



# SCReeD Dataset Declaration Form

Dataset name:\* Myths and Facts Data on Child Development

Dataset version:\* 1

Dataset URL:\* <https://github.com/CSCRC-SCREED/CSU-Myths-and-Facts>

Creation date:\* 1st June, 2023

Last update:\* 6th of August, 2024

Author(s):\* Mehedi Tajrian, Azizur Rahman, Ashad Kabir, Rafiqul Islam

Author(s) affiliation(s): School of Computing, Mathematics and Engineering, CSU

Author contact(s):\* mtajrian@csu.edu.au, azrahman@csu.edu.au, akabir@csu.edu.au, mislam@csu.edu.au

Keywords:\* age, brain, behaviour, care, child, cognitive, development, hear, impact, language, learn, make, talk, parent, play, social, world.

Description/background:\* The data is about the facts and myths related to child development. It comprises information from social media, such as blogs. It includes mainly the text data. However, we cleaned and processed it into qualitative and quantitative data.

Dataset funding: N/A

Attribute details:\* Word Count, Average Word Length, Sentence Count, Stop Words Count, Noun Count, Verb Count, Adjective Count, Adverb Count, Pronoun Count, Noun/Verb Ratio, Adjective/Adverb Ratio, Sentiment Polarity, Sentiment Subjectivity, Word2Vec Average, Unique Word Count, Lexical Diversity, Punctuation Count, Uppercase Word Count, Average Sentence Length, Exclamation Mark Count [Click or tap here to enter text.](#)

Intended target (if specified): Data science, Child research, Misinformation classification, NLP, Data mining.

Format:\* CSV

License:\* Open

Standard compliance: [Click or tap here to enter text.](#)

Type:\* Qualitative, quantitative.

Size:\* 267KB

Availability: Click or tap here to enter text.

Data status: Clean data

Data provenance:\* Originated from the raw data that collected from blogs/SM.

Source computing infrastructure:\* Live system

Accompanying program(s)/script(s): Click or tap here to enter text.

Software installer or VM for replication: Click or tap here to enter text.

Generated or captured via:\* Manually

Category/categories:\* ML training data

Published in: Click or tap here to enter text.

Open research question(s) (if any): Click or tap here to enter text.

Potential use case(s) or application area(s): Click or tap here to enter text.

Data access control:\* Global access

Data retention period:\* 5 years

Data validation/checksum:\* CSU team

GDPR compliance:\* Yes

Consent: N/A

Ethics approval: N/A

Ethics considerations: Consulted with ethics office and they advised that it is not required.

*\* denotes mandatory field*

☒ I confirm that I have read, understood, and agreed to the submission guidelines, policies, and submission declaration.

☒ I confirm that the contributors of the dataset have no conflict of interest to declare.

☒ I agree to take public responsibility for my dataset's contents.

A handwritten signature in black ink, appearing to be 'Nahla' with a long, sweeping underline.

Signature of Corresponding Author  
(signed on behalf of all contributors)

Date: 25th July, 2024