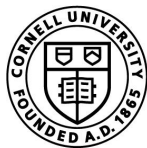# Automate or Assist? The Role of Computational Models in Identifying Gendered Discourse in US Capital Trial Transcripts

Andrea W Wen-Yi,   Kathryn Adamson,   Nathalie Greenfield,   Rachel Goldberg,

Sandra Babcock,   David Mimno,   Allison Koenecke

Content warning: There is violent description and emotionally intense language in this talk.

Racial bias in language has been widely acknowledged as a concern in capital trials.

Capital trials: trials of death sentence.

The New York Times

Texas Man's Death Sentence Thrown Out Over Racist Testimony

By Matthew Haag

Oct. 4, 2017

How about gender bias language?

Gender bias language: language that invoke gender stereotypes

How about gender bias language?

Nathalie Greenfield, one of Ms. Andrew's lawyers, said gender stereotypes infected the trial and poisoned the jury.

"Every single day the state was presenting gendered evidence about her appearance, about her clothing, about her sexual practices, about her skills as a mother," she said. "We've got someone who is at risk of execution for not conforming to gender stereotypes."

"Gender bias is normalized and tolerated to an extent that racial bias no longer is in the administration of the death penalty," said Sandra Babcock, a law professor at Cornell who represents Ms. Andrew in a related case. "Women on trial for capital murder have been subjected to similar shaming tactics for hundreds of years."

source: "Did Prosecutors' Sex Shaming Help Send Brenda Andrew to Death Row?", New York Times, 2024

Is there a more objective evidence?

Can you measure gender bias language in trial transcripts?

Goal: Annotating four types of gender bias language in trial transcripts.

- **Emotions** *"[The defendant] is an actress who feigns grief and hysteria when it suits her. "*
- **Betrayal of Gender Norms** *"[The defendant] is a cunning, manipulative woman motivated by power and control."*
- **Hypersexualization** "*[The defendant] uses people to satisfy her sexual desires, and in furtherance of her plot to kill for money."*
- **Bad Mother** *"[The defendant]'s house is so disgusting for her kids to live in."*

Barbara

I got a degree in geology. When I came back here with my husband, I had done plenty of field mapping and working with GIS, and so I got a job with a soil scientist here who helps people try to get septic permits – and environmental field assessments, watershed mapping, wetland delineation.

Henry

What kind of insight has that given you into the area, given your particular professional angle?

Well, the one thing is that it's very low. The land is so low and the water table is so high. And it's tough for people to get septic permits. As environmental standards have gotten more stringent, some local people that have inherited land are finding that they can't get a permit, or they have but they have to install a very, very expensive pretreatment system, which is too much to pay for. And so I've seen it be a challenge for local people who are trying to stay where they grew up. And so that's a big limiting factor right now for development, which is okay. But it's interesting. Sometimes it's tough when people can't build the house on their land that they inherited or can't do anything with their land, but I understand the limitations also. It's critical to maintain the water quality.

There are people who have even purchased land that they were told could get a system or thought could get a system or just sort of in an unofficial way looks like it. There are options and there are hard luck sort of considerations that the county tries to help people. But it's a balance between helping people who have these hardship situations and protecting the environment. And that's difficult as the person on the ground telling the landowner, "Sorry," and especially when it's not a developer. It's just a young couple or something.

Henry

So do you think that that's the major sort of limiting factor, environmental factor impacting development in Down East?

Q.4. Community and Environmental Change

Henry

Talk about changes that you've seen in this area.

Barbara

The biggest change that I'm aware of is the sense – I think that now there's the sense the commercial fishing as a livelihood is dying as a viable option. And I know – I think that that is a change since 30 years or 35 years ago. I don't remember that being sort of this sense of doom. For me that's the most profound. That's a huge thing that I think is very sad. And that has to do with – that's such a complex situation -- the water quality and also the overharvesting and tearing up of the bottom – all these different things that are factors.

Imported seafood is so much cheaper. And the little sea-side restaurants around here don't sell local shrimp on their shrimp burgers. The fancy restaurants have local seafood. Even if the stocks are fine, it's that you can't make a living spending backbreaking work getting paid – earning nothing, very little.

Henry

Yeah. So it's a combination of environmental factors and market forces?

Barbara

Or pressure on water quality and market forces. And then also the complicated relationship between the methods of trawling or methods that destroy habitat– just several factors. A lot of fishermen work on dredges and they go off to other jobs. It might be there would be less and less seafood for us. The fish houses have largely closed; there used to be quite a few more fish houses and places to sell the seafood that you would catch.

Henry

And what's contributed to the water quality issues? What's caused that problem?

Barbara

The biggest factor in my understanding is stormwater runoff. We don't have a lot of – considering the amount of wetlands we have and the amount of land – we don't have a lot of

impervious surfaces Down East. I think Down East is still in pretty good shape, as opposed to

Q.6. Barriers to and opportunities for achieving your vision

Image source: iStock.com/PhonlamaiPhoto. https://www.istockphoto.com/photo/machine-learning-concept-gm1096227330-294334525.

Three-phase study about employing language models in annotating gender bias language

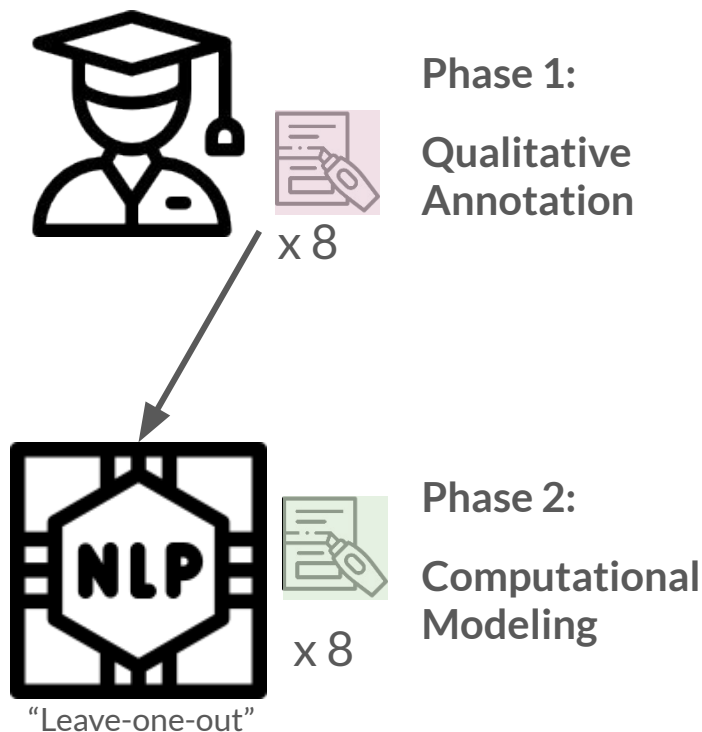# Three-phase study about employing language models in annotating gender bias language
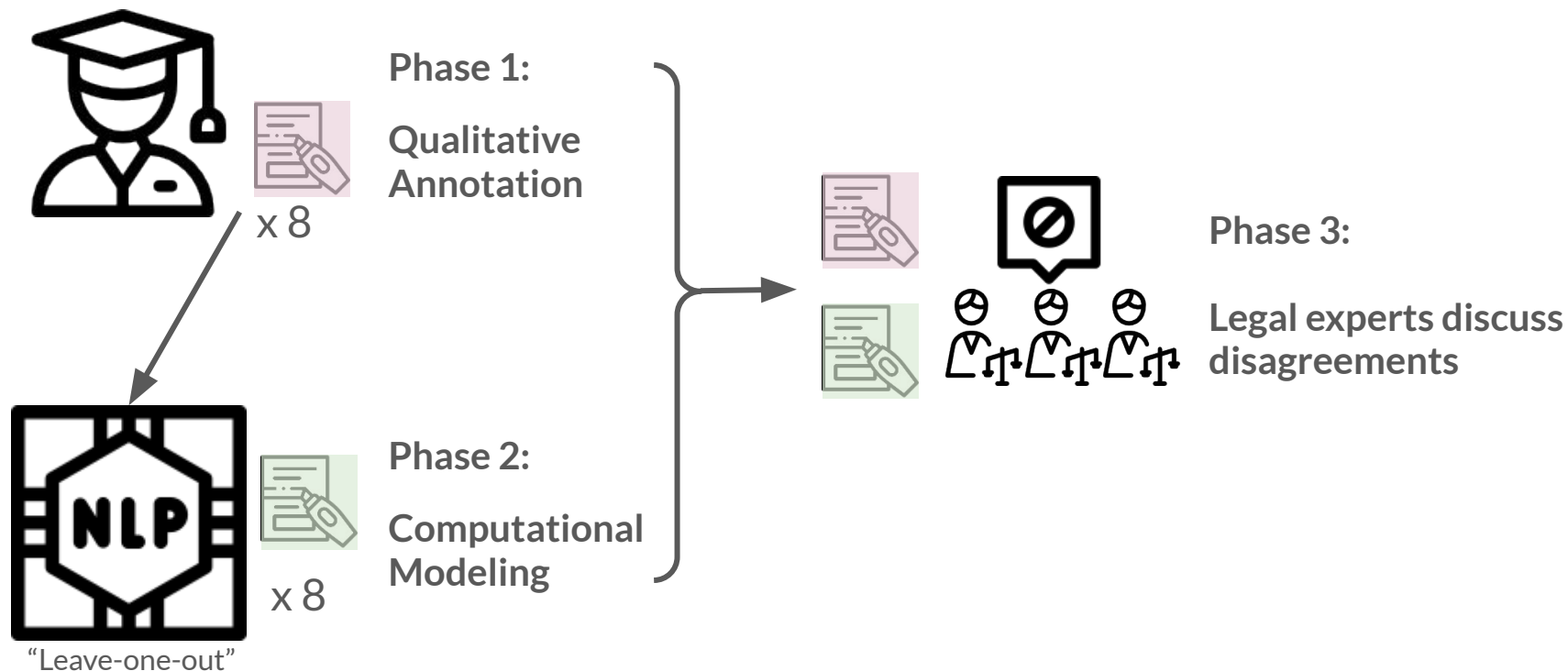

x 8

Phase 1:

Qualitative Annotation

Three-phase study about employing language models in annotating gender bias language

Phase 1:

Qualitative Annotation

x 8

Phase 2:

Computational Modeling

x 8

"Leave-one-out"

Three-phase study about employing language models in annotating gender bias language

x 8

"Leave-one-out"

Phase 1:

Qualitative Annotation

Phase 2:

Computational Modeling

Phase 3:

Legal experts discuss disagreements

The model-experts agreements vary across labels.

	Emotions	Betrayal of Gender Norms	Hyper-sexualization	Bad Mother
Experts agree with **annotators**				
Experts agree with **model**				
Experts have no **consensus**				

The model-experts agreements vary across labels.

	Emotions	Betrayal of Gender Norms	Hyper-sexualization	Bad Mother

Experts **agree with annotators**

Experts **agree with model**

Experts have **no consensus**

Emotions
- 39% (11)
- 57% (16)

Betrayal of Gender Norms
- 41% (13)
- 44% (14)
- 15% (5)

Hyper-sexualization
- 80% (24)
- 10%
- 10%

Bad Mother
- 69% (18)
- 12%
- 19% (5)
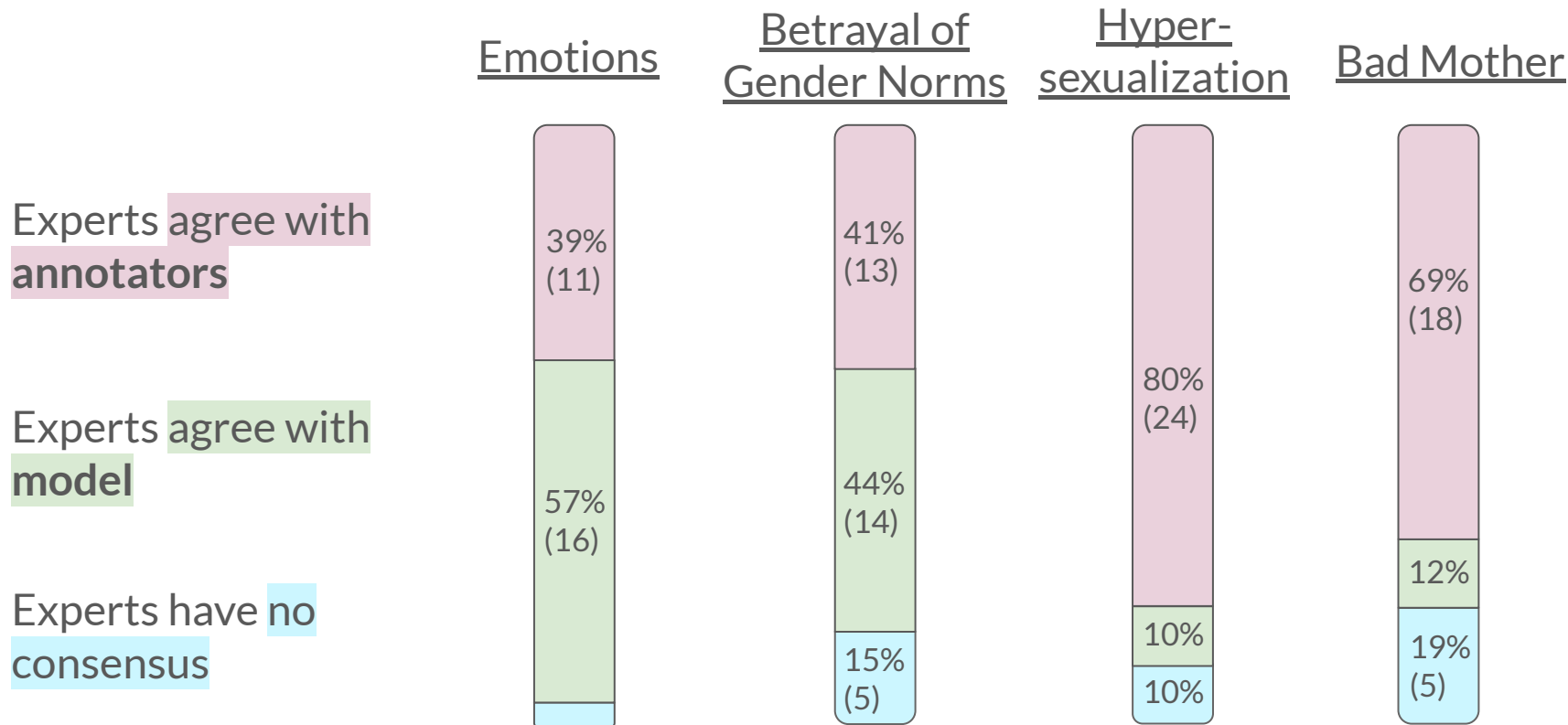
Problem 1: The model learn the wrong things...

Emotions

Betrayal of Gender Norms

Hyper-sexualization

Bad Mother

Experts agree with **annotators**

Experts agree with **model**

Experts have no consensus

Emotions
- 39% (11)
- 57% (16)

Betrayal of Gender Norms
- 41% (13)
- 44% (14)
- 15% (5)

Hyper-sexualization
- 80% (24)
- 10%
- 10%

Bad Mother
- 69% (18)
- 12%
- 19% (5)

16
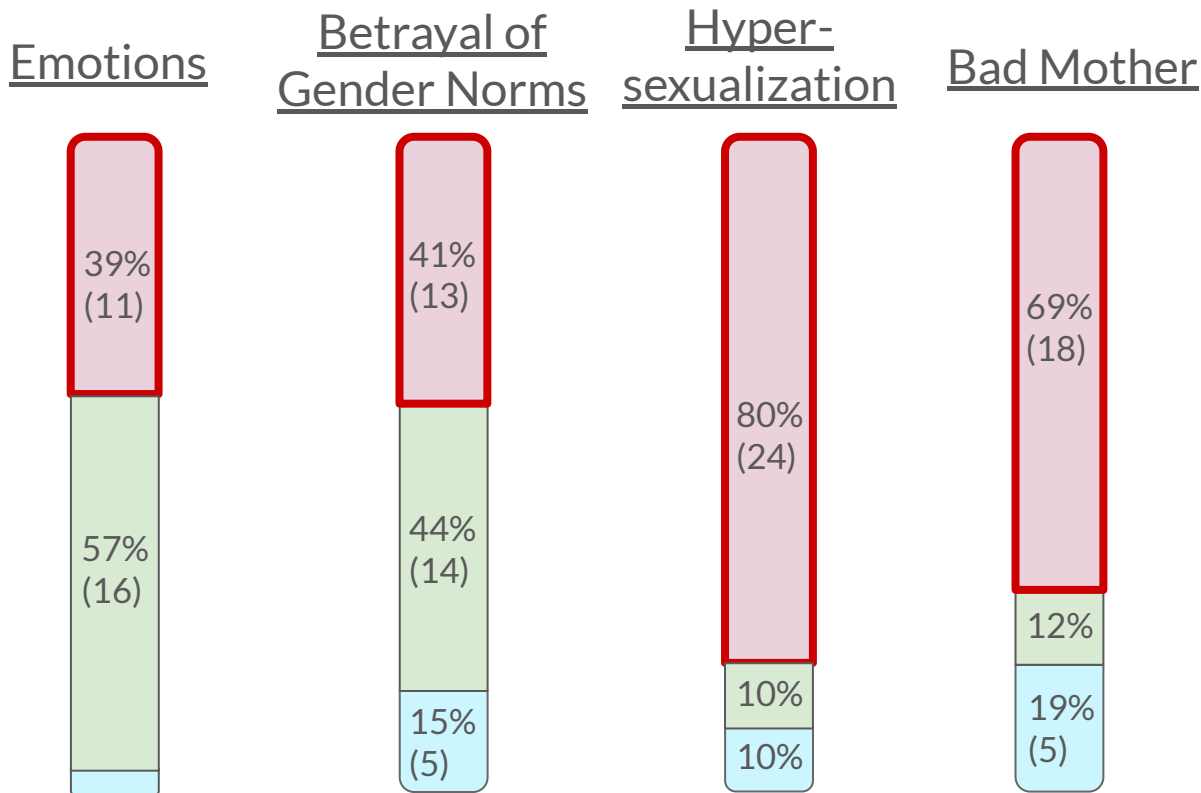
1a. The model can misidentify whether it is the defendant being discussed.
[Emotions]

" *[The defendant] said that [the victim] looked really mad, really upset.* "

1b. The model incorrectly picks up passages describing relationships but not **Hypersexualization**.

" *[The defendant] had a job for a period of time as a greeter at [a supermarket]. This was the end of May, into the beginning of July. So about a little more than a month.* *So that is where [the defendant] met [her second husband]. She took up with him, and they were married in December.* "

1c. The model incorrectly identifies passages describing defendants' experiences of violence from their parents for **Bad Mother**.

> " *You will hear that these parents battered each other and their children; that they neglected them; they abandoned them. [...] You will hear that* [the defendant] *from the beginning* **acted as buffer between her parents and her siblings,** *accepting blame and punishment,* **accepting the screaming, the beatings and the threats, to protect her younger siblings from their parents.** "

Problem 2: This annotation is hard for legal experts as well.

	Emotions	Betrayal of Gender Norms	Hyper-sexualization	Bad Mother
Experts agree with **annotators**	39% (11)	41% (13)	80% (24)	69% (18)
Experts agree with **model**	57% (16)	44% (14)	10%	12%
Experts have no consensus		15% (5)	10%	19% (5)
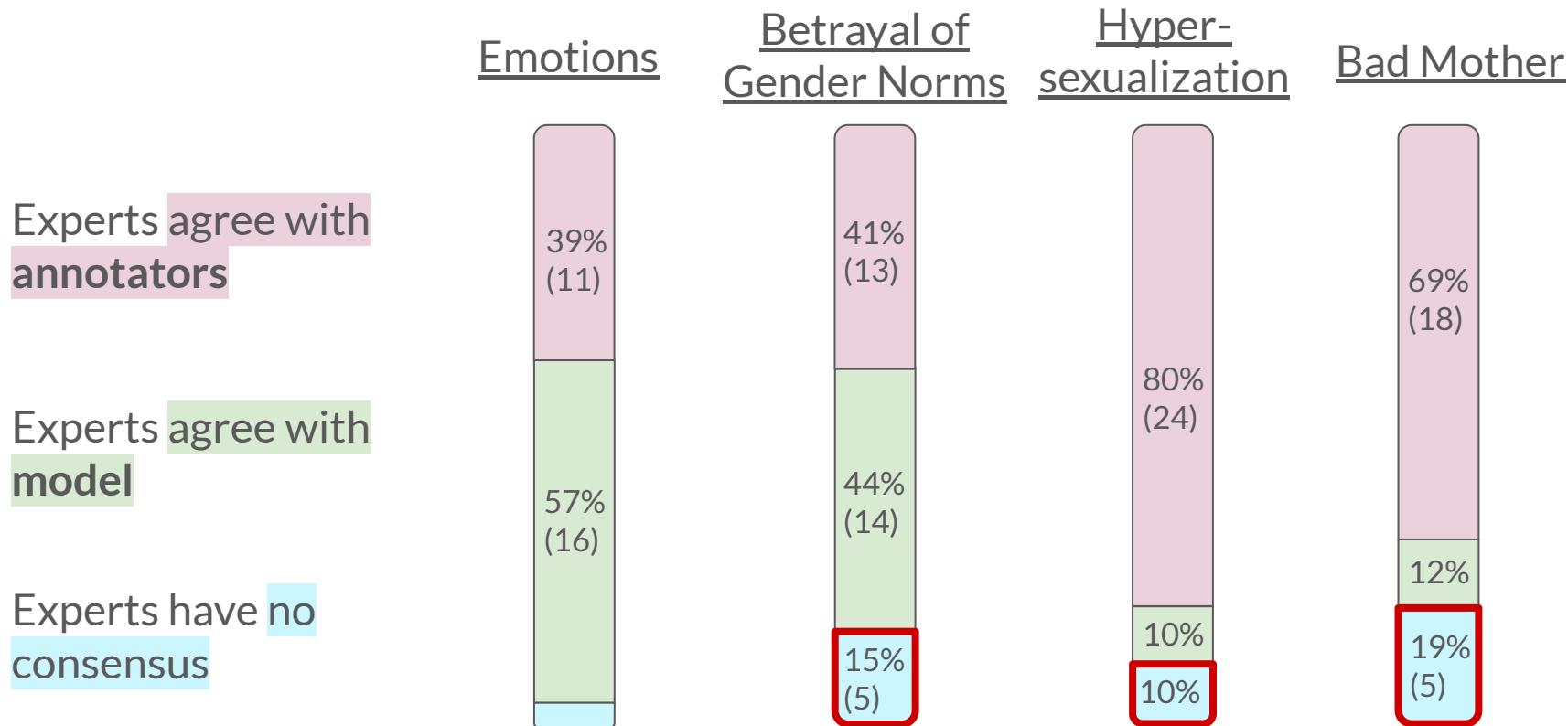
2a. The legal experts need to tell apart factual statements from value judgements. [Bad Mother]

Fact

[The defendant] had been trying to get [a man] to kill her daughter for at least months. It got to the point that he went to this child's father and said [the defendant] won't leave me alone about killing [her daughter].

Value

[The defendant] became part of the culture, drugs, thugs, and violence.
You heard about [the defendant's] children,** her son in prison for years and one of her daughters is in prison for five years. And her other daughter is with her grandmother. And **we can only hope that there may be some hope for that daughter.

2b. The legal experts need more contexts to gauge why a certain event is being discussed. [Betrayal of Gender Norms]

> " *Q. Now, was there anything characteristic about [the defendant] that would catch your attention?*
> *A. Whenever she would drive her car, she would always have her window down and she would always be smoking her cigarette. And just the mannerism that she would hold her cigarette to the side.*
> *Q. How did she do that?*
> *A. Just by flicking her cigarette.*
> *Q. Any particular style that you refer to it as?*
> *A. Well, as I said before, the kids told me it's like a Hollywood style.* "

2c. The legal experts need background knowledge to understand who is being discussed. [Hypersexualization]

> *Q: How long were you having sex with your brother's wife?*
> *A: From about five years.*
> *Q: So you had sex with your brother's wife for five years; how many times?*
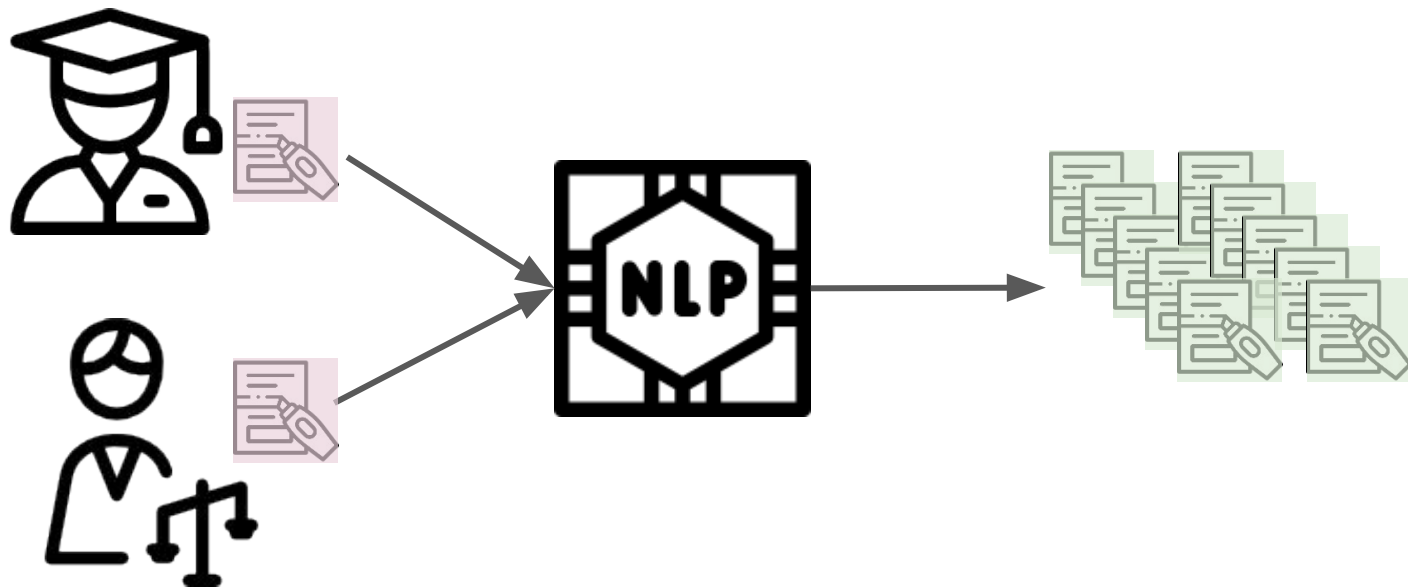> *A: Not so much.*
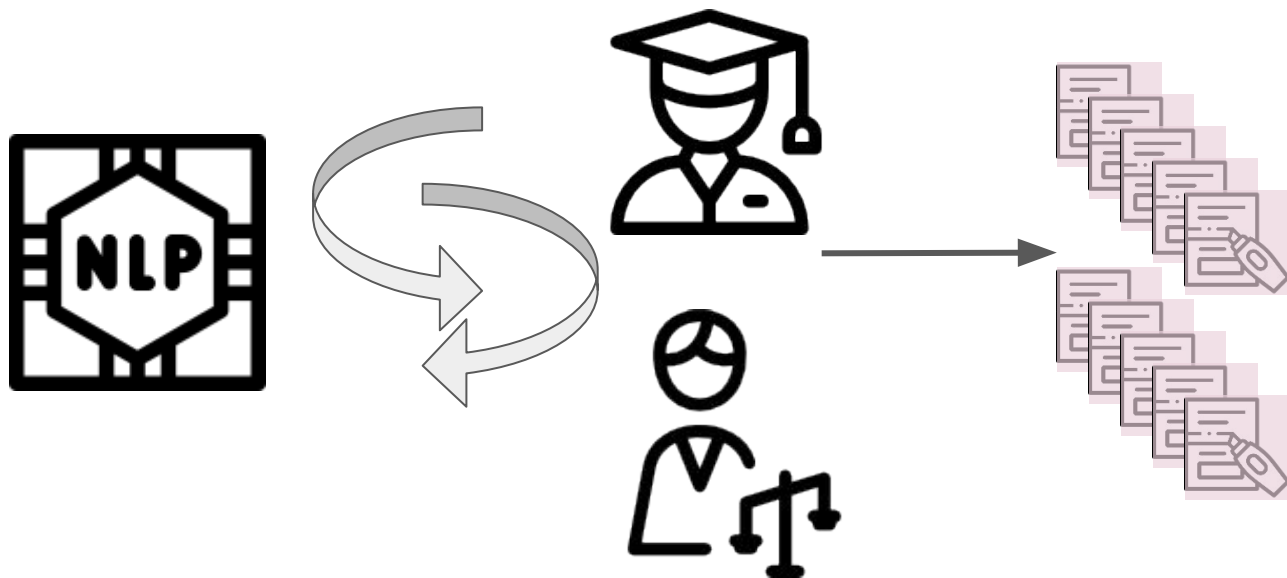> *Q: What does, "not so much" mean; 2, 20, 100?*

It is unrealistic to remove legal experts from the annotation process.

Phase 1:

Qualitative Annotation

Phase 2:

Computational Modeling

Phase 3:

Legal experts discuss disagreements
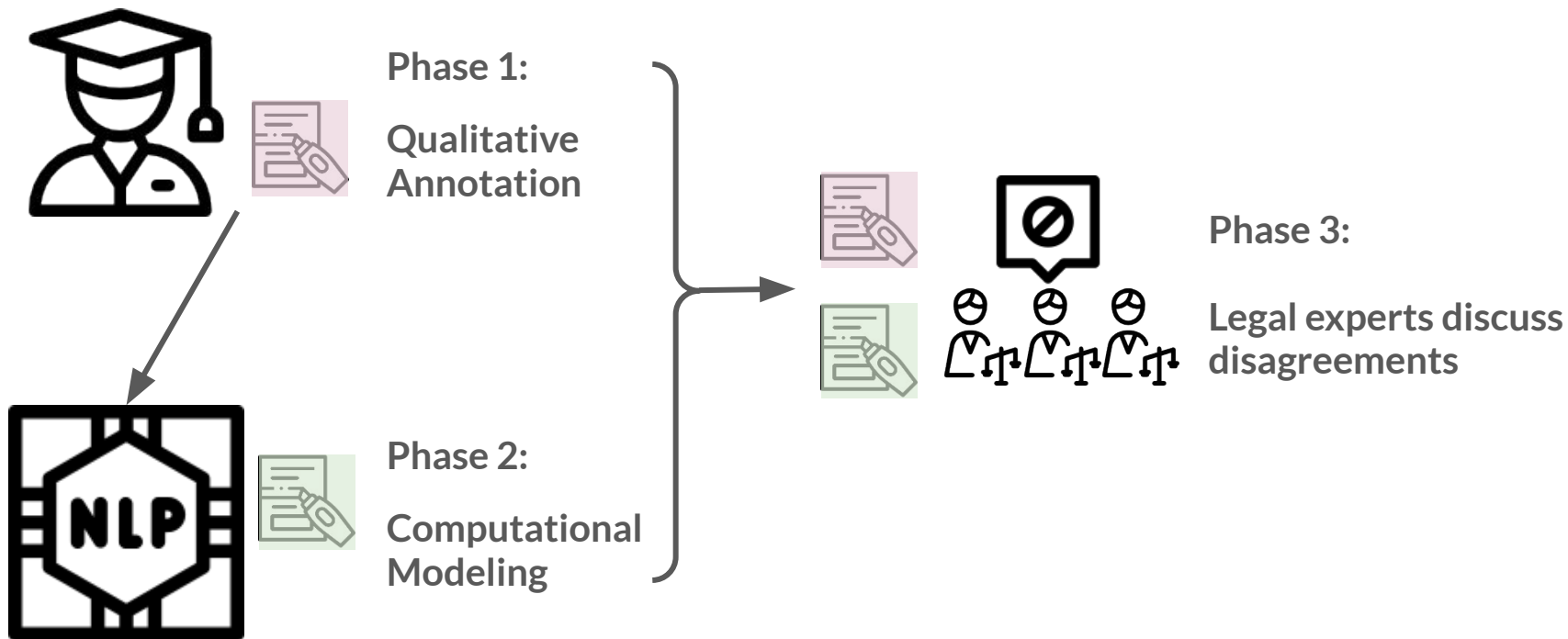
We are *not* replacing experts with a model...

Legal experts desire new perspectives that could be offered by language models

Legal experts desire new perspectives that could be offered by language models

"The model adds another layer of perspectives. I see [the disagreement between the model and annotator] as a benefit — it helps remind me that I may have my own biases.
Reading these tricky passages helps me identify my blind spots."

In assisting with annotating complex and nuanced concepts, the value of the computational model lies in the imperfect predictions it makes.

Showing "similar, but not quite" texts prompt legal experts to **reflect, refine, concretize, and build consensus** for the annotations.