# Introduction

Gender discrimination is a problem, especially in the technology industry. A recent study by XYZ looked at GitHub, one of the most popular software repositories with 14 million users, to see if the effect can be seen by comparing pull requests between male and female contributors. A pull request happens when someone proposes changes on a software code repository that is being hosted on GitHub. The request can either be accepted or denied by the owners of the repository. The researchers found that the approval rates of pull requests, when made without revealing ones gender, were comparable between both genders. The rate of approval for women fell significantly if the coder’s gender was identifiable upon requesting a pull whereas the rates for men did not change much.

Wikipedia, the largest encyclopedia mankind has ever created, also has a gender problem. Estimations about exact numbers are difficult to find but it is undisputed, that the vast majority of contributors and even editors are of male gender. Especially the latter is important since editors, specifically veteran editors, do have the powers to reject articles and restrict access of other contributors which has a severe impact on the acceptance rate regarding articles that are written about female figures. The resulting bias with respect to the variety of content is increasingly recognized leading to Wikipedia’s gender gap to be an ever increasing topic of interest, drawing the attention of scholars of new media and similar fields. Discussing the systemic bias on Wikipedia is important as the platform poses the most powerful source of information worldwide. It is available in 275 languages and is being accessed by millions of people regularly. Many assistive devices and applications such as Smartwatches, Google Assistant, Alexa, Siri or Cortana draw their knowledge directly from Wikipedia without informing their users about the source or authors. Through these developments, the world, as it is represented by Wikipedia, is widely being considered as natural.

In this work, we address gender bias discrepancies within the technology industry by providing facts on historical developments and current events. Additionally, we try to measure possible trends for the future of digital environments that have collaborative characteristics, draw implications for the quality of publicly available information and participation rates and provide suggestions that may have an impact on future developments.

# Considerations

Most studies on gender bias have been conducted either in Europe or in the United States. Only very few researchers actively work on that topic from other parts of the world. The participants of the studies conducted are more likely to be people from those geographical regions. Thus, their provided data and insights inherently carry strong biases and are not globally representative. A lot of studies in this area are user studies, which poses an issue in our specific case as it can be shown, that women are overall less likely to participate in user studies compared to men. As an example, this phenomenon has an impact on the data used in Figure 1 as the underlying studies could be skewed accordingly. It therefore must be expected that numbers of female editors in Wikipedia, as shown in Figure 1, are even lower in practice.

# Wikipedia

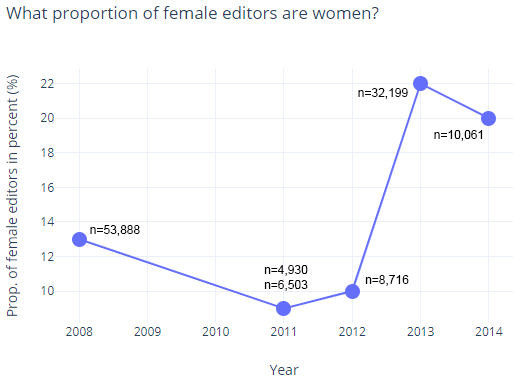


Figure 1: 6 Studies were conducted by the Wikimedia Foundation between 2008 and 2014 to measure the proportions between male and female editors that are active on Wikipedia. The corresponding number of participants is annotated for each study accordingly.

The Wikimedia Foundation conducts various user studies at annual intervals, and the data thus collected is made publicly available. Part of these studies is also to determine the ratio of female to male editors on Wikipedia. We analyzed the results of those studies to see how the ratios have changed over time. We found that women are strongly underrepresented on Wikipedia, even in the strongest estimate. However, a positive trend can also be observed from the year 2011 onward.