Damegender: Writing and Comparing Gender Detection Tools

David Arroyo Menéndez

June 8, 2020

Presentation

- Thesis Student: David Arroyo Menéndez
- Title: Writing and Comparing Gender Detection Tools from a Name
- Thesis Director: Jesús González Barahona
- Objectives:

Compare APIs to detect gender from a name

Apply machine learning to classify nicknames, diminutives or new names as male or female

Understand the impact about open data in census about names and gender versus datasets created by companies or free software communities.

Apply this tools to the scientific communities or communities based on Internet.

Activities

We have presented this working in progress in:

Scientific events:
- Madrilenian Software Research
- Group Retreat 2019 Workshop

Event to master students:
- Periodismo de Datos (Medialab Prado)

Industrial events:
- Python Barcelona
- Open South Code

Results

Software

Free Software released with GPLv3 integrated in the industry

- git clone https://github.com/davidam/damegender.git
- pip3 install damegender (more than 3000 downloads by month)

Publications

- Damegender: Writing and Comparing Gender Detection Tools (Preprint)
- Damegender Manual: Counting Males and Females in Internet Communities (On working)