AP2: Contentiousness Analysis Annotation Guidelines

1. **Introduction**

   The purpose of this contentiousness task is to analyze the degree of controversy within forum posts and their titles. A total of 500 Reddit posts gathered from the following 6 subreddits will be used: unpopularopinion, TrueUnpopularOpinion, rant, changemyview, PurplePillDebate, and the10thdentist due to the 100 request limit from each request from the Reddit API. To avoid just randomly selecting the newest posts from each subreddit which could lead to differing posts if someone were to recreate this task. We originally started with just unpopularopinion, TrueUnpopularOpinion, rant, changemyview and PurplePillDebate, but due to the Reddit API's 100 request limit and some posts having no text in the body, we added the10thdentist to fill in the remaining posts that we would we using. According to Merriam-webster.com, contentiousness is defined as anything that is " *likely to cause disagreement or argument*" and "*exhibiting an often perverse and wearisome tendency to quarrels and disputes*"
   "The following information must be detected and meaning must be disambiguated to grade contentiousness:

2. **Generalizations**
   Definition:
   A generalization is a broad statement or concept. It involves drawing conclusions or making assumptions about a subject without considering individual differences or exceptions. Generalizations oversimplify complex realities and overlook variations or nuances within the context, leading to potential inaccuracies or biases.

   Annotation Procedure:
   Document the identified generalizations along with specific examples from the posts. Note the frequency of occurrence and prominence of each generalization to assess its significance. Consider the context of each generalization and its relevance to the overall classification task. Mark it with the abbreviation [GRN-n], where n is the nth generalization in the post.

   [2.1]: "TITLE: The boy scouts never should have admitted girls
   Part of the Body: When you are young and its just boys around, the dynamic is totally different. [**GNR-1**] You start constructing things, competing with each other. You develop implicit honour rules and form brotherly bonds.

   The moment a girl joins the group the dynamic is suddenly different.[**GNR-2**] Suddenly the girl has lots of power as the only girl. Some boys stop being interested in the competitions and exploring and building, as they just want to compete for the girl. They suddenly care more about looking cool to the girl, and looking cool often means not engaging in things like building.

3. **Polarizing Language**
Definition:
Polarizing language refers to expressions, phrases, or statements that evoke strong emotions or opinions, often leading to division or controversy among readers. This includes language that is confrontational, inflammatory, or highly opinionated.

Annotation Procedure:
Annotators should highlight phrases or sentences that convey strong emotions, such as anger, frustration, enthusiasm, or disdain. Additionally highlight language that is confrontational or inflammatory in nature, such as personal attacks, insults, or derogatory remarks aimed at individuals or groups.

[**3.1**]: "TITLE : Dear Reddit fuck off [**PLZ**] already with the Talibangelical ads from u$er He Gets Us who cannot even be blocked
Body: That's some serious bullshit[**PLZ**] allowing the same people who want to strip men, women, and children of their rights to advertise their bullshit on our feeds and we cannot even block them. Fuck you[**PLZ**] for this.

4. **Reactionary Language**
Definition:
Reactionary language refers to expressions, phrases, or statements that resist or oppose societal, political, or cultural changes. This includes language that promotes nostalgia, advocates for the preservation of traditional values, expresses skepticism towards progressive ideas or anticipates disagreement or confrontation.

Annotation Procedure:
Annotators should highlight phrases or sentences that anticipate disagreement or confrontation, suggesting an expectation of opposing viewpoints or criticism. Highlight phrases or sentences that express resistance to societal or cultural changes, advocating for maintaining the status quo or reverting to previous practices.

[**4.1**]: "TITLE : Duolingo is a videogame. If you want to learn a language, you need grammar focused content.
Body: I get downvoted whenever I say this in any sub that isn't specifically dedicated to language learning, so I assume this is an unpopular opinion. [**RCL**]

There's more to language than just words. You need to learn grammar. Duolingo doesn't teach grammar. They did have some grammar hints for some languages a while ago, but I believe they did away with those too.

If you want some language learning related leisure activity, then sure. I like to do it too with no illusion I'm actually learning.

Grammar books are boring, I totally get it.**[RCL]** But they are teaching you what you need to know.

5. **Tone**
   Definition:
   Tone refers to the author's attitude or emotional expression conveyed through the language, style, and phrasing of the post. It encompasses emotions such as anger, frustration, sarcasm, enthusiasm, or neutrality.

   Annotation Procedure:
   After identifying the presence of labels such as [GNR, PLZ, RCL], annotators should contextualize the entire text and derive the overall emotion or attitude being conveyed to the reader. Consider the broader context of the post, including the content(GRL-n), language used (PLZ), and intended audience, and assess the emotional expression conveyed through the language, style, and phrasing of the post to interpret the tone accurately.

   Tone Categories:
   Classify the tone into one of the following categories based on the overall emotional expression**:**
   1. **Positive**: Indicates a supportive or optimistic tone towards the topic discussed.
   2. **Negative**: Signifies a critical, confrontational, or pessimistic tone towards the topic.
   3. **Neutral**: Denotes a factual, objective, or unemotional tone without expressing strong opinions or emotions.
   4. **Mixed**: Represents a combination of positive and negative elements in the tone, indicating ambivalence or complexity.

   [5.1]: "TITLE : Stop projecting your feelings about your good parents onto people who had bad parents.
   BODY: How could you not like your parents? They're your parents!" Like, shut up. Not everyone had the exact same experience as you. People are allowed to dislike their parents if they were/are shitty, terrible, abusive people. **[TNE = Negative]**

6. **Combining the Labels**
   This section is just to show readers how a post would be annotated using the listed connotations above in unison.

   [6.1]: "TITLE : CMV: Dating for 2+ years and still not having a ring is completely normal and not a red flag**[PLZ]**

Body: So i'm not sure if this is just a thing on tiktok, or if others also believe this but i constantly see tiktoks where a girl will mention her and her significant other have been dating for 3+ years and there will be a flood of comments saying "wow and you still have no ring that's a red flag run sis**[PLZ]**". I believe marriage is a life long commitment and knowing someone for one year is not enough time to fully know them or make the biggest decision and marry them**[GNR-1]**. I know couples in my personal life who got married quickly just to divorce less than two years being married. I'm 20 and i've noticed a lot of people in my generation don't believe in marriage**[RCL]**, I think there should be a talk to have a few months within dating to see if that is both of your guys end goal. And i know it's different for couples who are 30+ and dating**[RCL]** but I still see this getting thrown around at people in their early 20s. Another thing to mention is I know after 8+ years and still no ring then that is considered a red flag**[PLZ]**, but as mentioned earlier i think each person should be aware of what they want to get out of the relationship within the first few months of dating**[GNR-2]**. To summarize it up, get married at whatever time you feel is best and you fully know your partner. If you've been dating for 2 years i don't think you should automatically expect a ring. **[TNE = Neutral]**

7. **Contentiousness Grading**

For each Reddit post and its corresponding pair of title and main body, the annotators have to provide a contentiousness level rating according to the following three-level scale:

1. <u>Low Contentiousness</u> - The annotator believes that the post doesn't cause any controversy and is just a general statement that invites people to share their opinions and is open to them or has nothing within the context that would incite or agitate people.
2. <u>Medium Contentiousness</u> - The annotator believes that the post has some points within the context that will cause discourse. The post has a bias for one side of the argument which could be shown with a lot of generalizations and questions the other side with some reactionary language, but doesn't attack them which can be expressed from the tone.
3. <u>High Contentiousness</u> - The annotator believes the post is prone to causing large disagreements or arguments. The post makes a lot of generalizations that may be supported by polarizing language, and reactionary language that tries to impose on conflicting opinions, this can be supported by a negative tone but, a negative tone doesn't always indicate high contentiousness.

4. The hierarchical structure of the importance of the labels is as follows:
   [PLZ > RCL > GNR > TNE]
   The rationale is that while all these labels can provide insights into the contentiousness of a post, polarizing language, and reactionary language are more directly indicative of controversy, followed by generalizations and tone, which may influence but not always directly signify contentiousness.

[7.1] Analyzing example [6.1], Based on the annotations provided, the text exhibits a medium level of contentiousness. This is because the post makes use of generalizations, particularly regarding dating, marriage, and relationship timelines, which can oversimplify complex dynamics

and contribute to polarization. Additionally, while there is no explicit polarizing language, the discussion of marriage expectations and societal norms surrounding relationships could evoke strong opinions among readers. The acknowledgment of variations in perspectives based on age and societal norms suggests an awareness of differing viewpoints, indicating a level of complexity in the discourse. However, the overall tone of the text remains relatively neutral, with the author presenting personal observations and opinions in a reasoned and reflective manner, without resorting to confrontational language or strong emotional appeals. As such, while the text touches on contentious topics, it does so with a moderate level of measured introspection