Placing M-Phasis on the Plurality of Hate: A Feature-Based Corpus of Hate Online

Dana Ruiter1*, Liane Reiners2*, Ashwin Geet D'Sa3,
Thomas Kleinbauer1, Dominique Fohr3, Irina Illina3, Dietrich Klakow1,
Christian Schemer2, Angeliki Monnier4
{druiter, kleiba, dietrich}@lsv.uni-saarland.de
{liane.reiners, schemer}@uni-mainz.de
{ashwin-geet.dsa, dominique.fohr, irina.illina}@loria.fr
angeliki.monnier@univ-lorraine.fr
* Equal contribution.

Motivation and Concept
Most annotated hate speech (HS) corpora are binary. This ignores the complex and subjective nature of hate. M-PHASIS is a novel HS corpus, based on a collection of user comments which...
- are not sampled based on slur-words.
- capture implicit hate.
- match migration-related regular expression keywords.
- are derived from a diverse set of outlets.
- are in French and German.
- are annotated based on features.
- descriptors of various types of hate.
- are collected with context.

Data Collection
- Period: January–May 2020
- Content: News articles and comments from German and French mainstream and fringe news outlets.

<table>
<thead>
<tr>
<th>Outlet (fr)</th>
<th>#Art.</th>
<th>#Com.</th>
<th>Outlet (de)</th>
<th>#Art.</th>
<th>#Com.</th>
</tr>
</thead>
<tbody>
<tr>
<td>France Info (fr)</td>
<td>20</td>
<td>618</td>
<td>Tagesschau (de)</td>
<td>13</td>
<td>1,020</td>
</tr>
<tr>
<td>Figaro (fr)</td>
<td>19</td>
<td>1,056</td>
<td>Welt (de)</td>
<td>74</td>
<td>736</td>
</tr>
<tr>
<td>Le Monde (fr)</td>
<td>16</td>
<td>554</td>
<td>Zeit (de)</td>
<td>15</td>
<td>999</td>
</tr>
<tr>
<td>Valeurs Actuelles (fr)</td>
<td>30</td>
<td>614</td>
<td>Focus (de)</td>
<td>13</td>
<td>962</td>
</tr>
<tr>
<td>AgoraVox (av)</td>
<td>11</td>
<td>369</td>
<td>Compact (co)</td>
<td>12</td>
<td>282</td>
</tr>
<tr>
<td>Riposte Laïque (fr)</td>
<td>35</td>
<td>1,435</td>
<td>Epoch Times (et)</td>
<td>2</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Junge Freiheit (fr)</td>
<td>55</td>
<td>1,374</td>
</tr>
<tr>
<td></td>
<td>131</td>
<td>4,646</td>
<td></td>
<td>187</td>
<td>4,821</td>
</tr>
</tbody>
</table>

Overview of number of news articles (#Art.) and comments (#Com.) collected from both mainstream (top) and fringe (bottom) media outlets in French (fr) and German (de) for inclusion in the dataset.

Target Analysis

Agent Evaluation Victim

Agent-Evaluation-Victim sankey for the most common agent, politician, in the German portion of the dataset. We show the 5 most common evaluations/victims respectively.

→ Decision makers most negatively evaluated.

An annotation pipeline for HS features in user comments across five modules (c_ne, c_pe, c_act, c_contr, c_emo). When a comment fulfills the requirement (y) of a module (e.g., contains negative evaluation for c_ne), follow-up categories (gray boxes) are annotated, otherwise (n) we skip to the next module.

Contributions
- Creation of feature-based HS corpus on French and German news outlet comment.
- Resulting annotations with high levels of inter-annotator agreement (0.77 ≤ κ ≤ 1).
- Target analysis showing...
  - politicians more criticised than migrants.
  - French action recommendations more radical.
- Domain analysis showing...
  - clear divide between mainstream vs. fringe media comments.
- Error analysis and discussion showing limitations of dataset (e.g., some annotation irregularities).

https://github.com/uds-lsv/mphasis