

# NLP Fairness for India

**Data Card Authors:** Shaily Bhatt, Sunipa Dev, Partha Talukdar, Shachi Dave, Vinodkumar Prabhakaran

This dataset is being released for reproducing the results of the paper “[Re-contextualizing Fairness in NLP: The Case of India](#)”. The aim is to demonstrate existent bias in corpora and models along the dimensions of region, and religion, in the Indian context. The dataset contains tuples of the form (identity term, attribute) (for eg: (gujarati, entrepreneur)). These tuples are then annotated by human-raters for whether the attribute is commonly associated with the identity term as a stereotype. The tuples were created with a combination of dictionary driven (relying on previous literature for list of characteristics and identity terms) and corpora driven (filtering based on occurrence in [IndicCorp](#)-en) approaches. Along with the tuples we also release the list of identity terms and proxy identity terms (first or last names with prototypical gender and region associations) as obtained from Wikipedia and list of templates used to perform the analysis of NLP models in the paper.

## Data Card

DATASET TEAM(S)	DATASET CONTACT	DATASET AUTHORS																								
Google Research India - NLU team Responsible AI and Human Centered Technology (RAI-HCT)	<ul style="list-style-type: none"><li>Shaily Bhatt: <a href="mailto:shailybhatt@google.com">shailybhatt@google.com</a></li><li>Sunipa Dev: <a href="mailto:sunipadev@google.com">sunipadev@google.com</a></li><li>Partha Talukdar: <a href="mailto:partha@google.com">partha@google.com</a></li><li>Shachi Dave: <a href="mailto:shachi@google.com">shachi@google.com</a></li><li>Vinodkumar Prabhakaran: <a href="mailto:vinodkpg@google.com">vinodkpg@google.com</a></li></ul>	<ul style="list-style-type: none"><li>Shaily Bhatt, Pre-Doctoral Researcher, Google, 2022</li><li>Sunipa Dev, Research Scientist, Google 2022</li><li>Partha Talukdar, Research Scientist, Google 2022</li><li>Shachi Dave, Software Engineer, Google 2022</li><li>Vinodkumar Prabhakaran, Research Scientist, Google 2022</li></ul>																								
PRIMARY DATA MODALITY	DATASET SNAPSHOT	DESCRIPTION OF CONTENT																								
Image Data Text Data <b>Tabular Data</b> Audio Data Video Data Time Series Graph Data Geospatial Data Multimodal (Please specify) Others (please specify) Unknown	<table><tr><td>Size of dataset</td><td>208 KB</td></tr><tr><td>Number of Instances</td><td>3852</td></tr><tr><td>Number of Fields</td><td>6</td></tr><tr><td><b>Field 1.</b> Identity term</td><td>Identity term for the tuple</td></tr><tr><td><b>Field 2.</b> Token</td><td>Attribute token of the tuple</td></tr><tr><td><b>Field 3.</b> Stereotypical</td><td>Number of annotators that labeled the attribute token to be considered stereotypically associated with the identity term in the society.</td></tr><tr><td><b>Field 4.</b> Non Stereotypical</td><td>Number of annotators that labeled the attribute token to not be considered stereotypically associated with the identity term in the society.</td></tr><tr><td><b>Field 5.</b> Not sure</td><td>Number of annotators unaware or unsure of any such association between the identity term and token</td></tr><tr><td><b>Field 6.</b> Total</td><td>Total number of annotations for the tuple</td></tr></table>	Size of dataset	208 KB	Number of Instances	3852	Number of Fields	6	<b>Field 1.</b> Identity term	Identity term for the tuple	<b>Field 2.</b> Token	Attribute token of the tuple	<b>Field 3.</b> Stereotypical	Number of annotators that labeled the attribute token to be considered stereotypically associated with the identity term in the society.	<b>Field 4.</b> Non Stereotypical	Number of annotators that labeled the attribute token to not be considered stereotypically associated with the identity term in the society.	<b>Field 5.</b> Not sure	Number of annotators unaware or unsure of any such association between the identity term and token	<b>Field 6.</b> Total	Total number of annotations for the tuple	<p>The dataset contains tuples of the form (identity term, attribute) (for eg: (tamil, mathematician)).</p> <p>These tuples are annotated by human-raters in two phases. The annotators were asked to label whether the attribute token is associated with the identity term as stereotypical in the society. Due to the phased nature of annotation, each tuple is expected to have annotation from either 3 or 6 raters. In this phased annotation, we aim for getting high precision in indicating whether a tuple is a commonly held stereotype. For this, first the tuple is shown to 3 raters. If <math>\geq 2</math> of the 3 label it as a stereotype, it is shown to 3 more raters for increasing confidence in the label.</p> <p>The tuples were created with a combination of dictionary driven (relying on previous literature for list of attributes and identity terms) and corpora driven (filtering based on occurrence in <a href="#">IndicCorp</a>-en) approaches.</p> <p>Along with the tuples we also release the list of identity terms and proxy identity terms (first or last names with prototypical gender and region associations) as obtained from Wikipedia and list of templates used to perform the analysis of NLP models in the paper.</p>						
Size of dataset	208 KB																									
Number of Instances	3852																									
Number of Fields	6																									
<b>Field 1.</b> Identity term	Identity term for the tuple																									
<b>Field 2.</b> Token	Attribute token of the tuple																									
<b>Field 3.</b> Stereotypical	Number of annotators that labeled the attribute token to be considered stereotypically associated with the identity term in the society.																									
<b>Field 4.</b> Non Stereotypical	Number of annotators that labeled the attribute token to not be considered stereotypically associated with the identity term in the society.																									
<b>Field 5.</b> Not sure	Number of annotators unaware or unsure of any such association between the identity term and token																									
<b>Field 6.</b> Total	Total number of annotations for the tuple																									
DATASET SUBJECT	EXAMPLE: DATA POINT	DATA FIELDS																								
Sensitive Data about people Non-Sensitive Data about people Data about natural phenomena Data about places and objects Synthetically generated data Data about systems or products and their behaviors Unknown <b>Others*</b> (*Data about social phenomena)	<p>This example is an actual data point from the data. As suggested by the heading of each column, the first column is the identity term, the second column is the attribute token, and columns 3 and 4 indicate the number of annotators that found the tuple to be Stereotypical and Non stereotypical. Column 5 indicates the number of annotators unsure about any such association. Column 6 indicates total number of annotations; derivable as a sum of columns 3-5.</p> <p>E.g. of Data Point:</p> <table><tr><th colspan="6"></th></tr><tr><th>Identi ty term</th><th>Token</th><th>Stereot ytical</th><th>Non Stereot ytical</th><th>Not sure</th><th>Total</th></tr><tr><td>Tamil</td><td>Mathema tician</td><td>6</td><td>0</td><td>0</td><td>0</td></tr><tr><th colspan="6"></th></tr></table>							Identi ty term	Token	Stereot ytical	Non Stereot ytical	Not sure	Total	Tamil	Mathema tician	6	0	0	0							<ul style="list-style-type: none"><li><b>Field 1.</b> Identity term<ul style="list-style-type: none"><li>Identity term for the tuple in consideration</li></ul></li><li><b>Field 2.</b> Token<ul style="list-style-type: none"><li>Attribute token for the tuple under consideration</li></ul></li><li><b>Field 3.</b> Stereotypical<ul style="list-style-type: none"><li>Number of annotators that labeled the attribute token to be considered stereotypically associated with the identity term in the society.</li></ul></li><li><b>Field 4.</b> Non Stereotypical<ul style="list-style-type: none"><li>Number of annotators that labeled the attribute token to not be considered stereotypically associated with the identity term in the society.</li></ul></li><li><b>Field 5.</b> Not sure</li></ul>
Identi ty term	Token	Stereot ytical	Non Stereot ytical	Not sure	Total																					
Tamil	Mathema tician	6	0	0	0																					

	<div></div>	<ul style="list-style-type: none"><li>○ Number of annotators unaware or unsure of any such association between the identity term and token</li></ul> <ul style="list-style-type: none"><li>● <b>Field 6.</b> Total<ul style="list-style-type: none"><li>○ Total number of annotations for the tuple</li></ul></li></ul>
DATASET PURPOSE(S)	KEY DOMAINS OR APPLICATION(S)	PRIMARY MOTIVATION(S)
Monitoring <b>Research</b> Production Others (please specify)	Domains Natural Language Processing, Algorithmic Fairness  Problem Space Bias demonstration in NLP models and data	This dataset is aimed to be used to demonstrate the prevalence of Stereotypes in corpora and NLP models along the dimensions of region, and religion, specifically in the Indian context. The primary motivation of curating this data, is that no previous dataset for this purpose exists.
DATASET USAGE	INTENDED AND/OR SUITABLE USE CASE(S)	UNSUITABLE USE CASE(S)
Safe for production use <b>Safe for research use</b> Conditional use- some unsafe applications Only approved use Others (please specify)	<ul style="list-style-type: none"><li>● To demonstrate existence of bias i.e prevalence of stereotypes or fairness issues in NLP models and data</li></ul>	<ol style="list-style-type: none"><li>1. As a benchmark for assessing fairness or ensuring lack of fairness</li><li>2. As a sole resource for any bias mitigation in production systems</li><li>3. To train demographic predictors using lists of proxy identity terms obtained from wikipedia with their prototypical associations</li></ol>
SAFETY OF USE WITH OTHER DATA	ACCEPTABLE TRANSFORMATIONS	BEST PRACTICES FOR JOINING OR AGGREGATING WITH DATASET
<b>Safe to use with other data</b> Conditionally safe to use with other data Should not be used with other data Unknown Others* (Please specify)	<b>Joining with other datasets</b> <b>Subsampling and splitting</b> <b>Filtering</b> <b>Joining input sources</b> <b>Cleaning missing values</b> <b>Anomaly detection</b> <b>Grouping and summarizing</b> <b>Scaling and reducing</b> <b>Statistical transformations</b> <b>Redaction or Anonymization</b> Others (please specify)	N/A (we have not attempted to use this dataset with other datasets, but we do not anticipate any issues)
VERSION STATUS	DATASET VERSION	MAINTENANCE PLAN
Regularly Updated New versions of the dataset have been or will continue to be made available.  <b>Actively Maintained</b> No new versions will be made available, but this dataset will be actively maintained, including but not limited to updates to the data.  Limited Maintenance The data will not be updated, but any technical issues will be addressed.  Deprecated This dataset is obsolete or is no longer being maintained.	<b>Current Version</b> 1.0 <b>Last Updated</b> 05/2022 <b>Release Date</b> 10/2022	<ul style="list-style-type: none"><li>● We might add data translated to other languages when such resources become available</li><li>● We will address any issues that people might face in the dataset usage</li></ul>
ACCESS POLICY		
The data will be accessible under the Apache License 2.0		
DATA COLLECTION METHODS	DATA SOURCES	DATA COLLECTION

API Artificially Generated <b>Crowdsourced - Paid</b> Crowdsourced - Volunteer Vendor Collection Efforts Scraped or Crawled Survey, forms or polls <b>Taken from other existing datasets</b> Unknown To be determined Others (please specify)	Tuples for annotation: Taken from existing datasets <b>Date of Collection:</b> Nov 2021 - Mar 2022 <b>Data Modality:</b> Text Data <b>Process:</b> <ul style="list-style-type: none"><li>Attribute tokens were obtained from previous literature and datasets, such as papers including: <a href="#">Malik et al., 2021</a>, <a href="#">Nangia et al., 2020</a>, <a href="#">Nadeem et al., 2020</a>.</li><li>Identity terms were obtained from Wikipedia along three axes, namely, gender, region, and religion.</li><li>Tuples were formed by a cross-product of the attribute tokens and identity terms.</li><li>Tuples that were better suited candidates for annotation were filtered by looking up the co-occurrence counts (number of times the identity term and token co-occur in a sentence) in the <a href="#">IndicCorp</a> English corpus. Only tuples that frequently co-occurred were retained and sent for annotation.</li></ul> Annotations: Crowdsourced - Paid <b>Crowd Data Platform:</b> Crowd Data Platform facilitates the collection of large-scale human-generated or human-augmented datasets used by teams in Google and Alphabet working on machine learning (ML) or other data-intensive products and services. <b>Date of Collection:</b> Nov 2021 - May 2022 <b>Data Modality:</b> Text Data	Tuples for annotation: Taken from existing datasets <b>Collected and included</b> <ul style="list-style-type: none"><li>Identity_term: Identity term of the tuple in question</li><li>Token: Attribute token of the tuple</li></ul> Annotations: Crowdsourced - Paid <b>Collected and included</b> <ul style="list-style-type: none"><li>Stereotypical: Number of annotators that labeled the attribute token to be considered stereotypically associated with the identity term in the society.</li><li>Non Stereotypical: Number of annotators that labeled the attribute token to not be considered stereotypically associated with the identity term in the society.</li><li>Not sure: Number of annotators unaware or unsure of any such association between the identity term and token</li><li>Total: Total number of annotations for the tuple</li></ul> <b>Collected and excluded</b> <ul style="list-style-type: none"><li>none</li></ul>
<b>INCLUSION CRITERIA</b>	<b>EXCLUSION CRITERIA</b>	<b>DATA PROCESSING</b>
Tuples for annotation: Taken from existing datasets <ul style="list-style-type: none"><li>Attribute tokens were obtained from previous literature and datasets, such as papers including: <a href="#">Malik et al., 2021</a>, <a href="#">Nangia et al., 2020</a>, <a href="#">Nadeem et al., 2020</a>.</li><li>Identity terms were obtained from Wikipedia along two axes, namely, region, and religion.</li><li>Tuples were formed by a cross-product of the attribute tokens and identity terms.</li></ul>	Tuples for annotation: Taken from existing datasets <ul style="list-style-type: none"><li>Only tuples that did not frequently co-occur or did not co-occur at all in IndicCorp english corpus were excluded.</li><li>Additionally, tuples with tokens that appear with all identity terms of a given axis are also removed.</li></ul>	Tuples were formed by a cross-product of the attribute tokens and identity terms. Identity terms were obtained from Wikipedia and attribute tokens were obtained from previous literature and datasets, such as papers including: <a href="#">Malik et al., 2021</a> , <a href="#">Nangia et al., 2020</a> , <a href="#">Nadeem et al., 2020</a> .
<b>SENSITIVE DATA</b>	<b>FIELDS WITH SENSITIVE DATA</b>	<b>SECURITY AND PRIVACY HANDLING</b>
User Content User Metadata User Activity Data Identifiable Data S/PII Business Data Employee Data Pseudonymous Data Anonymous Data Health Data Children’s Data <b>None</b> Others* (*please specify)	NA	NA
<b>SENSITIVE HUMAN ATTRIBUTES</b>	<b>SOURCE(S) OF HUMAN ATTRIBUTES</b>	<b>RATIONALE FOR COLLECTING HUMAN ATTRIBUTES</b>

Race	<b>[Human Attribute]:</b> Source <b>[Language]:</b> The presence of co-occurrence of tuples was looked up in the IndicCorp english corpus. This corpus is a collection of data obtained from news, magazines, and other internet sources. This inherently reflects language use. <b>[Culture]:</b> Annotators were asked to label whether the attribute token of the tuple is commonly believed to be Stereotypically associated to the identity term of the tuple. This annotation inherently and intentionally captures the view of the society or the culture.	We intend to study whether commonly held Stereotypes in the society are prevalent in NLP corpora and data. For doing this, it is essential to first build such a list of Stereotypes that are commonly held by the society. This inherently represents the cultural view-point of the society, for our case specifically, the Indian society. Thus, the capturing of culture and language human-attributes was intentional and driven by the need of such computational resources that can drive research around bias and fairness issues in NLP models to ensure that the same stereotypes are not perpetuated through technology.																				
Gender																						
Ethnicity																						
Socio-economic status																						
Geography																						
Language																						
Sexual Orientation																						
Religion																						
Age																						
Culture																						
Disability																						
Experience or Seniority																						
Others (please specify)																						
	DISTRIBUTION OF HUMAN ATTRIBUTES																					
	<table><tr><th>Language</th><th>Number of tuples</th></tr><tr><td>Count</td><td>3852</td></tr></table> <p>Language distribution of the InFormal dataset.</p> <table><tr><th>Culture (Stereotypes)</th><th>Marked Stereotypical by 0 annotators</th><th>Marked Stereotypical by 1 annotators</th><th>Marked Stereotypical by 2 annotators</th><th>Marked Stereotypical by 3 annotators</th><th>Marked Stereotypical by 4 annotators</th><th>Marked Stereotypical by 5 annotators</th><th>Marked Stereotypical by 6 annotators</th></tr><tr><td>Count</td><td>2775</td><td>762</td><td>248</td><td>49</td><td>6</td><td>9</td><td>3</td></tr></table>		Language	Number of tuples	Count	3852	Culture (Stereotypes)	Marked Stereotypical by 0 annotators	Marked Stereotypical by 1 annotators	Marked Stereotypical by 2 annotators	Marked Stereotypical by 3 annotators	Marked Stereotypical by 4 annotators	Marked Stereotypical by 5 annotators	Marked Stereotypical by 6 annotators	Count	2775	762	248	49	6	9	3
	Language	Number of tuples																				
	Count	3852																				
	Culture (Stereotypes)	Marked Stereotypical by 0 annotators	Marked Stereotypical by 1 annotators	Marked Stereotypical by 2 annotators	Marked Stereotypical by 3 annotators	Marked Stereotypical by 4 annotators	Marked Stereotypical by 5 annotators	Marked Stereotypical by 6 annotators														
	Count	2775	762	248	49	6	9	3														
TRANSFORMATIONS APPLIED	FIELDS TRANSFORMED	LIBRARIES AND METHODS USED																				
Anomaly Detection	Cross Product of tokens and identity terms <ul style="list-style-type: none"><li>Tuple → token X identity terms<ul style="list-style-type: none"><li>Tuples created by pairing each identity term with each attribute token to create a tuples</li></ul></li></ul> Tuple Filtering <ul style="list-style-type: none"><li>Longer list of tuples → shorter list of tuples<ul style="list-style-type: none"><li>Tuples that do not frequently co-occur in the IndicCorp-en corpus are filtered out</li></ul></li></ul> Annotation Aggregation <ul style="list-style-type: none"><li>Annotation of individual annotators → Number of annotators that marked a Tuple Stereotype<ul style="list-style-type: none"><li>Tuples are bucketed into categories of how many annotators marked the tuple as Stereotype</li></ul></li></ul>	<ul style="list-style-type: none"><li>Cross product: python basic functions</li><li>Tuple filtering: python basic functions, NLTK for tokenization</li><li>Annotation aggregation: python basic functions</li></ul>																				
Cleaning Mismatched Values																						
Cleaning Missing Values																						
Converting Data Types																						
Data Aggregation																						
Dimensionality Reduction																						
Joining Input Sources																						
Redaction or Anonymization																						
Others*																						
(*Cross-product of tokens and identity terms, tuple filtering, annotation aggregation)																						
SAMPLING METHOD(S)	SAMPLING CHARACTERISTIC(S)	<ul style="list-style-type: none"><li>SAMPLING CRITERIA</li></ul>																				
Cluster Sampling	Frequency-based sampling <ul style="list-style-type: none"><li>Tuples are selected based on the frequency of co-occurrence in the <a href="#">IndicCorp</a> english corpus</li></ul>	Frequency based sampling <ul style="list-style-type: none"><li>Tuples where the identity term and attribute token never co-occur or occur with very low frequency in the <a href="#">IndicCorp</a>-english corpus are filtered out.</li><li>Tuples where the attribute token occurs with every identity term of that axis are also filtered out.</li></ul>																				
Haphazard Sampling																						
Multi-stage Sampling																						
Random Sampling																						
Retrospective Sampling																						
Stratified Sampling																						
Systematic Sampling																						
Weighted Sampling																						
Unknown																						
Unsampled																						
Others*																						
(*Frequency-based sampling)																						
ANNOTATION WORKFORCE TYPE	ANNOTATION CHARACTERISTICS	ANNOTATION DESCRIPTION																				



Annotation Target in Data Machine-generated Annotations Human Annotations - Expert <b>Human Annotations - Non-expert</b> Human Annotations - Employees Human Annotations - Contractors Human Annotations - Crowdsourcing Human Annotations - Outsourced / Managed Teams Unlabeled Others* (*Please specify)	Stereotype annotation Total number of annotations 12486 Avg. Annotations per example 3.24 Number of annotators per example 3 or 6	Stereotype annotation <ul style="list-style-type: none"> <li>Annotation was done in two phases.</li> <li>In the first phase the tuple is shown to 3 annotators for labeling whether it is a commonly held stereotype in the society.</li> <li>If 2 or more annotators label it as a Stereotype, this tuple is sent to a second phase of annotation where it is shown to 3 more annotators.</li> <li>The annotations are finally aggregated to the number of annotators that label the tuple Stereotype, Non Stereotype, or unsure/Not sure and Total annotations.</li> </ul>
	<b>ANNOTATOR BREAKDOWN</b>	<b>ANNOTATOR DESCRIPTION</b>
	Stereotype annotation Annotator type Paid - Non-expert Total unique annotators 6 Average cost/tuple 0.17 USD Expertise of annotators Trained for task	Stereotype annotation <ul style="list-style-type: none"> <li>We recruited 6 annotators with diverse gender and region identities(self-identification): 3 male, 3 female, 2 each from Northeast and Central India, and 1 each from West and South India.</li> <li>Multiple training sessions (virtual) were held where the task was described along with 3-4 examples.</li> <li>To test their understanding of the task, we conducted a pilot annotation where each annotation required a justification.</li> <li>The authors reviewed these justifications, and any misconceptions were clarified.</li> </ul>
<b>VALIDATION METHOD(S)</b>	<b>VALIDATION BREAKDOWN</b>	<b>DESCRIPTION OF VALIDATION</b>
<b>Data Type Validation</b> <b>Range and Constraint Validation</b> Code/cross-reference Validation Structured Validation Consistency Validation Not Validated Others* (*Please specify)	Data Type Validation Number of data points validated 3852 Fields Validated: All Range and Constraint Validation Number of data points validated 3852 Fields Validated: Stereotypical, Non Stereotypical, Not sure, Total	Data Type Validation The token and identity term columns are checked to be strings of text. The Stereotypical, Non Stereotypical, Not sure, Total columns are checked to be integers. This was checked using and corrected (if needed) using basic python functions. Range and Constraint Validation Stereotypical, Non Stereotypical, Not sure columns were checked to be integers within the range of 0 to 6. The total column was checked for being either 3 or 6. The total column was checked to be the sum of the Stereotypical, Non Stereotypical, Not sure columns. The validation was done in excel and using basic python functions.
	<b>VALIDATORS CHARACTERISTIC(S)</b>	<b>VALIDATORS DESCRIPTION(S)</b>
	N/A (automatic validation)	N/A (automatic validation)
<b>ML APPLICATION(S)</b>		
N/A The dataset was not used for any applications. No training or fine-tuning of systems was performed. The data was only used for diagnostic analysis of existing models and not used to create any new systems		

Terms of Art

Concepts and Definitions referenced in this Data Card

Axes of disparities (or axes for short)	Identity terms	Attribute Tokens (or tokens for short)
Definition: These are dimensions along which disparities or bias exist in Indian society.  Here, we refer to 2 axes: region and religion	Definition: These are words used to describe a group of people with a common trait or identity. In the context of this data we identify identity terms along region and religion. For region, these are the terms used to refer to the people belonging to the different States and Union Territories of India; For religion, these are terms used to refer to people following the	Definition: These are characteristics or attributes for which we aim to identify bias. These span categories like profession, adjectives, socio-economic status, subjects of study and so on.  For eg: doctor, teacher (profession), poor, powerful (socio-economic status), smart, handsome, ugly

	<p>six most populous religions in India.</p> <p>For eg: Gujaratis is a term used to describe the people of Gujarat. Hindus is a term used to describe people who follow the Hindu religion, and so on</p>	<p>(adjectives), computer science, mathematics (subjects of study) and so on.</p>
<b>Tuple</b>	<b>Stereotype/Stereotypical</b>	<b>Bias</b>
<p>Definition: A combination of one identity term and one attribute token.</p> <p>For eg: (Hindu, Priest); (Punjabi, Dance) etc.</p>	<p>Definition: In social psychology, a stereotype is a generalized belief about a particular category of people. It is an expectation that people might have about every person of a particular group.</p> <p>Source: <a href="#">Wikipedia</a></p>	<p>Definition: Preference or prevalence of certain associations (mostly harmful) appearing more frequently in a model or data.</p>

<b>Reflections on Data</b>	
<b>Limitations due to human annotation</b>	The current dataset is annotated by 6 annotators who were recruited so as to have diversity in (binary) gender and regional identities. However, we recognize that this may still not necessarily capture the wide variety of perspectives that exist on this incredibly subjective task. Future iterations of such data collection should take more participatory approaches and involve communities with lived experiences on the harms of bias in society.
<b>No ground truth on “Stereotype”</b>	We recognize that there is no “ground-truth” on labeling something as a “Stereotype”. This is an inherently subjective opinion, influenced by socio-cultural factors and personal experiences. Thus, we caution against using the data in this dataset to in any way classify tuples as “Stereotypical” vs “Non-stereotypical”.
<b>Stereotypes not captured by this dataset</b>	Our approach towards filtering the set of tuples for annotation based on co-occurrence limits our data to only capture those stereotypes that are often explicitly mentioned in text. For instance, there might exist stereotypes in society that are not captured in data and hence will not be captured by our dataset.
<b>Caution against calling models “fair” based on evaluation on this dataset</b>	This dataset is insufficient to capture all the undesirable biases in Indian society. Our dataset covers only 3 axes of disparities, and reflects the judgments of a small number of annotators. Hence, they should be used only for diagnostic and research purposes, and not as benchmarks to prove lack of bias.