The data was anonymized, and only paraphrased quotes of social media data are used. Multiple coauthors hold different LGBTQ+ identities, enabling us to incorporate the sensitivity and contextualize our findings in the light of our lived experiences.

**LGBTQ+** is an acronym used to refer to individuals who use these sexual or gender labels as personally meaningful for them.
MENTAL HEALTH OF LGBTQ+ INDIVIDUALS

- **3 times** more likely than others to experience a mental health condition.
- **2.5 times** more likely to experience depression.
- **3 times** more likely to report suicidality and attempted suicide.

Mustanski et al., 2010
LGBTQ+ MINORITY STRESS

- Stigma, prejudice, and discrimination create hostile and stressful social environment
- Minority Stress Theory (proposed by Ilan H. Meyer in 1995)

_Minority stress theory_ proposes that LGBTQ+ mental health disparities can be explained largely by stressors induced by a hostile, homophobic, and transphobic culture, which often leads to harassment, mistreatment, and victimization. This may ultimately impact access to care.
SOCIAL MEDIA

- Widespread prevalence of social media use among LGBTQ+
- Stigma, self-disclosure, and support-seeking on social media
- Anonymity, Audience, and Sense-of-Belonging in Community
OUR WORK

... a **codebook** to identify minority stressors in social media per Meyer’s minority stress theory.

... a **machine learning classifier** to identify social media posts on minority stress

... a **lexicon** of social media language markers contextualized in minority stress theory
LGBT is not now, or ever will be, accepting pedophiles.

Hello all people. There has been recently a lot of posts, caused by some rather questionable images claiming that LGBT is adding a P for pedophiles, or the MAP stuff. This could not be further from the truth. This is a hoax being spread by 4chan, as you can see on the snopes page here. However, due to this hoax, we have been receiving multiple posts a day, either claiming we are, asking are we, or simply repeating our denial. To this effect we are putting up this sticky, and asking people not to keep posting such, further posts will be removed and referred to this sticky. thank you.
DATA AND STUDY

- 12,645 posts on r/lgbt subreddit
  - January 2016 to May 2018

- **Codebook** on minority stress experiences on social media
  - 350 randomly sampled posts were qualitatively inspected

- Machine learning **classifier** on minority stress experiences on social media
  - Used to classify the entire sample of data

- **Linguistic Cues** associated with the language of minority stress
CODEBOOK: ANNOTATION APPROACH

Drawn from Meyer’s Minority Stress Theory:

- Difference in time and context
- Differences in identity
- Differences in data and medium of communication

Prejudice Events | Perceived Stigma | Internalized LGBTphobia
Prejudice Events

Explicit actions of rejection toward an individual related to their LGBTQ+ identity.

- Verbal and Physical Violence
- Legal Discrimination
- Anxiety / stress about discriminatory political powers
PERCEIVED STIGMA

Internalized fears and anxiety of expected rejection from others

- Individuals constantly monitoring their behavior: how one dresses, speaks, walks, and talks in contrast to expected social norms.
- Lying to cover up identity
- Thinking others do not or will not “accept” them
- Feelings of not “fitting in”
INTERNALIZED LGBT-PHOBIA

Internalization of negative societal attitudes and negative feelings that come from self about self.

- Self-hate for being a part of the LGBTQ+ community.
- Feelings of negativity rooted from LGBTQ+ identity.
- Feelings of isolation or of ending up alone.
- Hate/distaste for others within LGBTQ+ community.
OBSERVATIONS

- We see a wide range of minority stressors
- Some minority stressors are perpetuated by some subsets of the LGBTQ+ population to other subsets,
- **Sense of belonging** within the **community**
Machine Learning Classifier
MACHINE LEARNING FEATURES

- Latent Semantics (Word Embeddings)
- Psycholinguistic Attributes (LIWC)
- Hate Lexicon
- Open Vocabulary ($n$-grams)
- Sentiment
- Mental Health Concerns (DASS: Depression, Anxiety, Stress, and Suicidal Ideation)
# MACHINE LEARNING PERFORMANCE

<table>
<thead>
<tr>
<th>Model</th>
<th>Pr.</th>
<th>Rc.</th>
<th>F1</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>0.70</td>
<td>0.49</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>Logistic Reg.</td>
<td>0.73</td>
<td>0.72</td>
<td>0.72</td>
<td>0.76</td>
</tr>
<tr>
<td>SVM (Linear)</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td>0.77</td>
</tr>
<tr>
<td>Random F.</td>
<td>0.76</td>
<td>0.67</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>0.73</td>
<td>0.72</td>
<td>0.72</td>
<td>0.77</td>
</tr>
<tr>
<td>MLP</td>
<td>0.75</td>
<td>0.74</td>
<td>0.75</td>
<td>0.80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>F1</th>
<th>Acc.</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>W2V</td>
<td>0.67</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>.+LIWC</td>
<td>0.71</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>.+.+HateLex</td>
<td>0.72</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>.+.+.+Ngrams</td>
<td>0.73</td>
<td>0.72</td>
<td>0.78</td>
</tr>
<tr>
<td>.+.+.+.+Senti</td>
<td>0.74</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>.+.+.+.+.+DASS</td>
<td>0.75</td>
<td>0.75</td>
<td>0.80</td>
</tr>
</tbody>
</table>
ANALYZING LANGUAGE CUES

- Use Minority Stress Classifier to infer the presence of minority stress in the entire dataset
- Obtain **discriminating keywords** of presence / absence of minority stress expressions using SAGE\(^1\)

---

Keywords related to sharing and disclosing to those who are close to the discloser (e.g., *tell mom*, *telling people*, *started talking*)

“I’m a coward to be unable to tell my mom to her face that I’m gay.”
Keywords that express self-experiences and life events (e.g., *started feeling, days later*, etc.)

“*started feeling these issues when I was 16 or 17, and because of my parents, I was like, “No I won’t do this”, and shoved them back down again.*”
Keywords relating to friendliness and sense of community (e.g., *lgbtq community, lgbt friendly*, etc.)

“I thank the LGBTQ+ community for being together, through friends and social media.”
# ERROR ANALYSIS OF MINORITY STRESS CLASSIFIER

<table>
<thead>
<tr>
<th>Precision</th>
<th>Recall</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>0.74</td>
<td>0.80</td>
</tr>
</tbody>
</table>

- Identify similar posts but contrasting outcomes on minority stress expressions
- Compare and contrast the similar posts
ERROR ANALYSIS: FALSE POSITIVES

- Non Self-Experiences
- Past Experiences
- Seeking Relationship Advices

“My best friend is a trans man who has been constantly abused [...] is now taking anxiety medication and have managed to survive thus far. Does someone know of any safe spaces for trans men who are recovering from abuse or from suicide?”
ERROR ANALYSIS: FALSE NEGATIVES

- Asking Question(s)
- Lack of context and non-explicit description of self-experience
- Typos and variants of slangs

Guys! I’m in need of a chest binder. I’m in extreme dysphoria that badly affects my daily life.
Discussion
ETHICS

- These automated approaches may reinforce stressors such as prejudice and stigma.
- Minority stress classifier could be misused by bad-actors.
- An individual may not be willing to have their minority stress be identified.
- Ethical complexity associated with using such classifier in non-anonymized platforms (Twitter, Facebook).
ETHICAL IMPLICATIONS

- An **opt-in approach**, where members of an online community are given the choice whether or not to participate in a mental health intervention (that uses their data).

- It is only meant to **complement** and **assist** human intervention.

- Any decision and interventions using our classification approach require **careful** and **in-depth supplemental ethical analysis**.
IMPLICATIONS

Social Media Interventions for LGBTQ+ individuals:

- Moderation and Support Matching Efforts
- Therapeutic Efforts (e.g., Journaling tools)
IMPLICATIONS

Public Health Implications:

- Clinical and Therapeutic Practice
- Minority Stress Research
TAKEAWAYS

- **Codebook** on minority stress in social media language

- A machine learning classifier to scalably identify minority stress in social media language

- **Lexicon** of minority stress language

  - Posts with minority stress contain personal and self-experiences

  - Posts with no minority stress are about raising awareness, and less-personal and more-general issues

Thank You

@kous2v | koustuv.com
koustuv.saha@gatech.edu