**ReadMe - Goodreads Corpus – Curated without full text reviews (anonymized)**

This datafile is organized by curations, that is to say, each row in the file refers to one curation made in the corpus of online book reviews we put together.

**Variable names:**

- ID: a number identifying a particular curation
- Link: identifier for title-page on the Goodreads website (identifying the page for the book that was reviewed, not for the specific review itself)
- Round: the round in the annotation process in which this review was annotated. The rounds ran from round 7 to 15. Before round 7 we used a different annotation software that did not allow curation. Round 0809 and round 1314 were merged as the meetings in between these rounds were canceled.
- Title: the title of the book that was reviewed
- Author: the author of the book that was reviewed
- Review: a number identifying a particular review
- Mode:
    - Absorption (the reviewer mentioned being absorbed in the book they are reviewing)
    - AbsorptionMention (the reviewer made a general statement on absorption)
    - NA (the review did not contain any absorption)
- Presence:
    - Present (the reviewer mentioned that absorption was present)
    - Negated (the reviewer mentioned that absorption was negated)
    - NA (the review did not contain any absorption)
- SWAS_Category (see the annotation guidelines for descriptions of the higher order categories):
    - Attention
    - Emotional_Engagement
    - Mental_Imagery
    - Transportation
    - Impact
    - NA
- SWAS_Tag (see the annotation guidelines for descriptions of the fine-grained tags)
    - A1 Altered_sense_of_time
    - A2 Concentration
    - A3 General_sense_of_absorption
    - A4 No_distractions
    - A5 Forgetting_surroundings
    - A6 Anticipation
    - A7 Inability_to_stop_reading
    - EE1 Perspective_taking
    - EE2 Sympathy
    - EE3 Emotional_connection
    - EE4 Empathy
    - EE5 Compassion_for_story_events
    - EE6 Anger

    - EE7 Fear
    - EE8 Emotional_familiarity
    - EE9 Wishful_identification
    - EE10 Emotional_understanding
    - EE11 Parasocial_Response
    - EE12 Participatory_Response
    - MS1 Imagery_of_character
    - MS2 Imagery_of_story_events
    - MS3 Imagery_of_story_world
    - MS4 Realness
    - T1 Presence
    - T2 Merge_of_fiction_and_reality
    - T3 Proximity_of_story_world
    - T4 Deictic_shift
    - T5 Part_of_the_story_world
    - T6 Return_deictic_shift
    - T7 Travel_in_story_world
    - IM1 Effortless_Engagement
    - IM2 Wish_to_reread
    - IM3 Anticipation_Bookseries
    - IM4 Addiction
    - IM5 Lingering_Story_Feelings
    - NA
- Statement: curated text segment from the review or NA (in case a review did not contain any absorption)
- Onset: start of the curation (in number of characters) or NA (in case a review did not contain any absorption)
    - 0 is considered to be the index of the first character of a review
- Offset: end of the curation (in number of characters) NA (in case a review did not contain any absorption)
    - 0 is considered to be the index of the first character of a review
- Genre_1: genre designation for this book with most votes on Goodreads
- Genre_2: genre designation for this book with second most votes on Goodreads
- Genre_3: genre designation for this book with third most votes on Goodreads
- Length_characters: length of the review in characters
- Length_tokens: length of the review in tokens
- Rating: rating given by reviewer for this book following Goodreads' rating system
- Rating_stars: rating given by reviewer for this book using star-rating system from 1 to 5