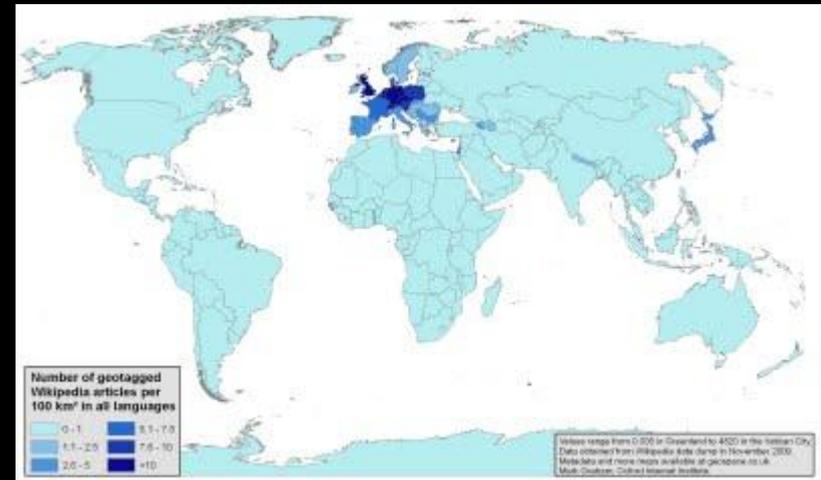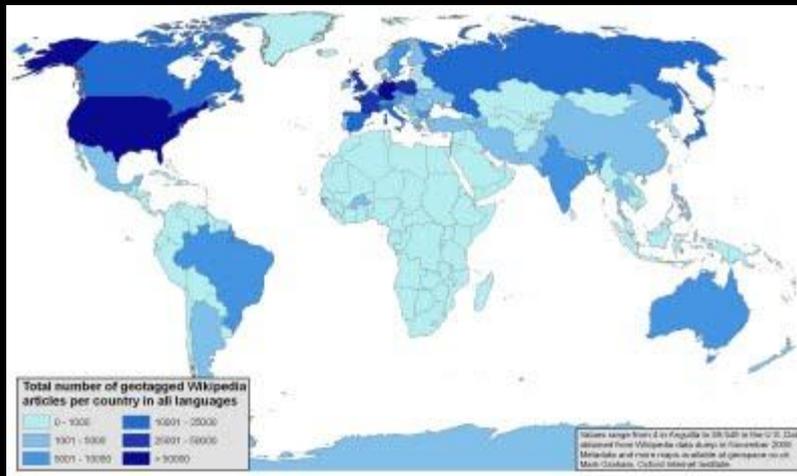
(clockwise) Wikipedia articles
   (a) by country;
   (b) normalized for area;
   (c) normalized for populace

What do you notice?
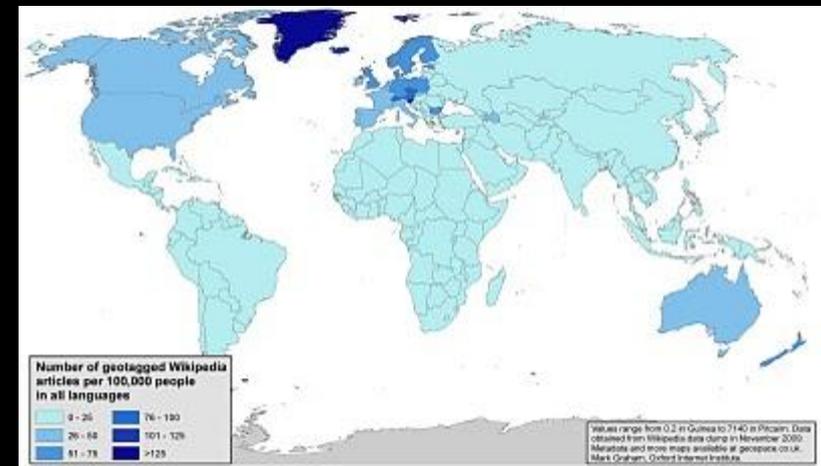
who writes wikipedia?

- analysis done by Aaron Swartz (2006) looking at *total number of characters added* suggest most substantive content added by individuals, often doing very few edits
  - people (like you?) make a small number of substantive contributions
- analyses counting *total number of edits* suggest most work done by a dedicated pool of self-identified "Wikipedians"
  - people like me make lots of edits reviewing, "cleaning up", and organizing content

what it looks like: editor experience

- ● 8.5-15% of editors are women [various]
  - ○ Internet trolling of women/POC/minorities
- ● Wikipedians are less "agreeable" & less "open"; cultural expression & tone - conflict, tendentiousness; 24% report harassment [Amichai-Hamburger 2008; WES 2011]
- ● editor frustration because of deleted/reverted work, or lengthy processes to justify / educate
- ● Note: WikiMedia Editors Survey 2011 looked at nationality/language; education; motivation & experiences; gender. NOT ethnicity, religion...

why does systemic bias persist?

- editing comes from editors: average Wikipedian
    - white, male [90%], 32yo, college-educated or student, English-speaking, US or European; single (60%); not a parent (80%) [WES 2011]
- Wikipedia culture established that replicates the "average Wikipedian"
- English sourcing
- digital divide

Editor Age Groups
- 12-17 13%
- 18-21 14%
- 22-29 26%
- 30-39 19%
- 40+ 28%

Editor Education Level Completed
- PhD 8%
- Primary 9%
- Masters 18%
- Secondary 30%
- Undergrad 35%

Editor Gender
- Female 9%
- Transgender/transsexual 1%
- Male 90%

Source: Wikimedia Foundation 2011 Editor Survey, April 2011
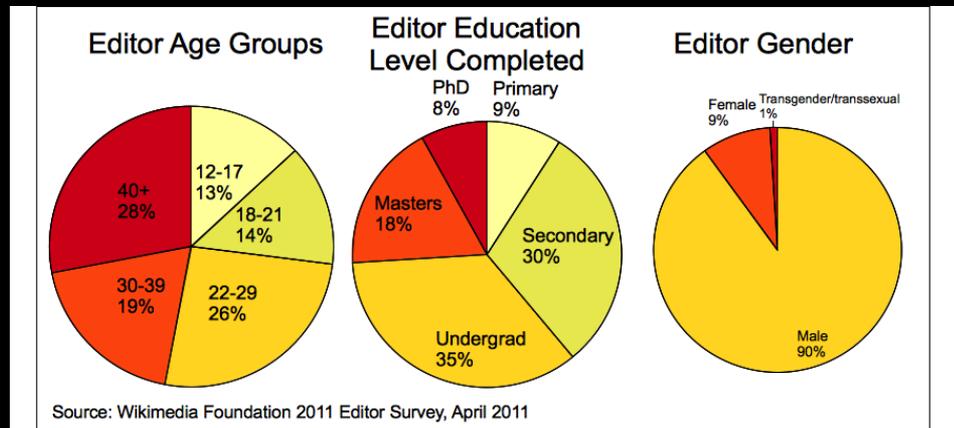
so why do we care?

- #6th ranked site on Internet
- #1 general reference on Internet -- the go-to for wide variety of answers for children, students, academics, press, etc.
- 2.7Bn pageviews in US in 2011

- the good news is that while Wikipedia's failings replicate those of the larger society, and in particular those of Internet society, *Wikipedia is easier to change!*

part 2
see for yourself

see for yourself

The best way to explore biases in Wikipedia is to pick some topics, and explore Wikipedia's coverage on them. An approach:

- Pick a topic of personal interest to you
- Identify your assumptions, or questions, about Wikipedia's coverage of that topic.
- Identify some ways to assess your assumptions / test your question.
- Did you find what you expected? Why or why not?
- What next?

see for yourself

The best way to explore biases in Wikipedia is to pick some topics, and explore Wikipedia's coverage on them. An approach:

- Pick a topic of personal interest to you
- Identify your assumptions, or questions, about Wikipedia's coverage of that topic.
- Identify some ways to assess your assumptions / test your question.
- Did you find what you expected? How was it different if not? Why?
- What next?

see for yourself

- ...
- Did you find what you expected? How was it different if not? Why?
  - This is where you attempt to pull it all together: Assess your own biases in formulating the question; assess what you found; figure out if it's real or an artifact of your method; if it's real, try to describe WHY it is.
- What next?
  - Applied research means your question is aimed at figuring out how to make something better. So if you found a bias, how can you (and others) remedy it?

see for yourself

- What are good ways to assess your assumptions / test your question?

Content coverage:

  - *Qualitative article assessment*
  - *Quantitative article assessment*
    - compare Wikipedia's coverage to another encyclopedia's coverage; OR
    - compare Wikipedia's coverage of this topic to Wikipedia's coverage of another topic.

Are there other ways to explore Wikipedia's coverage of that topic?

see for yourself

Some ways to compare articles on Wikipedia:

- Look at length of article
- Consider depth of coverage
- Is it appropriately cited and referenced?
- Is there currently vandalism, hatespeech, or obvious inaccuracies?
- Is it tagged for improvement?
- Count & consider the CATEGORIES at the bottom of the page
- Look at other navigational aids: navigation boxes, "see also" links, Wikipedia "awards"
- Click "What links here" -- is it linked as much as it should be?

see for yourself

Some ways to compare articles on Wikipedia:
- View the HISTORY to see the edit history.
  - Read the edit summaries, and compare some of the edits to see histories of vandalism, changes in content, edit-wars. What did the article look like to begin with? When was it first created?
- Look at TALK page (and any talk archives)
  - How many people are talking? What are they saying? Was the article challenged for notability or bias? What was the quality of discourse?
- If any challenges or disputes, read those.

see for yourself

Some possible things to compare:

- language coverage: An article on the English-language (http://en.wikipedia.org) versus one of the non-English WPs.
  - Compare a "local" subject in both versions
  - Compare a "universal" subject in both versions
- geographic topics: An article about a US city, landmark, historical event, or political leader; compared with a comparable article in an African country.

see for yourself

Some possible things to compare:

- gender biases: Pick a topic of mid-level importance in women's history or a cultural issue of interest to women; compare with similar topic for men.
- class & cultural biases: Pick a topic of mid-level importance to a minority or disenfranchised class; compare with similar topic

see for yourself

Some possible things to compare:

- subject biases: Pick a topic of mid-level importance in pop culture or sports versus a similar academic topic.

see for yourself

Other exercises:

- Browse biographical categories in Wikipedia. Look at a random sampling of articles. Are the articles representative of the gender, ethnic, or other diversity of the field? Why or why not?
Example: Biologists by field of research

- Consider the Wikipedia "Notability" policy. What is it trying to accomplish and is it successful? Does it encode systemic bias, and if so, does that advance or hinder the goal of the policy? How could the policy be rewritten to address your concerns, if any?
http://en.wikipedia.org/wiki/Wikipedia:Notability

references & further reading

- Wikipedia, "Systemic Bias" (encyclopedia article)
- Wikipedia, WikiProject Countering Systemic Bias
- Wikipedia Editors Survey, 2011
- Halavais & Lackaff, 2008, "An Analysis of Topical Coverage of Wikipedia"
- Graham, 2009, "Mapping the Geographies of Wikipedia Content"
- Aaron Swartz, 2006, "Who Writes Wikipedia?" at http://www.aaronsw.com/weblog/whowriteswikipedia (see also related essays)
- Lots of coverage in press & blogs. See, e.g., Cohen, NYT, 2011/01/31; Sue Gardner, 2011/02/19.

references & further reading

- "100 longest Wikipedia articles", http://www.buzzfeed.com/fjelstud/the-100-longest-entries-on-wikipedia
- Wikigroaning: Comparing two similarly- titled articles on subjects of widely disparate importance: http://www.somethingawful.com/d/news/wikigroaning.php

part 3
How to stop worrying
and learn to love editing Wikipedia

how to stop worrying &
learn to love editing Wikipedia

- **"so-fix-it"**
  - become an editor (and more)
  - small changes, big changes, writing articles
- **understand Wikipedia culture**
  - (a) effectively represent your subjects
  - (b) change Wikipedia culture for the better
- **change Wikipedia demographics**
  - underrepresented interests, skills, and identities

How to make an article bullet-proof: 1
Become an editor

- establish yourself as an editor
  - small edits are fine! typos, refs, facts
- follow the core Wikipedia principles
  - notable subject; neutral point of view; no original research; verifiable sources; you don't own it.
  - editorial behavior: assume good faith; talk about content not editors; avoid conflicts of interest.
- when challenged
  - stay civil, and
  - seek allies & outside opinion: friendly objective editors/admins for 2d opinions

How to make an article bullet-proof: 2
Prepare your content

- Pick a topic
- Review any notability or style criteria for your subject (e.g., biographical; entertainment)
- Find a model: similar topic, handled well
  - copy the source if you are new to Wikipedia!
- Start a draft off your userpage:
  - User:*Lquilter*/*Subject (draft)*
- Establish notability: Gather 3+ independent refs with detailed coverage of your subject; OR meet other notability criteria
- Seed wikilinks in relevant articles (e.g., awards, lists, subjects)

How to make an article bullet-proof: 3
Write your article

- ○ Assert notability claim in first sentence / paragraph
- ○ Include relevant categories
- ○ Add references and a references section
- ○ Add a section for awards / 3rd-party recognition
- ○ Add a section for notable works, and in comments or talk describe why the works are notable. Add wikilinks on the notable works if they are independently notable.

How to bulletproof other projects too

- ○ Pay attention to *who's participating*
  - ■ Get the numbers
  - ■ Try to assess why
  - ■ Systemic biases should be addressed; do not let "reverse racism", "colorblind", "open-to-all" narratives dissuade you.
- ○ When projects are ostensibly open, feel ENTITLED to ensure they are ACTUALLY open.
  - ■ use the mission statement!
- ○ Address non-representative participant outcomes OPENLY
  - ■ Talk about it.
  - ■ Do not let attempts to remedy systemic biases be derailed by accusations around intentional biases or lack thereof