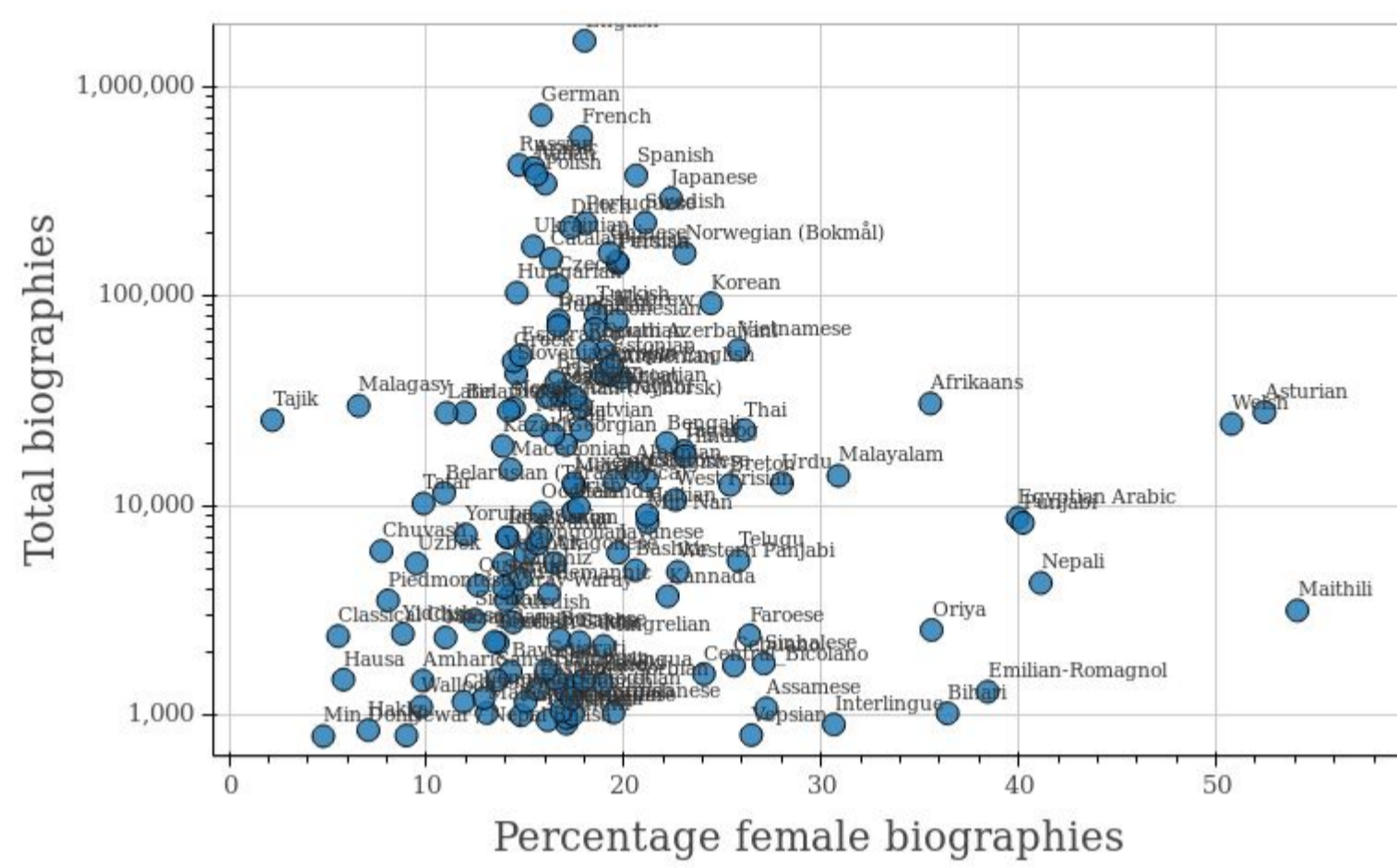


# Assessing the diversity in the content about humans in Wikimedia projects

Wikidata is a wonderful tool when looking for metadata about the content of Wikimedia projects. Over the years, several tools using Wikidata were created to understand the diversity of biographies in Wikimedia projects, mostly dedicated to the gender gap: [w.wiki/8zr](http://w.wiki/8zr)

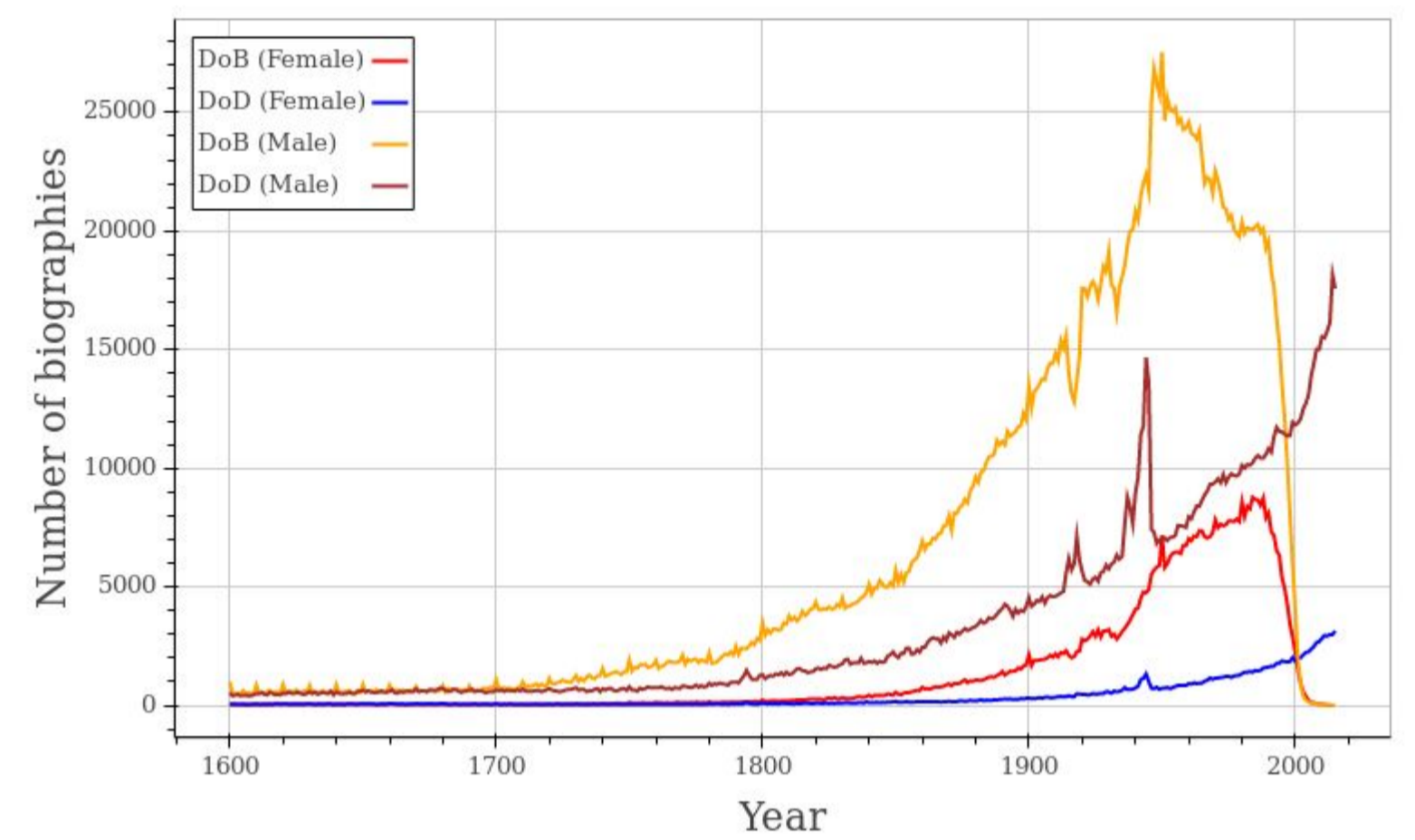
Two of these tools are:



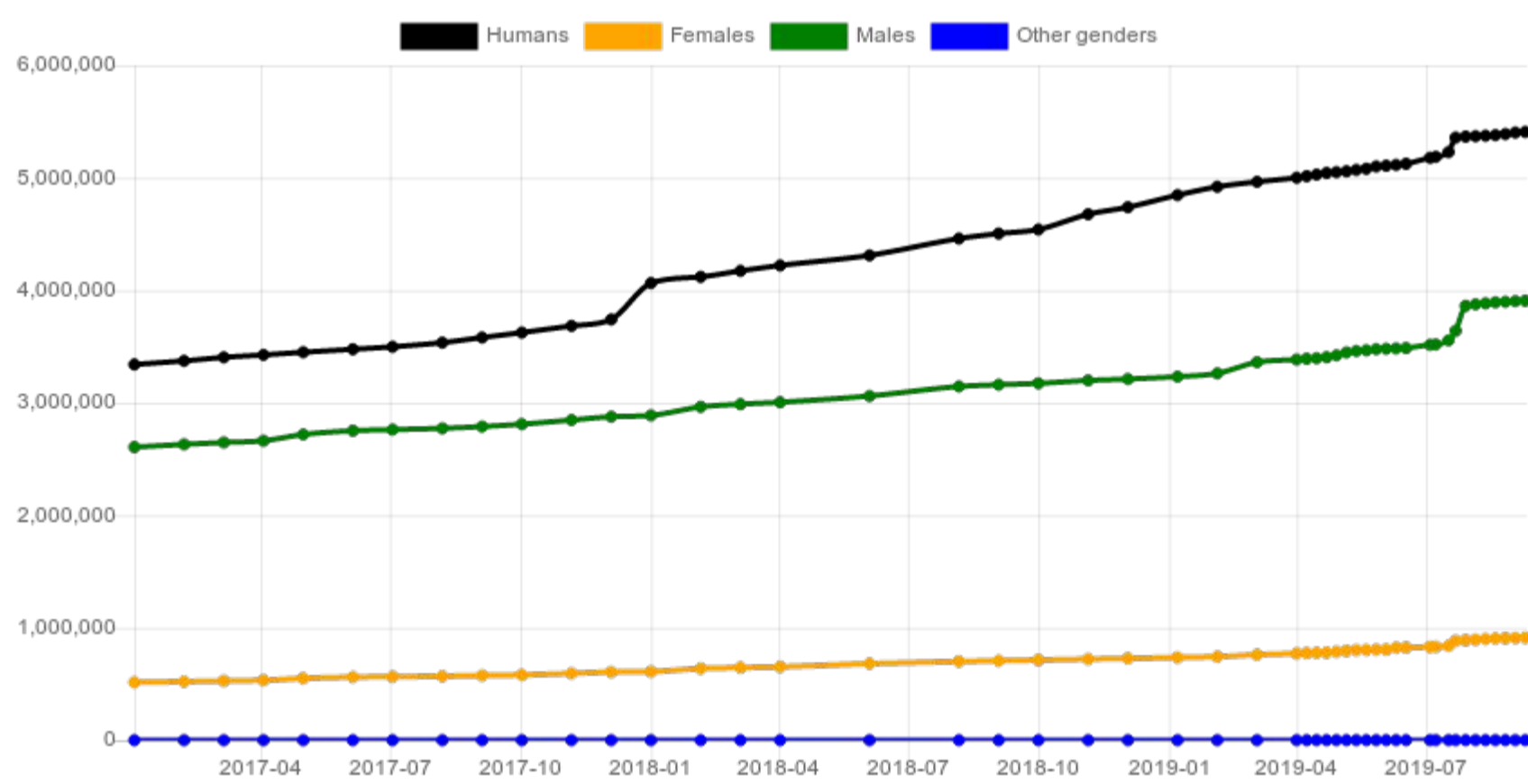
Number of biographies / percentage of female biographies in each Wikipedia edition

## 2014: Wikidata Human Gender Indicators (WHGI)

- [whgi.wmflabs.org](http://whgi.wmflabs.org)
- Author: Maximilian Klein
- MIT / CC BY-SA 4.0



Number of biographies in Wikipedia by year of birth and year of death of each person



Evolution of the number of items depicting humans in Wikidata

## 2017: Denelezh

- [www.denelezh.org](http://www.denelezh.org)
- Author: Envel Le Hir
- AGPL v3 / CC BY 4.0

Country	Total	Total with gender	Females	% Fem.	Gap	Males	% Mal.	Others	% Oth.
United States of America	341,326	341,323	65,004	19.045 %		276,024	80.869 %	295	0.086 %
United Kingdom	114,705	114,703	18,293	15.948 %		96,363	84.011 %	47	0.041 %
France	55,052	55,050	8,971	16.296 %		46,067	83.682 %	12	0.022 %
Canada	52,986	52,985	10,780	20.345 %		42,171	79.590 %	34	0.064 %
Germany	52,485	52,482	9,034	17.214 %		43,435	82.762 %	13	0.025 %
United Kingdom of Gre...	50,643	50,643	4,526	8.937 %		46,112	91.053 %	5	0.010 %
Australia	41,425	41,424	8,848	21.360 %		32,553	78.585 %	23	0.056 %
India	31,832	31,832	6,180	19.414 %		25,633	80.526 %	19	0.060 %
Japan	28,130	28,130	6,669	23.708 %		21,448	76.246 %	13	0.046 %
Italy	25,799	25,799	4,319	16.741 %		21,475	83.240 %	5	0.019 %
Russia	21,799	21,799	4,139	18.987 %		17,660	81.013 %	0	0.000 %
Spain	21,773	21,773	3,894	17.885 %		17,867	82.060 %	12	0.055 %
Norway	17,711	17,711	3,535	19.959 %		14,175	80.035 %	1	0.006 %
Sweden	17,452	17,452	4,391	25.160 %		13,054	74.799 %	7	0.040 %
Soviet Union	17,248	17,248	3,322	19.260 %		13,926	80.740 %	0	0.000 %
New Zealand	16,966	16,966	3,563	21.001 %		13,396	78.958 %	7	0.041 %

Gender Gap by country of citizenship in the English Wikipedia

## 2020: a tool to rule them all?

In order to share our efforts, we want to merge our tools. The goal is to assess the diversity in the content about humans in Wikimedia projects. We want to:

- keep the existing features of both tools,
- build an architecture to easily generate new reports
  - about gender gap
  - about diversity in general

Task on Phabricator: [T230184](https://phabricator.wikimedia.org/T230184)

### Contact

Envel Le Hir  
[envel@lehir.net](mailto:envel@lehir.net)  
[www.lehir.net](http://www.lehir.net)

