

# SEX-DISAGGREGATION OF SOCIAL MEDIA POSTS

*Understanding sustainable development through a gendered lens*

Global Pulse collaborated with Data2X and the University of Leiden to develop and prototype a tool to infer the sex of users. The tool automates the process of looking up public information from Twitter profiles, in particular the user name and profile picture. Using open source software, the tool analyses user names from a built-in database of predefined names (from sources such as official statistics) that contain gender information. User name alone may sometimes not be enough to discern sex, in which case the tool analyses profile photos, using face recognition software. The tool was tested on more than 50 million Twitter accounts from around the world to understand the different concerns and priorities of women and men on topics related to sustainable development.

## WHY SEX-DISAGGREGATION?

Data are key for informing decision-making and improving accountability. The availability of new data sources provides new opportunities to drive sustainable development and humanitarian action. Social media, in particular, provide new avenues for monitoring public perceptions and measuring development priorities and impact.

Disaggregation is central to the success of working with data. It can provide insights into the differences and inequalities between people on the basis of characteristics such as income, gender, age, race, ethnicity, migratory status, disability, or geographic location.<sup>1</sup> Sex-disaggregation can play an important role in providing information about the disparities between men and women, boys and girls.

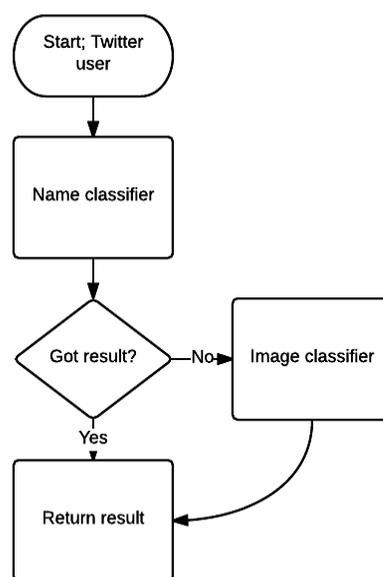
However, data from open social media channels such as Twitter may not always provide an indication of a person's sex. Together with Data2X and Leiden University's Centre for Innovation, Global Pulse developed open source software for sex-disaggregation of social media posts.

## WHAT ARE THE OBJECTIVES?

To develop open source software that can infer the male/female ratio of social media users and thus enable real-time analysis of differences in the life-experiences of girls, boys, women and men.

## HOW DOES THE TOOL WORK?

A key criterion in the development of this tool was to ensure that the approach can be applicable at global scale and across different languages. The tool disaggregates social media posts using a number of automated "classifiers," in a waterfall approach—



starting with the classifiers with the highest success rate and, when results are unknown or indecisive, moving on to another classifier. The validity of individual classifiers and overall results were assessed through a crowdsourced verification mechanism.

In the course of developing this tool, the classifier, which yielded the greatest success was name classification, where a

user's name is compared to a 'name dictionary' showing whether the content of the name was more likely to indicate a female or male.

The results could be further improved by using country-specific name dictionaries. This would require a more complex process of first determining the home country of a specific user. Since exact geo-location is often omitted in tweets, the approach would adopt a separate script for classifying the geo-location of a user, after which a country-specific dictionary could be used. If a country-specific dictionary is absent or results are indecisive, disaggregation will take place by relating a user's name either to language-specific or, subsequently, an overall overview of a number of dictionaries. The name classification process and script

<sup>1</sup> Zero draft: Transforming our world by 2030: A new agenda for global action

used in this tool built upon the code of 'Gender Computer' developed by TU Eindhoven.<sup>2</sup>

In addition to name classification, image recognition of a user's profile picture demonstrated good results for sex-disaggregation.

In this method, the script classifies a user's profile picture with the free-to-use tool Face++. However, if multiple persons are identified in the same photo, the results can be inconclusive. For the purposes of this prototype, the face that the algorithm concluded to be the most prominent in the picture was selected.

The script for sex-disaggregation of social media accounts is open-source and readily available.<sup>3</sup> Additionally, an online version of the tool itself has been made available for determining the sex of a person on the basis of their Twitter user name, first name, or an image URL.

## ACCURACY

To test the accuracy of the waterfall method of first deploying the name classifier followed by image recognition, a public website was created. The website allows users to manually determine whether a certain Twitter account is male or female.

The accuracy of the hybrid classification approach (the waterfall combination of name and image recognition) was compared with the crowdsourced results, which were assumed to be correct. The automated classification approach accurately determined sex in 74% of cases.

Crowd-Sourced as	Classified by Hybrid Classifier				
	Male	Female	Unknown	Unisex	Total
Male	736	28	53	3	820
Female	50	477	43	2	572
Other	206	44	42	8	300
Male	Precision = <b>74%</b>		Recall = <b>90%</b>	Accuracy = $(736+477+42)/(820+572+300)$ = <b>74%</b>	
Female	Precision = <b>87%</b>		Recall = <b>83%</b>		

<sup>2</sup> The code of Gender Computer (TU Eindhoven) can be accessed here: <https://github.com/tue-mdse/genderComputer>. For the overall script, we updated several of the dictionaries with new names and included additional country specific dictionaries.

<sup>3</sup> GitHub repository: [https://github.com/LU-C4i/gender\\_classifier](https://github.com/LU-C4i/gender_classifier)

## DATA PRIVACY

The methodology uses publicly available data from Twitter profiles. Moreover, only gender markers such as the name and profile pictures of users were applied to building the tool. Users for whom the name and profile picture were insufficient to allow the classifier to detect sex were categorized as unknown.

## POTENTIAL APPLICATIONS: A CASE STUDY

The tool could potentially be used for any study of tweets and other types of social media expression wherein the name and/or profile picture of users are available.

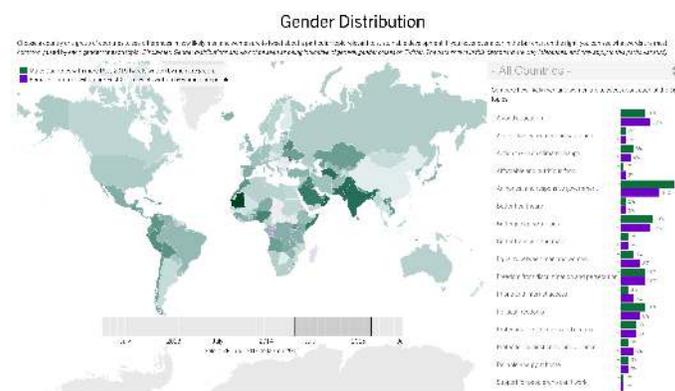


Figure 1: Screenshot from online dashboard showing how women and men tweet about global development topics.

For example, Global Pulse used the sex-disaggregation tool to improve an existing real-time online dashboard showing the volume of tweets around priority topics related to sustainable development. By filtering through 500 million daily tweets from over 50 million accounts for 25,000 keywords relevant to global development topics, this interactive dashboard showed which countries tweeted most about which topics between May 2012 and July 2015.<sup>4</sup>

In order to further refine the dashboard, the gender classification script was run over the entire dataset.

Once disaggregated by sex, the dashboard revealed new insights, highlighting the different concerns and priorities of women and men. For example, in Nepal, the sex-disaggregated data showed that women tweeted more on topics related to 'equality between men and women' and 'freedom from discrimination'. The second most tweeted topic was that of 'a good education'. In comparison, men discussed most about 'protecting forests, rivers and oceans', followed by 'a good education'. In the second quarter of 2015, discussions were dominated by 'support for people who cannot work'—a topic rarely mentioned at other times—and 'an honest and responsive government'. These discussions were prompted

<sup>4</sup> The project was initially developed by Global Pulse in collaboration with the UN Millennium Campaign and DataSift: <http://post2015.unglobalpulse.net/>

following the earthquake that hit Nepal on 25 April of that year. The above topics were widely mentioned by both men and women.

## CONCLUSIONS

The plight of women and men differ in many ways and one step towards understanding those differences is sex-disaggregation of available data. By embarking on the development of this prototype tool, and testing it on Global Pulse’s dashboard of global development tweets, the tool shows several early examples of the insights that can be gleaned about the differences between how men and women discuss global development on social media.

This tool may be the first open-source solution for inferring users’ sex from tweets (although many commercial tools exist). It is recommended that work on this open-source methodology be improved to allow application at a global level. Specifically, efforts should be made to include high-quality name data from more countries.

Gender disaggregation of social media content is just one important step. Future efforts to classify social media content by other markers—for example, inferring the age of social media users—could help further support and measure the achievement of the Sustainable Development Goals.

### HOW TO CITE THIS DOCUMENT:

UN Global Pulse, “Sex Disaggregation of Social Media Posts,” Tool Series, no. 3, 2016.