Task:

- 1. Categorize the different texts in one of the following category
 - o Question
 - Complaint
 - o Praise
 - Suggestion
 - o None
- 2. Categorize the different texts according to their main subject
 - Builds (What armor? What weapons? Which stats? ect...)
 - PvP related discussions (Arenas, OPRs, races, wars)
 - PvE related discussions (Expeditions/Dungeons, farming, crafting)
 - Hardware-related (pc, graphic requirements, graphic problems)
 - No categorization possible
- 3. Categorize the different texts according to their general feelings
 - o Positive (good, excellent, incredible, amazing)
 - Negative (bad, boring, not-working)
 - Neutral
 - o Cannot categorize

Tries per tasks

Table of annotation success rate:

- 1. Category is okay = 1
- 2. Category could be okay, but not main element of the text = 0.5
- 3. Category would've been good, if prompt allowed it: = 0.25
- 4. Category is wrong/Category is not part of the choice and is wrong = 0
- 1. All categories have been identifies = 1
- 2. 50% + of categories have been correctly identified = 0.75
- 3. Some categories (-50%) have been correctly identified = 0.5
- 4. One category has been identified/Suggested alternative could be ok = 0.25
- 5. No category has been correctly identified/No input/Suggestion category is wrong = 0
- 1. feeling correctly identified = 1
- 2. Two answers, but one is correct: 0.5
- 3. feeling not correctly identified/not answered/wrongly classified = 0