

Stories and personal experiences in the COVID-19 Discourse

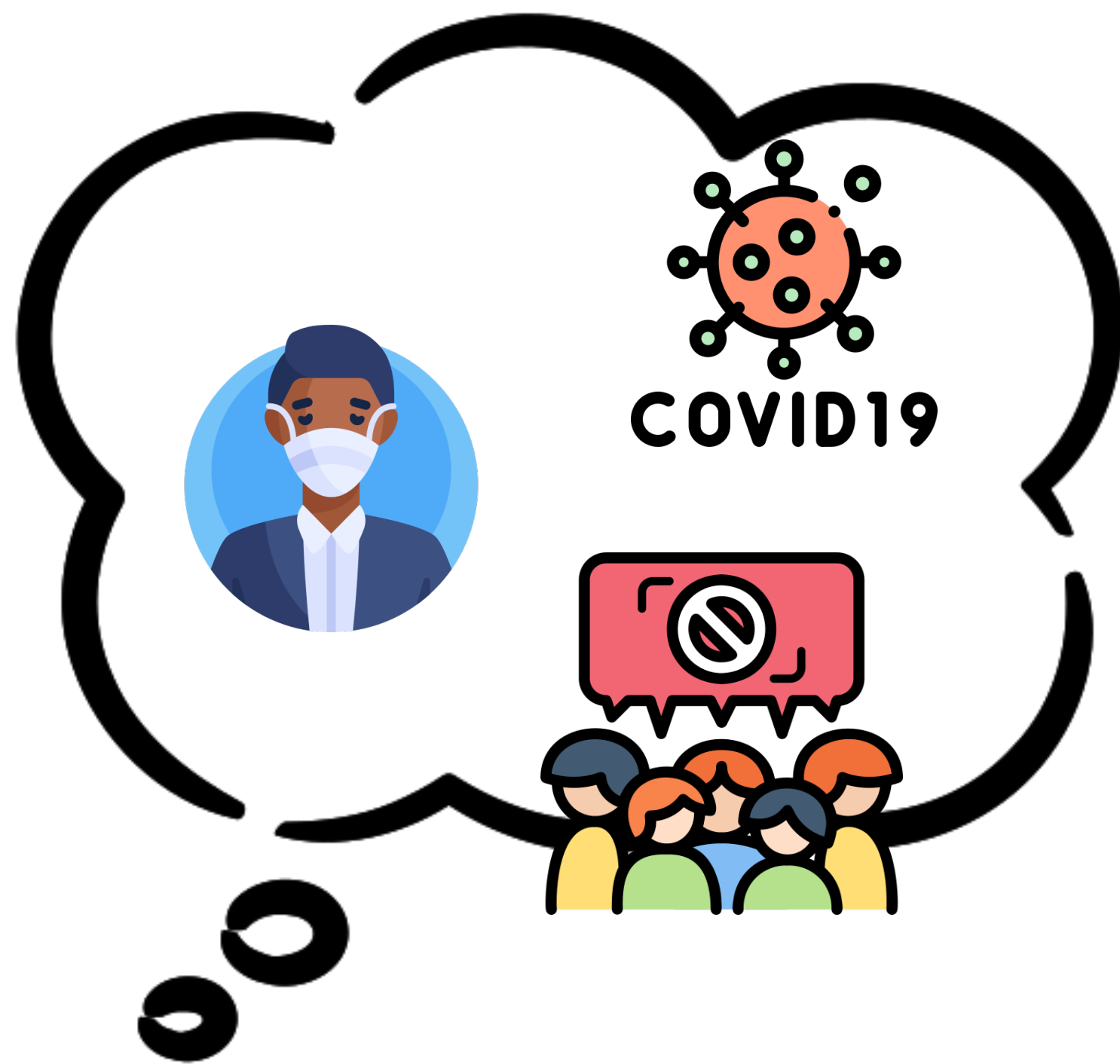
Neele Falk (*) and Gabriella Lapesa (**)

*Institute for Natural Language Processing, University of Stuttgart, Germany

**GESIS – Leibniz Institute for Social Sciences and Heinrich-Heine University of Düsseldorf, Germany

Goal

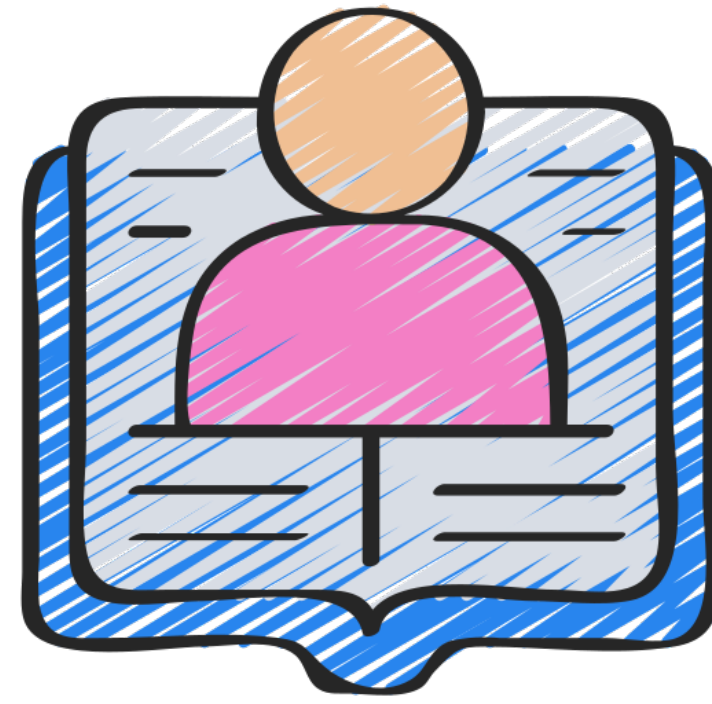
Understand storytelling in COVID-19 Discourse



- understanding personal experiences during crisis
--> driving force for political discourse
- reveal shared concerns
- identify potential actionable measures

How? Develop NLP pipeline to automatically analyse large-scale online discussion.

Why storytelling in online discussions?

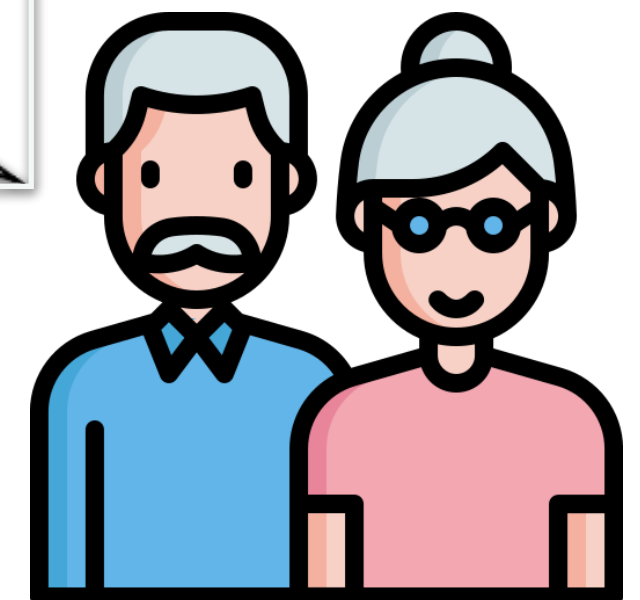


- storytelling = natural and human way to justify your positions
- we use storytelling in discourse to establish our collective identity
- to trigger empathy and perspective-taking

How does it look like?

i'm a post-grad student in melbourne. i can't access any archives within my state, my thesis has just stalled. i can't work because of the lockdown. i can't even go for a walk with friends because i'm trapped within a 5km radius of my house."

"maybe if we had actually done something as a nation when any of the other black people were murdered by police we wouldn't be having protests during a pan- demic. my wife is high risk (autoimmune) and as much as we would like to we are not out protesting (both white btw)."



Research Questions

How pervasive is storytelling in COVID-19 Discourse?

Which types of stories do speakers share?

Are they personal (about themselves) or more general (represent collectives)?

What is the function of the stories within the post they occur?

Do they provide clarification, propose a solution, disclose harm, or establish the speaker's credibility?

What are these stories about?

What are their topics?

What data are we using?

source



500 documents
2300 narrative/experience
spans annotated with
functions,
multiple domains, not
COVID-19

target



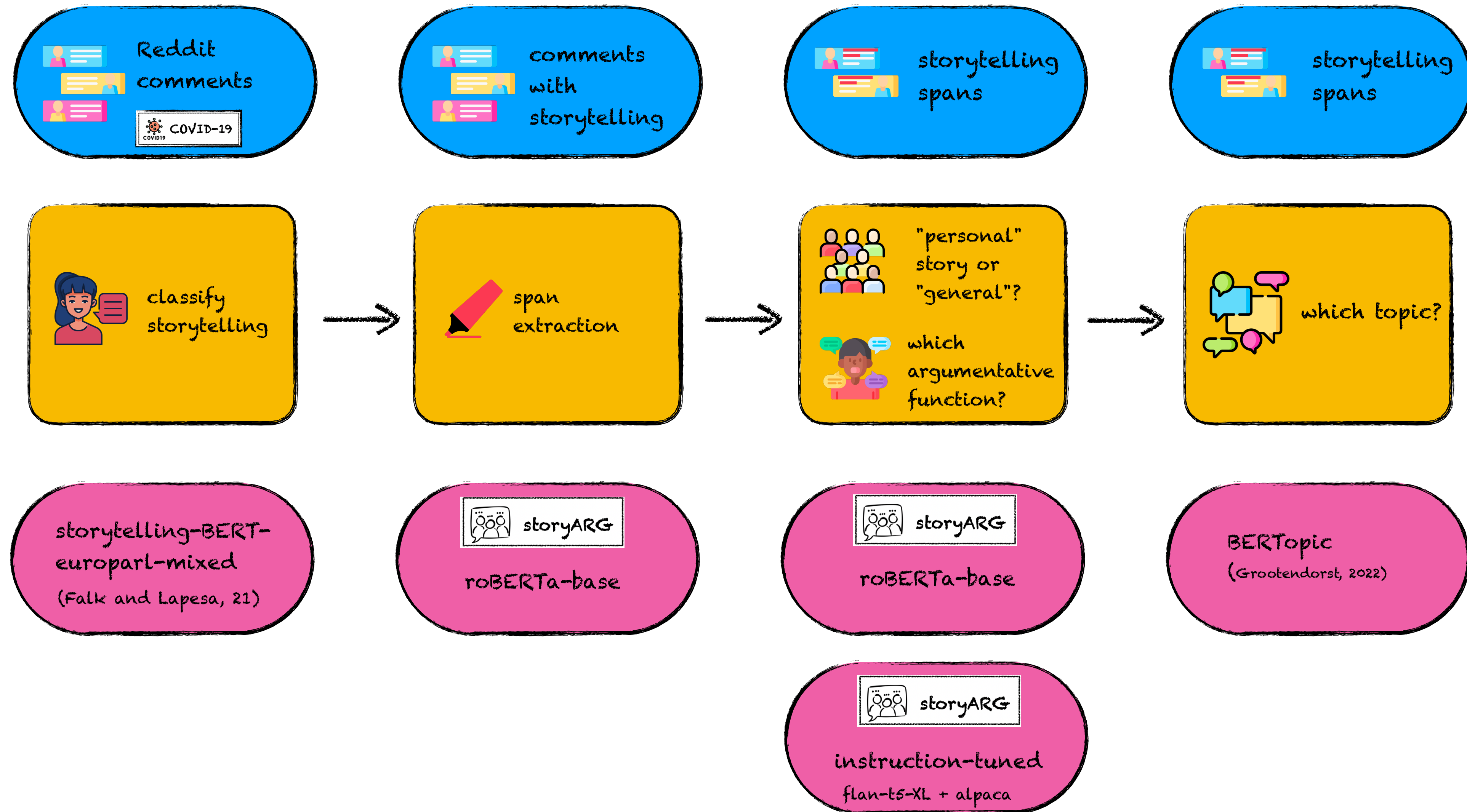
340k posts
validation set: 700
posts
reddit comments
related to COVID-19

challenges

out-domain
small training data

How do we retrieve our final discourse?

NLP Pipeline for retrieving the COVID-19 "storytelling" discourse



Evaluation

Which models are best at predicting argumentative function?

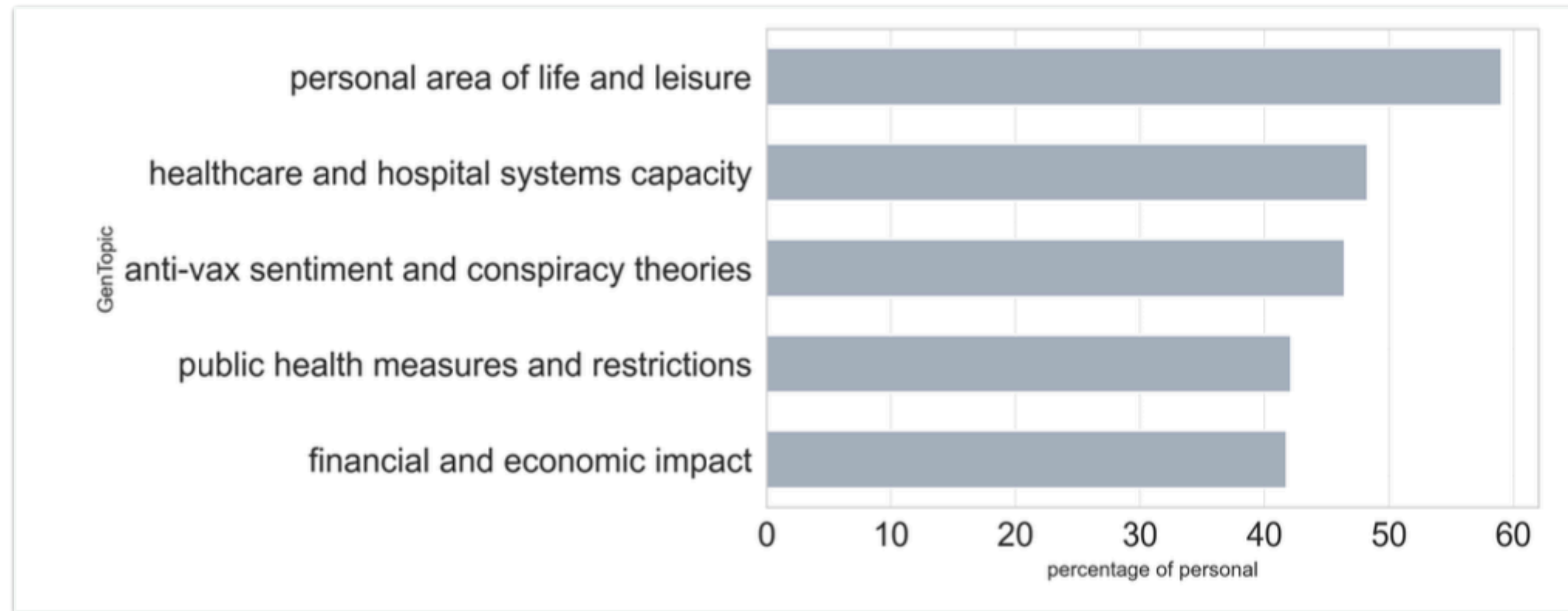
Instruction	Input	Target
Does the following personal experience or story express the a disclosure of harm ? A disclosure of harm is defined as follows: A negative experience is reported that was either made by the discourse participants themselves or that they can testify to and casts the experiencer as a victim. The experience highlights injustice or disadvantage. Answer with yes or no.	Yes, we do have to face negative (even aggressive) reactions on a regular basis.	yes

Roberta, fine-tuned on StoryARG → best on StoryARG

General-purpose instruction-tuned (zero shot)

Model 2, instruction-tuned on StoryARG → best on COVID-19

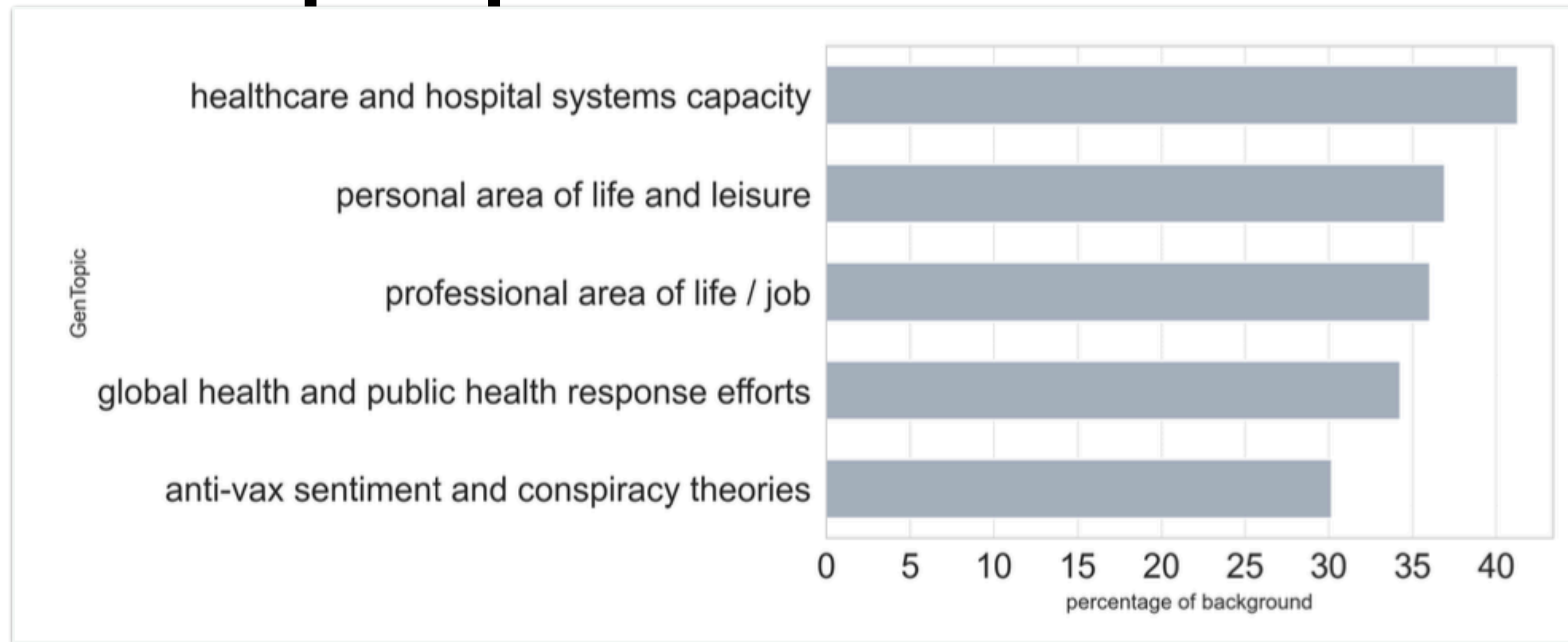
Where do we see more personal stories?



Social distancing results in highly personal experiences, particularly in areas where social interaction plays an important role.

"[...] in december if theres no vaccine for covid, ill move it again same wedding, just different day no biggie, [...], the peace of mind of not making any of my loved ones sick is absolutely worth it"

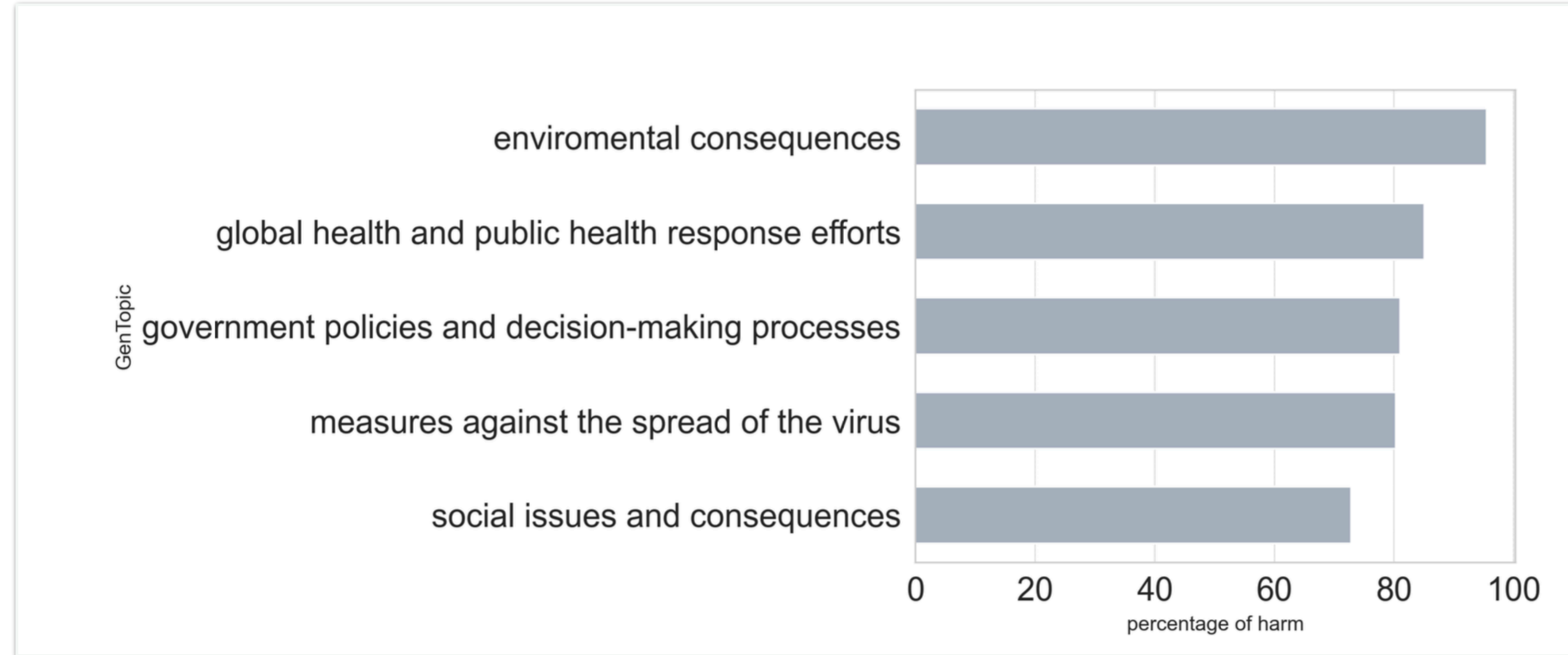
Where do people establish their background?



In controversial topics, "establish background" is often used to distinguish oneself from a group that is being questioned.

"i ain't no anti-vaxer (i have all my shots) and i know the short-term side effects are basically non-existent in some of the ones we working on but we really can't be that sure about them long term side effects"

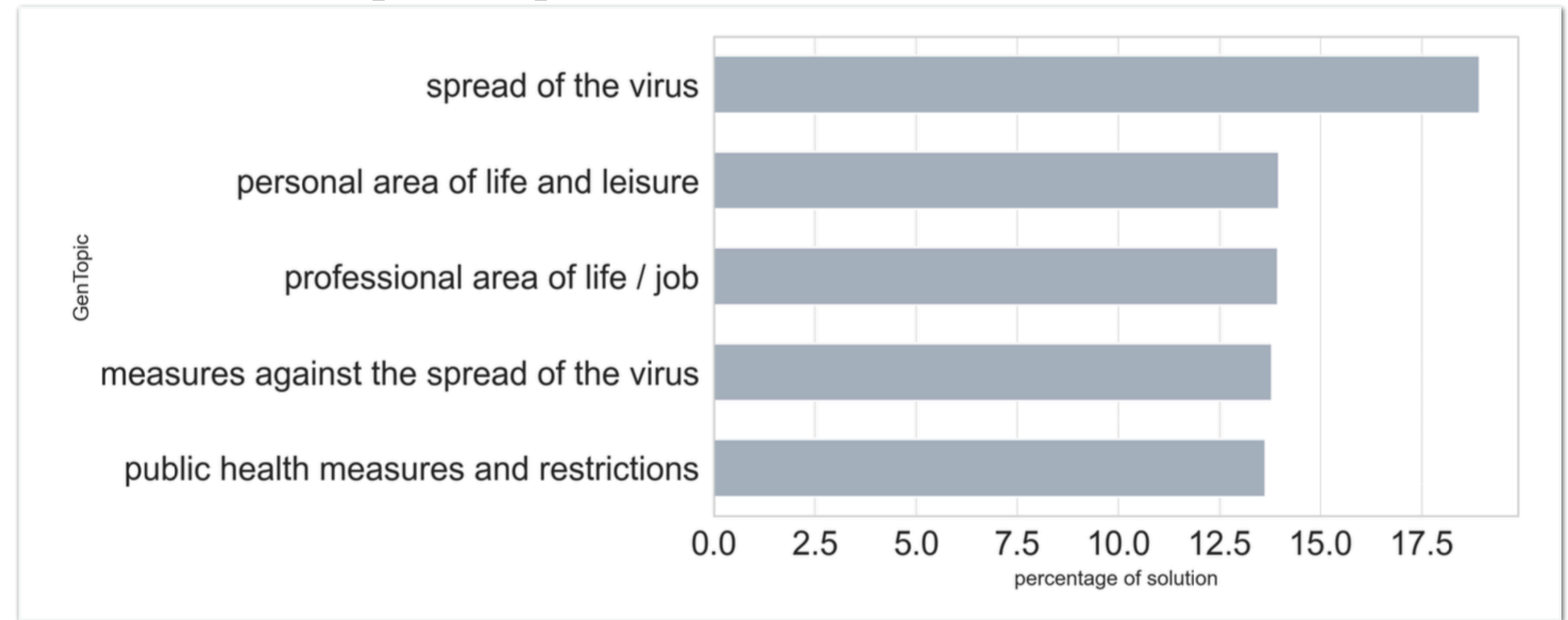
Where do we see disclosures of harm?



Issues that concern society and social problems often involve a high number of disclosures of harm.

"it's always insane to me how we live in a system where tons of people simply live paycheck to paycheck, can't afford health insurance,[...]"

Where do we see proposals for solutions?



Proposed solutions are mainly made in the work/school sector or in the concrete handling of measures to stop the spread of the virus.

"[...]have 1/5 of the population attending one day a week they would meet in class rooms with desks spread far enough as part for social distancing, while wearing masks, and have the teachers float from room to room[...]"

Conclusion

- NLP pipeline to extract and classify storytelling in COVID-19 discourse
- challenging: argumentative function, instruction-finetuning can help with domain gap
- identify specific trends regarding certain types of stories