My first idea involves far-right extremism as the general topic.

2 place you expect to find data:

GW’s Program on Extremism offers lots of data as well as organizations like the ADL which has a heat map and some other interesting resources tracking extremist violence.

3 questions you have:

I saw the Program on Extremism keeps updated info about the number of federal cases against individuals involved in the Capitol Hill Siege. One question I have involves the penalties and the average sentencing. This could be interesting to compare with data they have about sentencing for individuals travelling abroad to join terror groups.

I remember the president of GW’s Program on Extremism said that people who travel to Ukraine and train with far-right paramilitary groups don't face legal consequences the same way those who travel overseas to train with groups like ISIS might. Another question I have involves the number of people traveling to Ukraine to receive military training and the legal consequences for doing so.

After seeing the ADL heat map I’d like to look into where the most violence has been measured as well.

I would expect to find a lapse in legal penalties for far right groups vs other extremist organizations.

2.

Another topic I find interesting is disinformation. Word-mining seems very appealing in this context as it relates to key terms and phrases employed to bypass censors or bayesian filters that flag certain words.

2 place you expect to find data:

I don't know how I'd find data for that outside of manually collecting online communications and organizing them into key terms and phrases. Groups like Check my Ads offer data on the ad networks which producers of disinformation use to make money.

3 questions you have:

I'd love to look at what the diction of disinformation looks like, what are some key terms and how might it vary by topic or subsection?

I’d also like to look into the presence of bot accounts online within this context to get a sense of their prevalence.

Online Ad networks interest me for this topic. What do the networks that make disinformation profitable look like mapped out across the various sites they appear on?

I expect to find a common vocabulary among disinformation accounts and an ad network which connects various publishers of misinformation.