

ANNOTATION MANUAL: Online risks

GENERAL GUIDELINES

- When the conversation does not include enough context for the online risk, and we are not sure whether to annotate, we do not annotate.
- We annotate also negative responses if they fit within the guidelines (e. g., “wanna go for a smoke?” “nope, i don’t smoke”)

RESPONSES AND LONG CONVERSATIONS

We code responses when they are related to the person and explicitly tied to the risk.

Example:

- “I am so anxious that I cannot attend school. I need to see my doctor soon.” [MENTAL HEALTH]
- “I am sure it will be better, your doctor can prescribe you stronger meds and soon the anxiety will be gone” [MENTAL HEALTH] X “Every problem has a solution, don’t worry.” [NO TAG]

1. | (15) Aggression, violence, harassment, hate speech and conversations with elements of aggression

Does the line of chat include:

exposure to vulgar / aggressive conversation* | aggressive contents / insults / threats / defamation* | referring to or inviting to join violent / aggressive behavior (cyberaggression, harassment, violence) | hate speech, xenophobia, racism, discrimination against a nation / ethnicity / colour of skin / religion / sexuality / weight

*only when directed towards concrete people or groups

- We do not annotate vulgarisms that are not directed (e. g., just saying vulgarisms).
- We annotate vulgarisms even if they are in a friendly context (e. g., calling girl-friends bitches) towards people
- We also include aggression towards groups and public figures (politicians, celebrities)
- Implicit racism is also coded, as well as aggression related to dehumanization and dishonesty.

Example: | **referring to violent/aggressive behavior** | “...so when i came back from work she was already there... and then i hit her and we fought...”

Example: | **racism** | „Yeah, that neighborhood is so full of gypsies that it’s dark even during the day there“

2. | (16) Mental health problems and self-harm

Does the line of chat include:

referring to / complaining about / experiencing long-term mental health problems as depression, anxiety, phobias, paranoias, insomnia, eating disorders (anorexia, bulimia, binge eating), self-harming, suicidal ideation | referring to therapy, medication, psychiatric hospital | describing experiences with eating disorders, instructions for not eating and drastic losing weight, sending pro-ana contents

- We do not annotate short term feelings or affects (bad mood, sadness). It must be obvious from the conversations that the symptoms are long-term.
- Our goal is to detect discussing mental health problems, not diagnose and assume problems of our participants - therefore we code explicit mentions of mental health problems and not general discussions of bad mood etc.

Example: | **referring to self-harm/suicidal ideation** | “Ana... She was really mean. I had a panic-self-hate attack with like... umm... not positive consequences... and I wanted to overdose.”

Example: | **referring to therapy** | “My psychiatrist is so stupid, I need to find a new one, he kicked me out and doesn’t wanna prescribe me any meds!”

3. | (17) Alcohol and drugs a drogy

Does the line of chat include:

referring to own or someone else’s experience with alcohol or drugs (cigarettes, nicotine packs, tobacco, hookah, marihuana, abusing medication ...) | making plans to drink alcohol or take drugs | seeking drugs | supporting/justifying alcohol and drugs | talking about the intention to try/use alcohol or drugs

Example: | **talking about the experience with drugs** | “Man, I’m so effin stoned 😊!”

Example: | **making plans to drink alcohol** | “Are 3 bottles of wine enough? 😊”

4. (18) Weight loss and dieting

Is the content of the line any of the following:

| weight loss support | sharing experiences with weight loss | positive evaluation of weight loss | how to diet/how to lose weight/how to be very thin | experience with dieting | expressing dissatisfaction with one's weight (subjectively perceived excess weight) | sharing/recommending/evaluating proana sites/groups/articles/content if not clearly related to PPP |

Child weight loss is not: does not include those who want to have muscles, pathologies associated with PPP [this is mental health]

Category includes both positive and negative phenomena associated with weight loss.

Example: | expressing dissatisfaction with one's weight | "I guess I'm fat. I'm depressed."

Example: | encouragement to lose weight | "You should lose weight, Kaya, you've got high blood pressure and cholesterol. The extra weight is already hurting you. You don't want to have a heart attack."

Example: "I can go 30 days without eating"

5. | (19) Sexual content

Does the line of chat include:

links to / talking about pornography | discussing sexual experience and sex life | flirting with sexually explicit content | sexting | sexual innuendo (even if meant as a joke between friends) | soliciting nudes or sexual information

Example: | **discussing sex life** | "You did not answer me last time... what's it like during sex from a boy's point of view? :P :D"

Example: | **sexual innuendo (even as a joke)** | "oh so you're at home? man, you can jerk off all day haha!"