ON THE STATE OF SOCIAL MEDIA DATA FOR MENTAL HEALTH RESEARCH

KEITH HARRIGIAN, CARLOS AGUIRRE, & MARK DREDZE
JOHNS HOPKINS UNIVERSITY

A FIELD AT A CROSSROADS

- Adoption of computational methods for mental health in the clinical setting remains limited, despite almost a decade of active research
- Several challenges plague data acquisition in this domain
  - Variable clinical presentation of psychiatric conditions
  - Sensitive nature of annotated data & robust privacy regulations
  - Proxy-based annotation mechanisms are necessary to achieve scale

To what extent have data-related challenges hindered research progress and slowed the transition of computational methods into the clinical setting?

METHODS

LITERATURE SEARCH & ANNOTATION SCHEMA

LITERATURE SEARCH

Term Lists

Depression | Suicide | Anxiety | Bipolar
Mood | PTSD | OCD | Addiction
ADHD | Eating | Panic | Mental Health
Borderline Personality | Schizophrenia

Social Media | Electronic Media

Machine Learning | Inference
Prediction | Detection

Query Pool

JN

JMIR Publications
Advancing Digital Health & Open Science

PubMed.gov

arXiv.org

Query Structure

+

+

SELECTION CRITERIA AND EXCLUSIONS

1. Must contain non-clinical electronic media (e.g., social media, SMS, search query text)

Electronic Health Records (EHR) & transcribed interviews

2. Must contain written language (i.e., text) within each unit of data

Search query volume, mobile activity, images, speech

3. Must contain dependent variable that captures or proxies a condition listed in DSM-5

Date of diagnosis, unlabeled data dumps

ANNOTATION SCHEMA

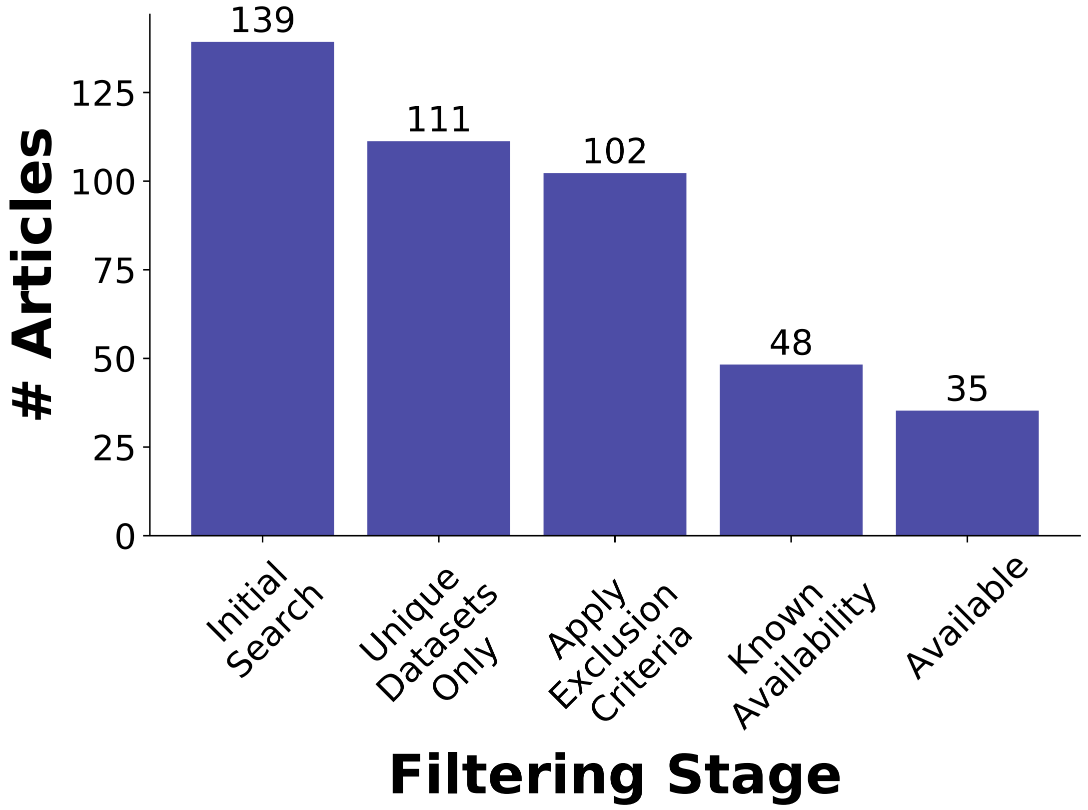
Field	Description	Example
Platform(s)	Electronic media source(s)	Twitter, SMS
Task(s)	Mental health condition(s) included as dependent variables	Depression, suicidal ideation, PTSD
Annotation Method(s)	Method for defining and annotating mental health variable(s)	Regular expressions, community participation, clinical diagnosis
Annotation Level	Resolution at which ground-truth annotations are made	Individual, document
Size	Number of data points at each annotation resolution for each task class	673 users, 576k comments
Language(s)	Primary language(s) of text in the dataset	English, Japanese, Portuguese
Availability	Whether the dataset can be shared and, if so, by what mechanism	Data usage agreement, IRB review, distribution prohibited

RESULTS

SUPPORTING DATA & ANALYSIS

RESULTS

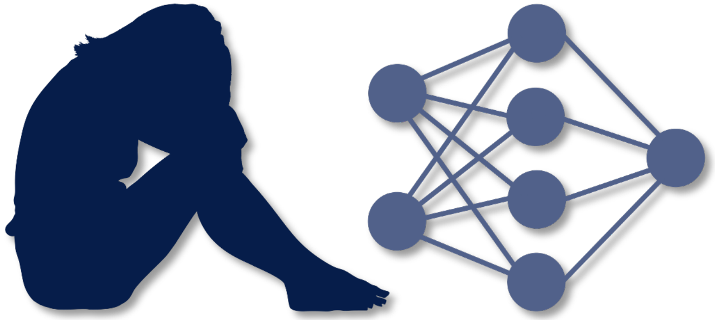
- Identified 102 unique datasets that meet our selection criteria
- Found an average of 12.75 new datasets released per year
- 2015 CLPsych Shared Task was the most reused resource[1]
- Unable to identify any accessible datasets with clinically-derived annotations



[1]"CLPsych 2015 shared task: Depression and PTSD on twitter." Coppersmith et al., 2015.

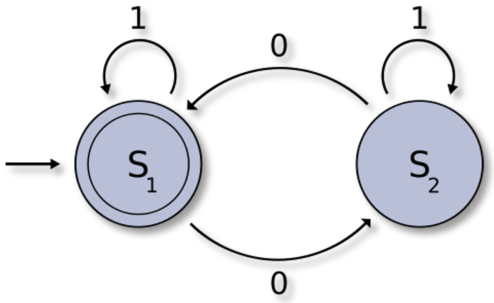
20 Platforms

Twitter and Reddit were most used; noticeable dearth of Facebook, Instagram, and YouTube



36 Modeling Tasks

Depression, suicidal ideation, PTSD, bipolar disorder, self harm, & eating disorders were most common



24 Annotation Mechanisms

Frequent use of regular expressions, clinical surveys, & community participation

6 Languages

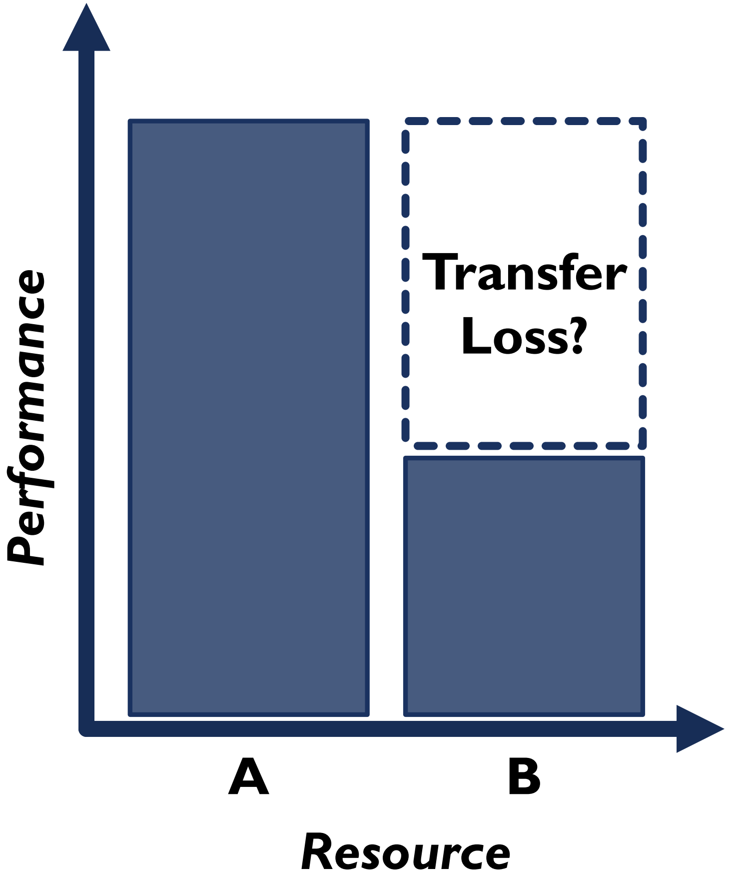
English, Chinese, Japanese, Korean, Spanish, Portuguese (though mostly English)

DISCUSSION

3 RECOMMENDATIONS FOR FUTURE DATASET CURATION

1. UNIFY TASK DEFINITIONS

- Over 20 unique annotation mechanisms identified (conservative estimate)
- Becomes difficult to contextualize algorithmic performance across studies
- Community needs **standardized** (and fair) benchmarks to inform interpretation of results as models are transitioned into the clinic



2. DEVELOP MECHANISMS FOR SHARING SENSITIVE DATA

- Recent research has called into question the strength of proxy-based annotations[1,2]
- Existing datasets with clinically-derived annotations are not currently shareable
- Leverage **privacy-preserving technology** to share patient-generated data or make data available via **secure computing environments**

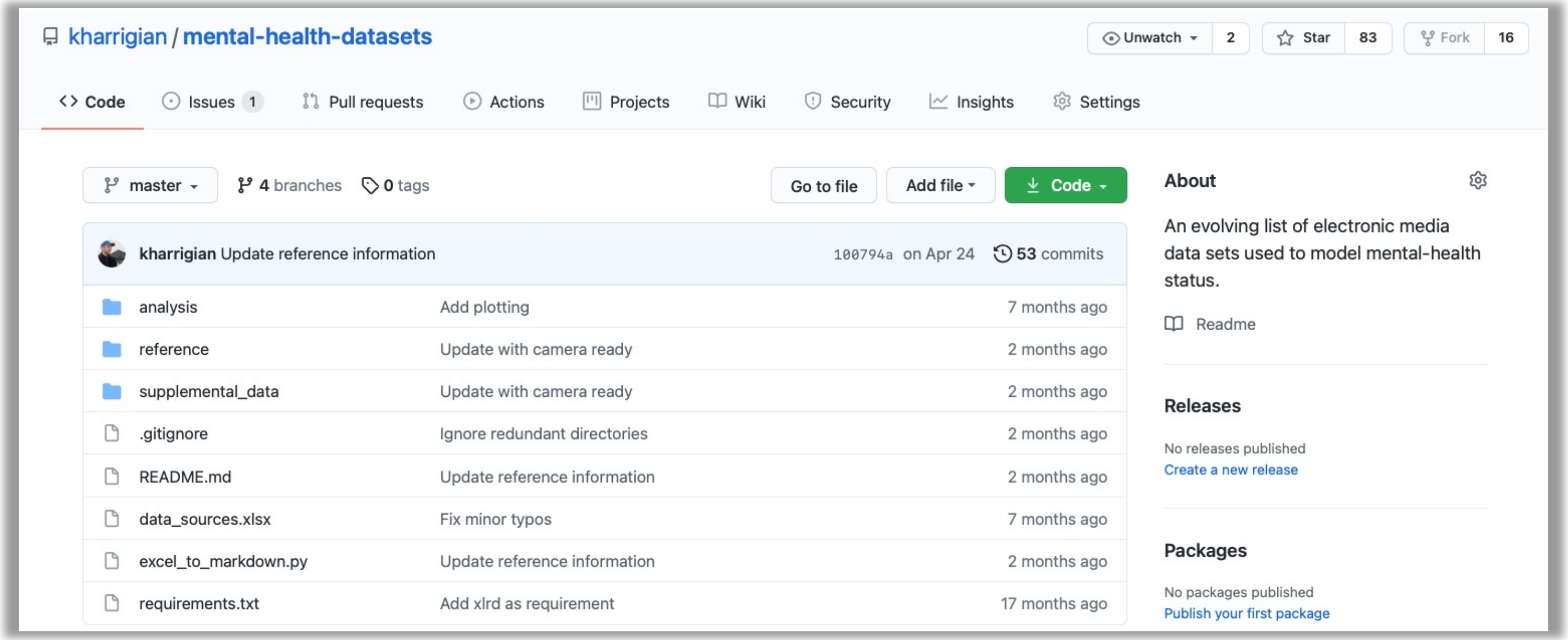
[1] "Methodological gaps in predicting mental health states from social media: Triangulating diagnostic signals." Ernala et al., 2019.
[2] "Do models of mental health based on social media generalize?" Harrigian et al., 2020.

3. CURATE DATASETS WITH POPULATION DIVERSITY IN MIND

- Several datasets sample demographically-matched or activity-matched individuals
- However, no dataset was specifically sampled to be representative of the general population
- Machine learning models underperform for people of color, even after addressing sample size issues[1]
- Leverage **self-disclosed demographics** when possible and start looking **beyond English**

[1] "Gender and racial fairness in depression research using social media." Aguirre et al., 2021.

MENTAL HEALTH DATASET DIRECTORY



Access and contribute to our directory of annotations at:

github.com/kharrigian/mental-health-datasets

THANK YOU FROM OUR TEAM

Keith Harrigian
PhD Student
Computer Science
kharrigian@jhu.edu

Carlos Aguirre
PhD Student
Computer Science
caguirr4@jhu.edu

Mark Dredze
John C. Malone Associate Professor
Computer Science
mdredze@cs.jhu.edu