One million posts

Annotation analysis (github issue #24, #25)

	Sentiment	Off-Topic	Inappropriate	Discriminating	Arguments Used	Personal Stories	Feedback
	3599	3599	3599	3599	3599	9336	6038
		Potentially undesirable content			Potentially desirable content		Neutral but requires a reaction by staff
Consolidated Annotator 1 Annotator 2 Annotator 3				1000 samples Source: random			
Annotator 1				1594 samples Source: 10 articles			
						3036 samples Source: 13 discussions	1953 samples Source: Replied by staff or pre-labeled
Annotator 2				513 samples Source: 10 articles			
						1077 samples Source: 13 discussions	486 samples Source: Replied by staff or pre-labeled
Annotator 4				492 samples Source: 10 articles			
					 	1624 samples Source: 13 discussions	

Comments

Annotation rounds

- There were 3 annotation rounds in total:
 - Round 1 was just a dry run whose results are not part of the dataset.
 - Round 2 consists of 1000 random samples.
 - Round 3 was a special round with the goal to increase the minorities in the "negative classes" and increase labels for the categories "feedback" and "personal stories"
- Annotator 3 vs 4
 - Annotator 3 was only present in round 2 whereas annotator 4 is only found in round 3.
- Overlapping annotator votes
 - In round 2 all three annotors have labeled the exact same 1000 posts. So the majority vote is taken as the final label.
 - In round 3 the annotators labeled on completely distinct sets. Except one single post (id 5444) is overlapping between annotator 1 and 2 concerning the additional feedback samples.
 However, both labeled unanimously with positive. The diagram accounts this label to annotator 1 for the sake of simplicity such that the total of the column is correct.

Round 3 analysis - "negative classes"

The 10 articles chosen for round 3 are:

Article ID dataset	Article title	Topic
1172	New York: Schwedin bekommt 18 Millionen nach sexueller Belästigung	Women / Feminism
1704	"Genderwahn": Ein Fall von Realitätsverlust	Women / Feminism
1831	Gewalt in der Partnerschaft: "Warum bleibst du bei ihm?"	Women / Feminism
1860	Vom Kinderarbeiter im Iran zum Zeltbewohner in Traiskirchen	Refugees
10425	Idomeni: Polizei setzte erneut Tränengas gegen Flüchtlinge ein	Refugees
10707	Junge Flüchtlinge: Die angeblichen Egoisten	Refugees
11004	Vergewaltigung am Praterstern: U-Häftling nach "Vorfall" verlegt	Refugees
9767	Israelischer Soldat soll verletzten Palästinenser mit Kopfschuss getötet haben	Israel and Palestine
10820	Israel lässt jüngste palästinensische Gefangene frei	Israel and Palestine
11105	Neues Gesetz soll israelisches Recht auf die Siedlungen ausdehnen	Israel and Palestine

Round 3 analysis - "personal stories"

The 13 discussions chosen for Personal stories are

Article ID dataset	Article title	
4	Welche Erfahrungen haben Sie als Linkshänder gemacht?	
5	Wie haben Sie das Jahr 1989 erlebt?	
6	Ihre schlimmsten Hotelerlebnisse	
8	Was wollten Sie als Kind werden?	
9	Wie war Ihr erstes Mal im World Wide Web?	
11	Wie gut leben Sie von Ihrem Job?	
1163	Der erste Urlaub alleine – wie war es für Sie?	
1855	Ihre Geschichte vom ersten Auto	
2123	Erzählen Sie uns von Ihren eindringlichsten Fernsehmomenten	
3581	Flucht aus der Heimat: Waren Sie oder Ihre Familie betroffen?	
5048	Welche Erfahrungen haben Sie bei der Arbeitssuche gemacht?	
5829	Plötzlich kein Lohn: Welche Erfahrung haben Sie gemacht?	
6551	Wenn Weihnachten traurig macht: Wie gehen Sie damit um?	