Use merge or similar more modern model for LDA (instead of doc2vec)

Get the latest version of hagenphase? should be able to use one of the compact models -> then do LDA

Concentrate on one single algorithm/technique. Pick something that works (the most modern one)

How do you structure your model?

New (might be a distraction): Do people have more mental illnesses or are they just more comfortable talking about it

How to measure: after topic modeling > go to literature (people’s abilities to talk about mental health) (Compounding variable)

Might be able to find more datasets on who are more comfortable talking about depression (self-reporting)

Go back into the citations, they'll have numbers

Do more research, look at it in the perspectives of time