**Fan fiction**

**Origin**

This datasets consists of a single, comma-separated, UTF8-encoded data table (*en.csv*). Each row represents metadata for a single text, drawn from a representative collection of English-language fanfiction that we have collected (in December 2018) from *Archive of our own* (AO3: <https://archiveofourown.org/>), a non-commercial platform, initiated by the Organization for Transformative Works. Fanfiction or transformative literature broadly refers to a rapidly expanding body of fictional narratives, typically produced by non-professional authors who self-identify as ‘fans’ of a particular oeuvre or individual work. Usually, these stories (or ‘fics’) are openly shared online with a larger fan community on platforms such as AO3. Fanfiction is the fastest growing form of writing in the world nowadays. When sharing their texts, fanfiction writers explicitly acknowledge taking inspiration from one or more cultural domains that are known as **fandoms** (e.g. Sherlock Holmes, Marvel Comics, Harry Potter). The reader should note, however, that many fanfics are explicitly erotic (and pronographic) in content, leading to unexpected settings for these stories and unexpectedly explicit interactions between characters.

**Columns**

This dataset comes with very rich metadata and for each fanfic and we have access to numerous (also author-provided) information fields. The entire dataset is limited to English-language fanfics and counts 24,143 individual items, which is still only a fraction of the entire archive.

* 'X' (int): row index column (*ignore*)
* 'work\_id' (int): a unique identifier
* 'title' (str): the fanfic’s title
* 'author' (str): the fanfic’s author (at least: the user name on AO3)
* 'rating' (str): a category label indicated for which audiences/age ranges the text is (not) suitable (e.g. “explicit”)
* 'category' (str): a category label for the sexual relationships discussed in the fanfic (e.g. M/M for homosexual encouters between two men etc.)
* 'fandom' (str): a label from which fandom(s) the fanfic drew inspiration (e.g. “Assassin’s Creed”); multiple fandoms are possible in the case of a “crossover” or “fusion fic”
* 'relationship': slash-separated character pairs, between whom a romantic relationship is discussed (e.g. “Harry Potter/Draco Malfoy”)
* 'character' (str): names of characters occurring in the story (e.g. Dan Howell, Phil Lester)
* 'additional.tags' (str): additional metalabels (mostly comma-separated)
* 'language' (str): the language (limited to English here)
* 'published' (date): date when the story was published
* 'status' (str): status indication of the work (e.g. “completed”)
* 'status.date' (date): date when the status was last updated
* 'words' (str): integer word token count for the document
* 'Chapters' (int): the number of chapters which the work contains
* 'comments' (int): the number of comments which were posted on the fanfic since it was first published.
* 'kudos' (int): the number of “likes” which the fanfic received since its first publication
* 'bookmarks' (int): the number of “bookmarks” (i.e. an expression of an interest to read in the future) which the fanfic received since its first publication
* 'hits' (int): the number of times the work retrieved from the website.