Amharic Version: Annotation Guideline

<u>Bahir Dar University, Bahir Dar, Ethiopia</u>
<u>&</u>
<u>University of Humbug, Hamburg, Germany</u>

<u>መጠይቅ!</u>

የዚህ መጠይቅ ዋና አላማ በማህበራዊ ሚዲያ ላይ የሚለቀቁ የጥላቻ (hate) እና አፀያፊ (offensive) ንግግሮችን ለመለየት የሚያስችል በቴክኖሎጂ ምርምር የሚታገዝ መፍትሄ ለመስጠት ነው። ስለሆነም የአማርኛ ቋንቋን በመጠቀም የሚተላለፉ ጥላቻ እና አፀያፊ ጽሁፎችን የያዘውን በትክክል ለይቶ ለመክፈል ያለም ስራ ነው። በዚህ መጠይቅ ላይ የቀረበው የጽሁፍ መረጃ ከቲዉተር (Twitter) የተወሰደ ስለሆነ የጥላቻ (hate) እና አፀያፊ (offensive) ጽሁፎችን ይዟል። ስለሆነም እነዚህን ንግግሮች ማየት ካልፈለጉና በዚህ ስራ ለመሳተፊ ካልፈለጉ ጽሁፎቹን ሳይመለከቱ ውጣ የሚለውን መተግበሪያ ተጫነው ይውጡ።

- በመልእክት ላይ ያተኩሩ (በራስዎ አስተያየት ላይ ላለማተኮር ይምክሩ)።

- ሥራውን ለመጫረስ አትቸኩሉ እና ሁል ጊዜ ጥርጣሬ በሚፈጠርበት ጊዜ ተመራማሪዎቹን ያግኙ።

በመጠይቁ ለመሳተፍ ከወሰኑ መመሪያውን በአጽኖት ያንብቡና የደስራው ይግቡ።

ስለትብብርዎ በቅድሚያ እናመሰግናለን!!!

የቀረቡት ጽሁፎች ከሚከተሉት አራት (4) ክፍሎች ውስጥ አንዱን በመመምረጥ መልስዎን ይስጡ።

I. የጥላቻ ንግግር (Hate Speech) ማለት በቀረቡት ጽሁፎች ውስጥ አንድን ግለሰብ ወይም ተቁም የቡድን ማንነቱን የሚያንቋሽሽ (በብሄር፤ በጎሳ፤ በሃይማኖት፤ በባህል፤ በልምድ፤ በአካል ጉዳተኝነት፤ በጾታ፤ በፖለቲካ አመለካከት ወ.ዘ.ተ.) ንግግር ማለት ነው። ለምሳሌ፤ መድሎ፤ ማስፈራራት፤ ለጥፋት/ጥቃት ማነሳሳት፤ ማንኳሰስ፤ ንቀት፤ አጀሙር፤ ስድብ፤ ዘለፋ፤ የመሳሰሉ ነገሮች ማንነትን መሰረት አድርገው ሲቀርቡ ነው። አንድ ጽሁፈ ጥላቻና አፀያፊ ቅልቅል ንግግሮችን ከያዘ ጥላቻ ንግግሮች ውስጥ ይከታታል።

II. አፀያፊ (Offensive) ንግግር ማለት አንዲን ግለሰብ ወይም ተቁም እንዲናደድ፤ እንድበሳጭ፤ እንዲቀየም ወይም ሞራሉ እንድጎዳ የሚያደርግ ትገቢ ያልሆነ ንግግር ማለትም ስድብ፤ ስም ማጥፋት፤ የአሽሙር፡ የንቀት፡ ወ.ዘ.ተ. ንግግር የያዘ ሲሆን ነው።

III. መደበኛ (Normal) ንግግር ማለት አንድ ጽሁፍ ምንም አይነት የጥላቻም ሆነ አጸያፊ ንግግር ሳይዝ ሲቀርና በአንባቢው ሰው አስተያየት ምንም ምጥፎ ነገር ያልያዘ ንግግር ማለት ነው።

IV. ለመወሰን የሚያስቸገር (UnSure) ንግግር ማለት አንድ ጽሁፍ መደበኛ ወይም ጥላቻና አፀያፊ ይሁን አይሁን ለአንባቢው ሰው መለየት ሲያቅ ው ለመወሰን የሚያስቸግር ይበ" />ል።

<u>የአጀምር ጤና አደጋ እና ደህንነት:</u>

ጎጂ ይዘትን መግለጽ ሥነ ልቦናዊ ጫንቀት ሊሆን ይችላል። በሂደቱ ወቅት ጫንቀት ወይም ምቾት የሚሰማዉ ማንኛውም ገላጭ እረፍት እንዲያደርግ ወይም ስራውን እንዲያቆም እና እርዳታ እንዲፈልግ እንመክራለን። ቅድመ ጣልቃ ገብነትን ለመቁቁም ከሁሉ የተሻለው መንገድ ነው

Bahir Dar University, Bahir Dar, Ethiopia
&
University of Humbug, Hamburg, Germany

Questionnaire!

The main objective of this research is to identify hate and offensive speech on social media with the help of technology. It will help to tackle hateful or offensive content that are written in Amharic languages and posted on social media. The data is crawled from Twitter using Twitter API and don't reflect the researchers' opinions.

If you do not want to see such content and do not want to participate in the work, you can click the logout button and leave without looking at the content.

- Focus on the message conveyed in the tweets and try not to focus on your own opinion on the topic.
- Do not rush to finish the task and always reach out to the researchers with questions when in doubt.

If you want toparticipate in the task, please read the guideline carefully.

Thank you for your participation in advance

Please read each tweet/comment carefully and choose its approperate label.

Category:

- **Offensive speech** is any form of bad language expressions including rude, impolite, insulting, or belittling utterance intended to offend or harm an individual.
- **Hate speech** is language content that expresses hatred towards a particular group or individual based on their group identities such as race, ethnicity, religion, gender, disability, political affiliation, or other characteristics. It also includes threats of violence associating group identities.
- **Normal** is any form of expression that does not contain any bad language belonging to any of the above classifications.
- **Unsure** any tweet that is difficult to give a specific label, or if you are not quite sure about the content's label, label it as unsure.

Mental health risk and well-being:

Annotating harmful content can be psychologically distressing. We advise any annotator who feels anxious or uncomfortable during the process to take a break or stop the task and seek help. Early intervention is the best way to cope.