

## Part 2 Task A:

You'll now be an annotator (note: this part is done individually). A research group at a new social media company is working on compiling a dataset for studying toxicity in online discourse and researching potential models to detect it. They've collected a dataset of Wikipedia comments and now want to annotate them. You've been asked to annotate a small portion of the dataset contained below.

- If you were assigned Task A in Part 1, use your own instructions doc that your group came up with to label the examples.
- If you were \*not\* assigned Task A in Part 1, use the instructions doc that should have been emailed to you by a different group.

At the end, there's a brief reflection question about how you found the labeling process.

Enter your labels comments below. Make sure you following the labeling scheme in the instructions exactly (i.e., don't write a label that's not included in the instructions).

1. Stalking? Harassment? What do you call it when Chzz deletes my contributions to the Softimage|3D page out of spite? Hypocrisy and favoritism seem to run rampant on this website.
2. It is obvious from your biased and prejudiced edits of my wiki entries that you are a sold-out shill for the pharmaceutical industry. Unfortunately for me, you carry too much weight with Wikipedia management, and so your campaign of cover-up will prevail.
3. I mean seriously, who are you? Are you joking? You make a couple of posts saying that the categories do not belong but don't explain why. You got to be kidding me. Your posts make it seem like you are a little kid.
4. First off, I am not self-appointed. I had to go through training twice before the community appointed me as an administrator. I wasn't rude or overly aggressive with you, especially considering that you were willfully ignoring the advice of others who had previously tried to guide you. In fact my message on your talk page is to the point and non-threatening, so I don't even see what your problem is.
5. It was very constructive you are just very very stupid.
6. I have restored material that was removed and added a substantial reference within the text and also tried to add some perspective and also some copyediting. I would appreciate those of you who feel the information restored is unfounded taking a look at the references at the bottom of the page. With the opening of the KGB archives a flood of material has become available.

7. This is completely ridiculous. The extent to which Google is viewed as an authority here is threatening to make Wikipedia simply a rehash of topics indexed in Google.
8. No doubt you are another Wikipedia Democrat. Haven't you realized yet that the democrats are just as evil as the Republicans?
9. I could care less whether you're acting in good faith or not...the fact of the matter is that both you and Rubin are incompetent.
10. I am assessing this article to Start/Low. There is still room for expansion and many parts of the article need to be proof-red for grammar and punctuation errors. More references could also be used in some sections.
11. Why are you threatening me, buddy? I didn't do anything to you! And like I care about editing Wikipedia. Loser.
12. So when is someone going to warn YOU about your toxic attitude then? Tell me that. Tell me when someone's going to give you a block warning for the bullshit you've pulled on Wikipedia, like getting into arguments with people and issuing blocks to them when you don't like hearing the truth.

Briefly reflect on the labeling task, including how easy/difficult you found it, how much ambiguity you found in the instructions and the degree to which you needed to use personal judgement to resolve uncertainty.

MIT OpenCourseWare

<https://ocw.mit.edu>

RES.TLL-008 Social and Ethical Responsibilities of Computing (SERC)  
Fall 2021

For information about citing these materials or our Terms of Use, visit:

<https://ocw.mit.edu/terms>