Towards a diversity-minded Wikipedia

Fabian Flöck, Denny Vrandecic, Elena Simperl – KIT Karlsruhe
The loss of viewpoint diversity

Organizing for America | BarackObama.com
Organizing for America is the grassroots organization that supports President Obama's agenda for change. Visit BarackObama.com for blog posts, local events, ... www.barackobama.com/


Barack Obama - Wikipedia, the free encyclopedia
Barack Hussein Obama II is the 44th and current President of the United States, having taken office in 2009. He is the first African American to hold the ... en.wikipedia.org/wiki/Barack_Obama

Barack Obama Biography - Biography.com
17 Mar 2011 ... Learn more about the life and presidency of Barack Obama on Biography.com. Watch videos, see exclusive photos and more! ... www.biography.com/articles/barack-obama-12782369

Barack Obama's popularity soars - in Germany - The New York Times
6 Jan 2008 ... BERLIN — Barack Obama’s popularity extends far beyond the heart of Central Europe. Germany has swiftly developed a serious ... www.nytimes.com/2008/01/06/world/europe/06obama-berlin.4.9043329.html

YouTube - Barack Obama Schwäbisch - Rede Berlin - dodokam
11. Jan. 2009 ... I am sure Barack would least approve of all your bitching. .... Barack Obama Berlin Speech Part1by PieReportNews 119025 views ... www.youtube.com/watch?v=aF3qxtO7qzw
Wikipedia is biased

- To what extend does Wikipedia represent all relevant PoVs?

- Neutral-point-of-view policy

- Examples: Conservapedia, Croatian and Serbian editions
Leveraging diversity in Wikipedia

Use Case - In collaboration with Wikimedia Germany:
Displaying warnings when detecting patterns of bias in editing behavior

Challenge:
Identify, understand and predict socio-technical mechanisms leading to bias
Context: Slowing growth and content saturation

- Less new articles, less edit growth since 2007
- Consolidation of articles
- No room for easy contribution

Figures by Suh et al.
“The Singularity is not near: Slowing growth of Wikipedia”
Socio-technical mechanisms leading to bias

- Consensus: social proof and consolidation
- Ownership behavior
- Opinion camps and editor drop-out
- Boldness and useful conflicts
- Bureaucracy and implicit social norms
- Personal characteristics of authors
- WikiProjects and other groups
Building a bias prediction model

1. Identify typical patterns for a socio-technical mechanisms

2. Evaluate if these patterns are typical for articles marked as “biased”

3. Train machine learning algorithms to predict bias using the patterns
1. Finding patterns:
   o Ownership behavior
     • Typical patterns in “Maintained” articles
       ▪ High # of reverted newcomers
       ▪ High concentration of edits, reverts
       ▪ Core high-activity-editor group
       ▪ ...
     • Methods
       ▪ Statistical analysis on revision history, etc.
       ▪ SNA
   o .. And for all the listed mechanisms
Conclusion

- Building a prediction model as complete as possible
- Converge with other models developed for Wikipedia in RENDER

→ Build tools that help the shrinking number of editors to cope with the work overload of finding and curing bias