Don't be abusive Before submitting a response, make sure the response does not spread any hate, bigotry, prejudice or illegal content. We want to maintain the conversations peaceful and not to degenerate into a conflict. We are talking about people not categories.	
Think about the objectives Before writing a response, think about the effect it may create and the one you want to obtain. Paying attention to the objectives will help use proper words.	
Call for influential users Enlisting influential supporters (civic leaders, politicians, subject experts) will help bring attention and deepen the effect to counter narrative.	
Use credible evidence The information in hate speech may be confusing and misleading. Ask for clarification when necessary. Counter it with credible evidence and use reliable sources.	
Think about the tone We can demonstrate understanding and support to those who might be attacked. Be careful of using sarcasm, humour, parody and satire. We can use them, if we are able to master it as they run the danger of being antagonistic.	

Table 1: Counter-narrative guidelines adopted by the three NGOs.

Hate Speech	Counter-Narrative
Every Muslim is a potential terrorist.	Every Muslim is also a potential peacemaker, doctor, philanthropist... What's your point?
I musulmani violentano le nostre donne e vanno castrati.	Buonasera, se ho capito bene, lei mi sta dicendo che tutti gli uomini adulti di fede islamica presenti in Italia, hanno violentato, violentano o violenteranno le donne italiane? Può portare dei dati a supporto della sua tesi?
Muslims rape our women and must be neutered.	*Good evening, if I understood correctly, you're telling me that all adult men of Islamic faith present in Italy, raped, rape or are going to rape Italian women? Can you bring in data in support of your statement?*
Le voile est contraire à la laïcité.	Bien au contraire la laïcité permet à tout citoyen de vivre librement sa confession.
The veil is contrary to secularism.	*On the contrary, secularism allows every citizen to freely profess his faith.*

Table 2: Example pairs for the three languages, along with English translations.

2018). Other examples of data augmentation strategies are back translation (Sennrich et al., 2016) and gold standard repetition (Chatterjee et al., 2017) that have been used in sequence-to-sequence Machine Translation. In all these tasks, adding the synthetic pairs to the original data always results in significant improvements in the performance.

In line with the idea of artificially augmenting pairs, and since in our dataset we have many responses for few hate speeches, we produced two manual paraphrases of each hate speech and paired them with the counter-narratives of the original one. Therefore we increased the number of our pairs by three times in each language.

Counter-narrative type annotation. In this task, we asked the annotators to label each counter-narrative with types. Based on the counter-narrative classes proposed by (Benesch et al., 2016; Mathew et al., 2018b), we defined the following set of types: PRESENTATION OF FACTS, POINTING OUT HYPOCRISY OR CONTRADICTION, WARNING OF CONSEQUENCES, AFFILIATION, POSITIVE TONE, NEGATIVE TONE, HUMOR, COUNTER-QUESTIONS, OTHER. With respect to the original guidelines, we added a new type of counter-narrative called COUNTER-QUESTIONS to cover expressions/replies using a question that can be thought-provoking or asking for more evidence from the hate speaker. In fact, a preliminary analysis showed that this category is quite frequent among operator responses. Finally, each counter-narrative can be labeled with more than one type, thus making the annotation more fine-grained.

Two annotators per language annotated all the counter-narratives independently. A reconciliation