

# Building Language Models for Social Good

*Case Studies on Homelessness, and Suicide Prevention*

USC Center for AI in Society

*Swabha Swayamdipta*

*Assistant Professor, USC Viterbi CS*

*Feb 06, 2026*

*The Webb Schools Tech and Innovation Seminar*

USC Viterbi



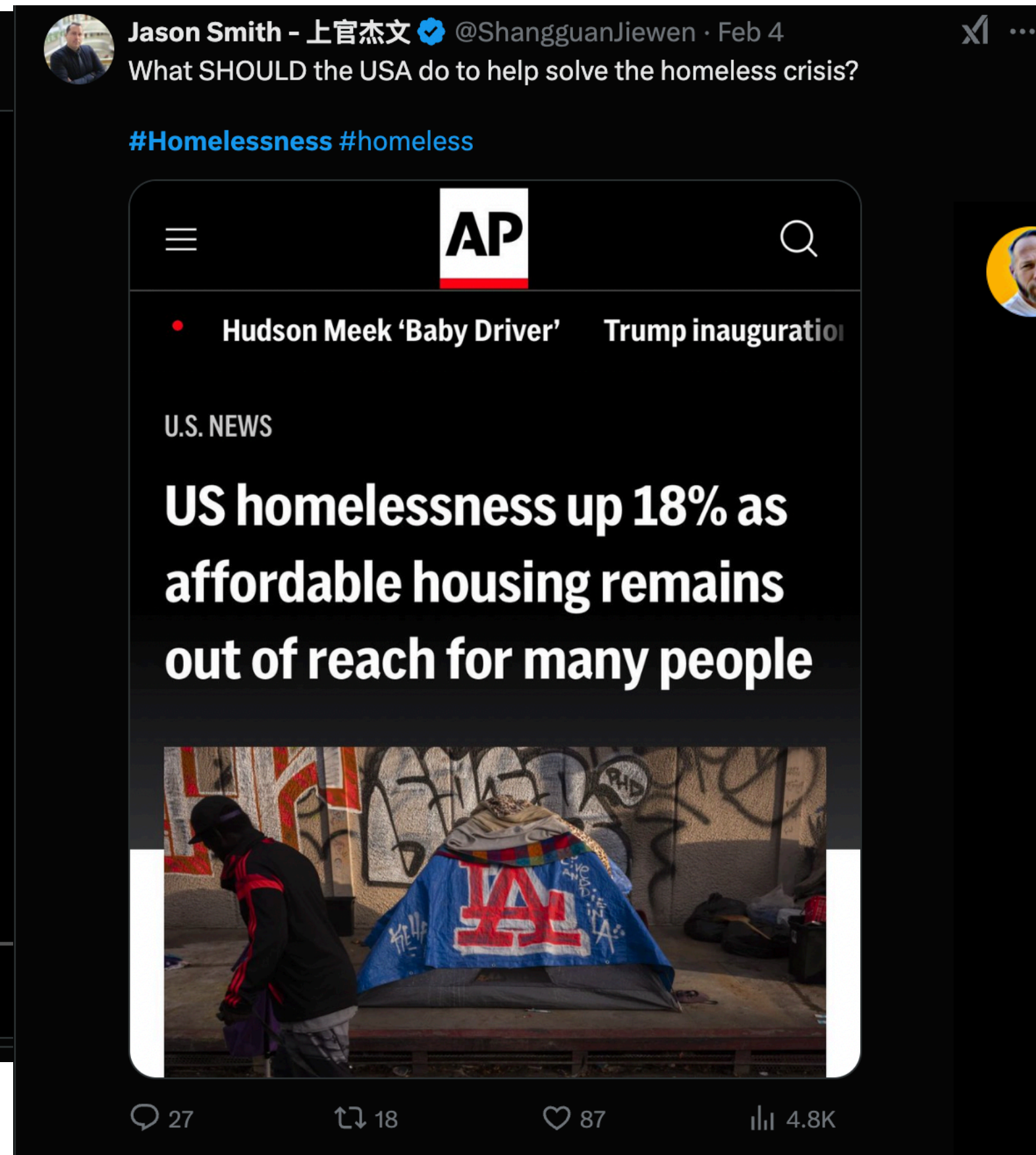
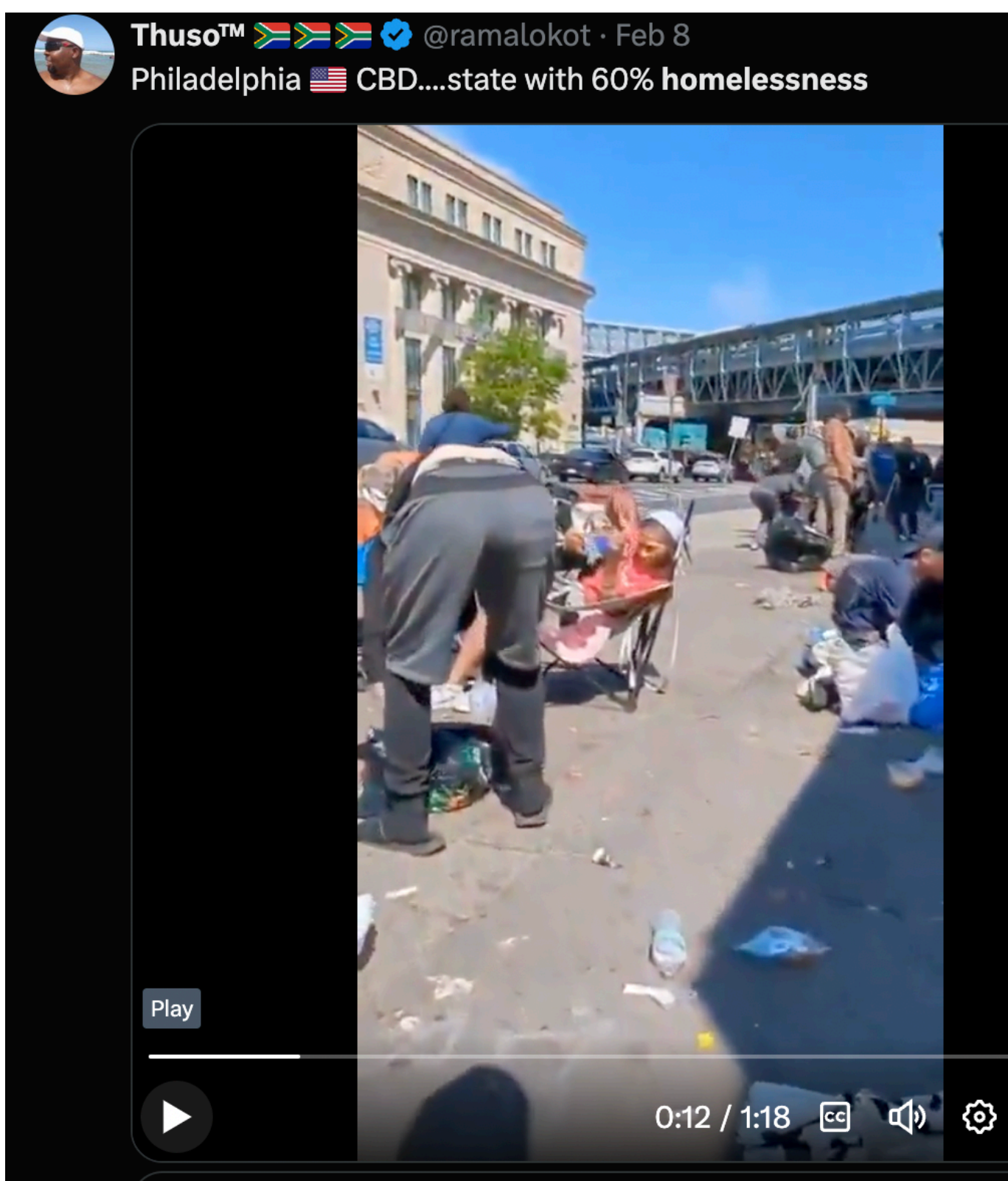
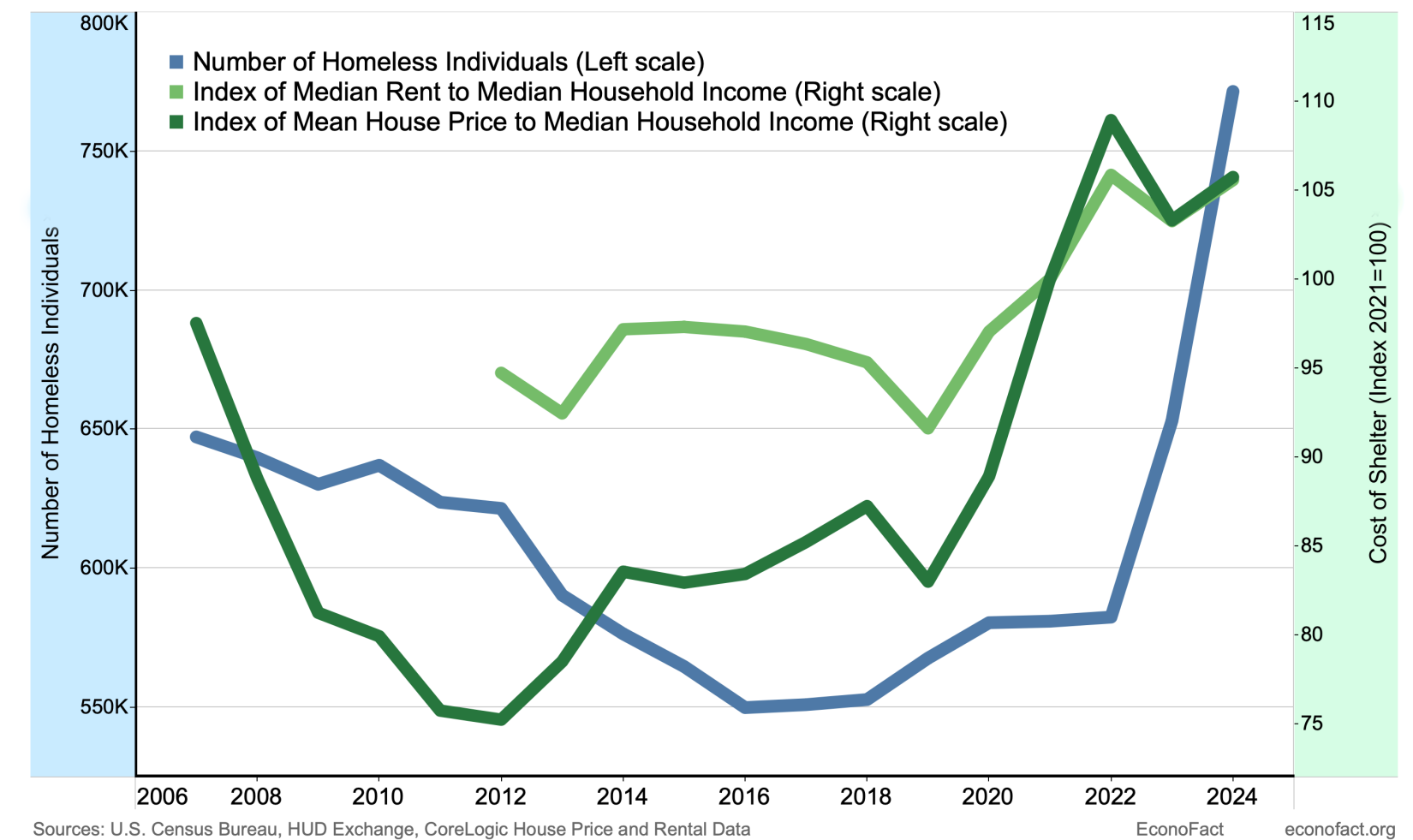
**Warning: Some content in this talk might be  
offensive / disturbing to the audience**





**Record-High Homeless Counts.** A record-high 653,104 people experienced homelessness on a single night in January 2023. This is more than a 12.1 percent increase over the previous year.

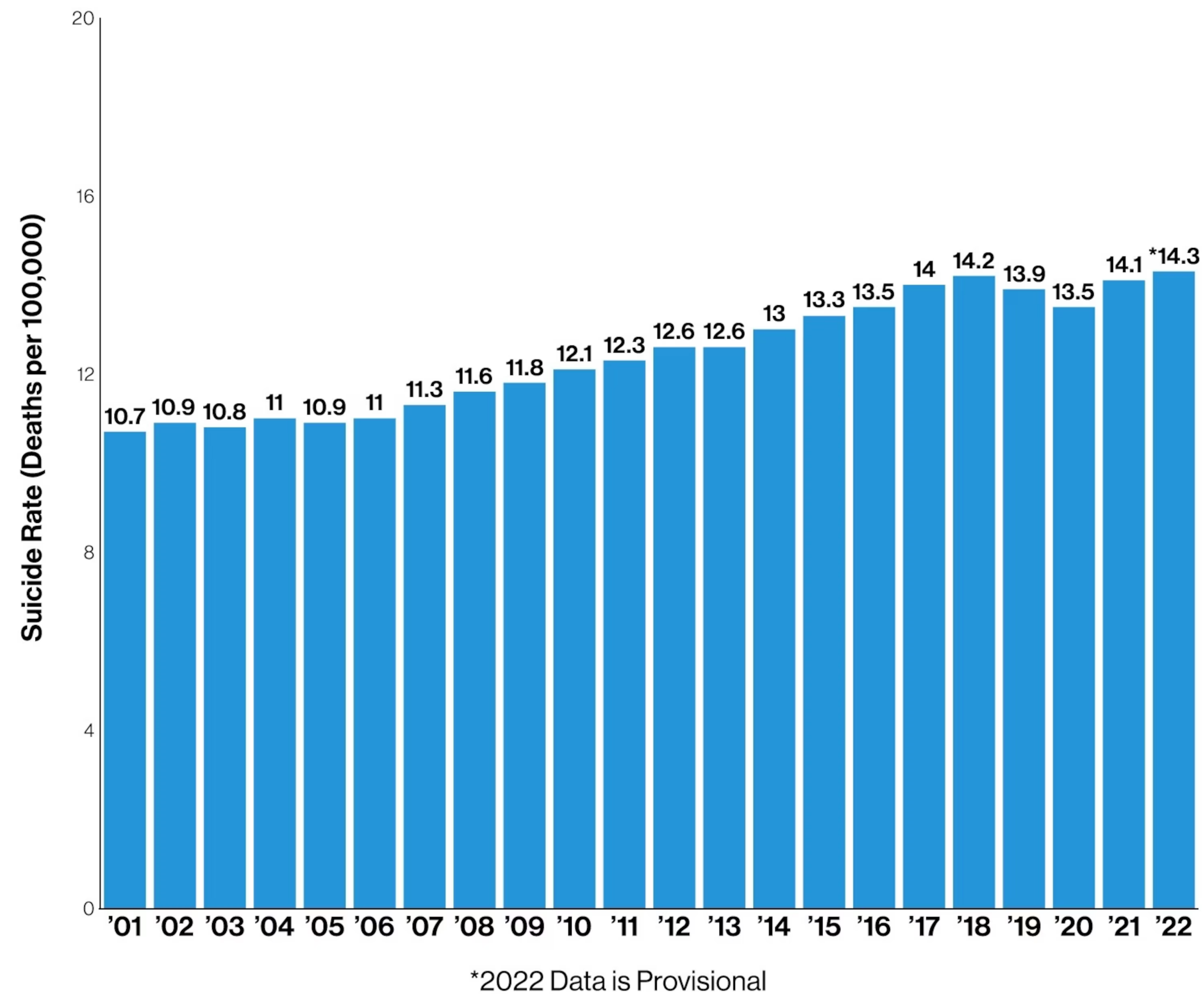
## HOMELESSNESS AND HOUSING COSTS





## Suicide Rate in the United States

2001-2022

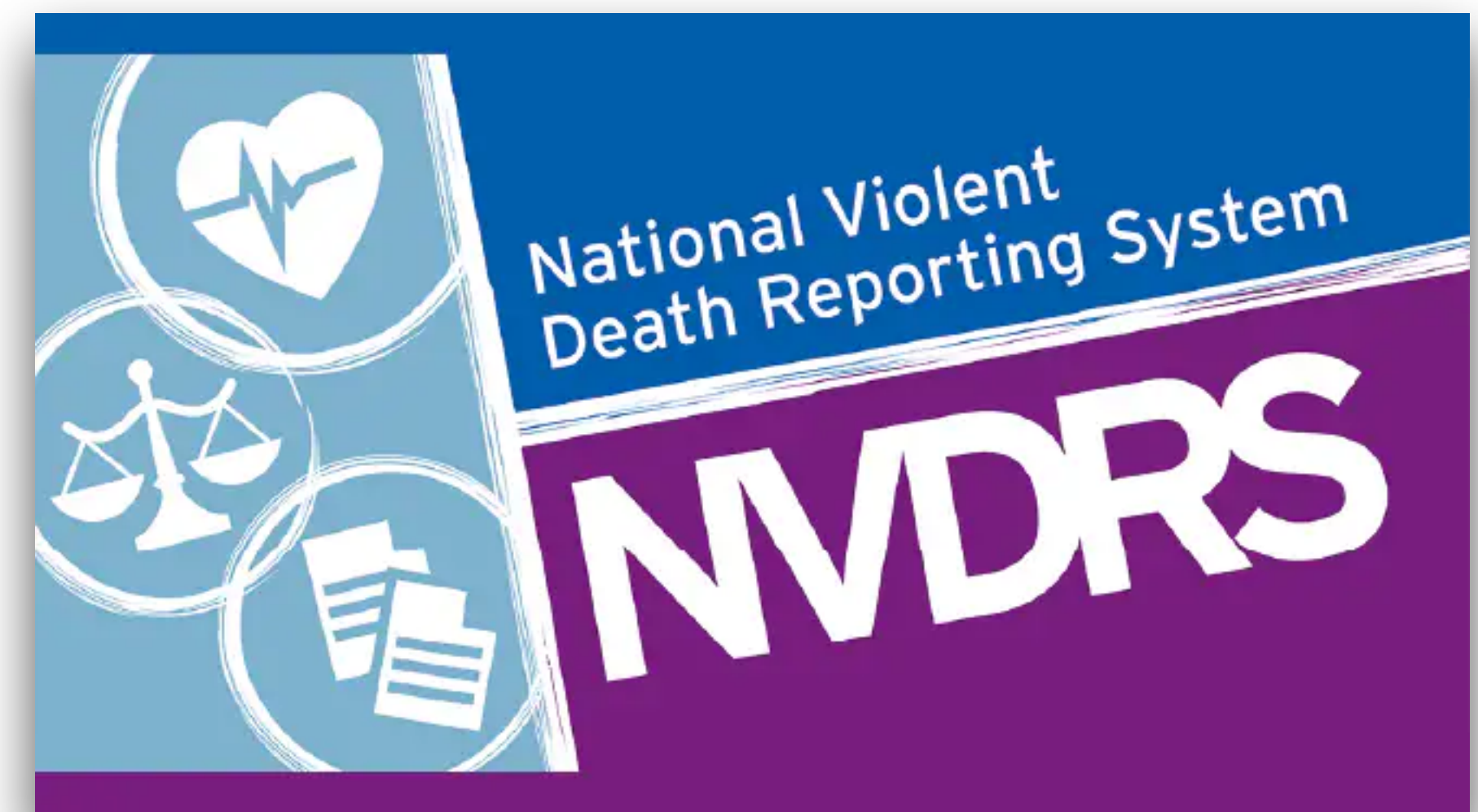


SOURCE: NATIONAL CENTER FOR HEALTH STATISTICS, NATIONAL VITAL STATISTICS SYSTEM

abc NEWS



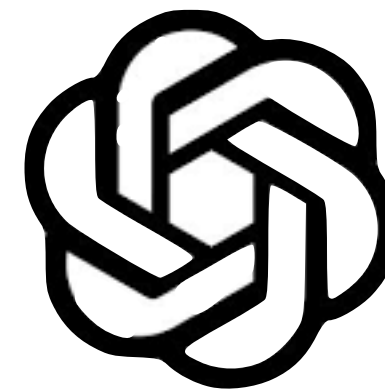
CDC Data on Suicide Reports Collected by Law Enforcement / Emergency Responders



Can AI assist in making sense of large amounts of sensitive information for better decision-making?

X

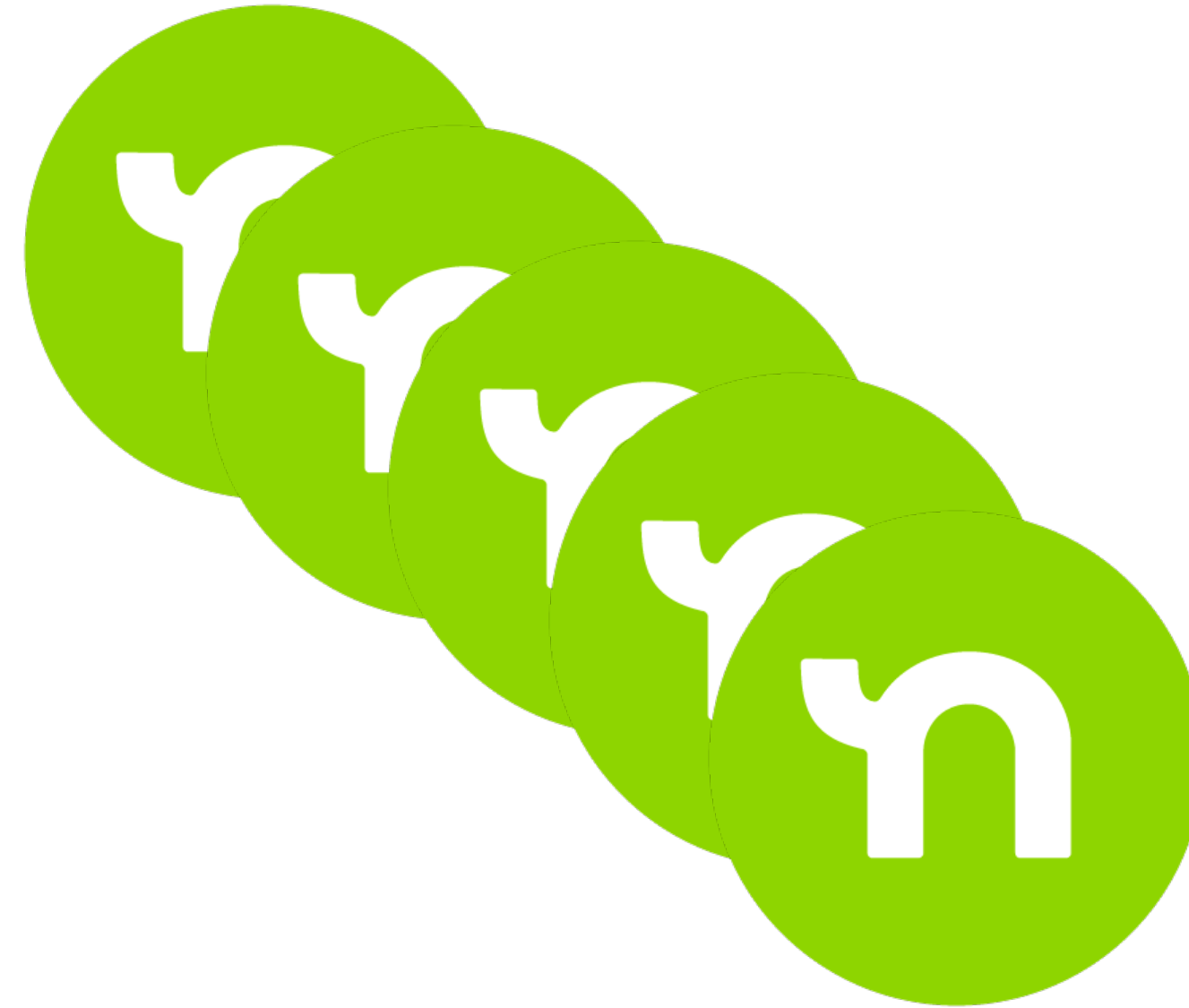
I was waiting on the corner of spring to cross the st Gamp; this homeless man started talking crazy to me and I was scared shitless then another homeless man comes up to him Gamp; said chill out she's my friend so the man left me alone. But Man I was so scared I was about to run 😭



What information does this post contain that could be useful to a social worker studying homelessness?

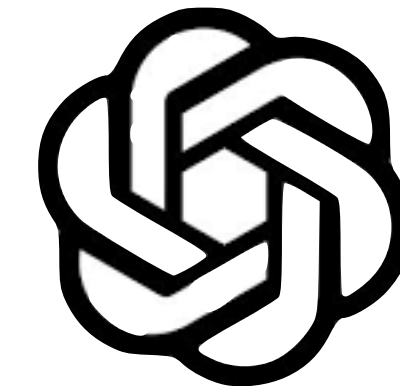
This message expresses fear and discomfort. The implication of the message is that the person felt threatened by the unpredictable behavior of the first homeless man, but was relieved when another homeless man intervened on their behalf. It also implies a certain level of fear or prejudice towards homeless people, associating them with danger or unpredictability.





n=5mil

What information do these posts contain that could be useful to a social worker studying homelessness?



This information can inform social work practice in areas such as:

- Crisis intervention and support
- Community safety and outreach
- Homelessness and social services
- Trauma-informed care and counseling



## OATH-Frames: Characterizing Online Attitudes Towards Homelessness via LLM Assistants

Jaspreet Ranjit♥ Brihi Joshi♥ Rebecca Dorn♥ Laura Petry♦ Olga Koumoundouros♦

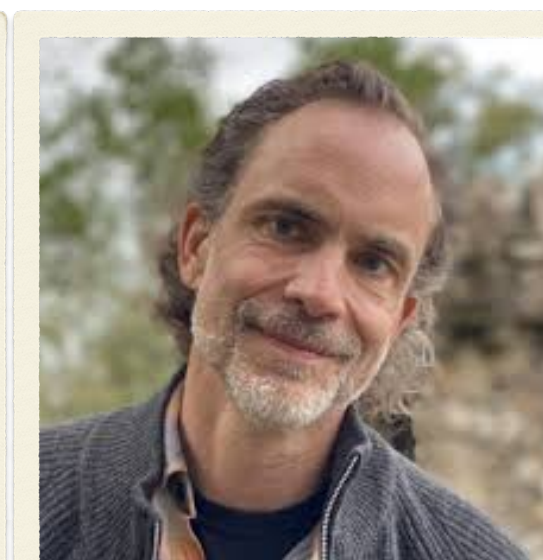
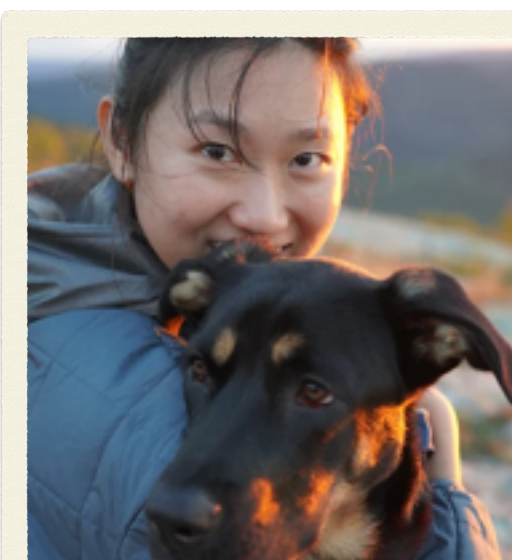
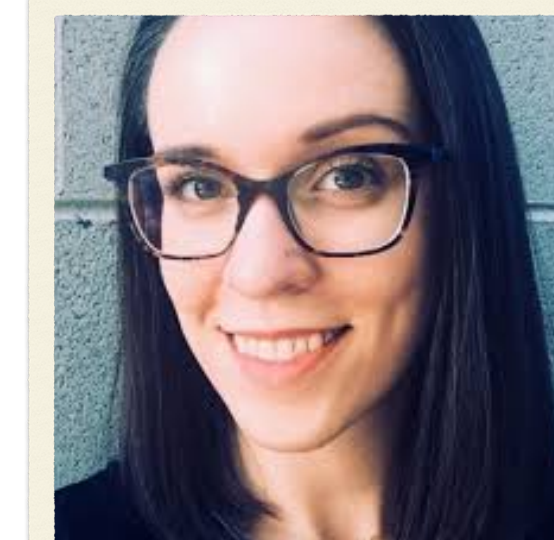
Jayne Bottarini♥ Peichen Liu♦ Eric Rice♦ Swabha Swayamdipta♥

♥Thomas Lord Dept. of Computer Science, University of Southern California

♦Suzanne-Dwork School of Social Work, University of Southern California



Outstanding  
Paper Award at  
EMNLP 2024





X

I think all homeless people should go to Ottawa and protest to our Government about the horrible conditions they live with. Being arrested might be better than their current conditions at least they'll have a bed and 3 meals per day. What do you think Evan???????

media portrayal

not in my backyard

harmful generalization

deserving/undeserving

personal interaction/observation

Perceptions

government critique

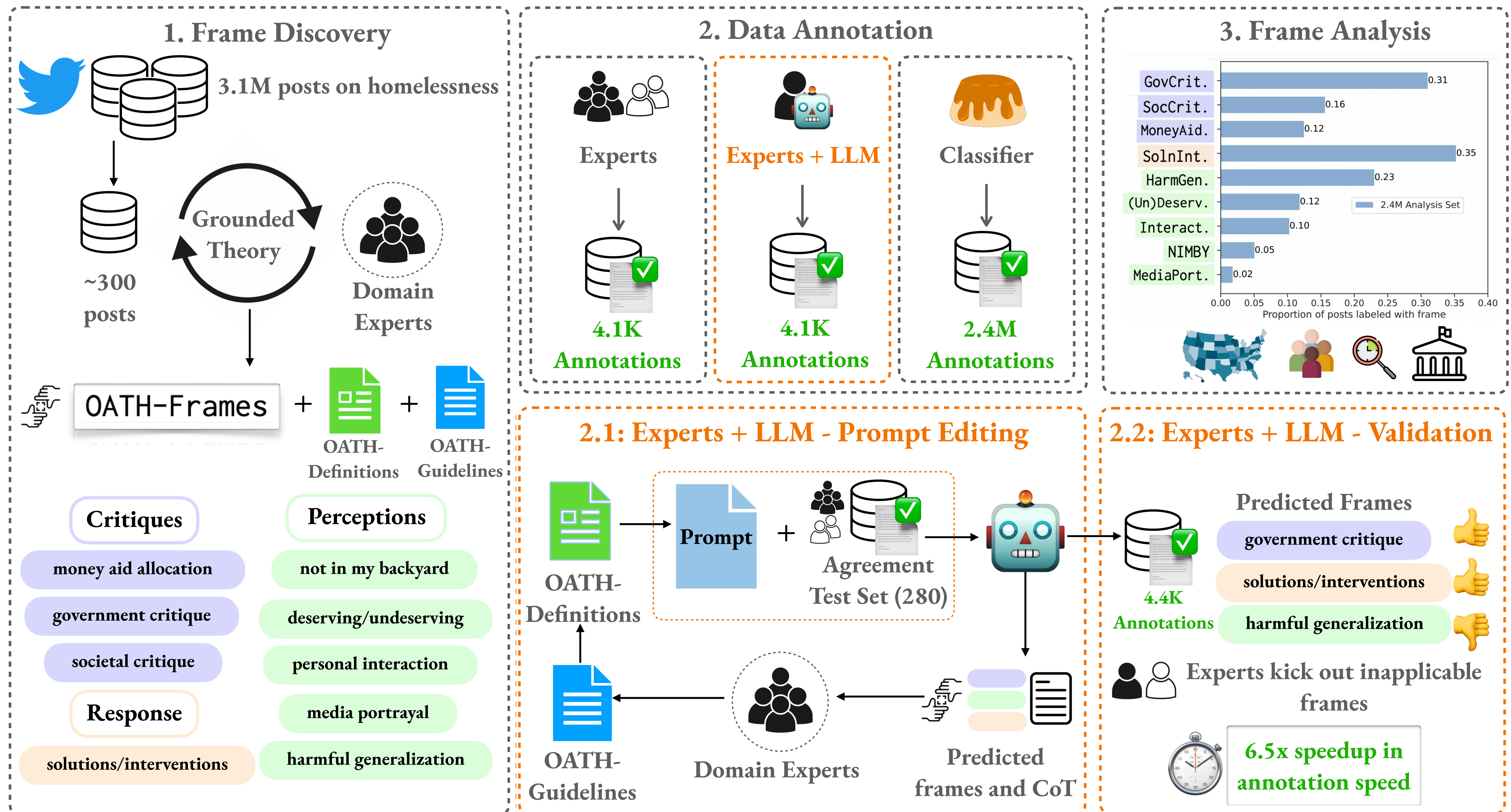
societal critique

money allocation

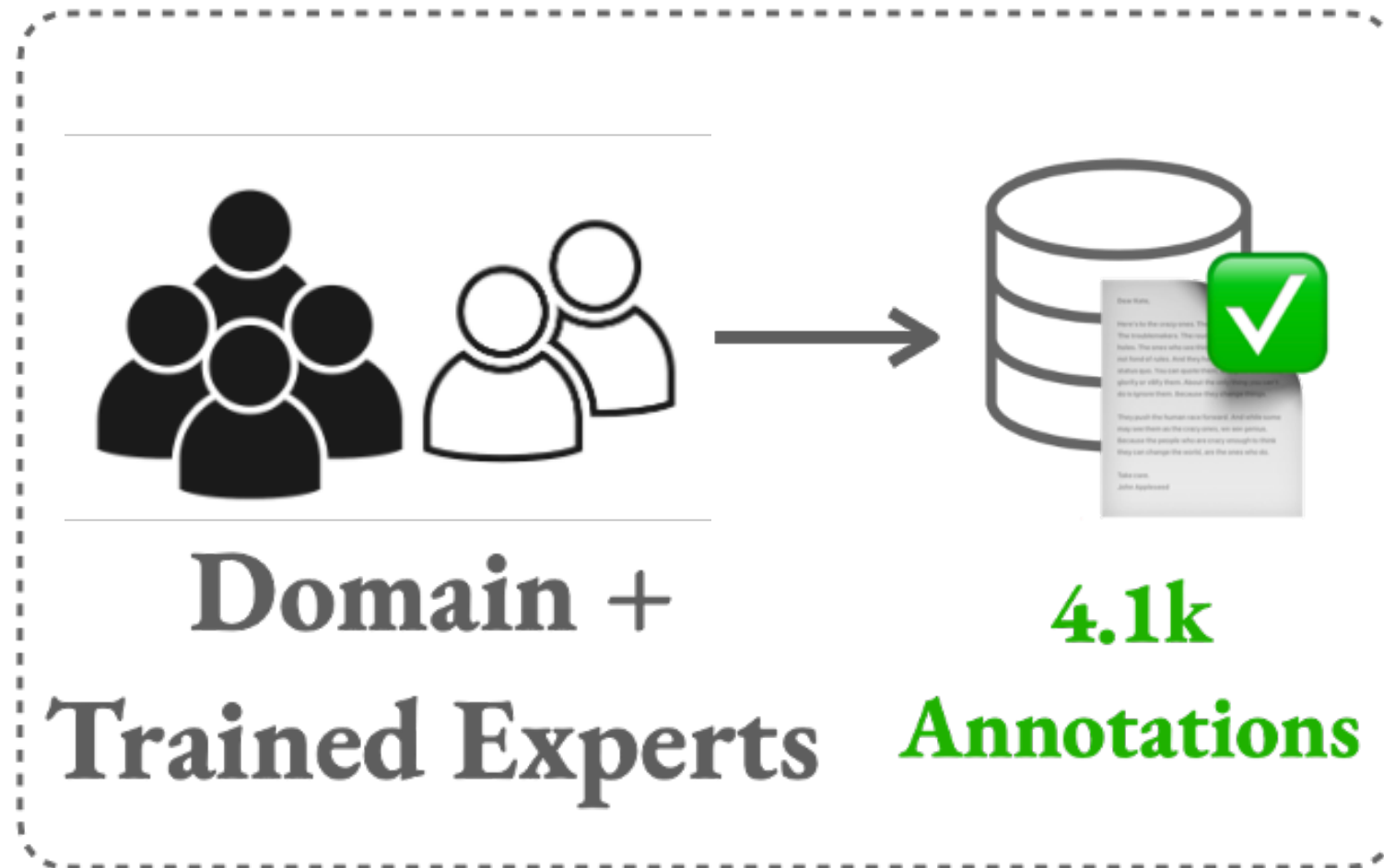
Critiques

solutions/interventions

Responses

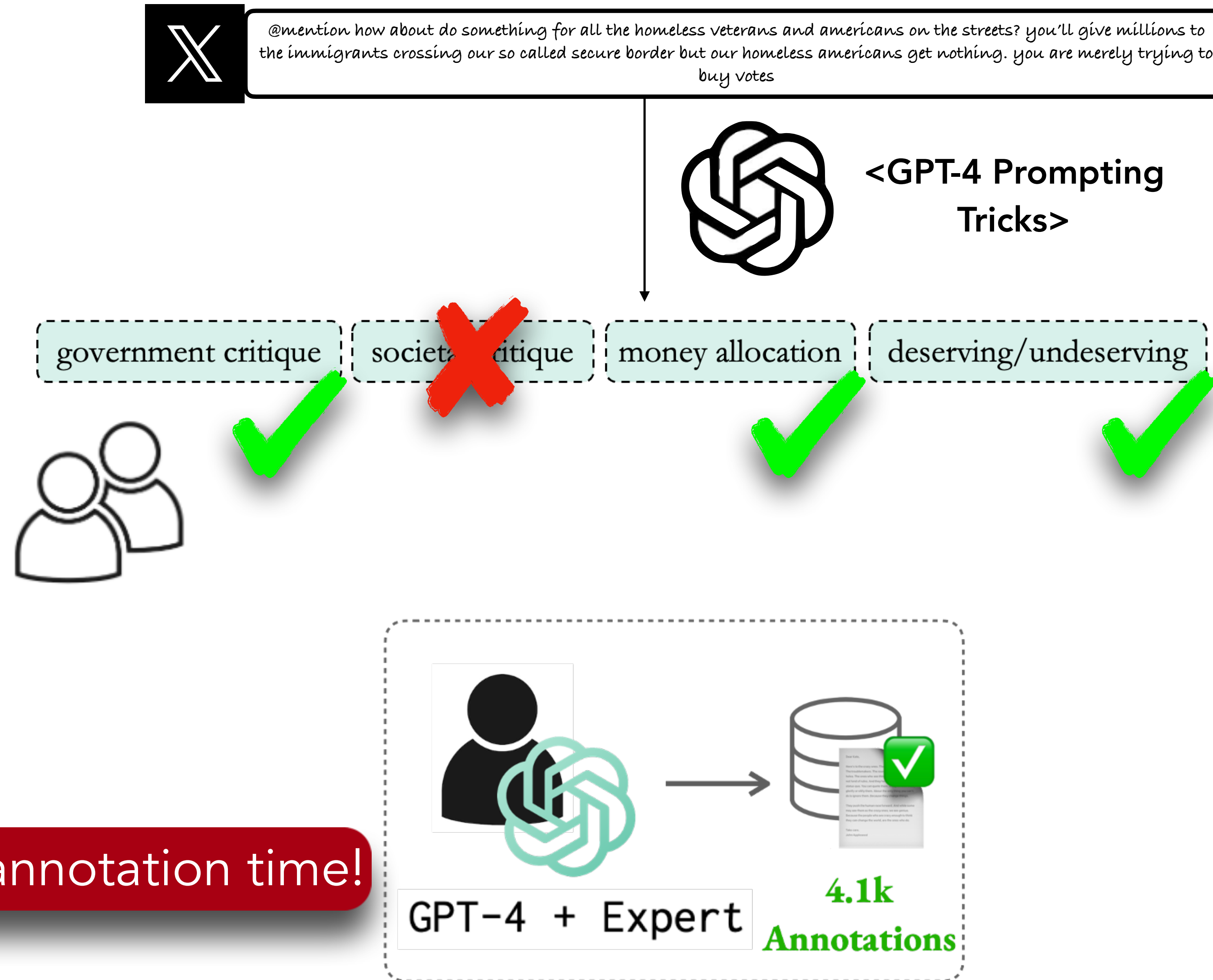


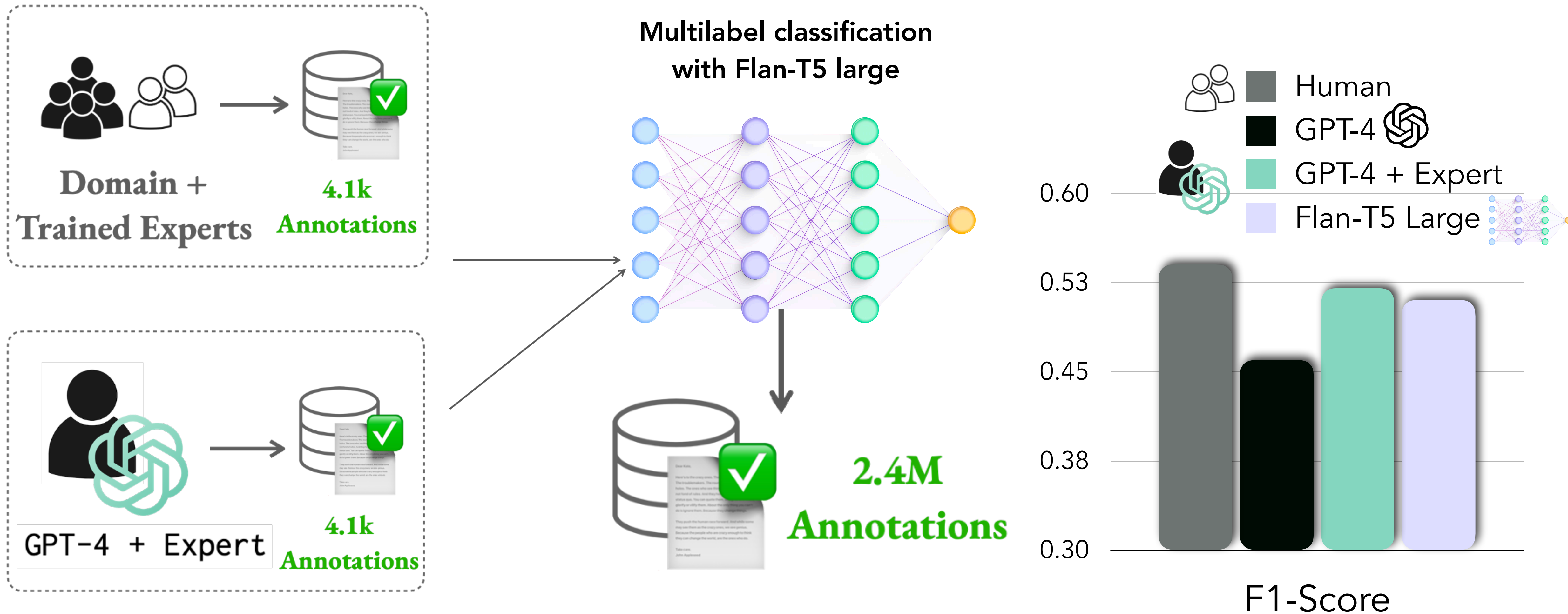




Laborious process which took more than 7-8 months!

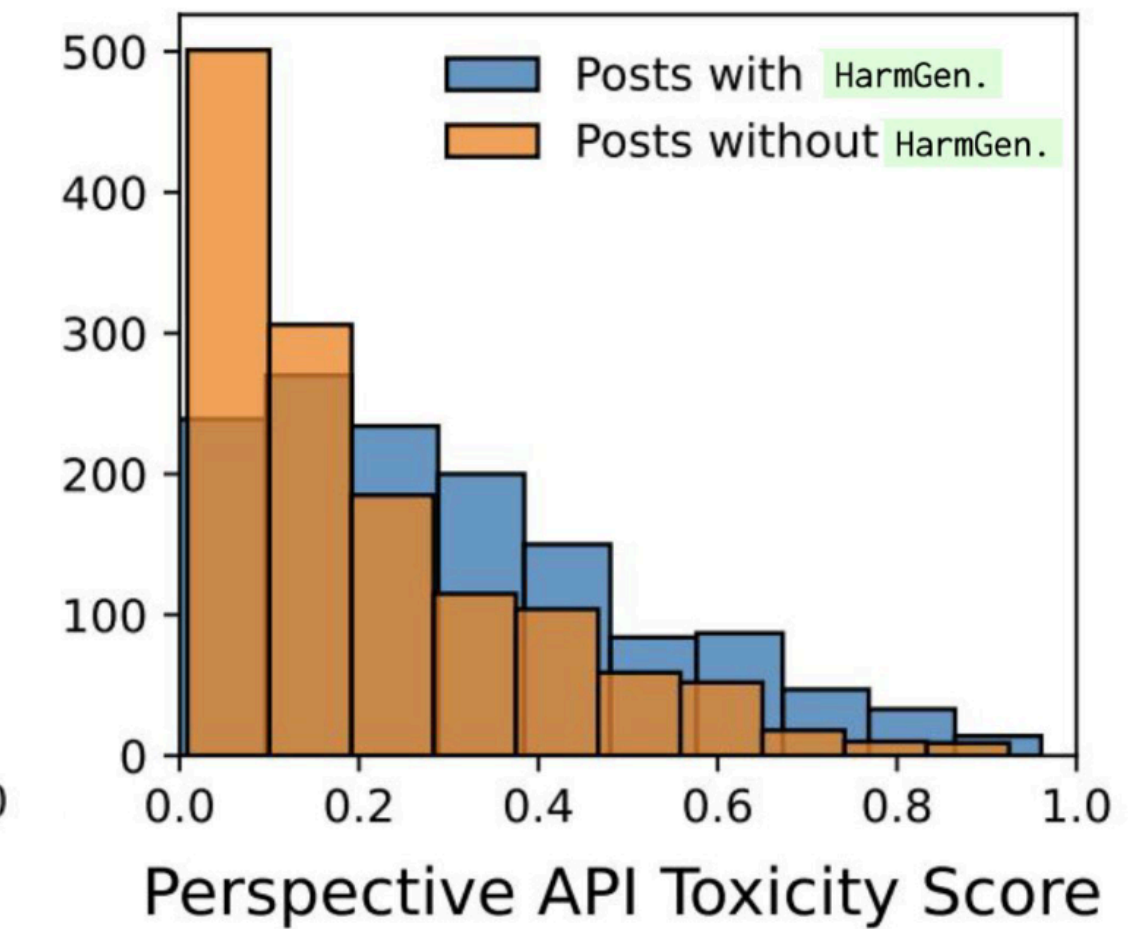
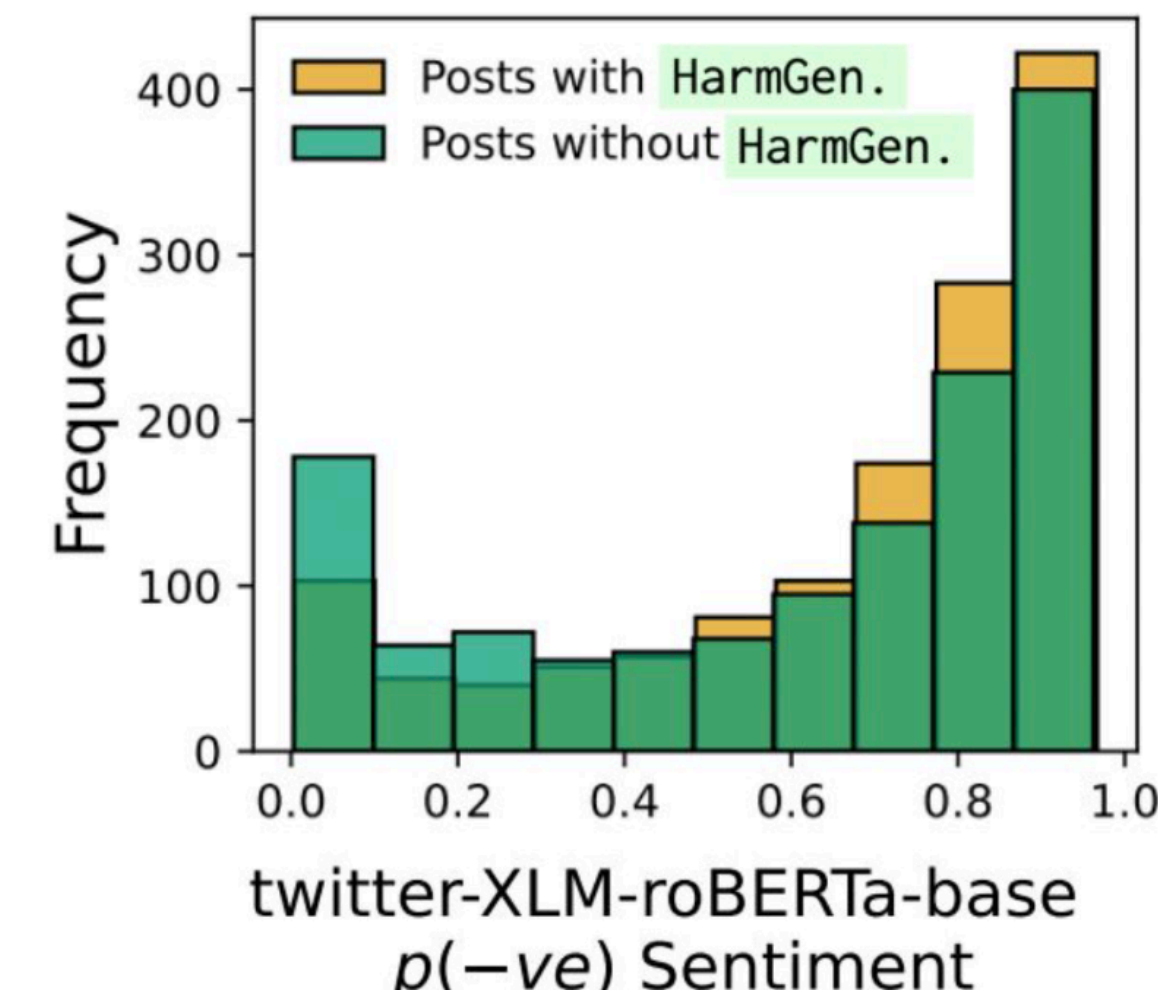
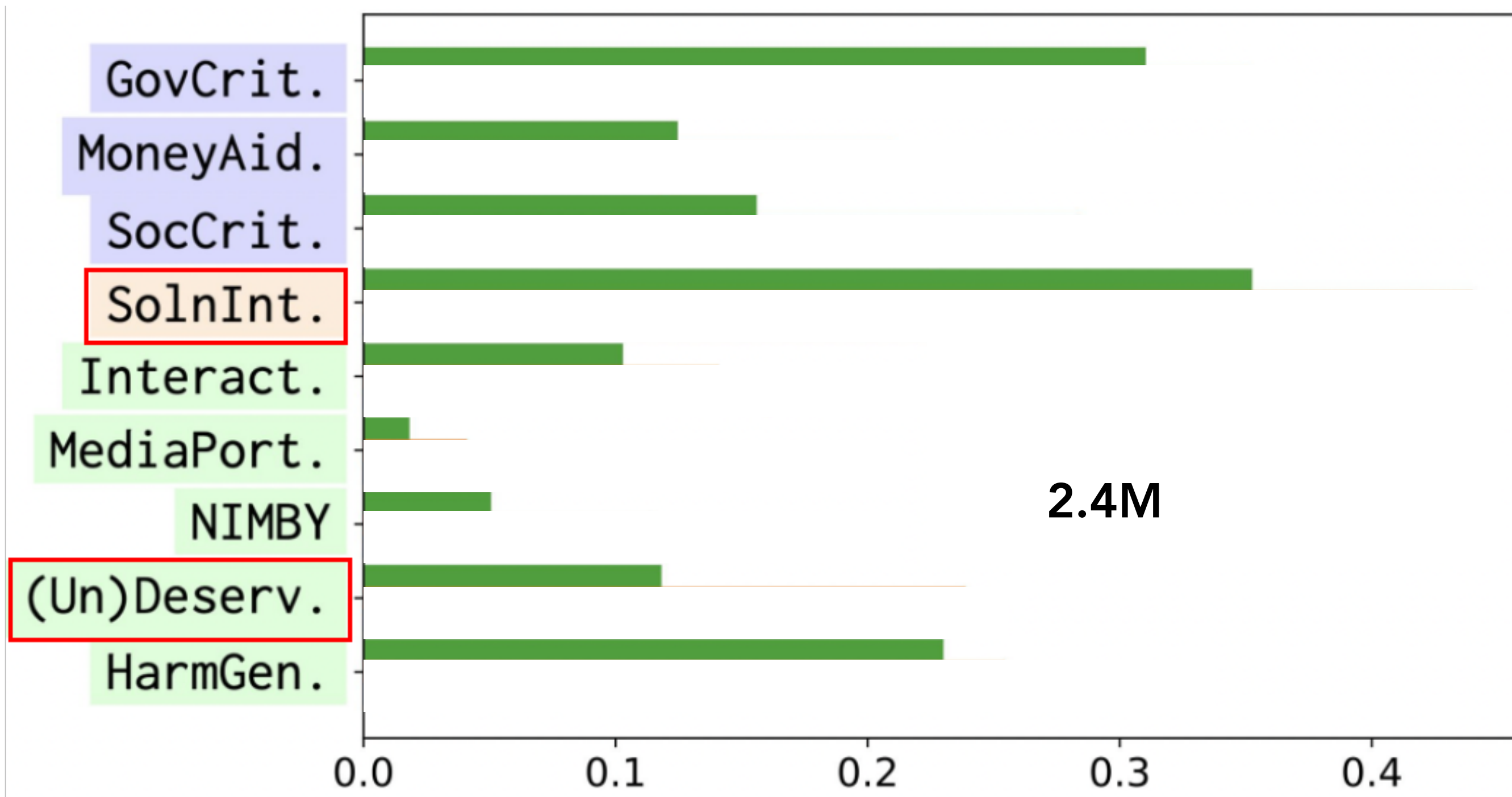
Results in 6.5x speedup in annotation time!



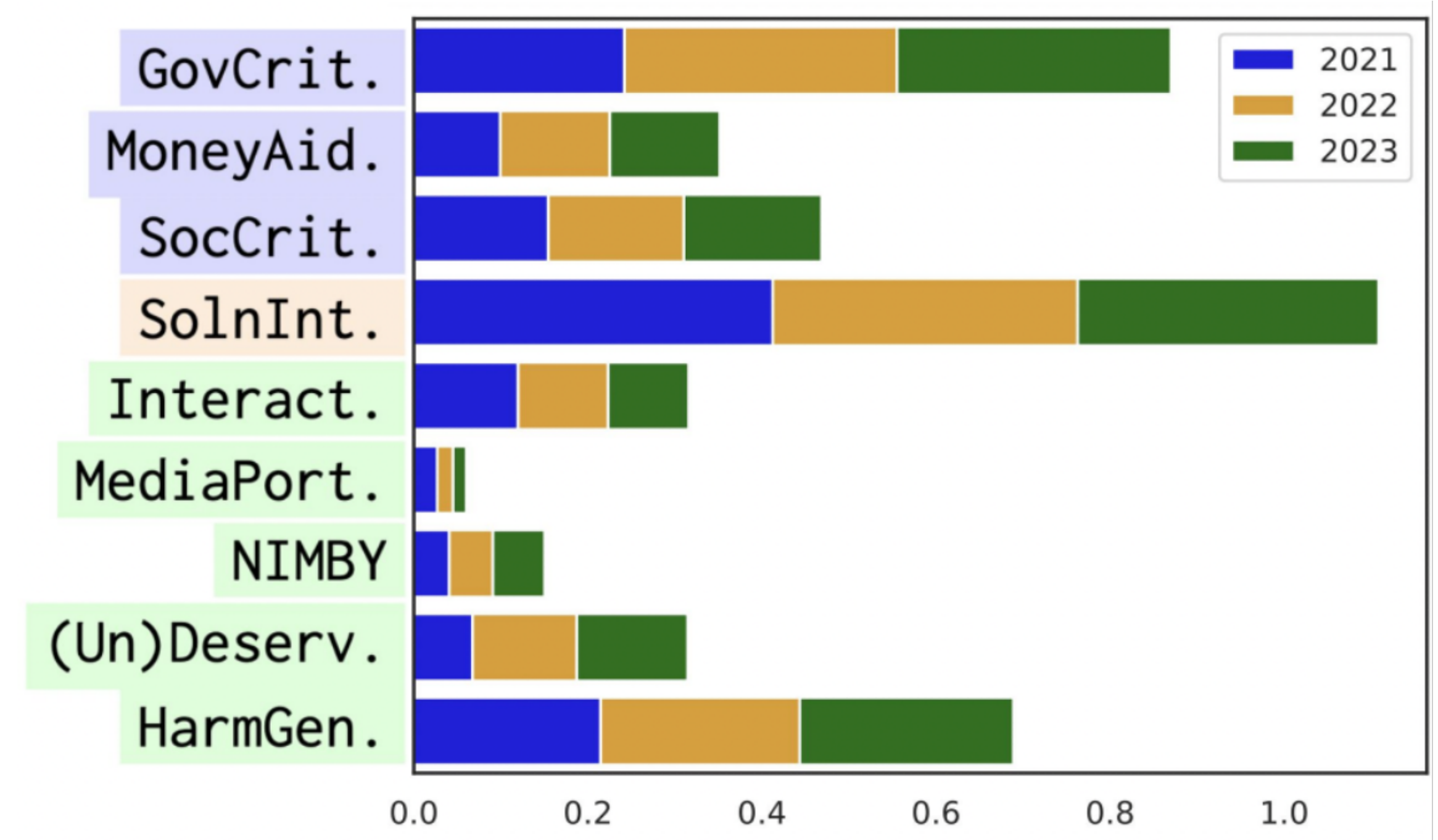
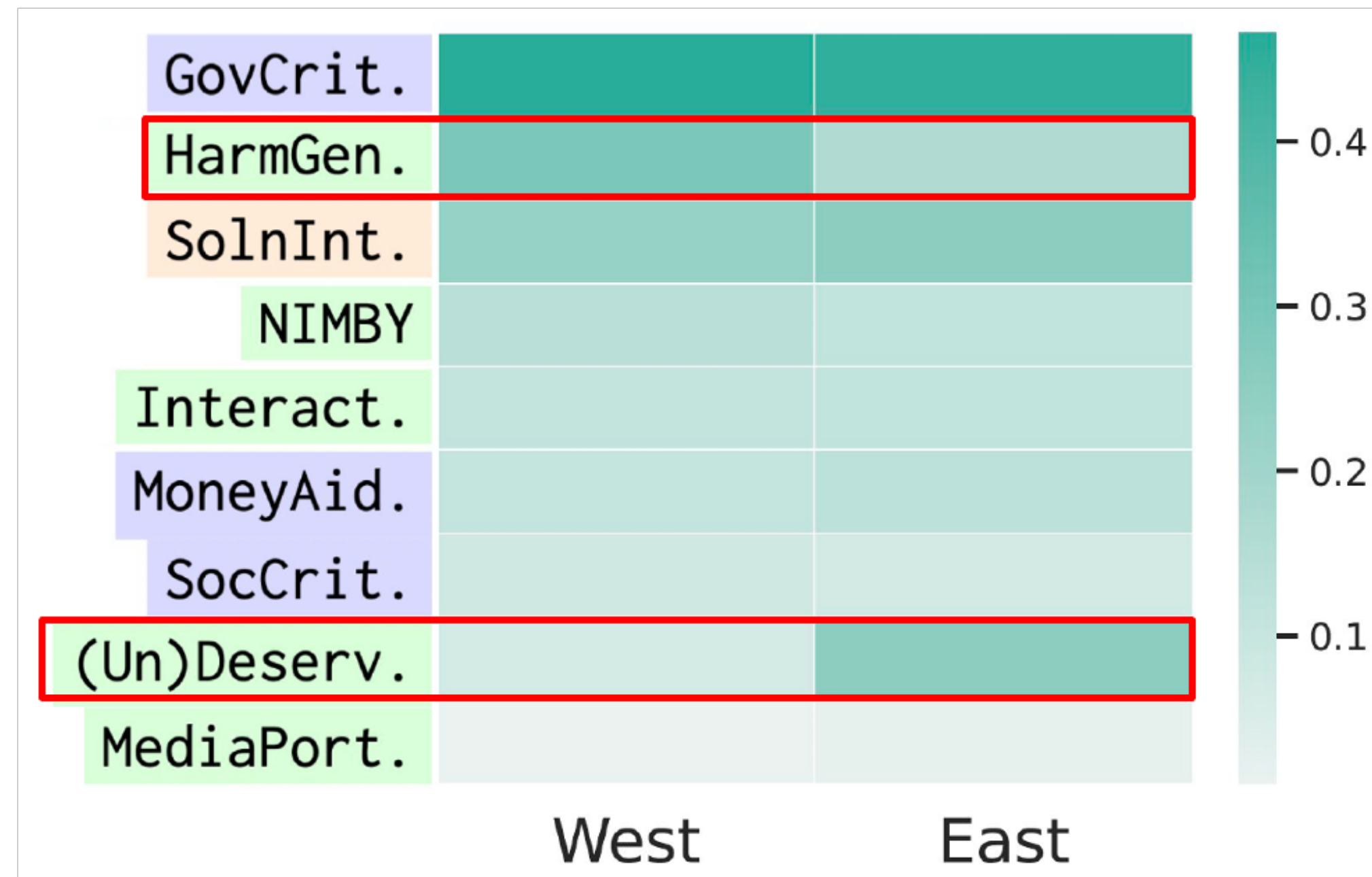


With some effort, language models can be used as assistants for doing a first round of annotations to determine pragmatic frames for complex social phenomena

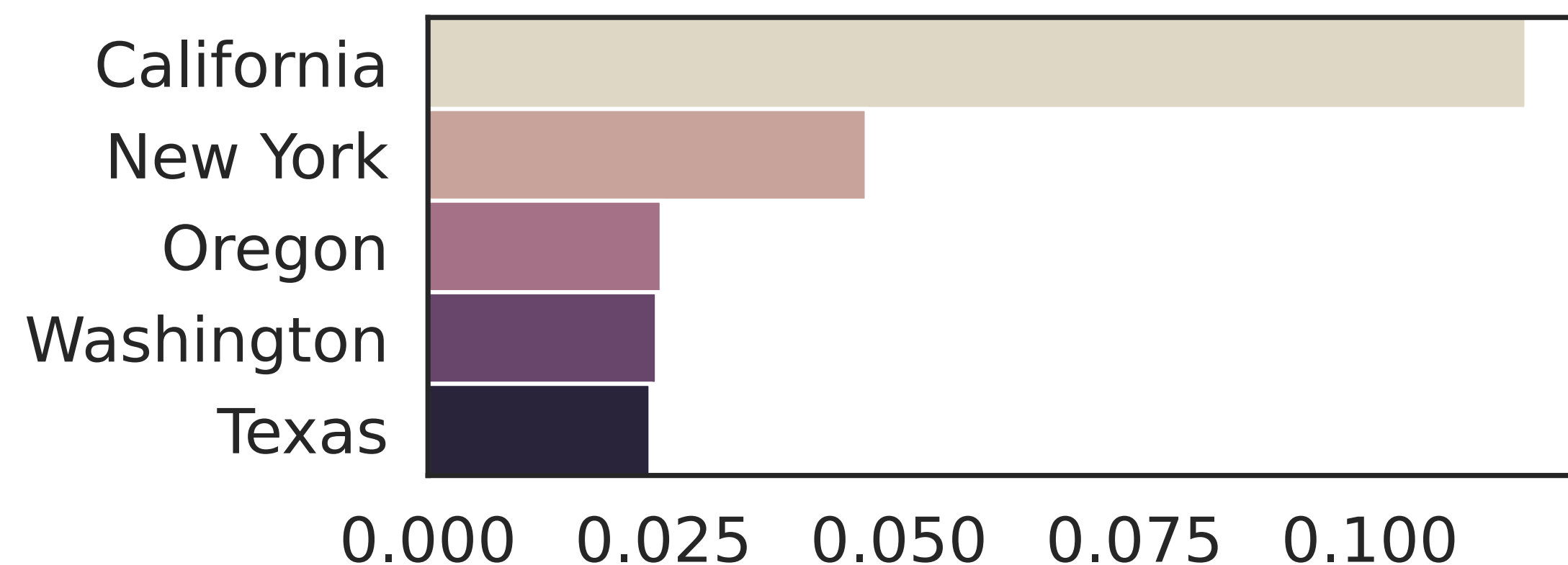




Analysis of predictions allows characterizing online conversations at scale, going beyond sentiment classification

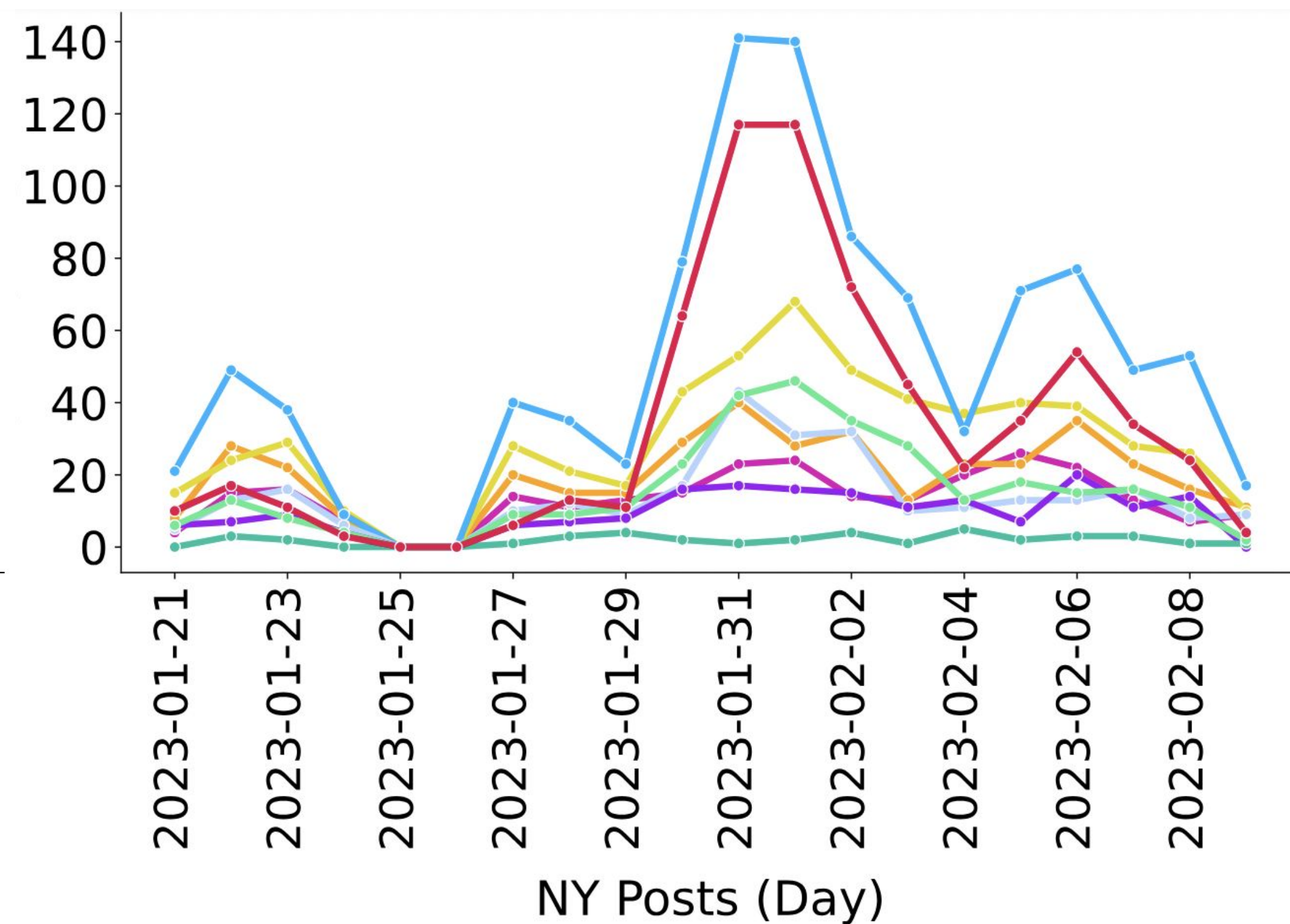
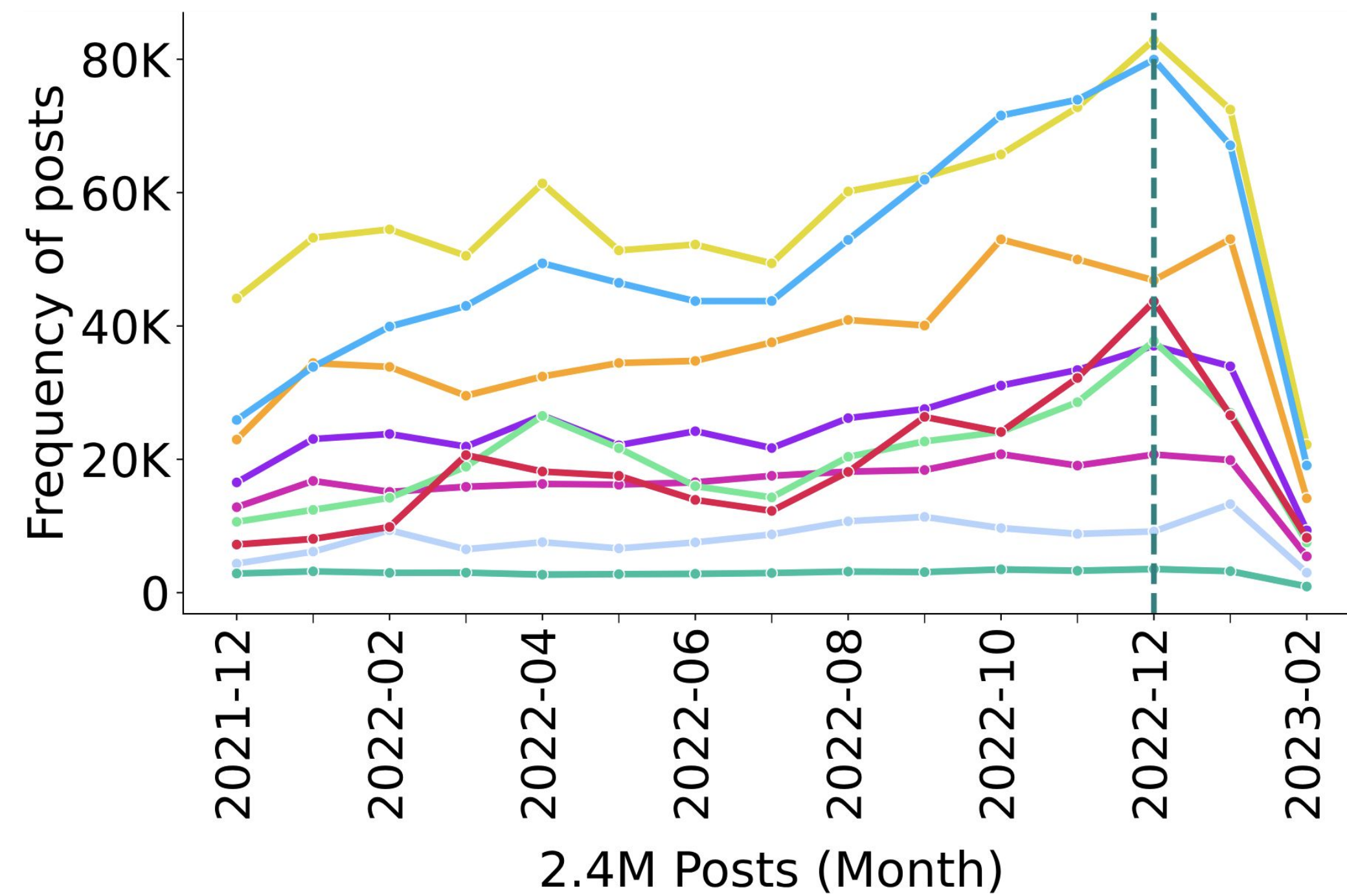


### Posts with mentions by state and labeled with the NIMBY frame



Analysis of predictions allows summarizing online conversations at scale





Trends in our predictions match real life events



Question

What are harmful stereotypes about people experiencing homelessness?

Clustering Method

Select a clustering method.

KMeans

Granularity

Start Date (YYYY-MM-DD)

2022-01-01

End Date (YYYY-MM-DD)

2023-01-01

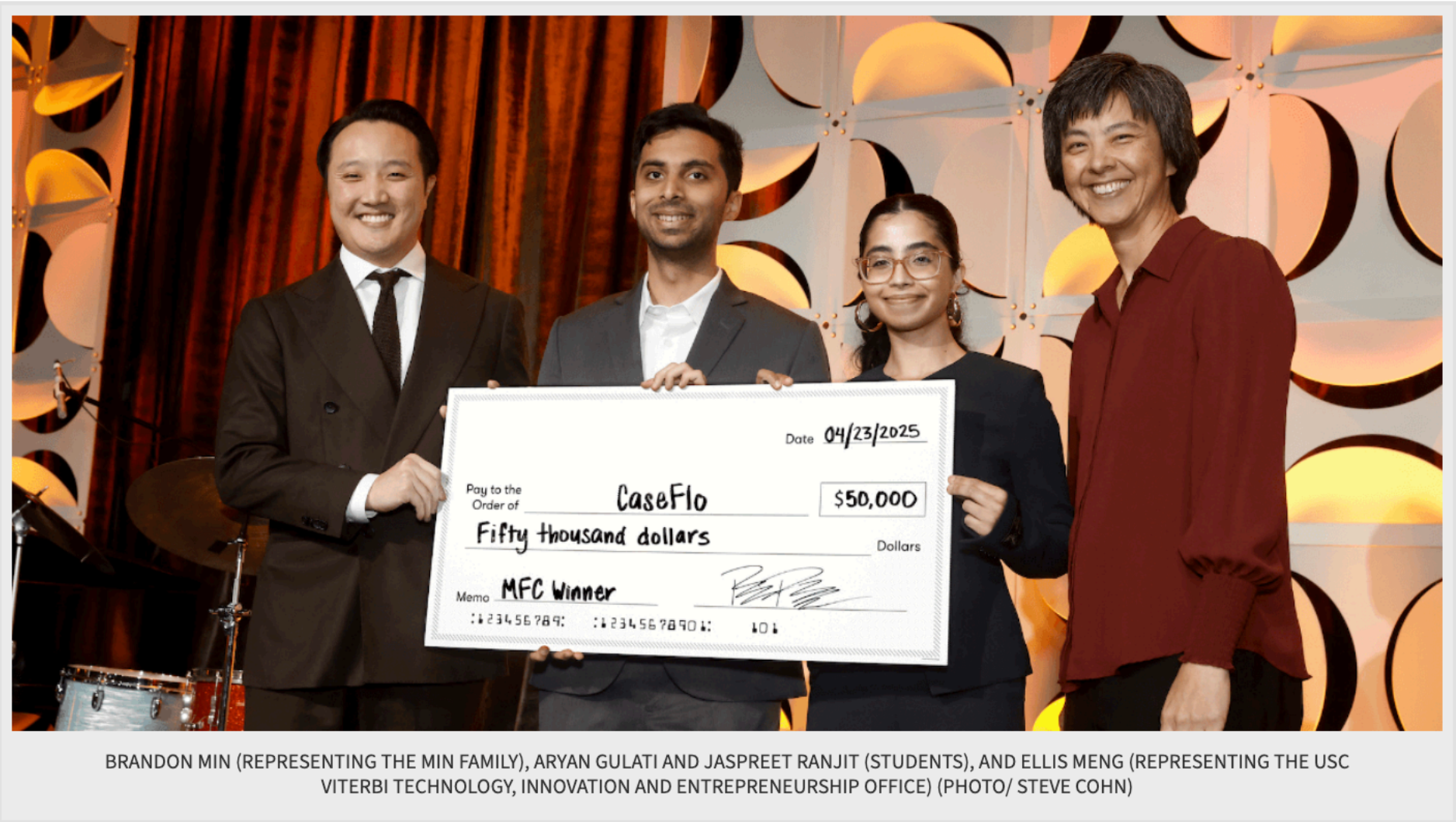
Submit

Dehumanizing Homeless Stereotypes

Description: Homeless individuals are dehumanized by stereotypes, allowing others to justify inhumane treatment.

Summary: This cluster of tweets expresses strong opposition to dehumanizing stereotypes surrounding homeless individuals. The authors, who claim to have represented or interacted with homeless people, condemn these stereotypes as "horrible and wrong". They argue that such stereotypes serve as a means for others to justify inhumane treatment of homeless people, treating them as "garbage". The overall sentiment is one of outrage and disgust at the dehumanizing effects of these stereotypes.

Representative Instances:  
["@5749000633872006893 @5848355391252785669 @2108533698217862798 @5833546692433784574 @6011456147601730776 @2832111684478264872 i haven't been homeless, but i've represented some homeless people. the stereotypes are horrible and wrong. they're just a way for others to dehumanize homeless people so they can justify treating them like garbage. it's sickening."]





# Uncovering Intervention Opportunities for Suicide Prevention with Language Model Assistants

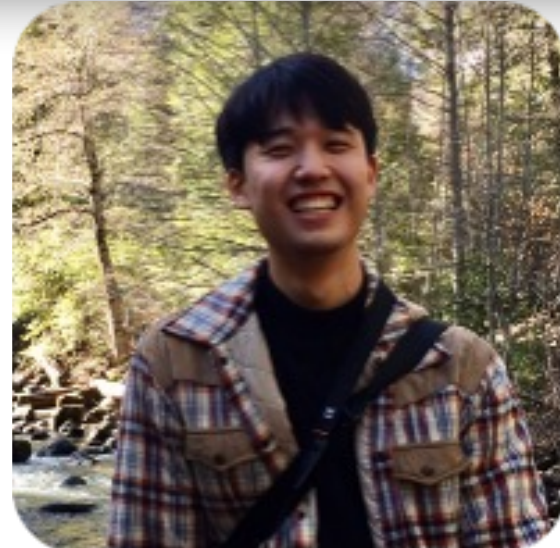
Jaspreet Ranjit<sup>♡</sup> Hyundong J. Cho<sup>♠</sup> Claire J. Smerdon<sup>♡</sup> Yoonsoo Nam<sup>♡</sup>  
Myles Phung<sup>♡</sup> Jonathan May<sup>♠</sup> John R. Blosnich<sup>♣</sup> Swabha Swayamdipta<sup>♡</sup>

<sup>♡</sup>Thomas Lord Dept. of Computer Science, University of Southern California

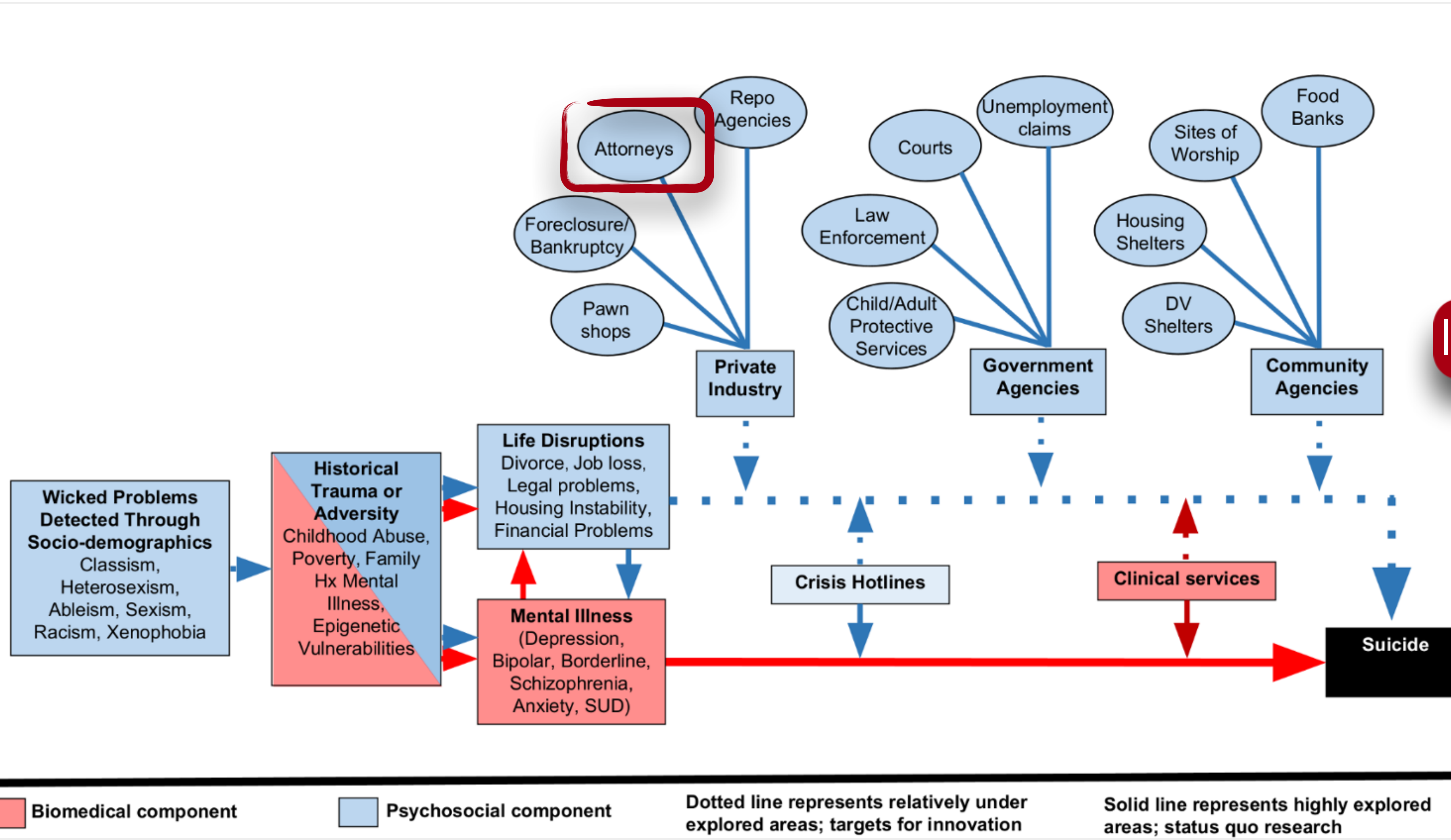
<sup>♠</sup>Information Sciences Institute, University of Southern California

<sup>♣</sup>Suzanne-Dwork School of Social Work, University of Southern California

{jranjit, hjcho, blosnich, swabhas}@usc.edu





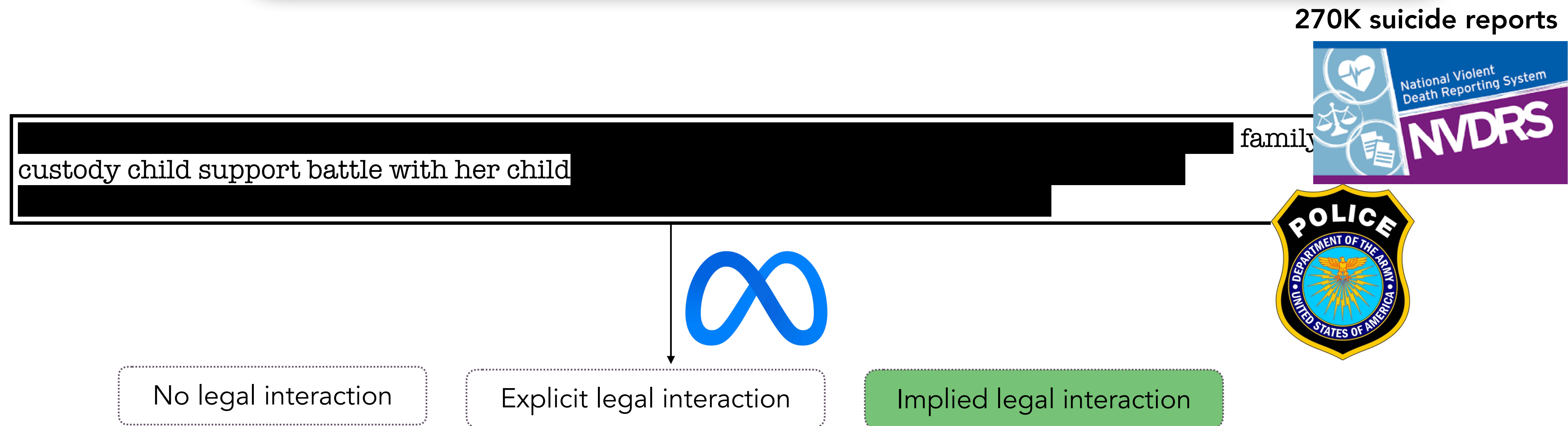


## Industries of Disruption

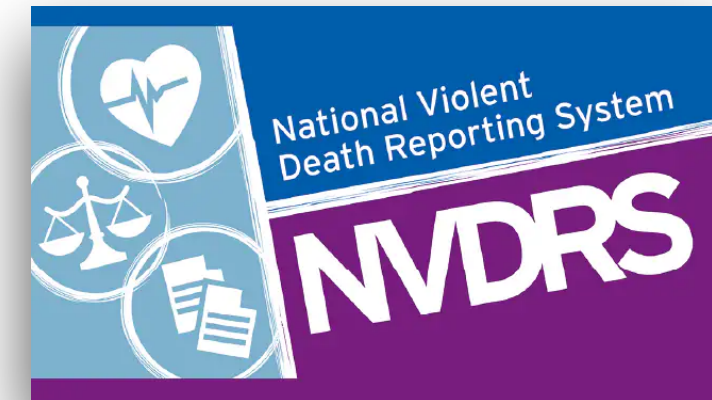


Heard you have been helping out our social work colleagues with language models. Could you help with analyzing suicide reports to explore novel factors for designing suicide interventions?

Can LMs assist experts in extracting new variables from NVDRS, potentially leading to new intervention opportunities?







22 year old white female hung herself in closet of her home. V was depressed over her job and family custody child support battle with her child V suffered from depression and was receiving treatment. V had previous attempt of suicide and a note was found at the scene

*Suicide report text redacted*

V (64 year old White male) had a long history of alcoholism and depression. V was taking antidepressants. V had been threatening suicide for three months. V had been court-ordered into a rehabilitation center but V was appealing this. On the night of the fatal incident, V's wife was sleeping on the couch since V had been

*Suicide report text redacted*

Legal and other non-clinical professionals could be trained to detect patterns and behaviors which may lead to suicide interventions

Victim, a 52 year old male, was found hanging from a tree in his home. V was unemployed and had a history of depression, anxiety, and alcoholism. Autopsy and forensic changes are advanced considering time of death is estimated to be 12-16 hours prior. Three empty bottles of beer were present along with 2 empty packets of Diphenhydramine. V and his wife are separated and are in the process of a divorce. Reportedly, the wife and the V had a disagreement in regard to financial matters (uncertain if financial problems are a factor). No note found.

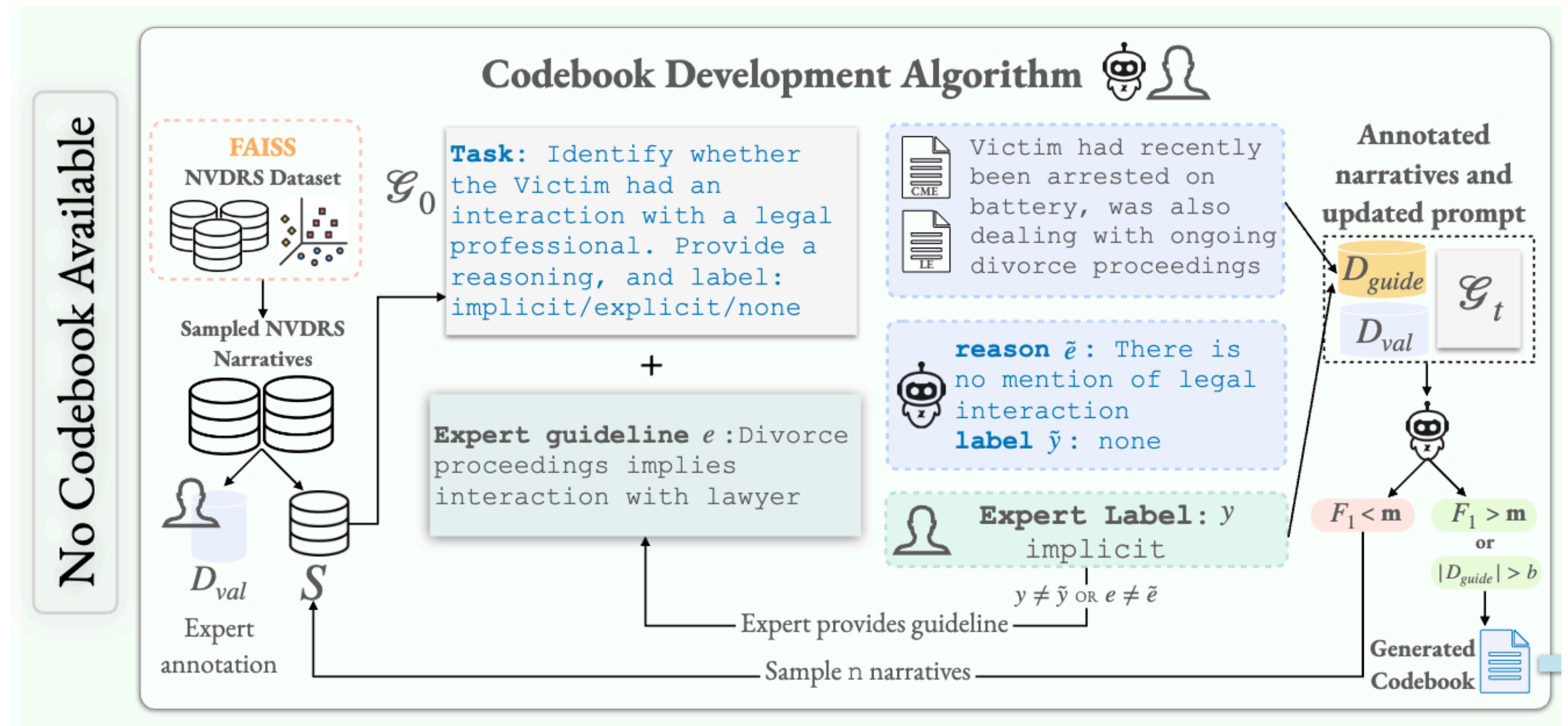
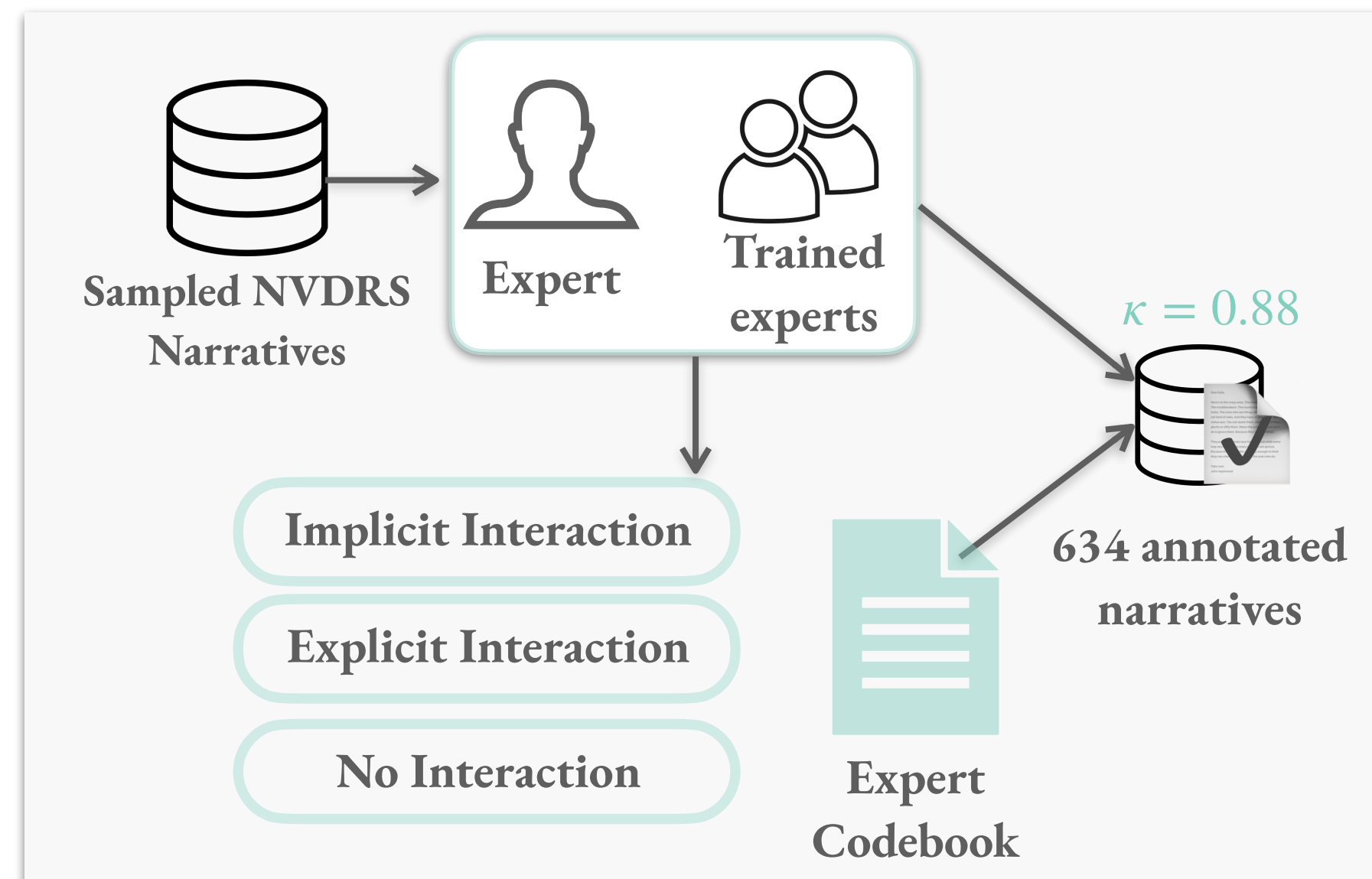
in the process of a divorce

financial problems

*Suicide report text redacted*

# LMs in Social Work Annotation

## Manual Codebook Development: Legal Interactions

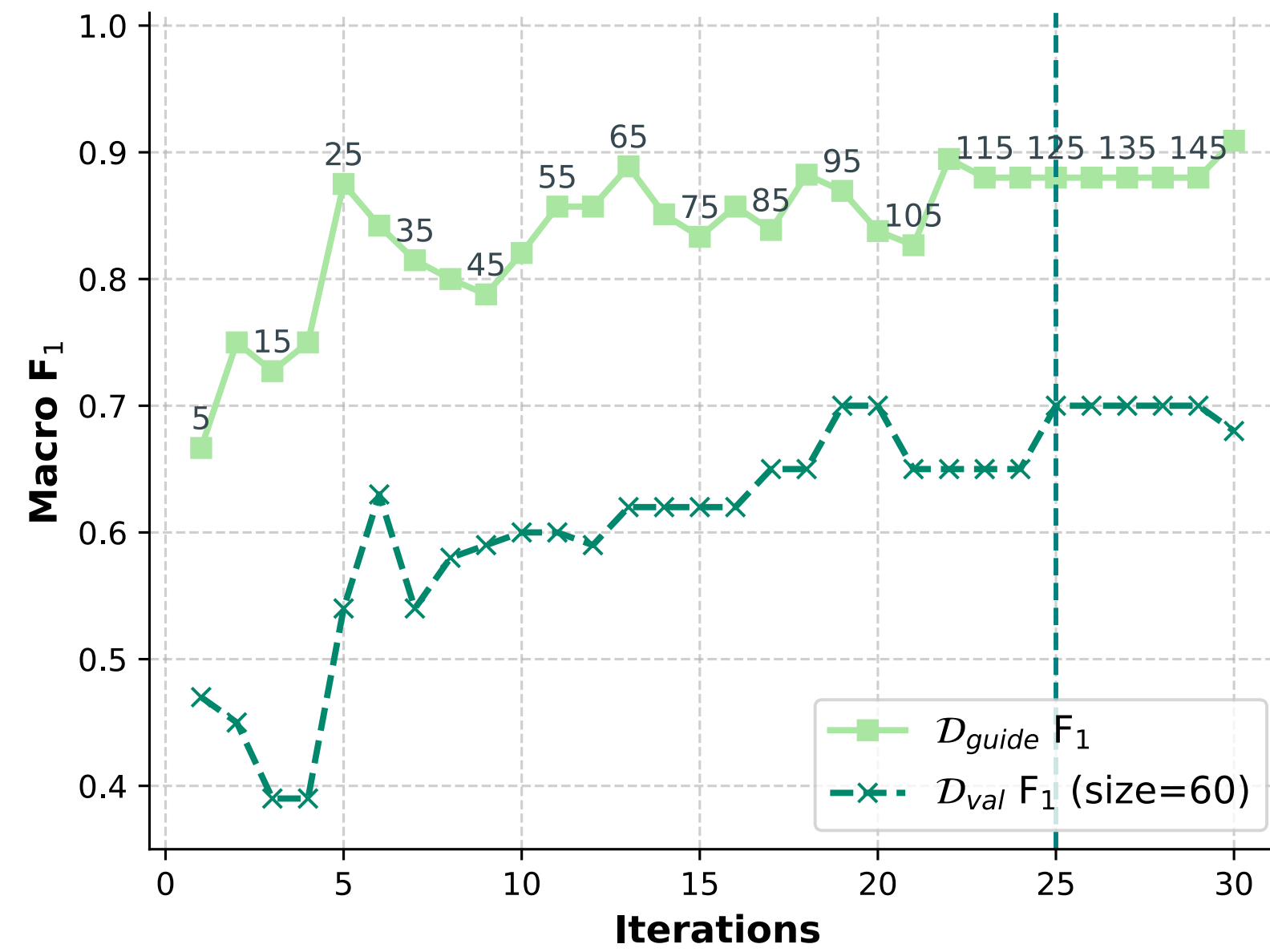


Speedup in process: weeks to hours!

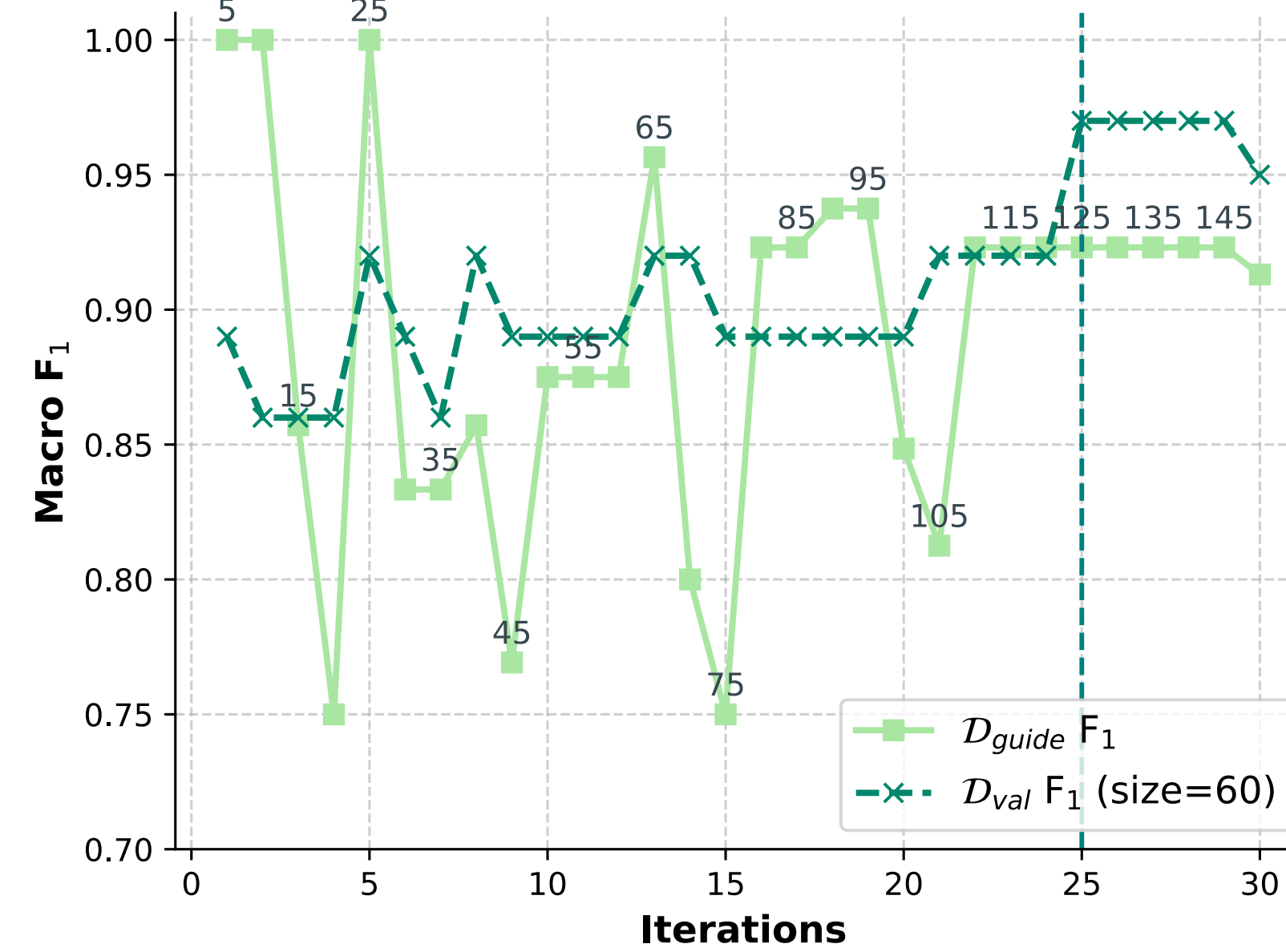
Instead of performance alone, speedup in manual labor at the same performance level should be an important metric in real-world social work annotation



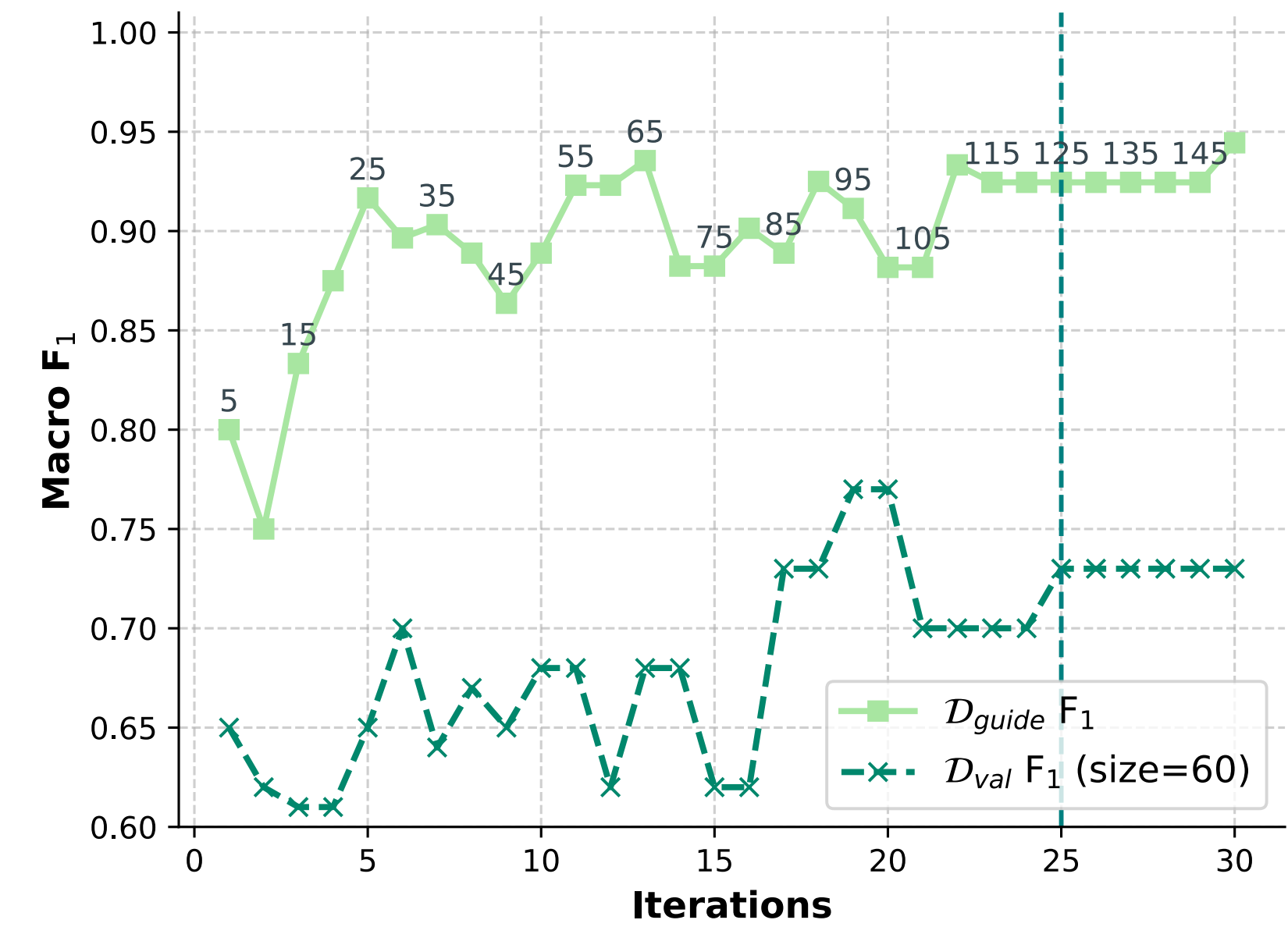
Implicit Interaction



Explicit Interaction



No Interaction

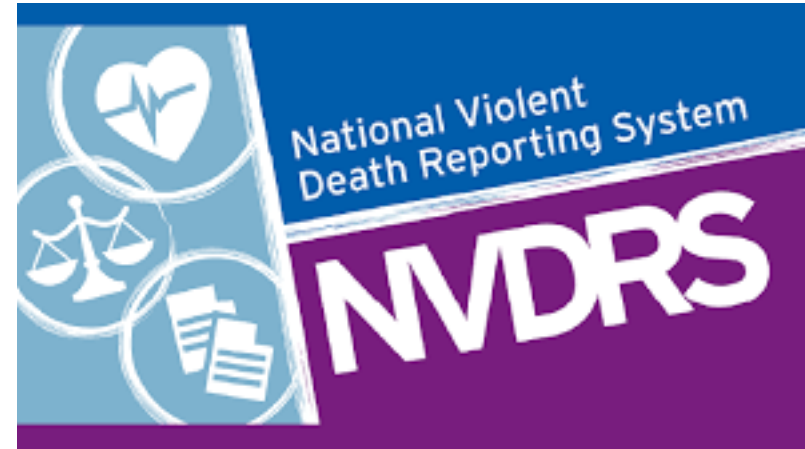


LMs provide better (more stable) predictions with longer iterations

Model	$\mathcal{G}_0$	$\mathcal{G}_{HitL}$	$\mathcal{G}_{expert}$
Llama-3-70B	0.57	0.80	0.78
Qwen2.5-32B	0.68	0.76	0.77
Qwen2.5-14B	0.63	0.75	0.77



270K suicide reports



10.4% of 270K suicide narratives had evidence of legal interactions!

## AI with Human Expert Guidance Can Greatly Alleviate Manual Labor in Sensitive Tasks

Our research was going at the pace of a bullock cart and you came at it with a jet plane!

USC Suzanne Dworak-Peck  
School of Social Work



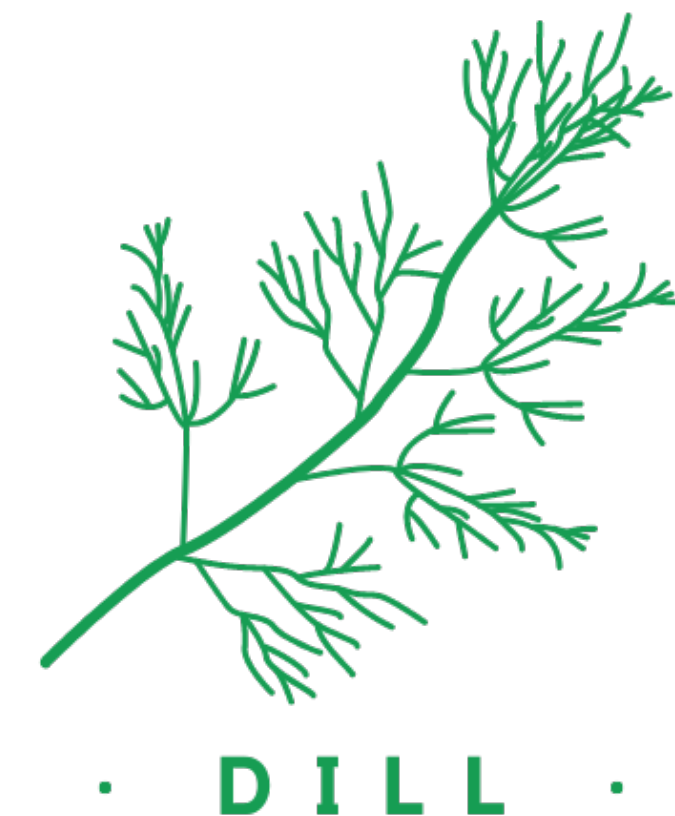
However, we need to take into account the possibility of unintended harms!





**USC** Viterbi  
Natural Language Processing

# Thank You!



[swabhas@usc.edu](mailto:swabhas@usc.edu)