

## Bachelor/Master Thesis

### Assessing Gender Inequality in User Access Behavior on Wikipedia

With more than 5 million articles connected by roughly 400 million links, Wikipedia—a free online encyclopedia—is a comprehensive source of information on the Web. The social significance of Wikipedia has been recognized by users from all over the world, some of which donate on a regular basis [1]. Moreover, Wikipedia receives a significant amount of Web traffic and is one of the most visited websites [2]. The diversity, amount of content and users, and often the collaborative nature of content production make it, however, difficult for Wikipedia not only to ensure efficient information access but also “equal” information access. For example, articles describing male and female personas. A recent study by Wagner et al. compared the gender inequality of Wikipedia articles with respect to their coverage, lexical, structural and visibility biases [3]. A natural continuation of this research is the assessment of gender inequalities from the Wikipedia user perspective and study their information access behavior i.e. *search* and *navigation* on *male* and *female* Wikipedia articles.

As part of the thesis, the plausibility of the following hypotheses has to be statistically tested using large-scale log data capturing *search* and *navigation* access to Wikipedia:

**Hypothesis 1:** Given women are in the network periphery and Wikipedia users tend to navigate to the periphery in general, women are predominantly accessed through navigation.

**Hypothesis 2:** Given men are in the network core and more popular in general, they are predominantly accessed through search.

The second part of this thesis aims to characterize Wikipedia gender inequalities in social media platforms, i.e., Twitter. Here the goal is to assess the posting and access behavior of social media users with respect to gender inequalities on Wikipedia.

**Hypothesis 3:** Men Wikipedia articles are more often shared/posted and accessed than female Wikipedia articles.

Ideally, the insight this works gives into the user behavior with respect to gender inequalities would help the Wikimedia Foundation in its efforts to counter systematic biases on Wikipedia [4].

#### Contact:

Dr. Dimitar Dimitrov  
 GESIS - Leibniz Institute for Social Sciences  
 Unter Sachsenhausen 6-8, 50667 Cologne  
 E-Mail: [dimitar.dimitrov@gesis.org](mailto:dimitar.dimitrov@gesis.org)  
 Phone: +49 (221) 47694-512

Prof. Dr. Stefan Dietze  
 Heinrich Heine University Düsseldorf - Data & Knowledge Engineering  
 Universitätsstr. 1 Building: 25.12 Floor/Room: 01.39, 40225 Düsseldorf  
 E-Mail: [stefan.dietze@hhu.de](mailto:stefan.dietze@hhu.de)  
 Phone: +49 (211) 81-13785

#### Introductory Readings:

[1] [https://wikimediafoundation.org/wiki/Fundraising\\_reports](https://wikimediafoundation.org/wiki/Fundraising_reports) Accessed: 2018-04-17

[2] <https://www.alexa.com/siteinfo/wikipedia.org> Accessed: 2018-01-12

[3] Wagner, Claudia, et al. "It's a man's Wikipedia? Assessing gender inequality in an online encyclopedia." *Ninth international AAAI conference on web and social media*. 2015.

[4] [https://en.wikipedia.org/wiki/Wikipedia:WikiProject\\_Countering\\_systemic\\_bias](https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Countering_systemic_bias) Accessed: 2018-01-12