| SeeGULL-T2I Visual Attributes **Data Card Authors:** Akshita Jha, Shachi Dave, Rida Qadri, Sarah Laszlo, Remi Denton, Vinodkumar Prabhakaran, Sunipa Dev | This dataset was created as part of the SeeGULL-T2I project. It consists of attributes along with human annotations about their visual nature. The dataset contains visual ratings for 1994 attributes present in SeeGULL. |
| --- | --- |
|  | |

| Data Card | | |
| --- | --- | --- |
| DATASET TEAM(S) | DATASET CONTACT | DATASET AUTHORS |
| Technology, AI, Society, and Culture (TASC) team, RAI-HCT  Google Research India - NLU team | * Sunipa Dev: [sunipadev@google.com](mailto:sunipadev@google.com) * Vinodkumar Prabhakaran: [vinodkpg@google.com](mailto:vinodkpg@google.com) | * [Akshita Jha,](http://who/akshitajha) PhD Student, Virginia Tech (Work done as a Google Student Researcher) * [Shachi Dave](http://who/shachi), Software Engineer, Google * [Vinodkumar Prabhakaran](http://who/vinodkpg), Research Scientist, Google * [Sunipa Dev](http://who/sunipadev), Research Scientist, Google |
| PRIMARY DATA MODALITY | DATASET SNAPSHOT | DESCRIPTION OF CONTENT |
| Image DataTextData**Tabular Data**Audio DataVideo DataTime SeriesGraph DataGeospatial DataMultimodal (Please specify)Others (please specify)Unknown | | Size of dataset |  | | --- | --- | | Number of Instances | 1994 | | Number of Fields | 5 | | ***Field 1***. Attribute | Attribute token | | ***Field 2***. Rating(Asia) | Rating given by annotator from Asia region. | | ***Field 3***. Rating(EMEA) | Rating given by annotator from EMEA region. | | ***Field 4***. Rating(North America) | Rating given by annotator from NA region. | |  |  | | ***Field 5.*** *Mean Rating* | Average score for the visual nature of the attribute based on human annotation of attribute terms on a Likert scale from 1 to 5. | | The dataset contains attribute terms like, sombrero, dark, brown, poor, and thin.  These attributes are annotated by human-raters who were asked to label whether the attribute could be visually depicted in the image based on a Likert Scale ranging from Strongly Agree (5) to Strongly Disagree (1). |
| DATASET SUBJECT | EXAMPLE: DATA POINT | DATA FIELDS |
| Sensitive Data about peopleNon-Sensitive Data about peopleData about natural phenomenaData about places and objectsSynthetically generated dataData about systems or products and their behaviorsUnknown**Others\*** (\*Data about social phenomena) | This example is an actual data point from the data. E.g. of Data Point:   | | Attribute | sombrero | | --- | --- | | Rating (Asia) | 5 | | Rating (EMEA) | 5 | | Rating (North America) | 5 | | Mean Rating | 5 | | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | * ***Field 1***. Attribute   + Attribute token in consideration * ***Field 2***. Rating (Asia)   + Rating given by annotator from Asia region. * ***Field 3***. Rating (EMEA)   + Rating given by annotator from EMEA region. * ***Field 4***. Rating (North America)   + Rating given by annotator from North American region. * ***Field 5.*** Mean Rating   + Mean Rating across all regions |
| DATASET PURPOSE(S) | KEY DOMAINS OR APPLICATION(S) | PRIMARY MOTIVATION(S) |
| Monitoring**Research**ProductionOthers (please specify) | Domains Vision, Natural Language Processing, Algorithmic Fairness Problem Space Demonstration of societal biases in Text-to-Image models and data. | This dataset is created to be a repository of attributes and the extent to which they can be visually represented in the image. Datasets like these will be instrumental in more effectively detecting stereotype harms in Text-to-image models. |
| DATASET USAGE | INTENDED AND/OR SUITABLE USE CASE(S) | UNSUITABLE USE CASE(S) |
| Safe for production use**Safe for research use**Conditional use- some unsafe applicationsOnly approved useOthers (please specify) | * To demonstrate the existence of bias i.e prevalence of stereotypes in the images generated from Text-to-Image models. | 1. As a benchmark for assessing fairness or ensuring lack of fairness 2. As a resource for any bias mitigation in production systems 3. To train demographic predictors using lists of proxy identity terms obtained from wikipedia with their prototypical associations |
| SAFETY OF USE WITH OTHER DATA | ACCEPTABLE TRANSFORMATIONS | BEST PRACTICES FOR JOINING OR AGGREGATING WITH DATASET |
| **Safe to use with other data**Conditionally safe to use with other dataShould not be used with other dataUnknownOthers\* (Please specify) | **Joining with other datasets**  **Subsampling and splitting**  **Filtering**  **Joining input sources**  **Cleaning missing values**  **Anomaly detection**  **Grouping and summarizing**  **Scaling and reducing**  **Statistical transformations**  **Redaction or Anonymization**  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. Actively Maintained 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. | **Current Version** 1.0  **Last Updated** 11/2023  **Release Date** 01/2024 | * We might add annotations for more attributes. * We will address any issues that people might face in the dataset usage. |
| ACCESS POLICY | RETENTION POLICY | WIPEOUT POLICY |
| The data will be accessible under the Apache License 2.0 | N/A | N/A |
| DATA COLLECTION METHODS | DATA SOURCES | DATA COLLECTION |
| APIArtificially Generated**Crowdsourced - Paid**Crowdsourced - VolunteerVendor Collection EffortsScraped or CrawledSurvey, forms or polls**Taken from other existing datasets**UnknownTo be determinedOthers (please specify) | Attributes for annotation: Taken from existing resources.**Sources**:  * Attribute tokens were obtained from previous literature [Jha et al, 2023 [1]](https://aclanthology.org/2023.acl-long.548/).   [1] Akshita Jha, Aida Mostafazadeh Davani, Chandan K Reddy, Shachi Dave, Vinodkumar Prabhakaran, and Sunipa Dev. 2023. [SeeGULL: A Stereotype Benchmark with Broad Geo-Cultural Coverage Leveraging Generative Models](https://aclanthology.org/2023.acl-long.548). In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 9851–9870, Toronto, Canada. Association for Computational Linguistics. | **Timeline:** Aug 2023 - Dec 2023 **Data Modality**: Text Data Annotations: Crowdsourced - Paid**Date of Collection:** Aug 2023 - Dec 2023**Instrumentation:** Google’s proprietary crowd work platform **Data Modality**: Text Data |
| INCLUSION CRITERIA | EXCLUSION CRITERIA | DATA PROCESSING |
| Tuples for annotation: Taken from existing datasets  * Attribute tokens were obtained from previous literature: [Jha et al, 2023.](https://aclanthology.org/2023.acl-long.548/) | Tuples for annotation: Taken from existing datasets  * All unique attributes present in SeeGULL were sent out for annotation. | Noisy attributes and non-alphabetic characters were removed from the data. |
| SENSITIVE DATA | FIELDS WITH SENSITIVE DATA | SECURITY AND PRIVACY HANDLING |
| User ContentUser MetadataUser Activity DataIdentifiable DataS/PIIBusiness DataEmployee DataPseudonymous DataAnonymous DataHealth DataChildren’s Data**None**Others\* (\*please specify) | NA | NA |
| TRANSFORMATIONS APPLIED |  | LIBRARIES AND METHODS USED |
| Anomaly DetectionCleaning Mismatched ValuesCleaning Missing ValuesConverting Data TypesData AggregationDimensionality ReductionJoining Input SourcesRedaction or Anonymization**Others\*** (\*Cross-product of tokens and identity terms, tuple filtering, annotation aggregation) |  | * Cross product: python basic functions * Tuple filtering: python basic functions, NLTK for tokenization * Annotation aggregation: python basic functions |
| ANNOTATION WORKFORCE TYPE | ANNOTATION CHARACTERISTICS | ANNOTATION DESCRIPTION |
| Annotation Target in DataMachine-generated AnnotationsHuman Annotations - Expert**Human Annotations - Non-expert**Human Annotations - EmployeesHuman Annotations - ContractorsHuman Annotations - CrowdsourcingHuman Annotations - Outsourced / Managed TeamsUnlabeledOthers\* (\*Please specify) | Visual Attribute annotation Number of annotators per example 3 | Visual Attribute annotation  * The attributes are annotated by human-raters who were asked to label whether the attribute could be visually depicted in the image based on a Likert Scale ranging from Strongly Agree (5) to Strongly Disagree (1) |
|  | ANNOTATOR BREAKDOWN | ANNOTATOR DESCRIPTION |
|  | Annotator type Paid - Non-expert  Total unique annotators  Total cost of annotation  Expertise of annotators Trained for task | * We recruited \_\_\_ annotators across all regions for annotating stereotypes. * To test their understanding of the task, we conducted a pilot annotation. |
| VALIDATION METHOD(S) | VALIDATION BREAKDOWN | DESCRIPTION OF VALIDATION |
| **Data Type Validation**Range and Constraint ValidationCode/cross-reference ValidationStructured ValidationConsistency ValidationNot ValidatedOthers\* (\*Please specify) | N/A | Data Type Validation  The attributes were checked to be strings of text. The rating were checked to be values corresponding to the Likert Scale. This was checked using and corrected (if needed) using basic python functions. |
|  | VALIDATORS CHARACTERISTIC(S) | VALIDATORS DESCRIPTION(S) |
|  | N/A (automatic validation) | N/A (automatic validation) |
| ML APPLICATION(S) |  |  |
| 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 ArtConcepts and Definitions referenced in this Data Card |
| --- |
| Attribute Tokens (or tokens for short) |
| Definition: These are characteristics or attributes for which we aim to rate the visualness. 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 (adjectives), computer science, mathematics (subjects of study) and so on. |
|  |

###### 

| Reflections on Data | | |
| --- | --- | --- |
| Limitations due to human annotation | Annotation about the visual nature of attributes is subjective. While we attempt to capture diversity in our annotator pool w.r.t. gender and geographical region, we recognize that it still does not capture all different opinions and perspectives. Future iterations of such data collection should take more participatory approaches and involve communities with lived experiences on the harms of bias in society. | |
| No ground truth on the visualness of attributes | We recognize that there is no “ground-truth” on labeling the visual nature of attributes”. This is an inherently subjective opinion that is influenced by personal experiences. Thus, we caution against using the data in this dataset in any way for classification of such attributes. | |
| Attributes not captured by this dataset | We use attributes from existing sources. This limits which attributes get annotated for their visual depiction, and there exist attributes not captured by our dataset. | |
|  | | |

###### 