Online-SoBA Testing Guide

Contextualization:

This project, *Online-SoBA (Online Social Behavior Analysis)*, is a semi-automatic system capable of analyzing short texts corresponding to comments written in Portuguese in reaction to '*posts*' on social media platforms (namely social networks and online newspapers), so that it is possible to identify behaviors that define social opinions at a given time.

To face this challenge, the texts contained in the NetLang Corpus are analyzed, extracting information on the topic and the polarity of each one.

The calculation of the polarity of a comment is carried out through the number of words with positive or negative connotation that constitutes it. Subsequently, according to the sentence structures that characterize these behaviors in natural language, the theme of the comment is assigned.

By analyzing the resulting graphs of this process, we can quickly make a visual deduction about how social opinion on an issue has evolved over time.

1. First Phase (Sources):

- Go to Online Social Behavior Analysis;
- On the top of the page, select *Source Analysis*, to view the distribution graph of comments, grouped by the respective source of extraction;
- Draw your conclusions by manipulating the graph (it is possible to drag/zoom with the tools in the lower left corner, as well as disable each source individually by clicking on the graph's legend);
- Click on *Contributing Files* and set the *Time Interval* between 2020/01 and 2020/08;
- Click on the name of the first file for a more detailed view of it (this is similar to using the *Search by File* mode).

2. Second Phase (Themes):

2.1. Predefined Themes:

- Select *Search by Theme* at the top of the page, to see information about the comments previously identified as containing a certain theme;
- From the dropdown menu, select *Sexualidade – Homosexualidade*;
- Explore the 3 graphs (*Polarities*, *Sources*, *Theme vs NetLang*) containing information on the topic under analysis and draw your conclusions (when in doubt, all graphs have tooltips with short descriptions of their content);
- Click *Contributing Files* and set a time range so that you can see which files contributed the most comments in that range.

2.2. Custom Themes:

- To change the search mode and look for a theme that does not exist yet, in the menu where you selected the theme you will click the search mode *Regex*;
- Enter the regular expression: *andr.*ventura*;
- Notice that the *Regex* search mode presents results in a similar way to the previous one but has the advantage of allowing you to define your own themes through regular expressions;
- Explore the different graphs and draw your own conclusions.

3. Third Phase (Files):

- Select at the top of the page *Search by File*, to see information about each file that is contained in the corpus;
- From the dropdown menu select *Sol_extraction_english_145.json*;
- Explore the 4 modes of viewing information about this file (note that clicking on a theme in the *Present Themes* graph is similar to using the *Search by Theme* mode);
- After you have made a quick analysis of the original post's content and seen some of the comments, ask yourself whether the main themes being detected by the program make sense in the given context.

4. Fourth Phase:

- Answer the questionnaire available at https://forms.gle/DKJdgVEYAWk337VL9