

Ethical Considerations When Utilizing Foreign Language Text for Natural Language Processing

Ethics of Big Data and Artificial Intelligence Conference (Nov. 3, 2022)
Robert Brathwaite, Ph.D. (Michigan State University)

Dictionary: Actions of Interest (All Data)

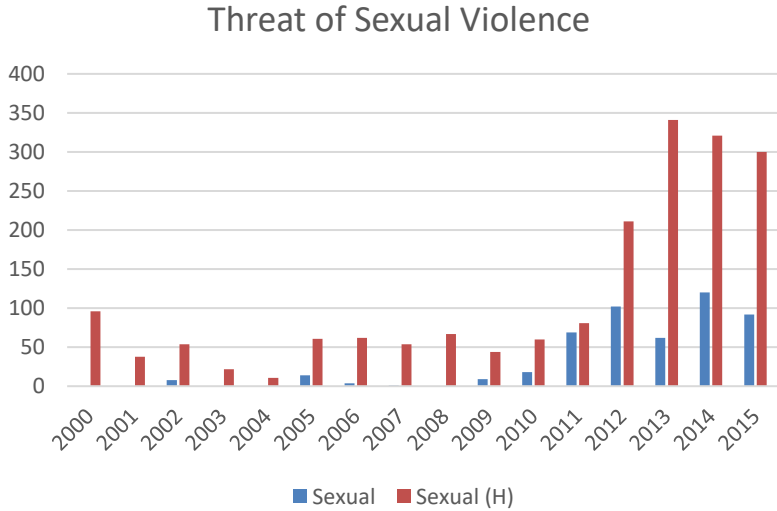
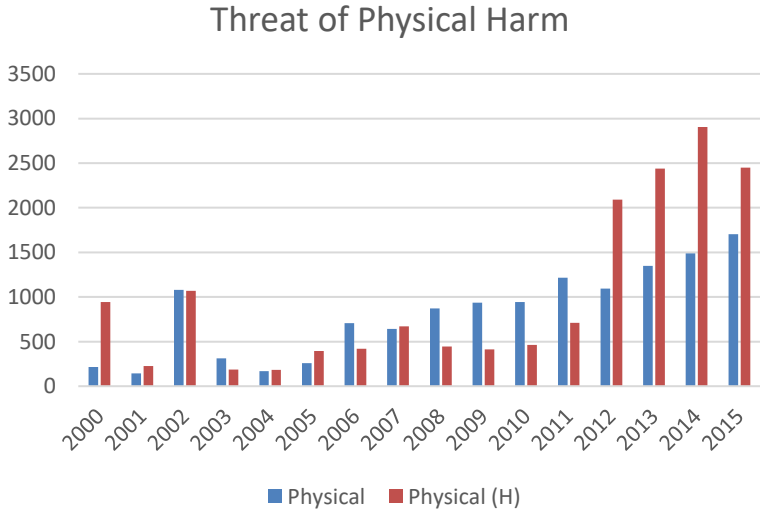
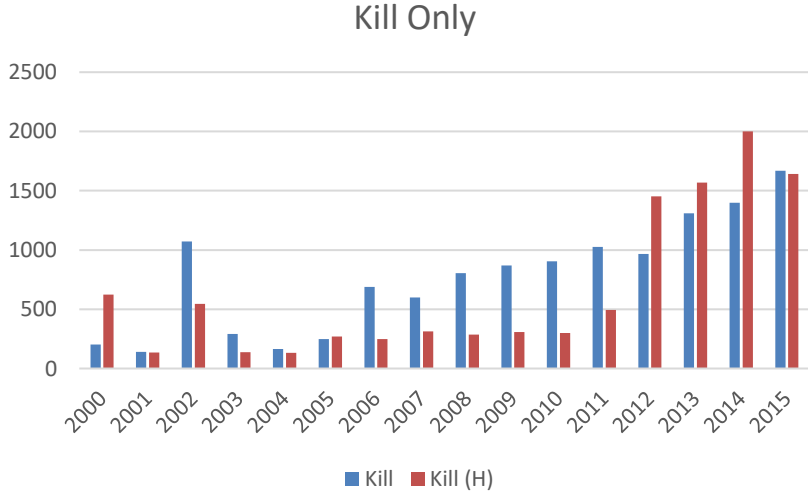
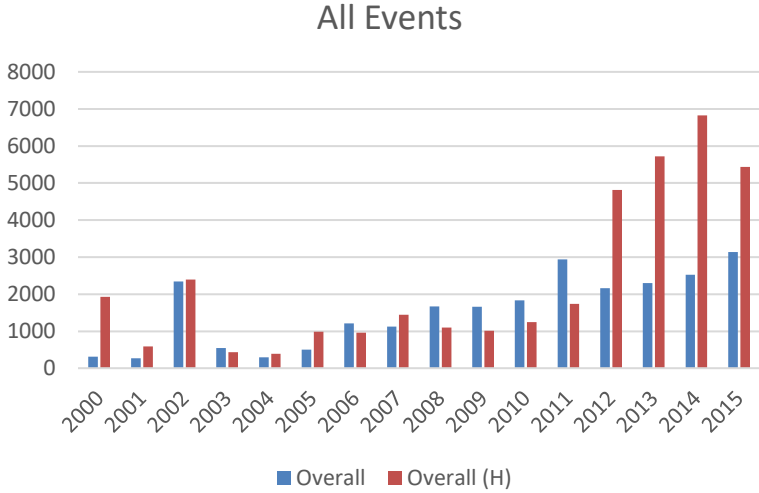
Threat to Physical Harm	Threats of Sexual Violence	Threats to Personal Property	Threat to Personal Dignity
kill	rape	arson	discriminate
murder	sexual violence	burn	strip
torture	sexual assault	fire	abuse
assault		destroy	slavery
beat		destroy property	atrocitiy
disappear		destruct	humilate
massacre		destruct property	
		rob	
		dacoity	
		banditiy	

Religious Violence Event-Data Example

“The law and order situation in *Manipur* had deteriorated due to the *killing* of a *Muslim* activist by *Assam Rifles* on *11th July, 2014*.”

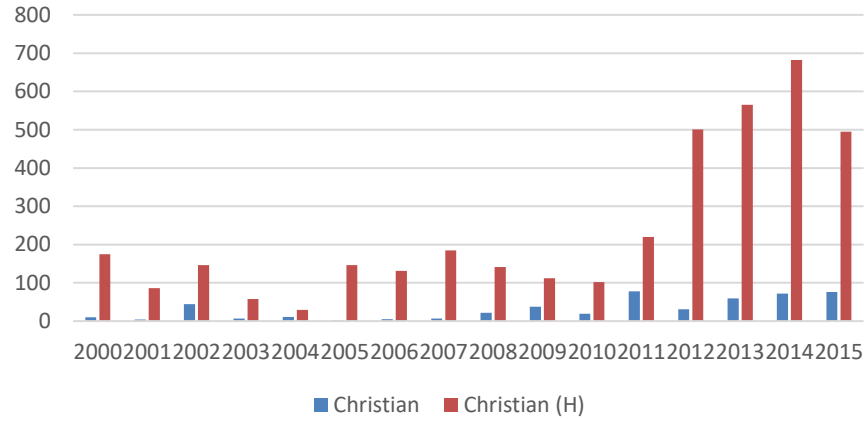
where *what* *whom*
who *when*

Event Count Comparison (India)

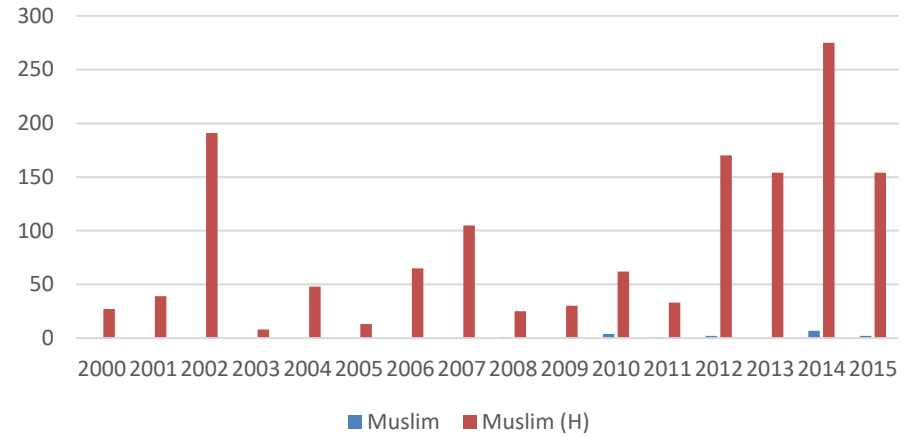


Victim Count Comparison (India)

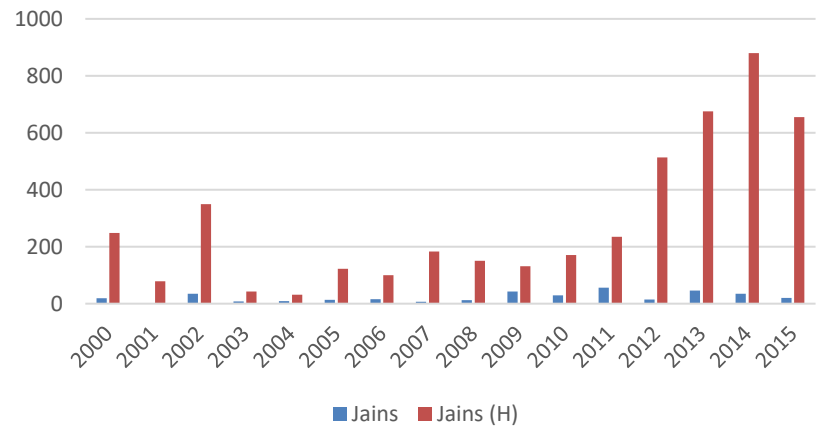
Christians



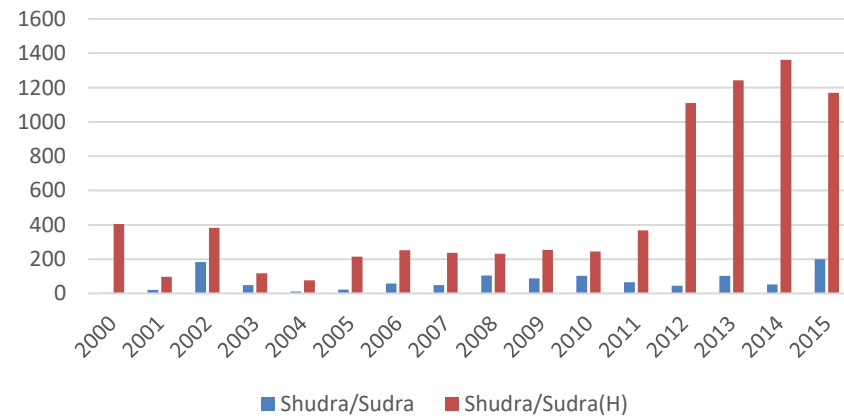
Muslim



Jains



Shudra/Sudra



Implications

- **Bias**
 - Media Sources
 - Language Use

- **Cultural Knowledge**
 - Research Results
 - Vocabulary Use

Thank You