

Anonymization Guide

Overall goal: Anonymize the essays such that no essay can be connected to its author, while trying to preserve the narrative sense of the essay.

Common fixes:

Anonymize: Use of author's own name, ID number, place of residence. Names of author's friends, roommates, and family members. Schools that the author attends. Ages and birthdates of authors or author's family members.

In-text fixes: Use asterisks and generalize the information using capital letters. Delete any personal header information and/or personal messages. Example:

Original text example:

(I'm so sorry that this essay is not good!!)

Aaron Yamada

ID# 12345678

Yo me llamo Aaron Yamada y soy de Morgan Hill, California.

Anonymized text example:

Yo me llamo *FIRST_NAME* *LAST_NAME* y soy de *CITY*, *STATE*.

Codes used in 2019-2020:

FIRST_NAME

LAST_NAME

CITY

STATE

UNIVERSITY

AGE

NUMBER

PLACE

BIRTH_DATE

EMAIL

ADDRESS