

## Social exclusion categories, their definition and the rubric for annotation

### 1. Social exclusion categories and definition

We define social inclusion in online communities as “*having an equitable opportunity to participate, by choice, in online activities of a community, without being disadvantaged on the basis of their identity and other factors (e.g., preferences)*”. Barriers for such equitable participation cause **social exclusion** through prejudiced attitudes and discriminatory behaviours within a community arising based on identity characteristics such as a person’s age, culture, beliefs, abilities, family composition, gender identity, sexual orientation, appearance, income and language use. In the table given below, we have identified attributes or characteristics that may deter individuals from having access to an equitable opportunity to participate in online platforms, particularly when these attributes are maligned (i.e., talked badly about in a derogatory way). For example, a comment stating that girls can’t write code would be offensive for women and they may not feel comfortable participating in an online environment (which has the potential to trigger social exclusion). Talking badly about disability (e.g. saying someone is retarded) will offend people with intellectual disability. The second column in the table below shows the name of the attribute, the third column shows the definition of these attributes and the last column sample textual comments belonging to a particular category (which can cause social exclusion).

No	Attribute/characteristic	Definition	Examples of socially-exclusionary comments
1	Gender	This represents the biological characteristics of a person (e.g. male or female) or the internal sense of self that a person identifies with (e.g. a man identifying as a woman or vice versa).	I don’t think <u>girls</u> in any support could fix that  JavaScript is the <u>ugly girl</u> freshman year with glasses and acne
2	Sexual orientation	This attribute represents the identity of an individual based on their sexual preference such as straight, lesbian, gay, bisexual, transgender, queer or intersex.	Notice how this question was closed off by trully proud <u>gay lord</u> american names.  You <u>gay</u> prick. Why the fuck did you revert my edit ..
3	Ethnicity	This attribute identifies the ethnic group an individual belongs to. Examples include Indian, American, and African.	Who the hell are you to talk? <u>An Arabian terrorist?</u> ”

4	Religion	This attribute identifies the religion one belongs to - e.g., Islam, Christianity and Hinduism.	Oh look a fucking <u>islamist</u> with admin powers fuck you asswipe  Go for it dipshit. You're just another <u>islamo christian</u> Terrorist.
5	Disability	This attribute identifies the disability (physical or mental) that impedes their ability to fully participate in the society (e.g., wheel-chair bound, dyslexia).	Oh my god the nazi moderator <u>retards</u> strike back  Stackoverflow is on that proverbial slippery slope. Yeah well fuck your mother you <u>retarded</u> <u>autist</u> you should kill yourself you fucking skinny ass nazi
6	Location	This attribute identifies a specific location a person is from (e.g., from Africa and Pakistan).	Fuck off already get the fuck out of my country Move to fucking <u>Pakistian</u>  why the fuckin <u>indian</u> guy edited my question  The above posted by a typical <u>American</u> retard
7	Employment status	This attribute identifies whether someone is employed or not, and if they work, whether they are a part-timer.	“Well, I think you’re ridiculous and borderline nazi. Go find a hobby. You seem like the type who is on SO way too much. <u>Maybe get a job</u> ”
8	Language use	This relates to the ability to speak a particular language such as the English language.	“pls use english not niggaenglish”  “Don’t mean to be a dick but you’re

			gonna have to learn to speak English better if you expect people be able to help you here.”
9	Age	This attribute identifies the age group an individual belongs to (e.g., young vs. old, and specific age-groups such as teenagers and millenials).	Go fuck Loy Wong you dirty <u>old</u> white asian basher
10	Software-specific	This relates to derogatory comments made on specific software, sometimes in comparison with competitors (Mac vs. PC, Chrome vs. Firefox, Java vs. Python)	<p>“<u>Windows is junk</u> and any company that uses windows for workstations are idiots and obviously don’t understand the higher long term IT costs on infrastructure and maintainance of running windows.”</p> <p>“I’m deleting my account on <u>Reddit</u>. <u>It’s full of crap</u>”</p> <p>“Hello it would be cool if you can add a method that returns the Auth ticket for the user in the current app. This method exists in steam user js but <u>javascript is a pain in the ass</u>. This method will be so useful for interacting with community servers.”</p>

11	Hardware-specific	This relates to derogatory comments made on specific hardware	“ <u>Stupid</u> <u>mac</u> <u>machine</u> ” “ <u>Useless dell pc</u> ”
----	-------------------	---	--

## 2. Rubric

Based on reading a comment an annotator must identify whether or not a specific attribute is present or not.

- A value of 1 indicates the presence of an attribute.
- A value of 0 indicates the absence of an attribute.

Having read a comment, an annotator will enter a value (1 or 0) for all the 11 attributes identified above.