

# Analyzing reddit Data to Answer Questions in Clinical Psychology

*Emma G. Preston, B.A.*  
*USC Clinical Science Doctoral Program*  
*September 10, 2021*

# Agenda

**1 Background on Psych Research**

**2 Stigma & Social Media**

**3 Reddit Data**

# Introduction

Who am I

2nd year student using social media data to answer clinical research questions in psychology

Clinical  
Science @  
USC

6 year program that combines research and clinical work. 5 years of research and clinical work at USC, 1 year of full time clinical work at a hospital

# “Clinical Psychological Research” Includes:

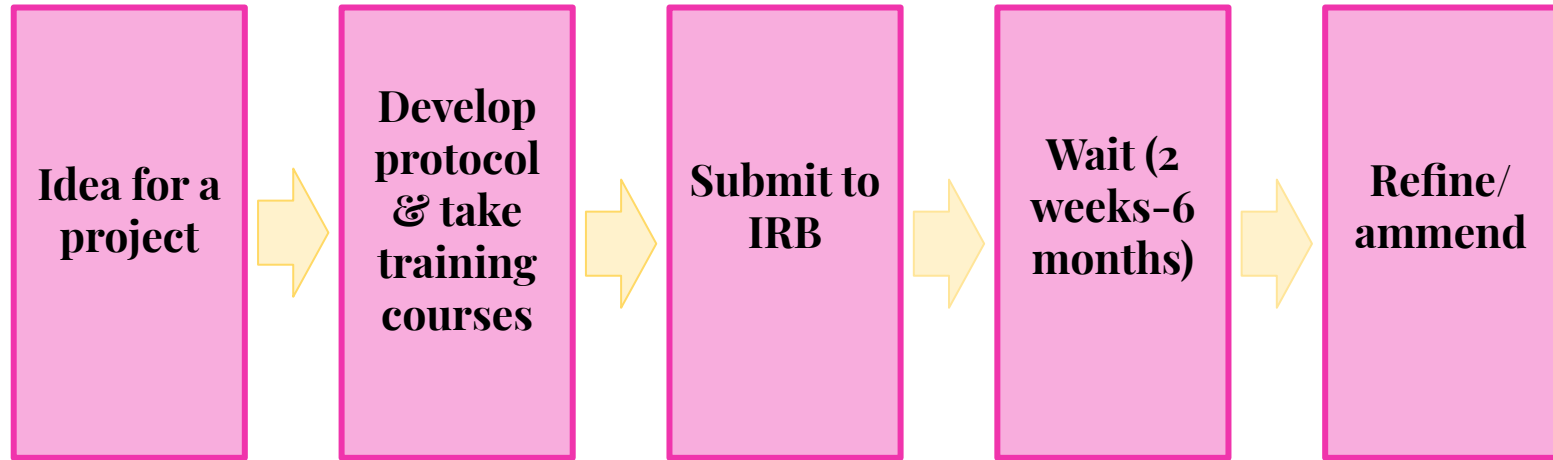
- Mood Disorders (e.g., depression, anxiety, bipolar, etc.)
- Cognitive Disorders (e.g., ADD/ADHD, learning disabilities, etc.)
- Developmental Disorders (e.g., Autism, Downs Syndrome, etc.)
- Personality Disorders (e.g., Borderline PD, Antisocial PD, etc.)
- Trauma Disorders (e.g., PTSD)
- Sleep/Wake Disorders (e.g., insomnia, sleep apnea, etc.)
- Sexual disorders (e.g., erectile dysfunction, female orgasm difficulties, etc.)
- Substance use disorders (e.g., alcoholism, drug addiction, etc.)
- Suicidality and self-injurious behaviors

# Research that is trying to understand and prevent human suffering

**Why do people  
suffer?**

**How do we prevent  
& treat suffering?**

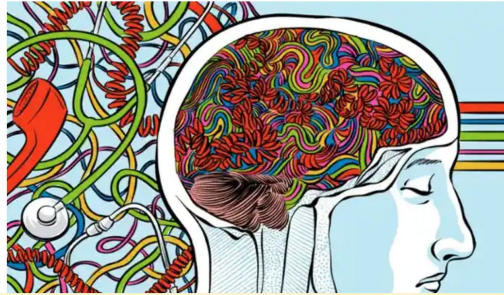
# Logistics in clinical psychology research: Internal review boards (IRBs) and human subjects research



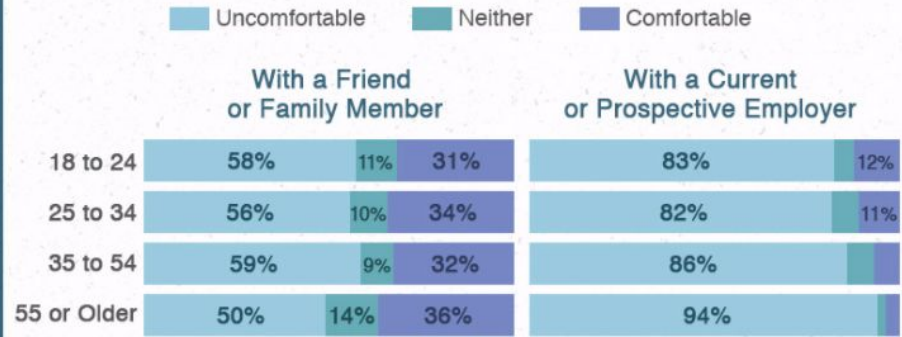
# Stigma

## Mental health stigma hasn't gone away

Stigma surrounding mental health comes in many forms, and it's important to understand what the differences are



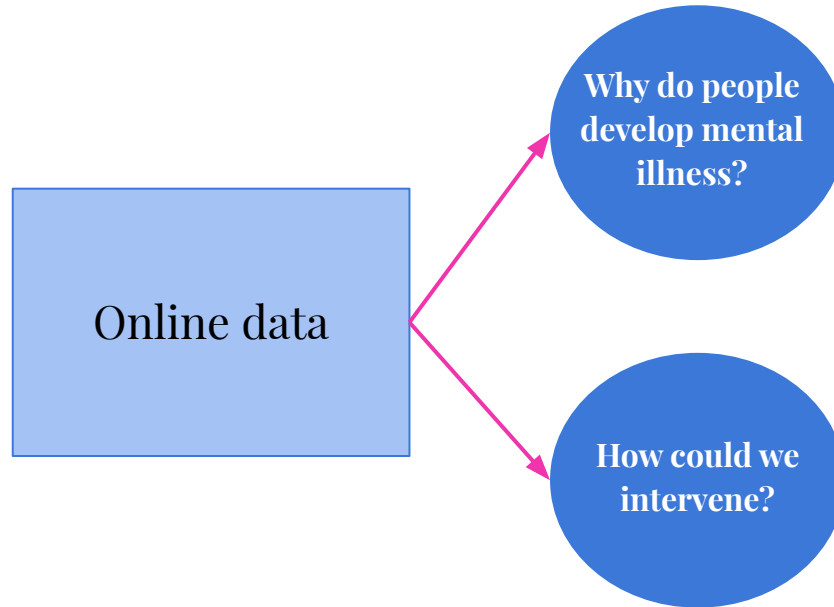
## LEVEL OF COMFORT DISCUSSING MENTAL HEALTH, BY AGE



44%

OF ADULTS WITH MENTAL ILLNESSES  
REPORTED BEING DISCRIMINATED AGAINST  
OR DISMISSED BY THEIR PHYSICIAN

People are often more comfortable talking about their mental health online than they are in a research lab or doctor's office







**reddit as a gold  
mine**

# “reddit: the front page of the internet”

- Chat-based
- Anonymous
- 430 million active users
- Minimal moderation & minimal community guidelines
- Subreddit = topic-specific chat

# Mental health subreddits

- r/depression: 788k subscribers
- r/anxiety: 482k subscribers
- r/ADHD: 1.3 million subscribers
- r/addiction: 59k subscribers
- r/suicidewatch: 302k subscribers
- r/offmychest: 2.5 million subscribers
- r/schizophrenia: 43k million subscribers
- r/raisedbynarcissists: 718k subscribers
- r/OCD: 129k subscribers

What clinical questions could we answer with  
reddit data?

# What happens on reddit influences real life

## Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity

**Munmun De Choudhury**  
School of Interactive Computing  
Georgia Institute of Technology Atlanta GA  
munmund@gatech.edu

**Sushovan De**  
Department of Computer  
Arizona State University  
sushovan@asu.edu

**JMIR Publications**  
Advancing Digital Health & Open Science

Articles ▾ Search articles

Journal of Medical Internet Research

Journal Information ▾

Browse Journal ▾

Sub

Published on 10.4.2018 in Vol 20, No 4 (2018): April

Preprints (earlier versions) of this paper are available at <https://preprints.jmir.org/preprint/8219>, first published June 15, 2017.



**Harnessing Reddit to Understand the Written-Communication Challenges Experienced by Individuals With Mental Health Disorders: Analysis of Texts From Mental Health Communities**

Albert Park<sup>1</sup> Mike Conway<sup>1</sup>

## Detecting anxiety on Reddit

**Judy Hanwen Shen**  
University of Toronto  
Toronto, Canada  
judyhanwen.shen@mail.utoronto.ca

**JMIR Publications**  
Advancing Digital Health & Open Science

Journal of Medical Internet Research

Published on 12.10.2020 in Vol 22, No 10

Preprints (earlier versions) of this paper are available at <https://preprints.jmir.org/preprint/20201>, first published October 12, 2020.



**Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study**

Daniel M Low<sup>1,2,3</sup> Laurie Rumker<sup>4,5,6,7</sup> Tanya Talkar<sup>1,8</sup> John Torous<sup>9</sup> Guillermo Cecchi<sup>10</sup> Satrajit S Ghosh<sup>1,2,11</sup>

## An Analysis of Domestic Abuse Discourse on Reddit

**Nicolas Schrading<sup>1</sup> Cecilia O. Alm<sup>2</sup> Ray Ptucha<sup>1</sup> Christopher M. Homan<sup>3</sup>**

<sup>1</sup> Kate Gleason College of Engineering, Rochester Institute of Technology

<sup>2</sup> College of Liberal Arts, Rochester Institute of Technology

<sup>3</sup> Golisano College of Computing and Information Sciences, Rochester Institute of Technology

{jxs8172<sup>§</sup>|coagla<sup>§</sup>|rwpeec<sup>§</sup>|cmh<sup>†</sup>}@{<sup>§</sup>rit.edu|<sup>†</sup>cs.rit.edu}

- How do LGBTQ+ individuals experience depression differently than non LGBTQ+ individuals?
- What kind of language is associated with psychosis? Can we predict psychosis from language?
- How do survivors of sexual trauma talk about experiences with mental health or medical professionals?

# Tip of the Iceberg: data science tools to answer psychological problems

**Topic  
Modeling**

“Understanding  
Weight Loss via  
Online  
Discussions”

-Liu & Yin, 2020

**Prediction**

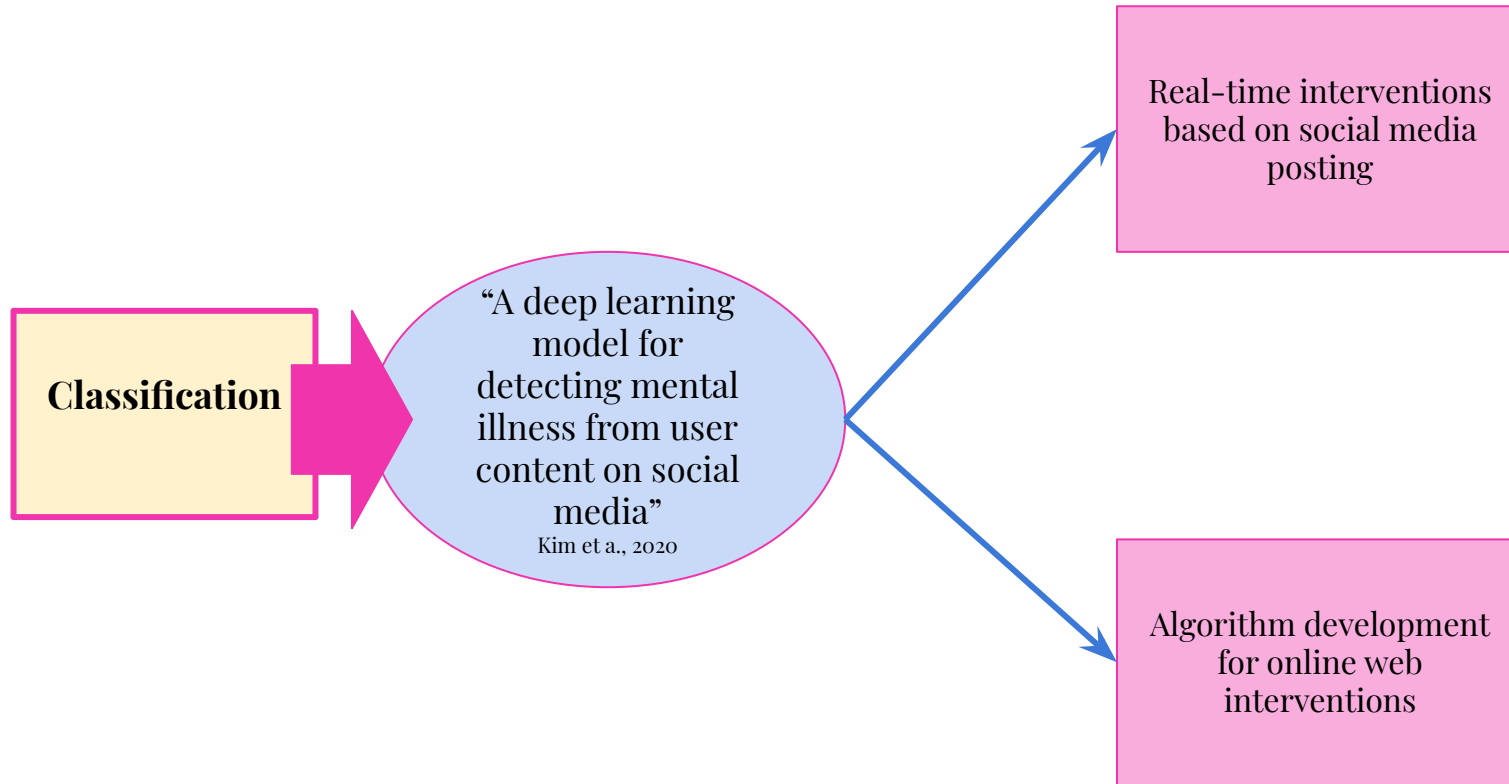
“Not Just  
Depressed: Bipolar  
Disorder  
Prediction on  
Reddit”

Skeulic et al., 2018

**Classification**

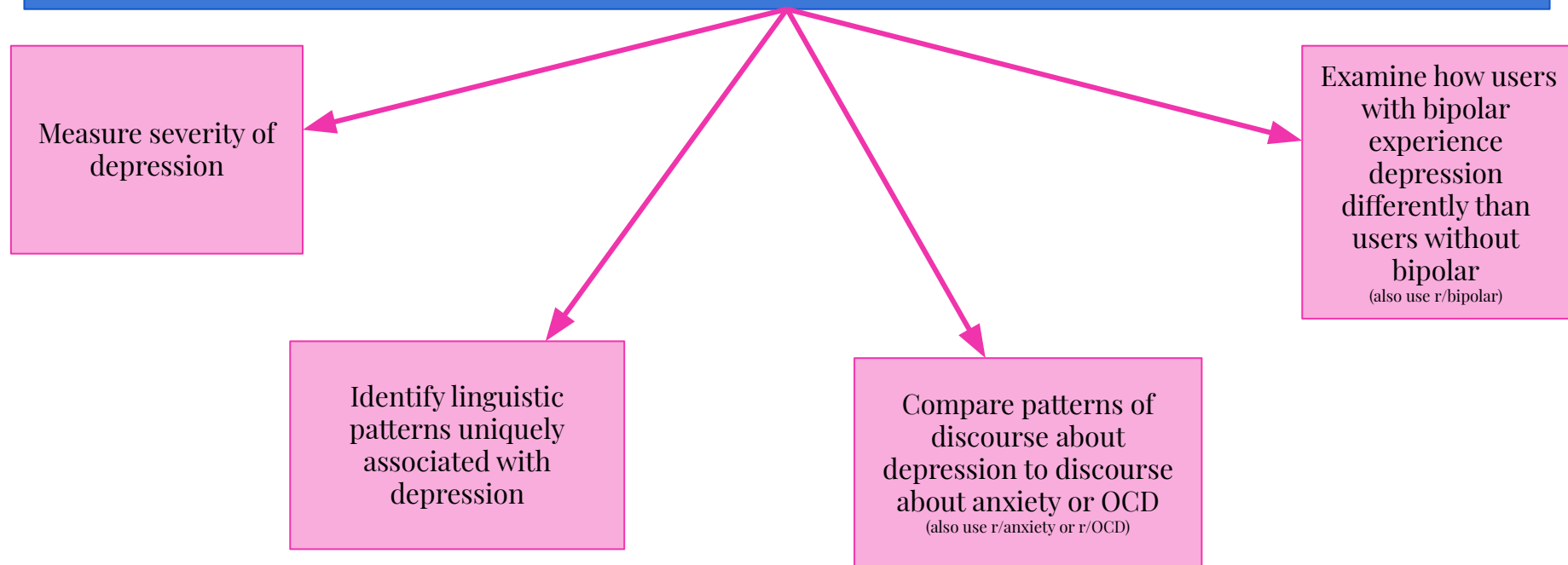
“A deep learning  
model for  
detecting mental  
illness from user  
content on social  
media”

Kim et al., 2020

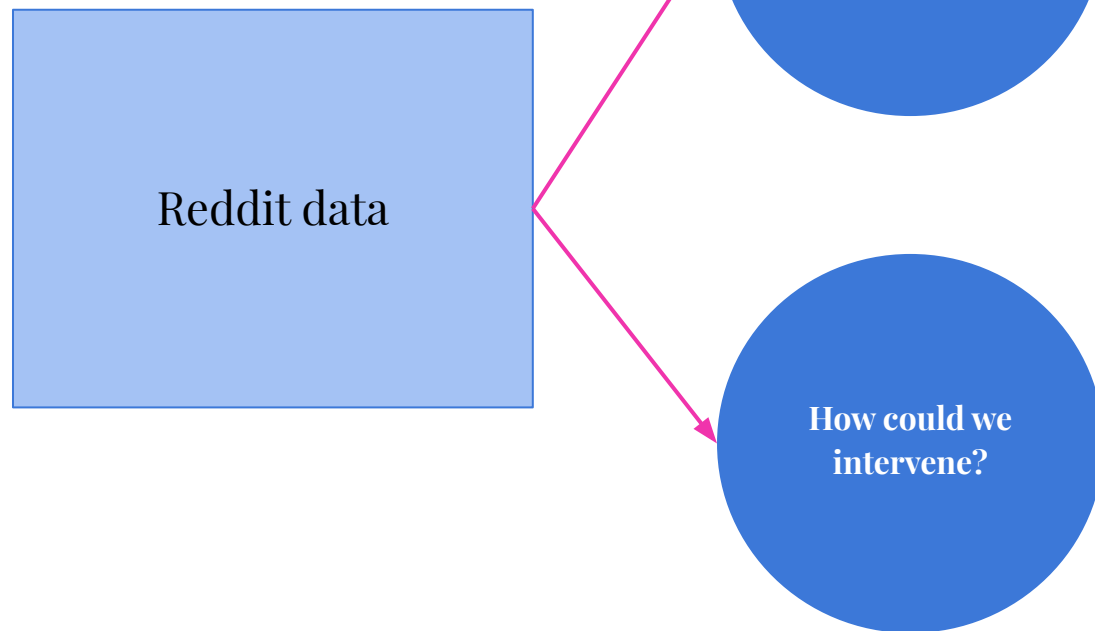




# r/depression: 788,000 subscribers



How can reddit data help us better understand, prevent, and treat human suffering?



# Questions?

*Feel free to email me at any time at [epreston@usc.edu](mailto:epreston@usc.edu)*