When Social Networks Write Science History
Priority Disputes on Wikipedia

Presenter: Arno Simons
Co-authors: Wolfgang Kircheis, Marion Schmidt, Martin Potthast, Benno Stein
Workshop: “Framing Innovation in a Networked World” (2–3 September 2021)
How are CRISPR-related priority disputes enacted on Wikipedia?
Three bits of sociology

“Ready-made science” vs. “science-in-the-making” (Latour 1987)
Wikipedia = a setting for observing the construction of facts

“Retrospective accounts” (Deuten & Rip 2000)
Interpretations of the past to orient present action towards the future
The neglected sibling of “prospective accounts”

Studying “controversies” online (Venturini 2012)
Cartography: Recursive adjustment of exploration and representation
Distant reading + close reading
Three groups of actors...
1) Our own team

Arno Simons | DZH | Political Science & STS
Wolfgang Kircheis | Leipzig U | Computer Science
Marion Schmidt | DZH | Science Studies & Bibliometrics
Benno Stein | BHU Weimar | Information Science
Martin Potthast | U Leipzig | Digital Humanities

Wikipedia Science Analytics
https://osf.io/ygur7
2) Our “CRISPR heroes”
2) Our “CRISPR heroes”

George Church
Jennifer Doudna
Feng Zhang
Emmanuelle Charpentier
3) Our Wikipedians
3) Our Wikipedians

**Age distribution**

- 13% of editors are under 17.
- 14% are in the group 18–21.
- 26% are 22–29.
- 19% are 30–39.
- 28% editors are aged 40+.

59% of the editors are aged 17 to 40.

**Gender**

The 2013 study The Wikipedia Gender Gap Revisited measured gender bias in survey completion and estimated that as of 2008, 84% of English Wikipedia editors were male. In the worldwide Wikipedia Editor Survey 2011 of all the Wikipedias, 91% of respondents were male.

**Nationality**

The greatest number, or plurality, of editors (20%) reside in the United States, followed by Germany (12%) and Russia (7%). The only country not in Europe or North America in the top 10 is India (3%).

**Language**

Most users primarily edit (76%) and read (49%) the English Wikipedia, followed by the German Wikipedia at 20% and 12%, and the Spanish Wikipedia at 12% and 6% respectively. More than half (51%) of editors contribute in two or more languages.

**Why contribute?**

- 71% of the editors contribute because they like the idea of volunteering to share knowledge.
- 69% believe that information should be freely available.
- 63% pointed out that contributing is fun.
- Only 7% edit Wikipedia for professional reasons.
Distant and close readings
Inspecting the editing landscape from far

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

“CRISPR”
(2068 revisions)

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

“CRISPR gene editing” (199 revisions)

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Only the “history” section
Inspecting the editing landscape from far

X-axis: Time (3-months slices)

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Y-axis (left): Number of character added/removed

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Only the “history” section

All sections taken together

Y-axis (right): Number of references
Inspecting the editing landscape from far

Launch & growth of history section

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

"Hot" phases

Only the “history” section

All sections taken together
Inspecting the editing landscape from far

Only the “history” section

Migration to new CRISPR entry

All sections taken together
Plotting heroes from far
Plotting heroes from far
Plotting heroes from far

"CRISPR gene editing"
Plotting heroes from far

\[
\text{age} = \frac{\text{num of revisions since first occurrence of name}}{\text{num of all revisions}}
\]
Plotting heroes from far

prominence = total num of occurrences of name across all revisions / num of revisions since first occurrence of name
Plotting heroes from far

```
controversiality = \frac{\text{sum of neg change ratios in name freq from rev to rev}}{\text{X-axis}}
```

---

**CRISPR** (C1)

- Mojica
- Koonin
- Sontheimer
- Mannaffini
- Gasiunas
- Brouns
- Yang
- Xu
- Liang
- Horvath
- Barrangou
- Ishino
- Church
- jinek
- Gootenberg
- Kawamura
- Mulepati
- Myhrvold
- Nishimasu
- Shibata
- DiCarlo
- Liao
- Schaefer
- Stemberg
- Yin

**CRISPR gene editing** (C2)

- Doudna
- Charpentier
- Zhang
- Yang
- Siksnys
- Jia
- Kuan

**Notes:**

- Endurance99
  - 0.1
  - 0.5
  - 1.0

---

**Prominence**

- 0.0
- 0.5
- 1.0
- 1.5
- 2.0
- 2.5
- 3.0
- 3.5
- 4.0

**Controversiality**

- 0.0
- 1.0
- 2.0
- 3.0
- 4.0

**Age**

- Older
- Younger
Plotting heroes from far

endurance = $\frac{\text{num of revisions containing name}}{\text{num of all revisions}}$

"CRISPR" (C1)

"CRISPR gene editing" (C2)

Prominence

Controversiality

Age
Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

The four scientists on our earlier slide take different positions
Plotting heroes from far

"CRISPR" (C1)

Less controversial but quite prominent

"CRISPR gene editing" (C2)
Plotting heroes from far

"CRISPR" (C1)

"CRISPR gene editing" (C2)

Quite controversial but less prominent
Reading from close

Three controversies

- Discovery of repeating DNA segments (CRISPR)
- Discovery that CRISPR is an immune defense system
- Development of CRISPR “genetic scissors”

Wikipedia editors negotiate their (different) accounts of “CRISPR Heroes”

- In and out of names and claims

Wikipedia’s rules and policies used for (de–)legitimation of edits

- WP:MEDRS (Reliable sources)
- WP:PROMO (Self promotion)
Conclusions

Wikipedia not just a provider but also a *producers* of knowledge

CRISPR entry/revision = a “retrospective account” in its own right

Wikipedia works as a filter for relevance

Claims, Names, Sources

Knowledge claims emerge from “socio-textual networking”

Words, Reputation, Policies

**Adaptive** tool development worked well for us

Flexibility to iterate between distant and close reading

Reflexivity: We also produced a retrospective account (of how Wikipedians account for CRISPR)
Thank you for listening!

Wikipedia Science Analytics

https://osf.io/vgur7

Arno Simons, Wolfgang Kircheis, Marion Schmidt, Martin Potthast, Benno Stein