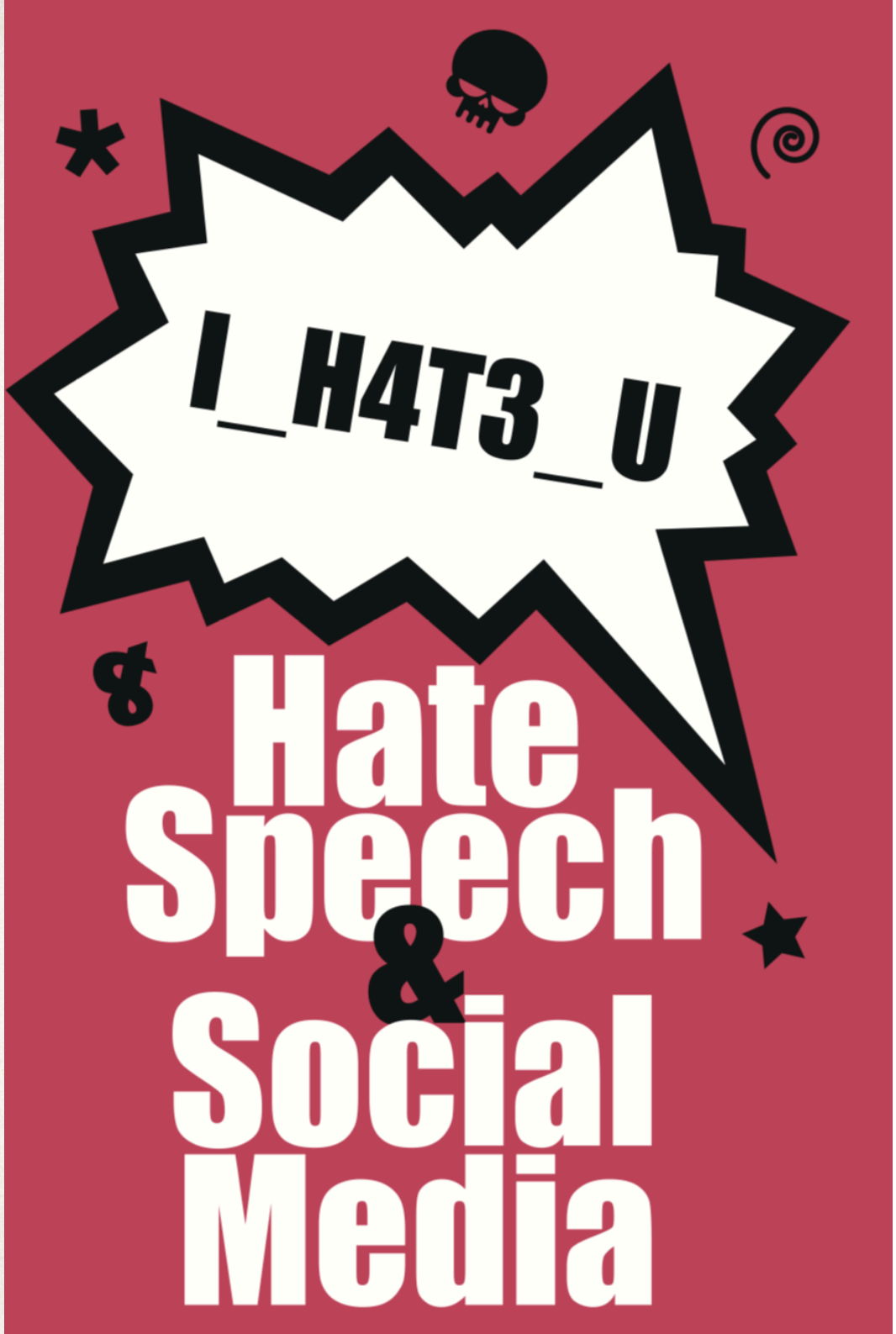
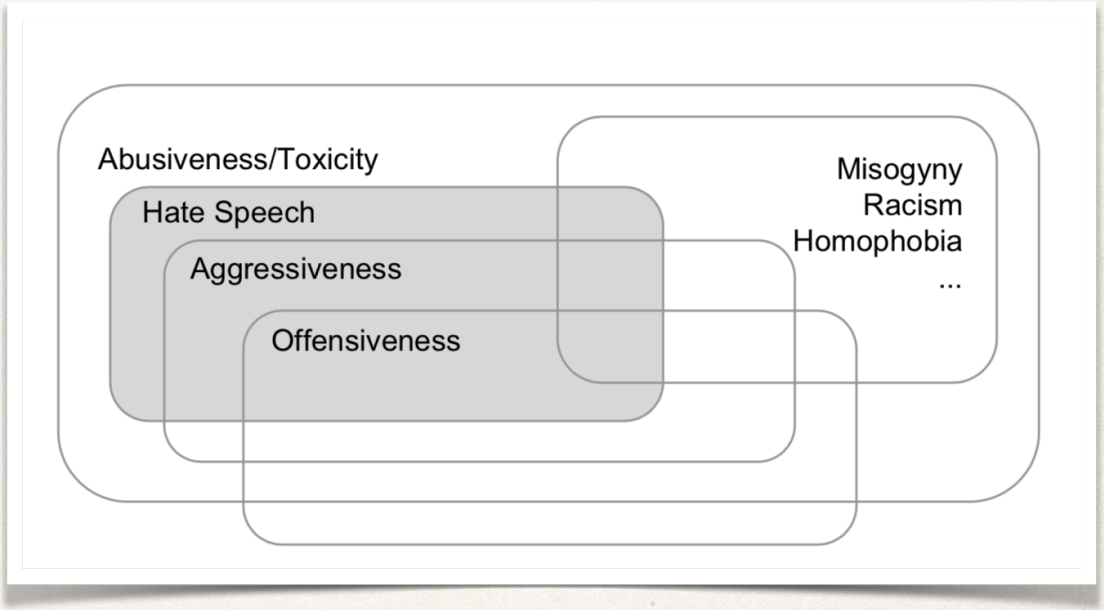
Abusive language / Hate speech et similia



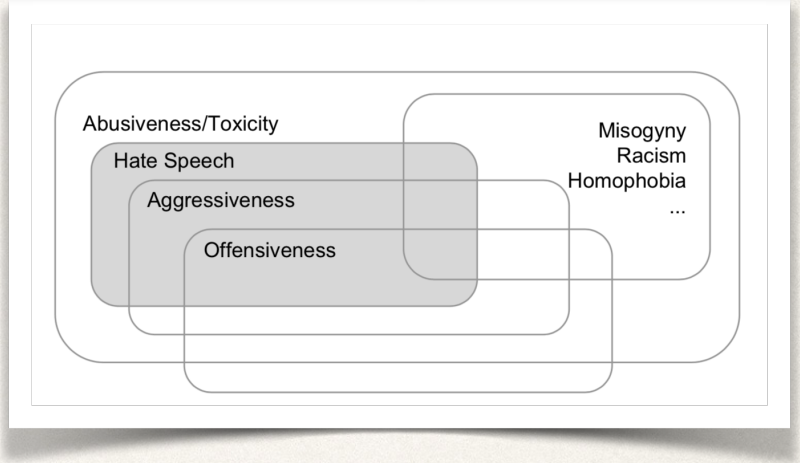
Hate speech and defining aspects



- Abusive language: umbrella term, which covers several related phenomena
  - The definition puzzle
- Relation between topical focuses in abusive language phenomena
- Offensiveness intersects with abusiveness, but includes also phenomena which are not abusive (cathartic swearing)
- Ontological modeling can help -> O'Dang! The Ontology of Dangerous Speech Messages.

Fabio Poletto, Valerio Basile, Manuela Sanguinetti, Cristina Bosco, and Viviana Patti. 2020. Resources and Benchmark Corpora for Hate Speech Detection: a Systematic Review. Journal of Language Resources and Evaluation. Springer.

Hate speech and defining aspects



- How to define the boundaries?
  - Free speech (sayable and unspeakable)
  - Toxic speech and blurred boundaries
  - Offensiveness (swearing)
- Hate speech: key aspects in many definitions:
  - Incitement to hatred towards a target belonging to a vulnerable category.
  - Expressions of hate consist of verbal attacks that incite violence and discrimination against individuals or entire groups characterized by a particular race or ethnicity (Africans, Roma, etc.), religion (Muslims, Jews, etc.), gender (women), or sexual orientation.
  - Minorities or (marginalized) groups with a long history of discrimination.

Debate on illegal hate speech

- Governments: laws
- Companies, social media platforms: content moderation
- Hate campaigns: not just a single post
- Peaks of escalation! Identifying the moments of change could be important! Longitudinal NLP could play a role (Maria Liakata)

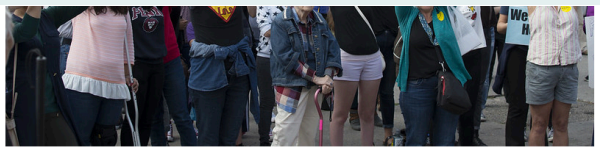
Subscribing to the code the ICT companies have agreed to commit to **remove the hateful contents within 24 hours** from the notification

SOCIAL MEDIA

New U.N. Report on Online Hate Speech

By **Evelyn Douek** Friday, October 25, 2019, 1:28 PM

> "Hate speech", a short-hand phrase that conventional international law does not define, has a double-edged ambiguity. Its vagueness, and the lack of consensus around its meaning, can be abused to enable infringements on a wide range of lawful expression. … Yet the phrase's weakness ("it's just speech") also seems to inhibit governments and companies from addressing genuine harms such as the kind that incites violence or discrimination against the vulnerable or the silencing of the marginalized.



Protest against Islamophobia and hate speech (Source: Flickr/Fibonacci Blue)

David Kaye, the United Nations special rapporteur on the promotion and protection of the freedom of opinion and expression, recommended in June 2018 that social media companies adopt international human rights law as the authoritative standard for their content moderation. Before Kaye's report, the idea was fairly out of the mainstream. But the ground has shifted. Since the release of the report, Twitter CEO Jack Dorsey responded to Kaye agreeing that Twitter's rules need to be rooted in human rights law, and Facebook has officially stated that its decisions—and those of its soon-to-be-established oversight board—will be informed by international human rights law as well.

Now Kaye has a new report, released Oct. 9—a timely evaluation of one of the biggest challenges in the regulation of online speech. Despite some tech companies expressing openness to Kaye's approach, in general these companies continue to manage "hate speech" on their platforms, as Kaye notes, "almost entirely without reference to the human rights implications of their products." And it remains unclear how these standards, developed for nation-states, can be put into practice in the very different context of private companies operating at mind-boggling scale and across a wide variety of contexts. These questions are the central concern of Kaye's latest report, which evaluates the human rights law that applies to regulation of online "hate speech."

Hate speech and defining aspects
Hate speech detection
Task definition

- Hate speech is commonly defined as any communication that is abusive, insulting, intimidating, harassing, and/or incites to violence, hatred, or discrimination, and it is directed against a person or a group belonging to a vulnerable category on the basis of some characteristics:
  - race, race, ethnic origin, religion, gender, age, physical condition, disability, sexual orientation, political conviction (Erjavec and Kovačič 2012)

- Binary classification task:

``decide whether a given message is a hateful speech utterance or a harmless one against a given target"

- Online hate is expressed in different forms depending on the subject against whom it is targeted
  - Target oriented nature of hate speech
    - Vulnerable category
    - Inciting hate or violence

Hate speech and defining aspects
Operational Definitions for Annotators

- Hate speech is commonly defined as any communication that is abusive, insulting, intimidating, harassing, and/or incites to violence, hatred, or discrimination, and it is directed against a person or a group belonging to a vulnerable category on the basis of some characteristics:

- race, race, ethnic origin, religion, gender, age, physical condition, disability, sexual orientation, political conviction.

  - Operational definition (Sanguinetti et. al., 2018)

2 Hate Speech

Labels: No – Yes

Two aspects are taken into account for its identification:

- the **target**, which must be a group identified as one of the three categories included in the search, or even an individual considered for its membership in that category (and not for its individual characteristics);
- the **action**, or more precisely the illocutionary force of the utterance: this means that we must deal with a message that spreads, incites, promotes or justifies hatred or violence towards the given target, or a message that aims at dehumanizing, delegitimating, hurting or intimidating the target.

  - Hate speech and Offensive language: different dimensions

Example
Developing corpora for hate speech

- Selecting data samples
  - Collect data from sources representative of the phenomena to be studied
  - Filter data by keywords and hashtags representing:
    - Hate speech targets > e.g. women, immigrants (Romas, Muslims, …)
    - Forms of hate speech > misogyny, racism, xenophobia, religious hate..
    - Monitoring potential victims of hate accounts, downloading the history of identified haters and filtering Twitter streams with keywords, i.e. words, hashtags and stems!

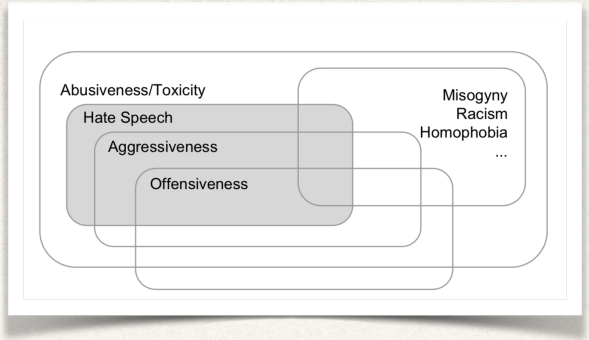
Developing corpora for hate speech

id_str	target	hate speech	aggressiveness	offensiveness		irony	stereotype	
782117718791221248	ethnic group	no	no	no	no	no	0	
782128837496745984	religion	no	no	no	no	no	0	
782142959789670401	ethnic group	no	no	no	no	no	0	
782145460664463360	Roma	no	no	no	no	no	0	
782165094318956548	ethnic group	no	weak	no	no	yes	0	
782195284105371648	Roma	yes	no	strong	no	yes	1	
782204731959734272	Roma	no	no	no	no	yes	0	
782241280659169281	Roma	yes	strong	weak	no	yes	3	
782268118194229248	Roma	no	no	no	no	no	0	
782349137257922560	Roma	no	no	no	no	no	0	
782462957842300930	ethnic group	no	no	no	no	no	0	
782508027815485442	Roma	no	no	no	no	yes	0	
782512181707440128	Roma	no	weak	no	no	no	0	
782559406311477248	Roma	yes	weak	no	no	yes	2	
782563896934666240	Roma	no	no	no	no	no	0	
782584588103278597	ethnic group	no	strong	strong	no	yes	0	
782588461006090240	religion	no	no	no	no	no	0	
782596951283933184	religion	yes	weak	weak	no	yes	3	
782614667759849472	ethnic group	yes	weak	no	yes	yes	3	
782627058115641345	religion	yes	weak	no	no	yes	3	
782640781290983424	ethnic group	no	no	no	no	no	0	
782686657732640768	religion	yes	strong	no	no	yes	3	
782787286857494528	ethnic group	no	no	no	no	no	0	
782838281444683776	ethnic group	no	no	no	no	no	0	
782838442044559361	ethnic group	yes	weak	weak	no	no	1	
782861476126162944	religion	no	no	no	no	no	0	

Example
Developing corpora for hate speech

- Human annotation
- Annotation scheme / Multilayered annotation schemes
  - Define a set of labels oriented to made explicit the presence of hate speech … in texts, starting from some operational definition
    - Coarse-grained: Hateful? Yes or no
    - Fine-grained: e.g. more relevant aspects characterizing hate (e.g. stereotypes, offensive attitude, aggressive attitude)
  - The entity towards which the hate is expressed (target)
  - The presence of figurative use of language: irony

Benchmark corpora & Multilingual Perspective



- Very high number of resources and benchmark corpora for many different languages developed in a very narrow time span

- This confirms the growing interest of the community around abusive language in social media (and HS detection in particular)

Name	Task	Sub-Tasks	Focus	Language	Size	Teams
SemEval 2019 task 5: HatEval [Basile *et al.*, 2019]	HS	✓	misogyny, racism	EN, ES	19,600	74
AMI at IberEval 2018 [Fersini *et al.*, 2018a]	HS	✓	misogyny	EN, ES	8,115	11
AMI at EVALITA 2018 [Fersini *et al.*, 2018b]	HS	✓	misogyny	EN, IT	10,000	16
HaSpeeDe at EVALITA 2018 [Bosco *et al.*, 2018]	HS	✓	racism, generic	IT	8,000	9
MEX-A3T at IberEval 2018 [Álvarez-Carmona *et al.*, 2018]	AG	-	-	ES	11,000	7
TRAC-1 [Kumar *et al.*, 2018]	AG	-	-	EN, HI	15,000	30
GermEval 2018 task 2 [Wiegand *et al.*, 2018b]	OF	✓	-	DE	8,541	20
SemEval 2019 task 6: OffensEval [Zampieri *et al.*, 2019]	OF	✓	-	EN	14,100	115

Table 2: Shared Tasks on Hate Speech detection (HS), aggressiveness (AG) and offensiveness (OF) identification as main task with presence of sub-tasks, specific focuses, size of datasets and number of participating teams.

Fabio Poletto, Valerio Basile, Manuela Sanguinetti, Cristina Bosco, Viviana Patti: Resources and benchmark corpora for hate speech detection. systematic review. Lang. Resour. Evaluation 55(2): 477-523 (2021)

Example
Hate Speech Corpus

- twita (http://twita.di.unito.it/) is a collection of texts from Twitter in Italian language that is continuously going on since 2012
- **Hate target: immigrants**
- Smaller datasets extracted from the main collection TWITA and filtered according to set of carefully selected keywords representing hate speech against migrants
    - An annotation scheme was designed for making explicit the main features of hate speech: stereotypes, aggressive attitude…
- HS as a complex and multi-layered concept
    - Multilayered annotation scheme
- Teams of annotators for applying the annotation on the datasets
- Guidelines: https://github.com/msang/hate-speech-corpus
- Crowdsourcing experiments for enlarging the datasets and collecting opinions of several people about what hate speech is

Manuela Sanguinetti, Fabio Poletto, Cristina Bosco, Viviana Patti, Marco Stranisci. An Italian Twitter Corpus of Hate Speech against Immigrants. *In Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan, May 7-12, 2018.*

Annotation Scheme

Hate speech

Yes: any expression showing both features outlined in our operational definition (action + target).

La prossima resistenza la dovremmo fare subito contro gli invasori islamici!
We should start fighting Islamic invaders right now!

Aggressiveness – focus on the speaker's intention

Weak: any expression that implies or legitimate discriminating attitudes, refers or hints to the target group as a potential threat, or claims that it enjoys some privileged treatment.

Nuova invasione di migranti in Europa.
New migrants invasion in Europe.

Strong: any expression that refers – implicitly or explicitly – to violent actions of any kind.

Cacciamo i rom dall'Italia!
Let's kick Roma people out of Italy!

Offensiveness – focus on the hurtful effect

Weak: any expression that portraits the target group with negative or unpleasant features.

Italiani sfrattati e immigrati viziati.
Italians [are] evicted and immigrants [are] spoiled.

Strong: any outrageous, degrading or overtly insulting expression addressed to the target group.

Zingari di merda!!!
Fucking gypsies!!!

Irony

Yes: broad term including nuances such as humour, sarcasm, satire.

Ora tutti questi falsi profughi li mandiamo a casa di Renzi??!
Now are we going to send all these fake refugees to Renzi's house??!

Stereotype

Yes: any implicit or explicit generic attribution of negative features to a whole target group, based on the alleged feature of some of its members.

Roma è in bancarotta ma regala 12 milioni ai rom.
Rome is out of money but gives away 12 millions [€] to Roma [people]

Intensity of hate speech

1: implicit incitement - attributes negative qualities to the target

I migranti sanno solo ostentare l'ozio. La gente è stufa.
Migrants can only show off their laziness. People are fed up.

2: implicit incitement - uses dehumanizing or discriminatory language

La polizia i controllori fermano solo italiani rom e immigrati non li avvicina nemmeno rischiano la vita.
Policemen and conductors only inspect Italians they don't even get close to Roma or immigrants they risk their lives

3: explicit incitement - generally justifies or promotes hatred or violence

Quella schifosa rom prende anche in giro. Speriamo che con i loro fuochi tossici si brucino e crepino tutti alla svelta. TOLLERANZA ZERO.
That filthy Roma woman is even mocking. I hope they are all burned down by their toxic fires and croak quickly. NO TOLERANCE.

4: explicit incitement - personally calls for openly violent actions

Hanno rotto il cazzo con tutti questi atti terroristi. Io sono pronto alla guerra.
They're pissing me off with all these terrorist attacks. I'm ready for war.

Semeval 2019- task 5: multilingual detection of hate against immigrants and women

Twitter data




Hate Targets

> Immigrants: especially raised by political changes occurred in the last few years. Governments and policy makers are currently trying to address the issue

> Women: hate against the female gender: long-time well-known form of discrimination

Cascade tasks

> Hateful tweets against targets: yes/no

> Finer-grained aspects of hateful tweets:

> Incitement is against specific individuals/groups of people

> Aggressive behavior

Multilingual tasks

> English & Spanish

https://competitions.codalab.org/competitions/19935

Annotation guidelines: https://github.com/msang/hateval/blob/master/annotation_guidelines.md

Semeval 2019- task 5: multilingual detection of hate against immigrants and women

Task A: Hateful? yes, no

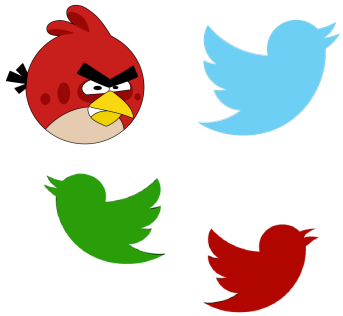
Predict whether a tweet in English or Spanish with target women or immigrants contains HS or not.

Hateful?

[hateful] @USER_345234 @USER_372644 Shut the fuck up. Your a fucking cunt. Shut the fuck up. Your a stupid cunt suck my dick

[non-hateful] Karma is a bitch.

Semeval 2019- task 5: multilingual detection of hate against immigrants and women



Task B: Is a hateful tweet also aggressive? yes, no
Is a hateful tweet against an individual or generic target? individual, generic

Target:
individual (including hateful messages purposely sent to a specific target)
generic (referring to hateful messages posted to many potential receivers)

Aggressive?

@USER_NAME Stupid ugly cunt who needs to die

Is the target individual or generic?

Women should kindly stay in the kitchen

Hate speech on line

Off line violence

Multilingual and multitarget perspective Hate against immigrants and women

- Training data

Women

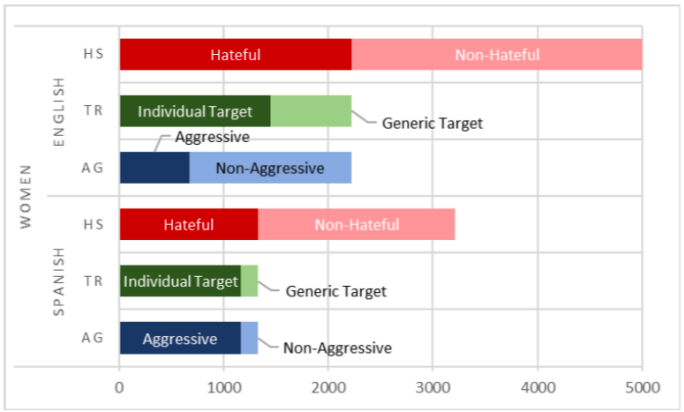


Figure 1: Distribution of the annotated categories in English and Spanish training and development set for the target women.

Immigrants

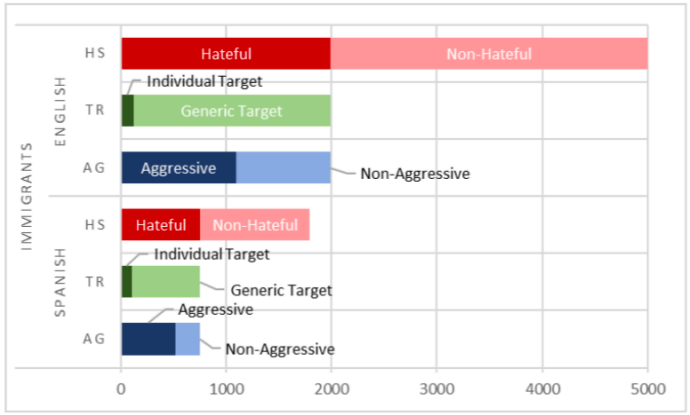


Figure 2: Distribution of the annotated categories in English and Spanish training and development set for the target immigrants.

Valerio Basile, Cristina Bosco, Elisabetta Fersini, Debora Nozza, Viviana Patti, Francisco Manuel Rangel Pardo, Paolo Rosso, Manuela Sanguinetti. 2019. SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter. SemEval@NAACL-HLT 2019: 54-63

Multilingual and multitarget perspective Hate against immigrants and women

- Women are frequent target online abuse in several countries:
- Some data-driven evidences:
    - Violation of the patriarchal norm
- In our corpora (not only Italian) we clearly find traces of the fact that often the trigger for the misogynistic attack, even online, is the so-called violation of the patriarchal norm by the single woman (Kate Manne, "Down girl")
- Online misogyny often targets a specific person and not a category.
    - Especially the woman who plays a social role traditionally considered male is the subject of online misogynistic attacks: woman in politics, the scientist, the leader of an ecological movement
- The verbal violence of the hater targets the single woman, accompanied by the desire to threaten her integrity, with a greater danger of interference with real life and a transition from verbal violence online to physical offline
- The case of hate speech towards migrants is different:
  ingroup / outgroup dynamics studied by social psychology come into play. We have seen in the data that this is reflected in a verbal expression where the category tends to be hit (the category of immigrants is inferior and constitutes a threat)
    - Hate speech and populist rhetoric (Comandini and Patti, 2019)

Gloria Comandini and Viviana Patti. 2019. An Impossible Dialogue! Nominal Utterances and Populist Rhetoric in an Italian Twitter Corpus of Hate Speech against Immigrants. In Proceedings of the Third Workshop on Abusive Language Online, pages 163–171, Florence, Italy. ACL.

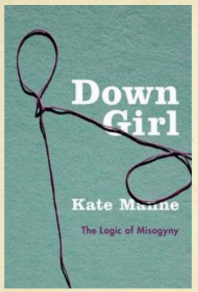
Same Target Different Topical Focus Misogyny vs sexism

- Philosophical debate on whether sexism and misogyny are distinct concepts (Manne, 2017), or whether misogynistic speech is hate speech (Richardson-Self, 2018)

- In (Pamungkas et al., 2020) cross-domain experiments provide some empirical results interesting for this debate:

  - Hate speech toward women has a stronger relation to misogyny than sexism.

  - This is in line with recent philosophical accounts theorizing misogyny and sexism as related but distinct mechanisms that enforce the norms of patriarchy (Manne, 2017), and arguing for considering misogynistic speech as a specific kind of hate speech (Richardson-Self, 2018).



Sexism and misogyny share a common purpose:

To maintain or restore a patriarchal social order

MISOGYNY (police force)	SEXISM (theory)
A property of social system The hostile reactions women face in navigating the social world. It helps us uphold sexist beliefs and enforces patriarchy It functions like a "police force", punishing women who deviate from feminine roles, policing women's subordination, enforcing male dominance It is moralistic	Patriarchal ideology, a set of beliefs, that justifies and rationalizes a patriarchal social order Think men are less suited for feminine roles. Belief women are inherently inferior It is scientific

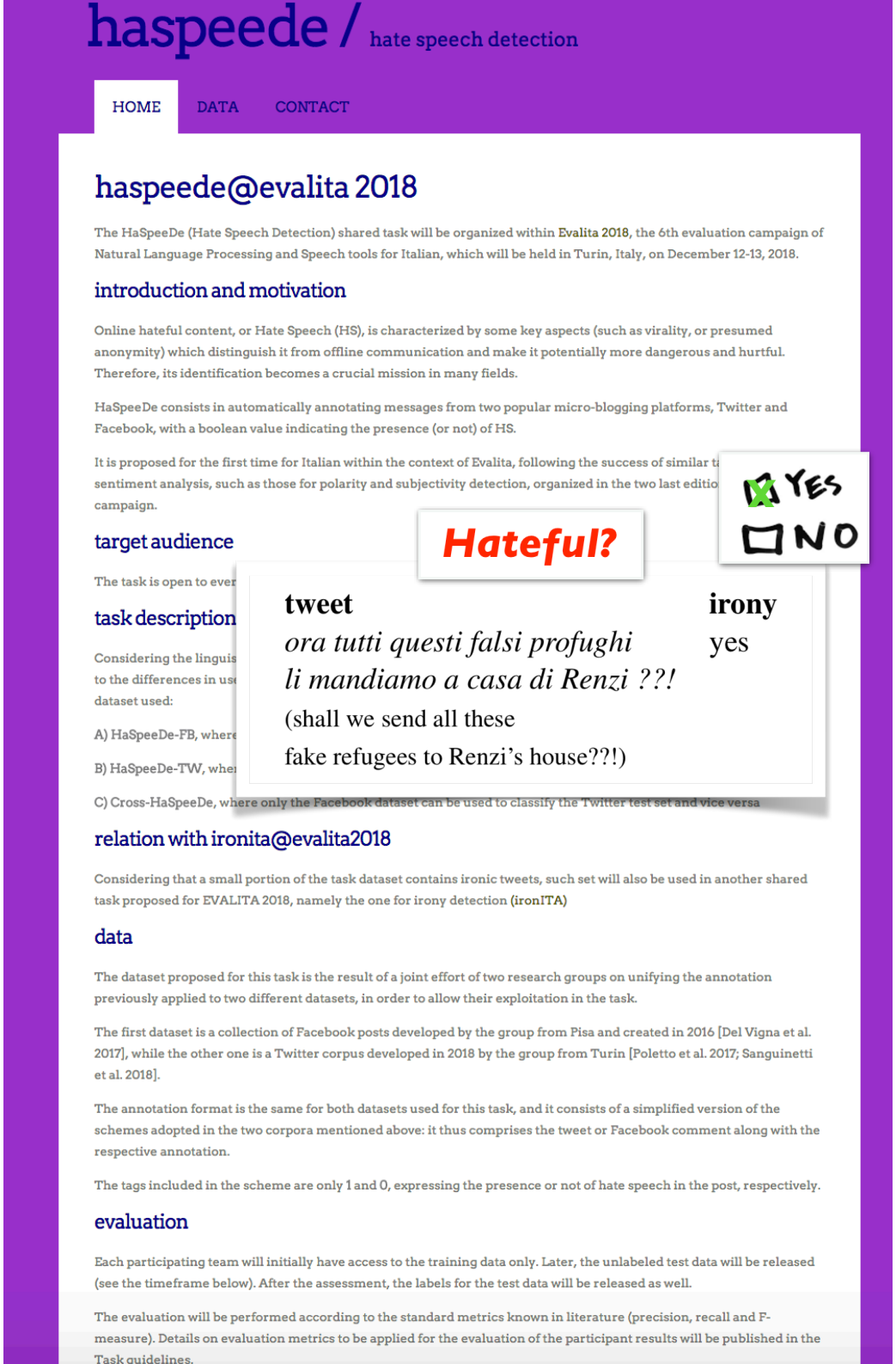
Is there a thing in the constitution about women shutting the hell up? COME ON! #shutthehellupwomen

@jackheathh I'm not sexist but women drivers are bad and when i mean bad I mean BAD

Endang Wahyu Pamungkas, Valerio Basile and Viviana Patti. 2020 Misogyny Detection in Twitter: a Multilingual and Cross-Domain Study, Information *Processing & Management, Elsevier.*

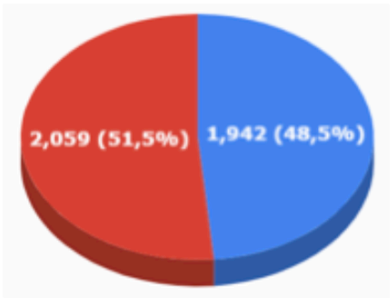
Hate Speech Detection HaSpeeDe 2018

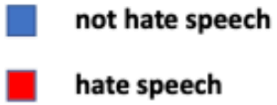
- HaSpeeDe (Hate Speech Detection) shared task at Evalita 2018
  - http://www.di.unito.it/~tutreeb/haspeede-evalita18/index.html
- Course-grained binary task: automatically annotate messages with a binary value (0 or 1) indicating the presence (or not) of hate speech
- Task main purpose: encourage and promote the participation of several research groups, making a shared dataset available in order to allow an advancement in the state of the art in this field for Italian tools as well.

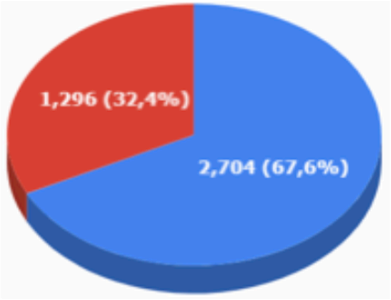


Hate Speech Detection HaSpeeDe 2018

- Two datasets released to participants:

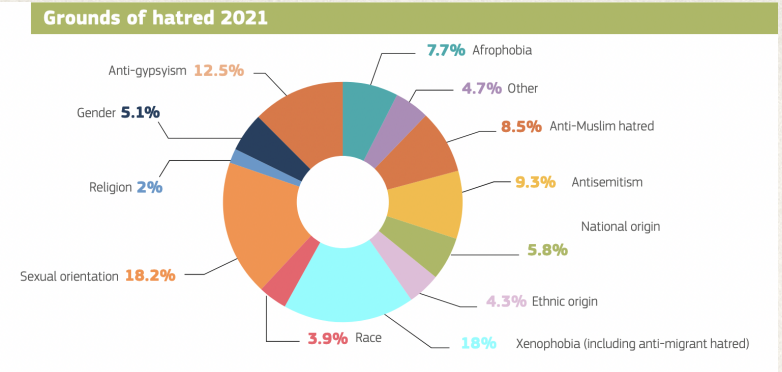






- Three main sub-tasks (depending on the dataset used):
  - HaSpeeDe-FB = Facebook Train+Test
  - HaSpeeDe-TW = Twitter Train+Test
  - Cross-HaSpeeDe:
    - a.Cross-HaSpeDe_FB = Facebook train + Twitter test
    - b.Cross-HaSpeDe_TW = Twitter train + Facebook test

New targets HODI



Home Task Description How to participate Important Dates Organizers and Contacts

About HODI

Welcome to the web page of HODI, the first
Italian at 8th evaluation campaign EVALITA

Despite advancements in LGBTQIA+ rights,
LGBTQIA+ individuals. Natural language pr
combatting online hate speech since it can
moderators' labor and mental stress. Despi
detection datasets and models, very few st
community.

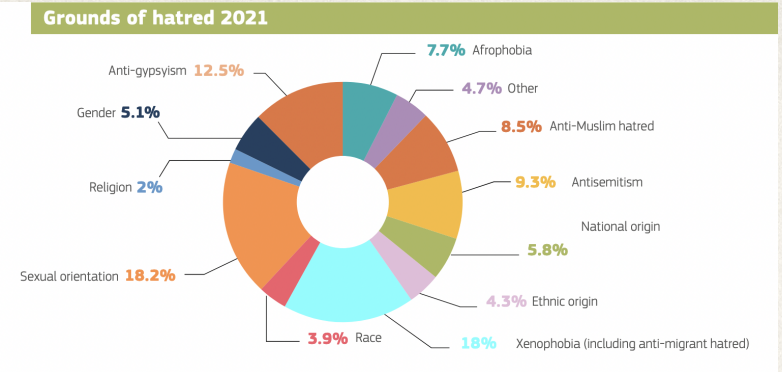
The **HODI** shared task will focus on identify

HODI is structured on two subtasks:

Subtask A
Hate speech detection: this is a bi- nary classification task, where systems are challenged to classify a
message is hateful or not against LGBTQIA+ community.

Subtask B
Explainability: once a message is classified as hateful, the objective is to identify the rationales of the classification model, i.e.,
those tokens in the sequence that contributed to the flagging of the message.

New targets



- Hate speech is commonly defined as any communication that is abusive, insulting, intimidating, harassing, and/or incites to violence, hatred, or discrimination, and it is directed against a person or a group belonging to a vulnerable category on the basis of some characteristics:
  - race, race, ethnic origin, **religion**, gender, **age, physical condition, disability,** sexual orientation, **political conviction**

haspeede3 / hate speech detection

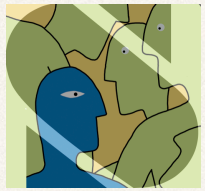
Political and Religious Hate Speech Detection

This Third Edition focuses on Hate Speech in Twitter proposing 2 tasks:

- **Task A**: Political Hate Speech Detection: a binary classification task aimed at determining whether the message contains Hate Speech or not
  - **Textual**: participants can only use the provided textual content of the tweets from PolicyCorpusXL for development.
  - **Contextual**: participants can employ for development the textual content of the tweets plus contextual information that will be given to them (i.e., metadata of the tweet and author, friends, retweets, and reply relations).
- **Task B**: Cross-domain Hate Speech Detection: a binary classification task with test data from different domains. The main objective is to explore cross-domain hate speech detection in two evaluation settings
  - **XPoliticalHate**: the test set will consist of tweets from PolicyCorpusXL (the same as in Task A), but se any kind of external data from other hate domains
  - e task here will be recognizing religious hate, therefore the test set will consist of tweets from the ReligiousHate corpus

Hidden bias? New Challenges!
From hate speech ,..
to microaggressions, stereotypes, unconscious bias

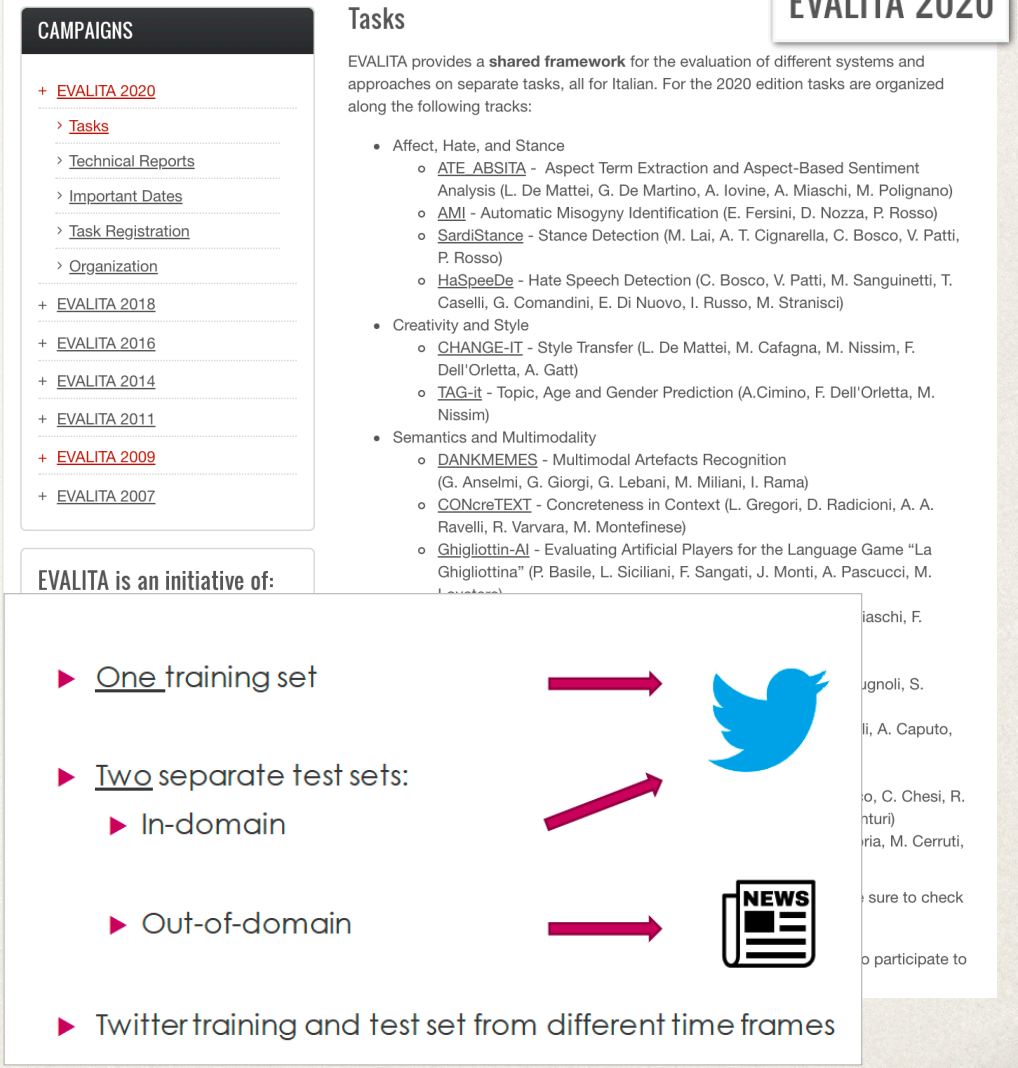




- Implicit hate
- Stereotypes behind hate
- Among the notions investigated for their relationships with toxic speech and related phenomena, stereotype is raising particular interest: it can play a leading role in novel detection tools and resource
- Stereotype has been studied in **social psychology**:
  inherently complex phenomenon that we can observe in a fine-grained way
  - We can distinguish stereotype from the evaluative and affective dimension of prejudice and from the specific forms of discredit it assumes
  - STudying European Racial Hoaxes and sterEOTYPES

HaSpeeDe 2 @ Evalita 2020
Stereotype Detection

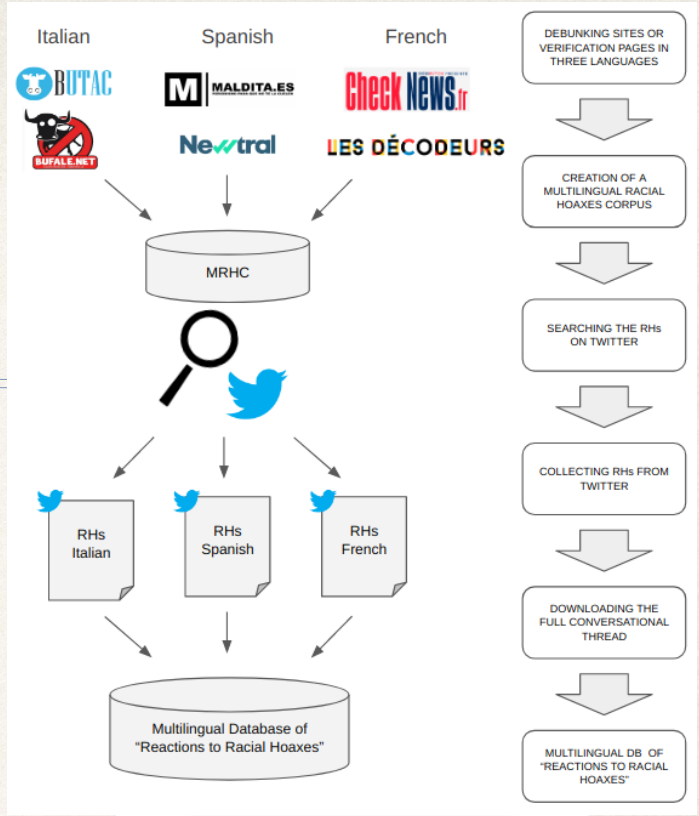
TASK A – MAIN	TASK B – PILOT 1	TASK C – PILOT 2
HATE SPEECH DETECTION	**STEREOTYPE DETECTION**	**IDENTIFICATION OF NOMINAL UTTERANCES**
Binary classification task: determine the presence	absence of hateful content towards a given target	***Binary classification task***: determine the presence	absence of stereotypes towards a given target	***Sequence labeling task***: identify Nominal Utterances in hateful data



Observation: in social media and newspapers' headlines, the most hateful parts are often verbless sentences or a verbless fragments, also known as Nominal Utterances (NUs)newspapers' headlines, the most hateful parts are often **verbless sentences or a verbless fragments**, also known as Nominal Utterances (NUs)
-> Walter Daelemans's Keynote

Stereotypes and hoaxes Conversational context collection/annotation

- https://www.irit.fr/sterheotypes/



Introduction

STERHEOTYPES is a three year (2021–2024) funded International project under the "Challenge for Europe" call by Carlsberg Foundation, Volkswagen Stiftung and Compagnia San Paolo, led by University of Bari 'Aldo Moro'.

The project focuses on '**racial hoaxes**', communicative acts created to circulate allegations of threats posed against someone's health or safety by individuals or groups, characterized by their race, ethnicity or religion. Aiming at understanding the social and **psychological processes** emerging from racial hoaxes in digital generations across three border Mediterranean European countries (**Italy, Spain and France**), it integrates different methodological approaches coming from psychology and **computational linguistics**.

Examples

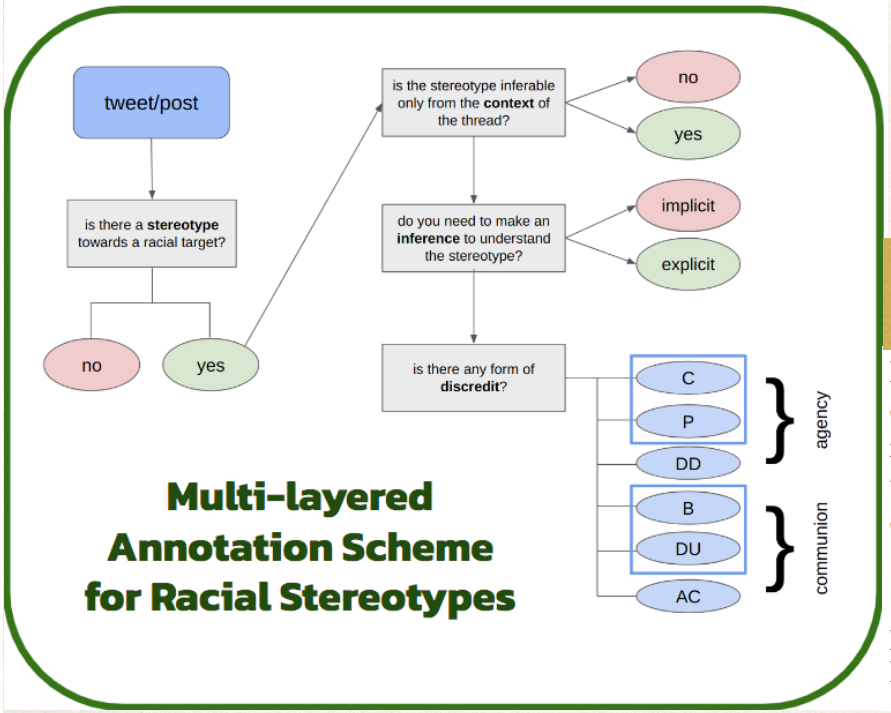
1. Immigrants out of control: they flee and injure an officer (**Security**)
2. Migrant with Covid repatriated. And now 100 agents are in quarantine (**Public Health**)
3. The electoral roll increases because the Government nationalizes 200,000 "illegals" (**Migration Control**)
4. A foreign minor, 4,700C per month, your grandmother, 426C pension per month (**Benefits**)
5. In Aubervilliers, the sheep ready to be slaughtered for #Eid on their way to the butcher. Mind boggling! #Ramadam (**Religion**)

Tom Bourgeade, Alessandra Teresa Cignarella, Simona Frenda, Mario Laurent, Wolfgang Schmeisser-Nieto, Farah Benamara, Cristina Bosco, Véronique Moriceau, Viviana Patti, Mariona Taulé: A Multilingual Dataset of Racial Stereotypes in Social Media Conversational Threads. EACL (Findings) 2023: 674-684

Stereotypes and hoaxes Conversational context collection/annotation

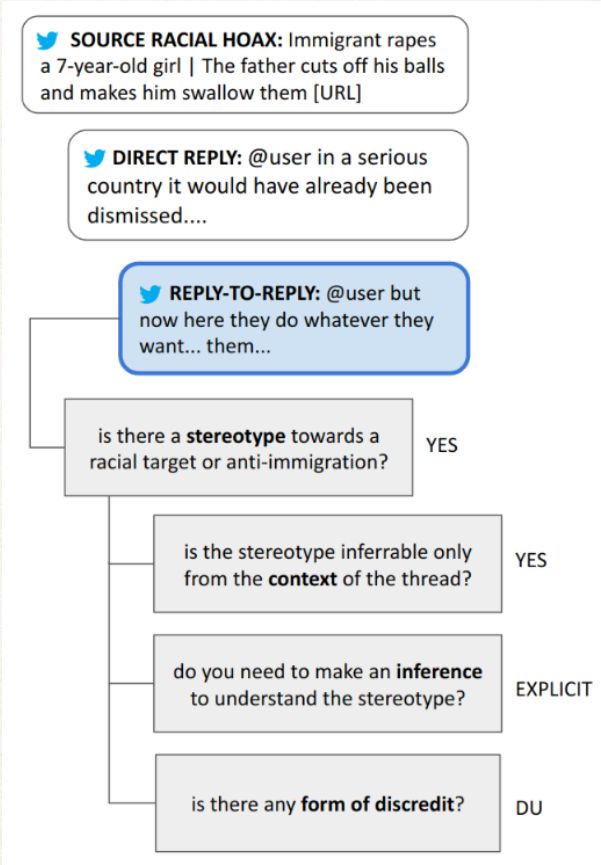
Conversational context

- https://www.irit.fr/sterheotypes/



Stereotype Content Model proposed by Fiske (1998)

Forms of Discredit. It encodes the precise form in which the text spreads a racial or anti-migration stereotype, attributing a type of behavior to the discriminated target. The values that can be applied are six: Affective Competence (AC), Attack to Benevolence (B), Competence (C), Dominance Down (DD), Dominance Up (DU) and Physical (P).



Tom Bourgeade, Alessandra Teresa Cignarella, Simona Frenda, Mario Laurent, Wolfgang Schmeisser-Nieto, Farah Benamara, Cristina Bosco, Véronique Moriceau, Viviana Patti, Mariona Taulé: A Multilingual Dataset of Racial Stereotypes in Social Media Conversational Threads. EACL (Findings) 2023: 674-684

PW1: Practical Lab, Dec 4

- Recognising canceling attitudes: an annotation task

Marco STRANISCI