

Hate speech tool for monitoring, analysing and tackling Anti-Muslim hatred online REC Action Grant (REC-DISC-AG-2016-04) 24 months (01.02.2018 - 31.01.2020)



This document was funded by the European Union's Rights, Equality and Citizenship Programme (2014-2020)

# Training seminar for stakeholders

















D21

December 2019

D21 – Training seminar for stakeholders		
WP Title	WP5 - Training, dissemination and sustainability strategy	
Due Date:	31/12/2019	
Submission Date:	20/12/2019	
Responsible Partner:	CCIF	
Version:	1.0	
Status:	Final	
Author(s):	Isis Koral (CCIF), Andrea Di Nicola, Daniela Andreatta, Gabriele Baratto and Elisa Martini (UNITRENTO), Serena Bressan (FBK), Jérôme Ferret and Mario Laurent (UT1-Capitole)	
Reviewer(s):	Isis Koral (CCIF), Andrea Di Nicola, Daniela Andreatta, Gabriele Baratto and Elisa Martini (UNITRENTO), Serena Bressan (FBK), Jérôme Ferret and Mario Laurent (UT1-Capitole)	
Deliverable Type:	R	
Dissemination Level:	PU	

### Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

The content of this document represents the views of the authors only and is their sole responsibility. The European Commission does not accept any responsibility for use that may be made of the information it contains.

# Table of contents

Executive summary	4
1. Introduction	5
2. Overview of the training seminar for academics and research organisations	6
Annex A - Agenda Hatemeter Training Seminar #2	10
Annex B - Presentations	11

# **Executive summary**

This document is the deliverable D21 "Training seminar for stakeholders" of project Hatemeter - Hate speech tool for monitoring, analysing and tackling Anti-Muslim hatred online, finalised within Activity 5.4 "Disseminating and communicating project information and results" of Workpackage (WP) no. 5 "Training, dissemination and sustainability strategy".

The training seminar for academics and research organisations has been held the **18<sup>th</sup> December 2019** at the **University of Toulouse**, Toulouse (France).

This deliverable provides a brief **overview of the Hatemeter project** and of the **Hatemeter Training deminar for stakeholders** (e.g., non-governmental organisations, law enforcement authorities, journalists, security companies, volunteers, students), with some **pictures** taken during the seminar. The documents contains the **agenda** of the event (Annex A) and the power point **presentations** (Annex B).

# 1. Introduction

Project "Hatemeter - Hate speech tool for monitoring, analysing and tackling anti-Muslim hatred online" aims at **systematising, augmenting and sharing knowledge of anti-Muslim hatred online,** and at increasing the efficiency and effectiveness of NGOs in **preventing and tackling Islamophobia** at the EU level, by developing and testing an **ICT tool (i.e., Hatemeter Platform)** that **automatically monitors and analyses Internet and social media data** on the phenomenon, and **produces computer-assisted responses and hints** to support **counter-narratives** and **awareness raising campaigns**.

More specifically, backed by a strong interdisciplinary effort (criminology, social sciences, computer sciences, statistics, and law), the Hatemeter Platform uses a combination of **natural language processing (NLP), machine learning, and big data analytics/visualization** to:

- A. identify and systematise in real-time actual "red flags" of anti-Muslim hate speech and/or possible related threats online (**Real-time Identification**);
- B. understand and assess the sets of features and patterns associated with trends of Islamophobia online (**In-depth Understanding**);
- C. develop an effective tactical/strategic planning against anti-Muslim hatred online through the adoption of the innovative Computer Assisted Persuasion (CAP) approach (**Tactical/Strategic Response**);
- D. produce an effective counter-narrative framework for preventing and tackling Islamophobia online and building knowledge-based and tailored awareness raising campaigns (**Counter-Narratives Production**).

The Hatemeter Platform has been **piloted and tested in three NGOs of EU Member States (MSs)** where the magnitude of the problem is considerable but no systematic responses have been implemented (**France, Italy and the United Kingdom**), thus enabling Project Hatemeter to address several objectives of the **Annual Colloquium on Fundamental Rights "Tolerance and respect: preventing and combating anti-Semitic and anti-Muslim hatred in Europe"** and the **European Agenda on Security** (2015), as well as the **priorities** of the REC call of proposals.

In order to strengthen **cooperation** between key actors and to ensure the **widest circulation** and **long term impact** of the project results on future research streams and operational strategies, the project favoured **capacity building and training** and the **sustainability and transferability** of the Hatemeter Platform among **other target stakeholder groups** (e.g., LEAs, journalists/media, etc.) across the EU and for **other forms of hate speech**, through the building of the **"EU laboratory on Internet and social media for countering online anti-Muslim hate speech"** (i.e., Hatemeter Lab).

# 2. Overview of the training seminar for academics and research organisations

The **Hatemeter Training seminar for stakeholders** has been held the 18th December 2019 at the "Salle des séminaires - Bibliothèque de l'Arsenal" Floor -1 of the Library of the Université Toulouse 1 Capitole 11 Rue des Puits Creuses, 31070 - Toulouse Cedex 07 (France).

The responsible partner of the event is **Collectif contre l'islamophobie en France** (CCIF) that organised the event together with the **University of Trento**, **Fondazione Bruno Kessler** and the **Université Toulouse 1 Capitole**.

The one-day training seminar has seen the participation of representatives of non-governmental organisations (NGOs) and civil society organisations (CSOs), anti-discrimination experts/professionals, civil servants, Muslim community leaders, journalists/media, law enforcement authorities (LEA), and security companies. The aim of the seminar was to supply detailed information on all the phases of the project, from the assessment of needs to reviewed state-of-the-art and the new knowledge and standards developed, to provide knowledge, competences and instruments (procedures, standards, guidelines), in order lay the ground for fully implementing an innovative approach to prevent and tackle Anti-Muslim hatred online (see also D19 - Training Module B; Hatemeter Lab meeting).

Fig. 1 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: Giulia Pirozzi (Amnesty International Italy).





Fig. 2 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: Isis Koral (CCIF).

Fig. 3 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: Bill Howe (Stop Hate UK).



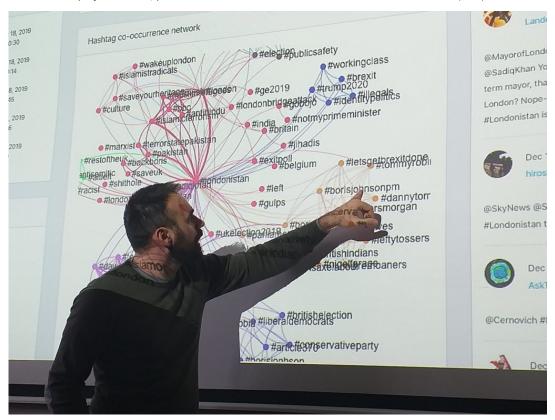


Fig. 4 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: Marco Guerini (FBK).

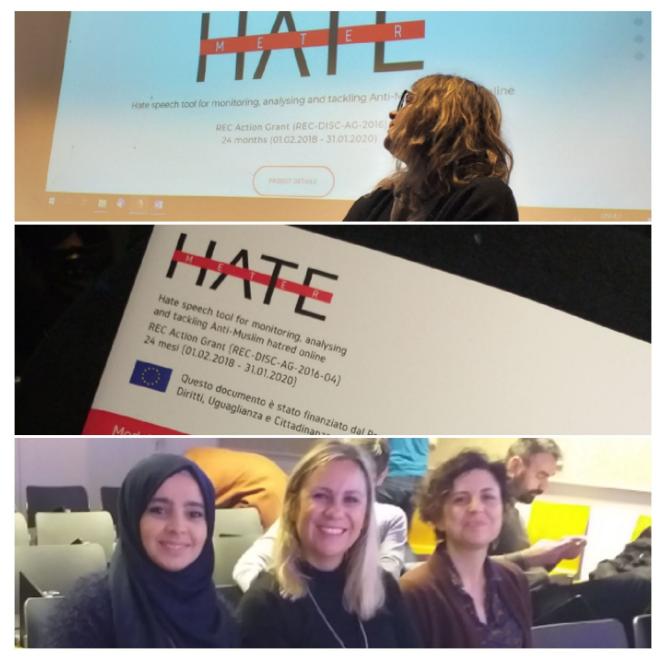
Fig. 5 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: Daniela Andreatta (University of Trento).



8

The participants at the seminar received the deliverable "**D19** – **Training Module B**" (published in English, French and Italian), designed to be used by stakeholders outside the Hatemeter project (i.e., NGO/CSO representatives; civil servants; Muslim community leaders; journalists/media; LEAs) as a **manual** to understand the Hatemeter Platform's (also referred to as the Hatemeter Tool) main **goals**, how it **works** and what it can **achieve**.

Fig. 6 – Overview of the project results, pilots and roundtable of the Hatemeter Lab: final pictures of the participants and deliverable.



The **feedback collected** during the seminar will be fundamental to assess the sustainability and dissemination of the results of the Hatemeter project, as well as to understand how the methodology and the Platform can been employed by experts on the topic in further works and research.

# Annex A

# Agenda Hatemeter Training Seminar #2

# 18 December 2019

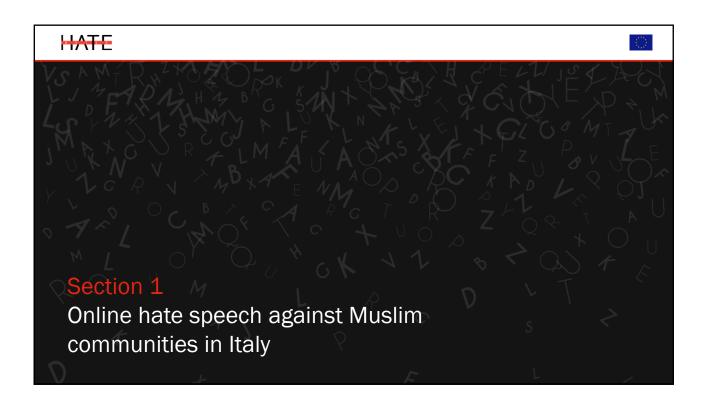
# Professionals

09.30 - 09.45	Registration of the participants
09.45 - 10.00	Welcome and presentation of the participants at the seminar (Université Toulouse 1 Capitole)
10.00 - 11.00	<ul> <li>Introduction of Hatemeter project: aim, methodology, expected results (University of Trento and Fondazione Bruno Kessler)</li> <li>1. The Italian Pilot: What is known about Islamophobia online // Effectiveness of the existing tools and techniques to fight against Islamophobia online (University of Trento)</li> <li>2. The French Pilot: What is known about Islamophobia online // Effectiveness of the existing tools and techniques to fight against Islamophobia (Université Toulouse 1 Capitole)</li> <li>3. The UK Pilot: What is known about Islamophobia online // Effectiveness of the existing tools and techniques to fight against Islamophobia (Teesside University)</li> </ul>
11.00 - 11.15	Coffee break
11.15 - 11.45	Presentation of the Hatemeter platform (Fondazione Bruno Kessler)
11.45 - 12.30	Presentation of the three Hatemeter pilots by NGOs representatives in Italy, France, and the UK (Amnesty International – Italy, Collectif Contre I'Islamophobie en France, and Stop Hate UK)
12.30 - 13.00	Conclusions on the qualitative and quantitative evaluation of the pilots (University of Trento)
13.00 - 14.00	Lunch
14.00 - 14.45	Work groups on the functioning of the Hatemeter platform and assessment of its efficacy and efficiency
14.45 - 15.30	Presentation of the results, suggestions, insights, recommendations by the work groups
15.30 - 16.30	Roundtable discussion on project Hatemeter: the efficacy and efficiency of the Hatemeter platform for NGOs, LEAs, policy makers, Muslim community leaders, ICT companies
16.30 - 16.45	Open questions and discussion
16.45 - 17.00	Seminar review and closing

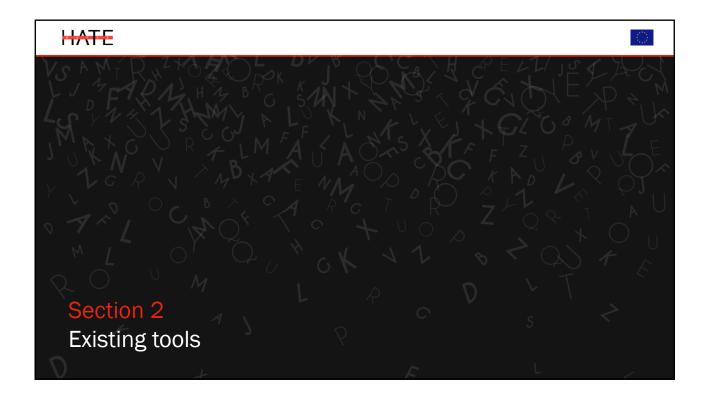
# Annex B

# Presentations





# HATE Online hate speech against Muslim communities in Italy State of the art Fake news and inflammatory statements against Muslims are spread on the Internet and through social media platforms. A recent study on Islamophobia in Italy found an increase in hate speech against Islam from neofascist and Catholic fundamentalist groups (Aglietti and Padovan, 2018). Muslims are the fourth most targeted group on Twitter (Vox, 2019). Hate speech is soaring and with xenophobia, Islamophobia, anti-Semitism and racism, particularly as a result of both terrorist attacks and the migration flows. Anti-Muslim hatred is not confined to the virtual world, since the Internet and social media promote religious hatred and xenophobia that lead to crimes in the real world. They also favour a rise in political extremism: far-right groups and radicalisation • of (young) Muslim in response to messages of exclusion. з

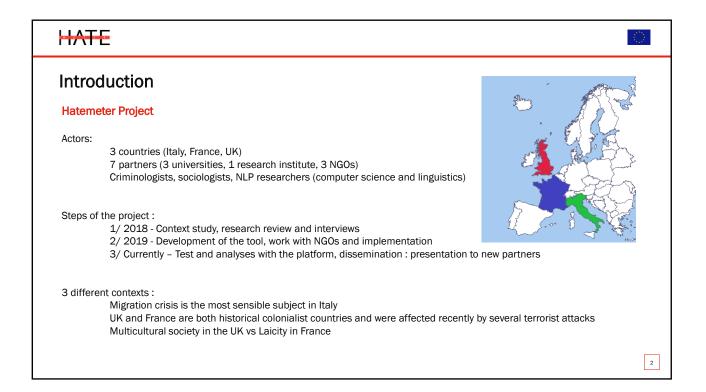


Is and the legal, political		Hate speech and/or Islamophobia
nd social techniques gainst hate speech nd/or Islamophobia in Italy	Monitoring tools	<ul> <li>Vox (2018). La mappa dell'intolleranza</li> <li>Amnesty International Italia (2018). Conta fino a 10. Barometro dell'odio in campagna elettorale. Roma: Amnesty International Italia</li> <li>Project React</li> <li>The European Observatory for Illegal Hate Speech Online and Fake News</li> <li>Hatebase</li> </ul>
	Educational tools	Project PRISM     Project 'PROXI'     Project 'BRICkS'     Project 'Media against Hate'     Project 'Words are Stones'     Project 'Silence Hate'
	Legal, political and social techniques	<ul> <li>Pugiotto. (2012). Le parole sono pietre? I discorsi di odio e la libertà di espressione nel diritto costituzionale</li> <li>Gometz (2017). L'odio proibito: la repressione giuridica dello hate speech</li> <li>Giovannetti and Minicucci (2015). L'hate speech nei new social media</li> <li>eMORE (n.d.). Monitoring and Reporting Online Hate Speech in Europe</li> <li>Italian legislation (e.g. Mancino Law, Legislative Decrees n. 215 and 216 of 9 July 2003; Article 33 of the Constitution; etc.)</li> <li>European Union legislation and directives (e.g. ECHR; Fundamental Rights of the EU; EU directives on Racial Equality; etc.)</li> </ul>

<ul> <li>Lack of clear legal de legal context) both at</li> <li>The legislative framew problematic when dis</li> </ul>	of religion and against online racisms and 'there have been only secondary and fragmented initiatives.' finitions of hate speech and hate crime problematise the fight against Islamophobia (which is not well delineated in the the national and at the European level. vork is challenging, as certain categories tend to be excluded from the Mancino Law of 1993. This becomes particularly criminatory actions against Muslims intersect with other factors, for example gender and sexuality. bre problematic as well as the prosecution of the 'hate preachers'.
<ul> <li>Previous research hig</li> <li>LEAs make available to Warnings to social me anonymity online is an</li> </ul>	hights that in Italy hate crime appears to be experienced less than hate speech bights that in Italy hate crime appears to be experienced less than hate speech cools and channels for online warnings but these are not very well known or used. Third Part reporting is not available. Adla supervisors are often inefficient as there are tolerant policies of self-regulation. According to the <b>postal police</b> , n illusion as investigatory activities lead to the identification of users; however, websites are often hosted on foreign therefore there are problems with both accessing data and shutting down some webpages. If hate speech
Lack of awareness raising	g campaigns

 $\langle \langle \hat{a} \rangle \rangle$ 







# Socio-legal background

### Our subject of study is mainly related or affected by 3 important pieces of law

### Law on « haine et incitation à la haine »

- Any hate speech, discrimination or violent against someone based on their ethnicity, religion, race or nationality is prohibited

- This also apply to writting discourses and to any public communication, including using electronic devices

"Ceux qui, par l'un des moyens énoncés à l'article 23, auront provoqué à la discrimination, à la haine ou à la violence à l'égard d'une personne ou d'un groupe de personnes à raison de leur origine ou de leur appartenance ou de leur non-appartenance à une ethnie, une nation, une race ou une religion déterminée, seront punis d'un an d'emprisonnement et de 45 000 euros d'amende ou de l'une de ces deux peines seulement."

"[...] soit par des discours, cris ou menaces proférés dans des lieux ou réunions publics, soit par des écrits, imprimés, dessins, gravures, peintures, emblèmes, images ou tout autre support de l'écrit, de la parole ou de l'image vendus ou distribués, mis en vente ou exposés dans des lieux ou réunions publics, soit par des placards ou des affiches exposés au regard du public, soit par tout moyen de communication au public par voie électronique [...]"

### Law on « statistiques éthniques »

- It's forbidden to gather any ethnic or religious data without special agreement

- No racial studies (can make the evaluation of discriminations harder and assumption of the far right activist stronger)

### Law on « laïcité » (laicity/securalism)

- Wearing any visible religious sign is prohibited to anyone working for the french state and in schools
- This law create debates and its interpretation by different political groups is still controversial

## **Epistemological issues**

### New laws and ambiguities

A new text of law was voted at Assemblée Nationale in July specifically addressing online hate speech. - incitation to hate, discriminate or commit violence, based on sex, sexual orientation, gender identity, disability, ethnicity, nationality, "race" or religion.

- social network has to delete it in less than 24hrs and replace it with a message indicating the removal.

During the debates at the Assembly, activists expressed some complaints on the text and asked to replace the term "Islamophobic" with "anti-Muslim".

- fear that "Islamophobia" could be use to prohibit any criticism of the religion itself.

- tension around the place of Muslims in French society.
- epistemological battles over terms qualifying the hatred of Muslims.

In 2017, the word "race" was removed from the constitution's Article 1 - viewed as a racist tool.

- question dividing researchers or citizens but the vote was unanimous.

→ racism based on biologic differences has been replaced by opposition between French culture and foreign culture, seen as invasive: "neo-racism" (Froio, 2018 ; Tahata, 2018).

HATE	
Other parameters	
Social network exchanges and media influence	
<ul> <li>Twitter mechanics also play a role in the form of hate speech.</li> <li>maximum size of tweets prevents the development of a constructed speech</li> <li>forces writers to be more concise and more divisive</li> <li>→ Polarization of the exchanges between support and anti-Muslims groups</li> <li>The effectiveness of a tweet seems linked to its controversial nature (Longhi, 2017)</li> </ul>	
False news reached more people and spread faster than the truth (Vosoughi, Roy and Aral, 2018)	
<ul> <li>« Islamophobic controversies often started from social media before getting debated in mainstream media and triggering reactions among politicians. » (Faytre, 2019)</li> <li>→ after a viral video posted on Twitter by a RN representant questionning place in society of women wearing th veil, continuous news channels in France have regularly debated this subject, during the month of October 2019 often receiving around 200,000 viewers → resulting in more online debates and hate speech</li> </ul>	ne
	<ul> <li>Social network exchanges and media influence</li> <li>Twitter mechanics also play a role in the form of hate speech.</li> <li>maximum size of tweets prevents the development of a constructed speech</li> <li>forces writers to be more concise and more divisive</li> <li>→ Polarization of the exchanges between support and anti-Muslims groups</li> <li>The effectiveness of a tweet seems linked to its controversial nature (Longhi, 2017)</li> <li>False news reached more people and spread faster than the truth (Vosoughi, Roy and Aral, 2018)</li> <li>« Islamophobic controversies often started from social media before getting debated in mainstream media and triggering reactions among politicians. » (Faytre, 2019)</li> <li>→ after a viral video posted on Twitter by a RN representant questionning place in society of women wearing th veil, continuous news channels in France have regularly debated this subject, during the month of October 201</li> </ul>

è	Sep 03 18:23:31	Q
	ouche Bien sûr qu'on a le dro hobe ! On en a même le devi	
que par	précaution. C'est une religion	n d'agression,
de haine	. #stopisiam	

5

# Section 2 Interviews and online observation

# HATE

### Interviews

### Main results from the in depth interviews

All the interviewees were Muslim women, activists in three different associations

According to them and our observation, **anti-Muslim** hatred takes **two** main forms: - **Hate speech**, based on **prejudices associated** with Muslims or people of North African descent and culture

- Cyberbullying campaigns targeting publicly known and recognised Muslim personalities or any person who have been exposed in the media as Muslim

Online harassment undertakes the following path:

- a religious sign visible in the media is associated with a willingness of its bearer to proselytise
- $\rightarrow$  proselytism for the Muslim religion is associated with extremist branches

 $\rightarrow$  these extremist branches are then associated with the jihadists  $\rightarrow$  and finally with ISIS and terrorist attacks

After being identified by a group, the **victims** can be **pursued relentlessly** by comments under their profiles, publications or by private messages.

They can receive « simple » insults to death threat.

### **Online observations**

### The three main groups producing anti-Muslim hate speech

 $\rightarrow$  Far right groups, called « Fachosphère », includes political parties, other official

groups as well as other individuals (Rassemblement National, FdeSouche, Génération Identitaire...)

Their discourse relates many societal problems to immigration and « choc des cultures » (cultural clash)

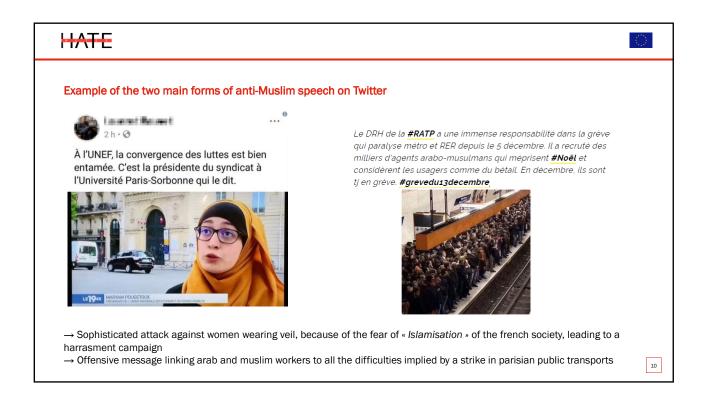
Some of them argue that they fight for the **common good**, including their target populations (Muslims), who would supposedly lives better if they leave France or if they adopt French culture.

### 

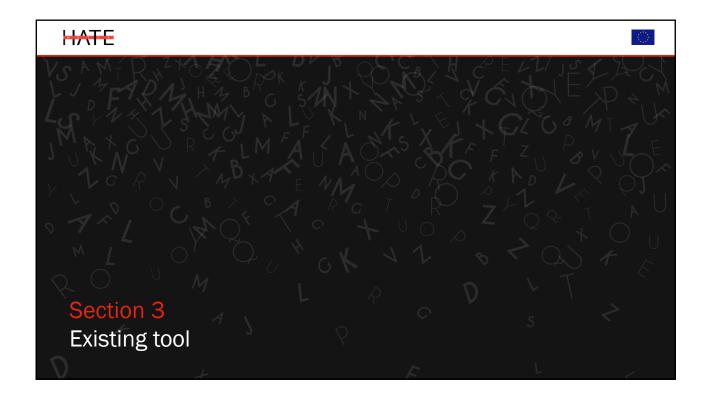
Their speeches defy the presence of Muslims in public space on the pretext that they must respect the historical right to govern and to speak in public held by Judeo-Christian white culture

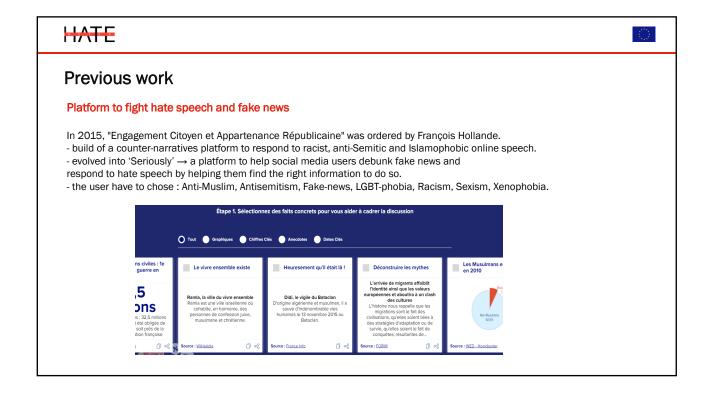
→ « Laïcistes » (secularist claimants), often politically affiliated with a left-wing party (intelectuals, journalists...) They claim an **extreme** interpretation of **secularism** such that no sign or religious claim should be tolerated in the public space (including the streets and media)

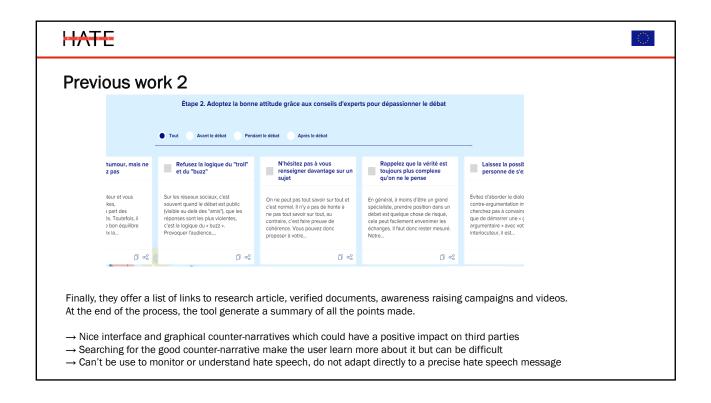
This group has the most **sophisticated** discourse in which hate speech is **subtle** when existent. They can launch **harrasment campaigns** aim to target Muslim personalities in order to **silence them.** 



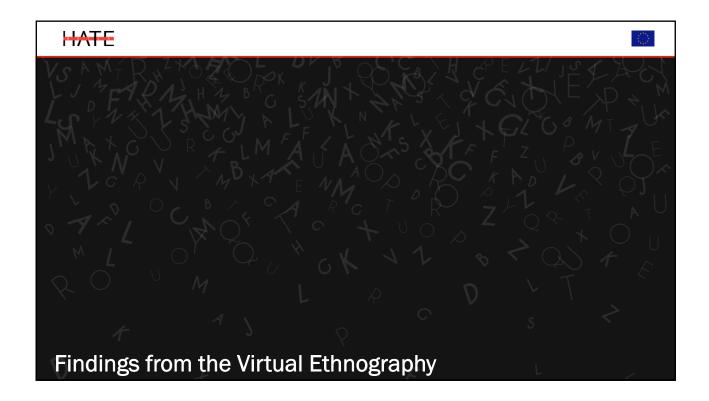
Republicains       20 ferr.         Republicains <td< th=""><th></th></td<>	
question client : bande de pourrie l argent n a pas d'odeur vous trahissez les valeurs de la république française honte a vous vous contribuez a l invasion islamiste vous finirez avec, cette racaille dans les fours en Pologne Magasin concerné : Reference et nom du produit : eivilite : Homme	

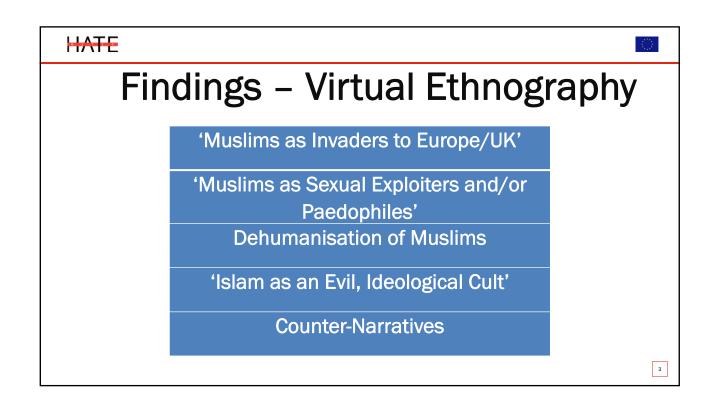


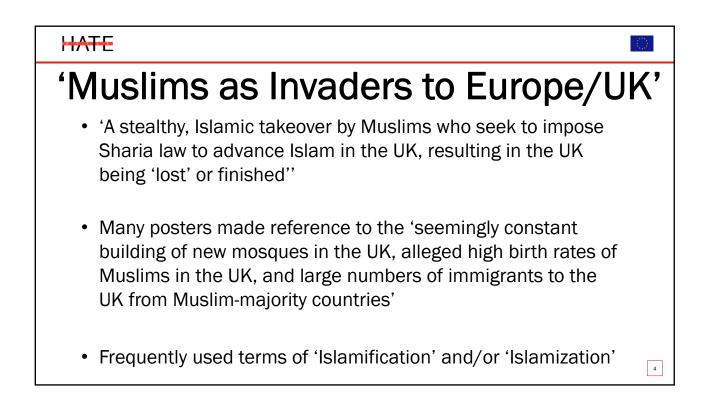


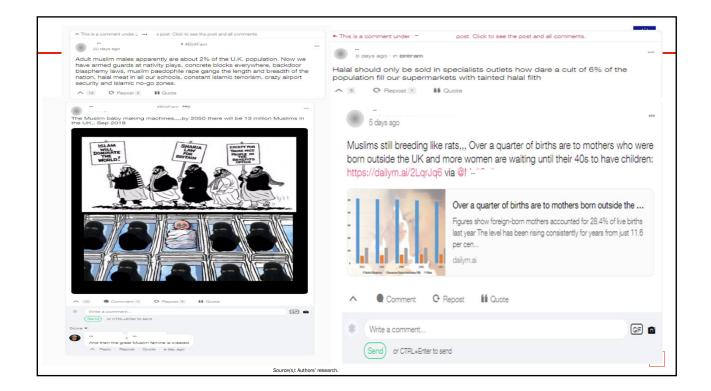


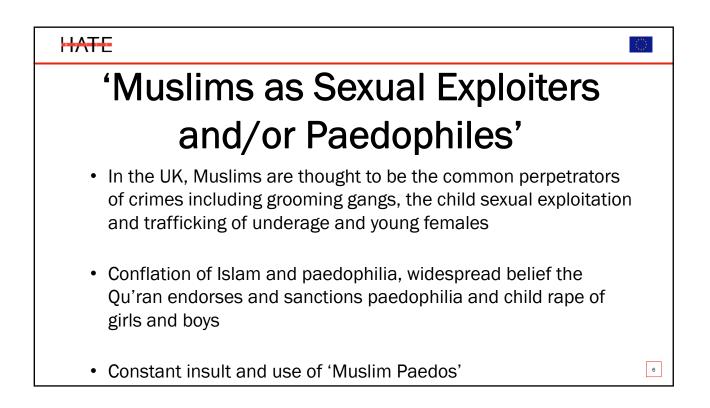




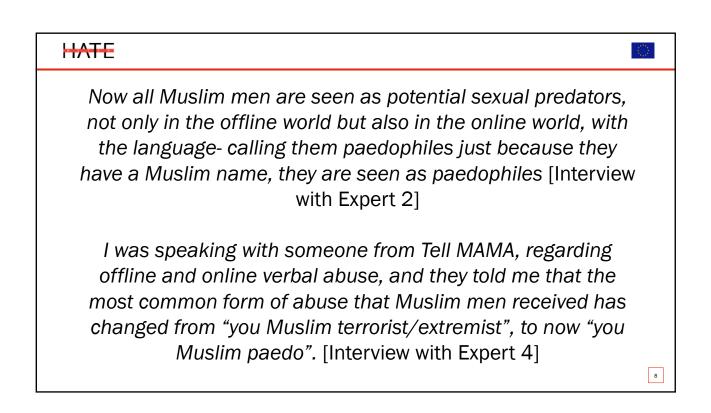






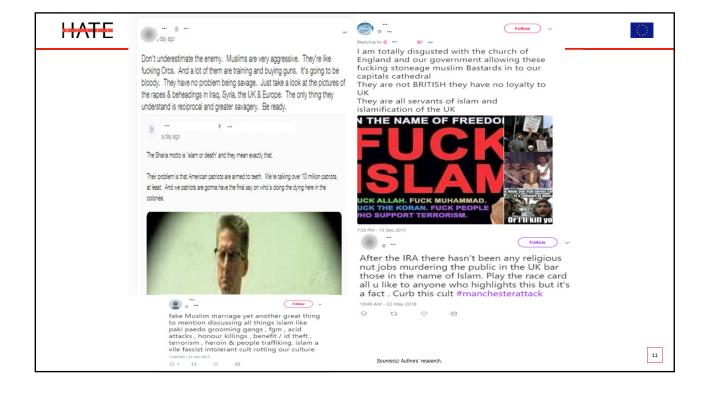




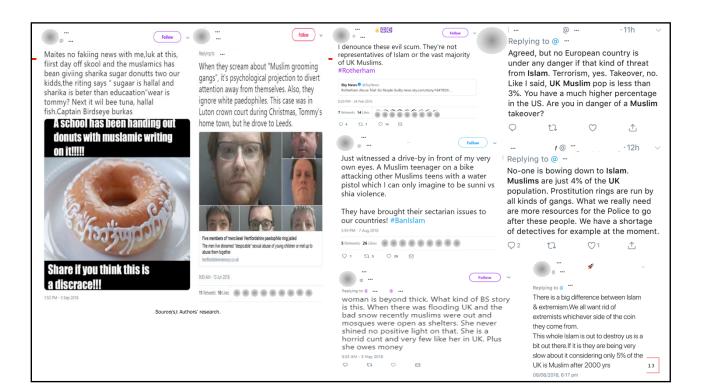


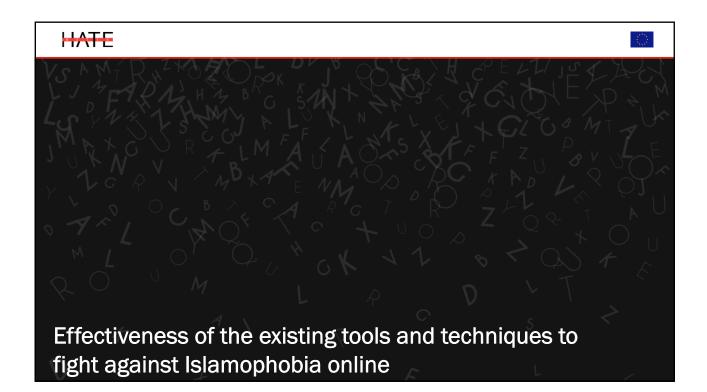


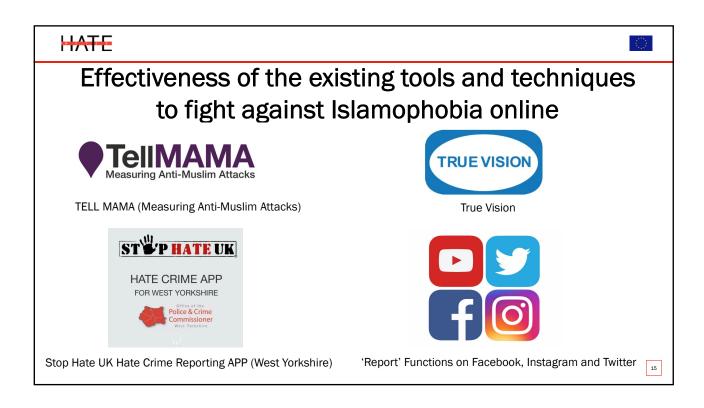
HATE	$\langle \bigcirc \rangle$
'Islam as a Cult, Not a Religion'	
<ul> <li>'Islam is not a religion, it is instead a monolithic 'death cult' that is intent on slaughtering non-Muslims in the U and wider Europe, taking over countries in the latter region'</li> </ul>	К
<ul> <li>'Muslims are mindless followers, intolerant evil, barbari prone to violence and possess paedophilic and rapist tendencies towards women and children'</li> </ul>	C,
<ul> <li>'Incompatibility of Islam and Muslims in the UK'</li> </ul>	10

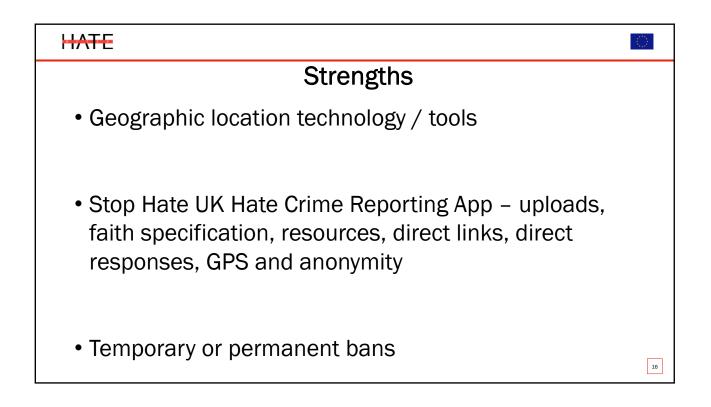


HATE	
Counter-Narratives	
<ul> <li>Instances of counter-narrative activity on various platforms dispelling negative myths and stereotypes about Islam and Muslims</li> </ul>	
<ul> <li>Mix of ordinary social media users and organisations</li> </ul>	
<ul> <li>Somewhat prevalent is the creation of parody accounts that use humour to negate and mock negative misconceptions of Islam by actively engaging with Islamophobic users in 'real-time'</li> </ul>	12









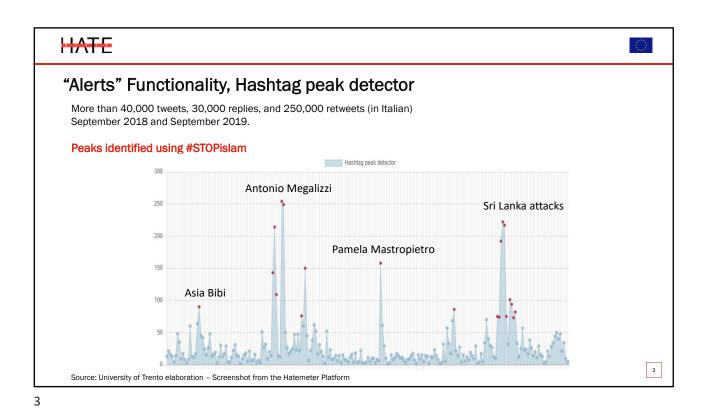
# Weaknesses

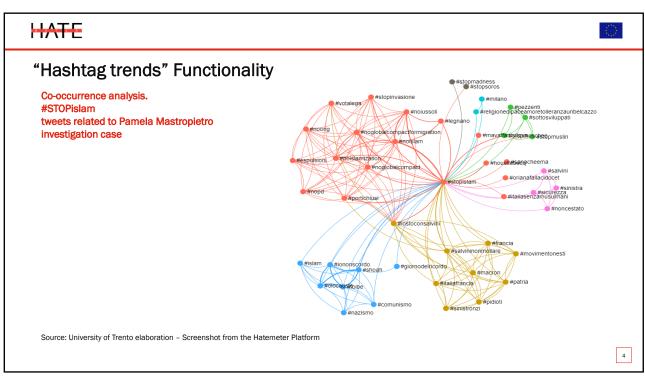
- Time-consuming
- 'Free speech' and censorship issues
- Challenges of identifying hate speech on platforms
- Creativity and ingenuity of perpetrators
- Difficulties for law enforcement

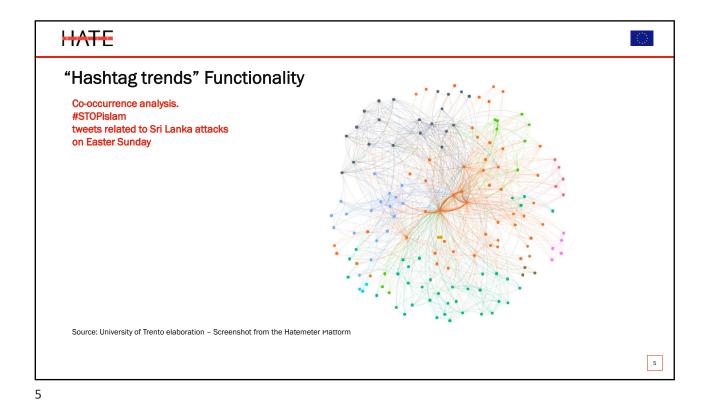
17

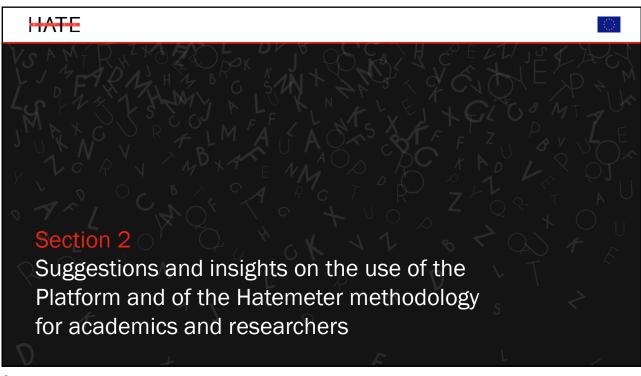


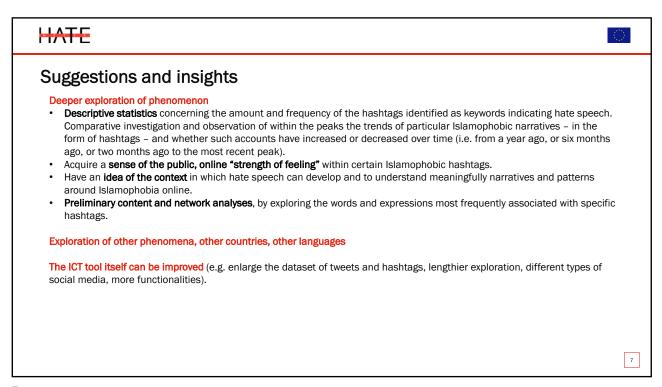








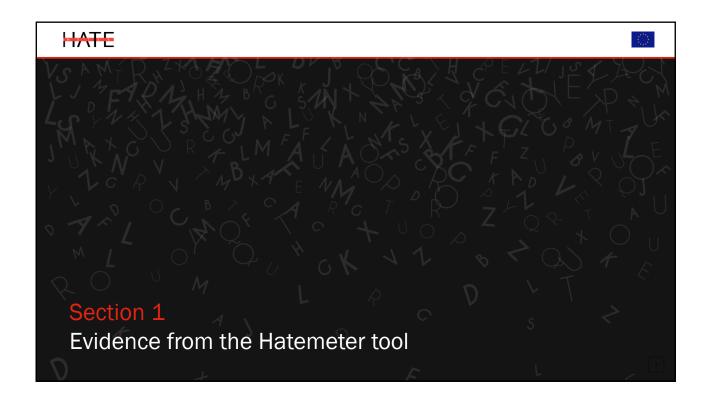












### One way to use the Hatemeter Platform

From large scale monitoring to qualitative analysis

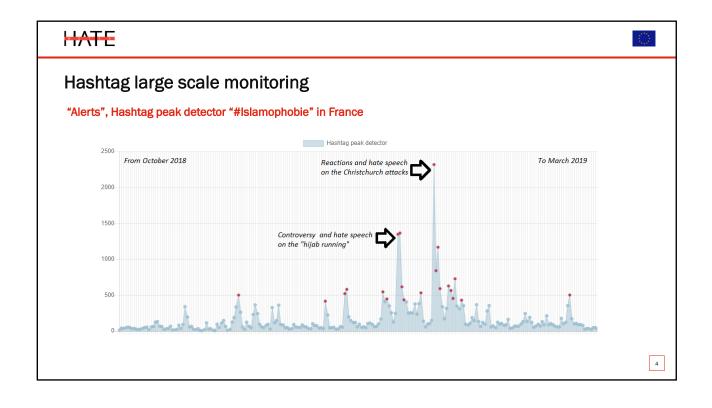
- Use the « Alerts » function on a Hashtag you want to study to detect a peak of islamophobia, and record the timespan concerned by this peak.

- Use the « Hashtag trends » function, chosing the same Hashtag and the previous time-span of interest.

 $\rightarrow$  read the most re-tweeted tweets in the list to understand which event produced the peak and how it was discussed on Twitter.

Use the « Hate Speakers » function on the **same Hashtag** and **same time-span** in order to know who are the biggest influencers.

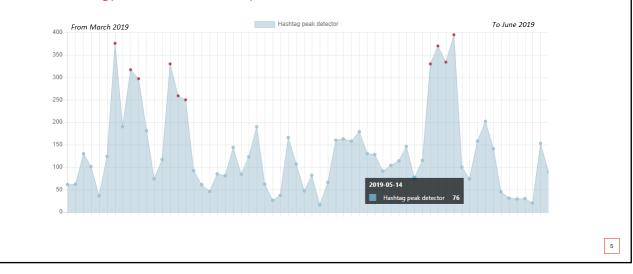
Understand how the users linked to this tweets are related to each other (Same group ? Same interest?) by looking at the hate speaker's cooccurrence graph and by clicking the name of the Hate Speaker to track back their actual Twitter account.

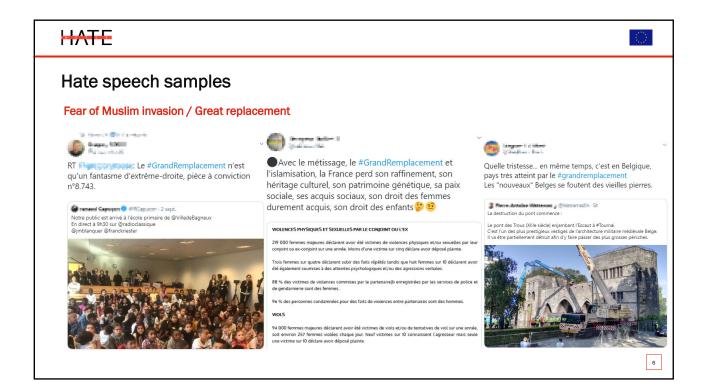


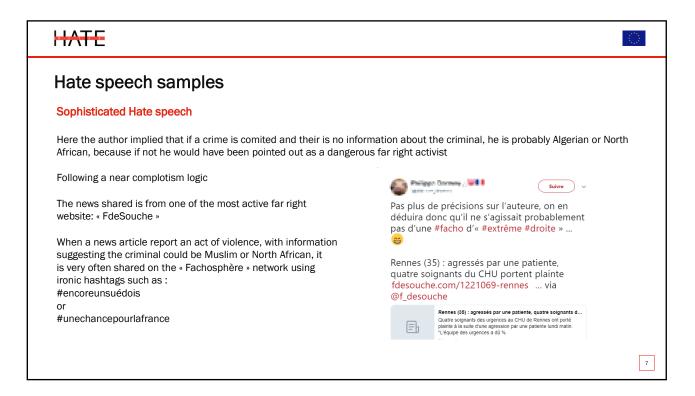
з

## Hashtag large scale monitoring

"Alerts", Hashtag peak detector "#GrandRemplacement" in France









#### Using the Hatemeter tool

#### Limits and benefits

It is difficult to capture every hate speech through the search for keywords or hashtags

- $\rightarrow$  Twitter users do not want to be legally concerned
- $\rightarrow$  They will often resort to a **subtle and sophisticated** speech

 $\rightarrow$  Their message can be claimed **politica**l. It can claim **common sense**, state a **"fact**", it can take the **form** of a **warning** against danger, or it can pretend to be a simple **informative** remark, which in fact is **hidding an** 

ideology

However, there is also a phenomenon well captured with our methodology :

- $\rightarrow$  Tweets with an **« innocent »** content followed by some hashtags
- $\rightarrow$  This tweets make **not much sense** unless we can understand what the hashtags refered to
- $\rightarrow$  Those relate to a specific **network** and gives them a political or ideological **affiliation**

Our methodology allows us to regularly **discover new hashtags** and **expand** our keyword and hashtag **samples** : "Cespaslislam", #reveilezvous, #resistantspourlafrance, #renaissancefrançaise, #remigration, etc

HATE			0
Different way to	use the tool		
Some Hashtags are most	ly used by the anti-racis	st activists	
#Islamophobie: Used to expo	ose anti-Muslim speech or	acts	
→ Usefull <b>to follow the</b> <b>news</b> on the subject → <b>Same Hashtags</b> can be <b>use ironically</b> with the opposite meaning	#IslamHorsDEurope Mar 25, 2019 17:34:09 #stopislam Mar 25, 2019 10:05:37 #Stopislamisme Mar 25, 2019 06:18:33		Après les tags naci et anti sitam retrouvés hier sur un stade de la Rochelle, ce matin c'est le chantier de la mosquée de Bergerac qui a été la cibie d'une tête de porc et du samp porcin. Un peu plus d'une semaine après «Ciritàticut» d'astampohobie
→ Victimisation	Keywords tags nazi chantier rochelle têteade porc bergerac	(Fislamophoble)	https://t.co/387/yms/NI https://t.co/X88/dX/Sin Mar 25 18/009 Mar 25 18/009
	matin		

9

#### **Counter-narratives**

#### Not efficient?

After the first phase of **counter-narratives collection** in France, we ran some **tests** with the Hatemeter tool to see if the counternarratives were **correctly assigned** to hate speech Tweets

In fact the **algorithm** was **working well** to assign a **pertinent** counter-speech to a hate speech but the **counter-speech** database **itself** was including some **inneficient** or sometimes **controversial** discourses:

- ungrammatical/bad written

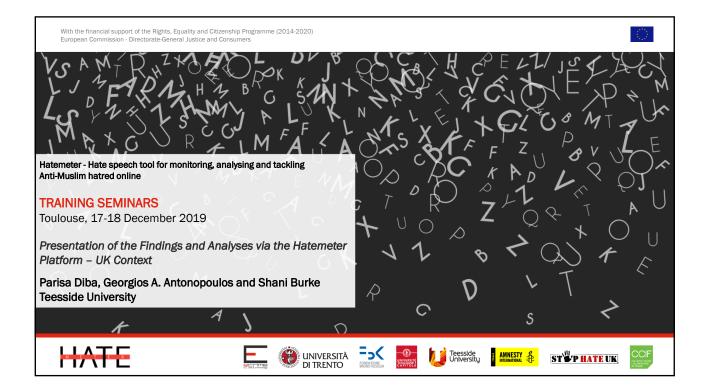
- no context/no content (messages like "not at all" or "you can't say that")

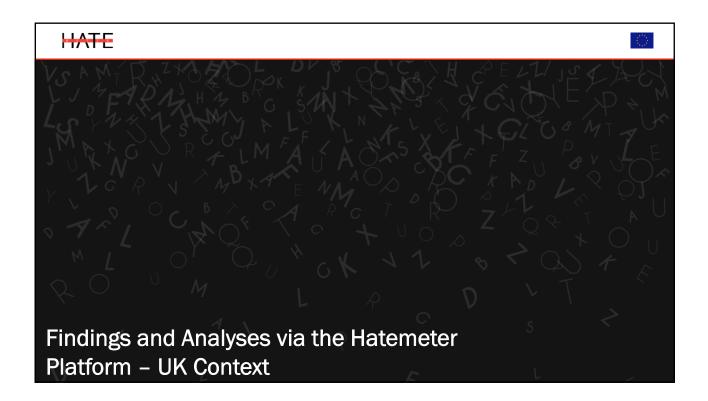
- too hostile or too emotional (sometimes insulting back in response to an insult message or comparison between religions to say than one is not better than the other)

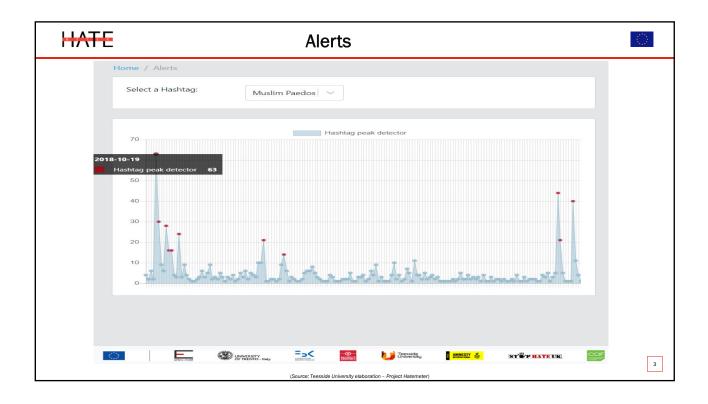
- too religious or controversial (for example: "Islam is universal" which can make haters against muslim even more angry, or "vieil is an obligation for muslim women" which seems a controversial assumption and not really a solid point to make hate speech decrease)

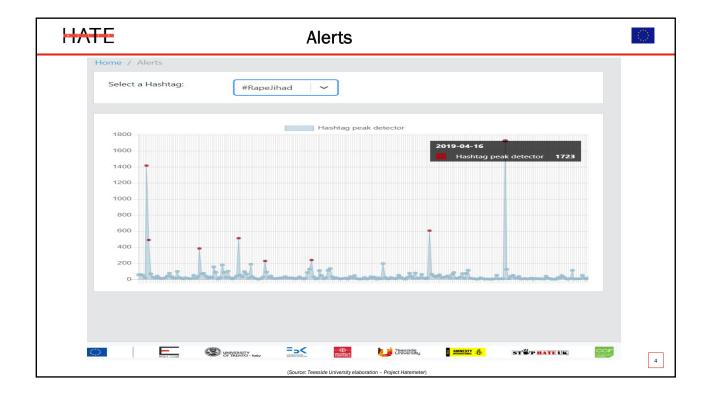
 $\rightarrow$  Operators have to be train/follow a formation to write counter-narratives Sometimes, being a victims of hate speech make it harder to respond efficiently to this hate speech

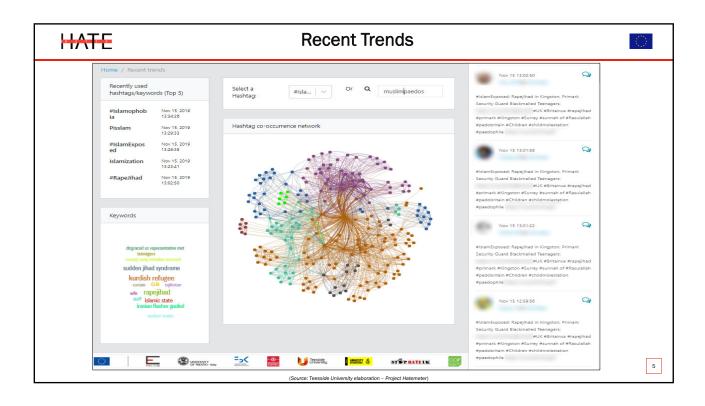
11

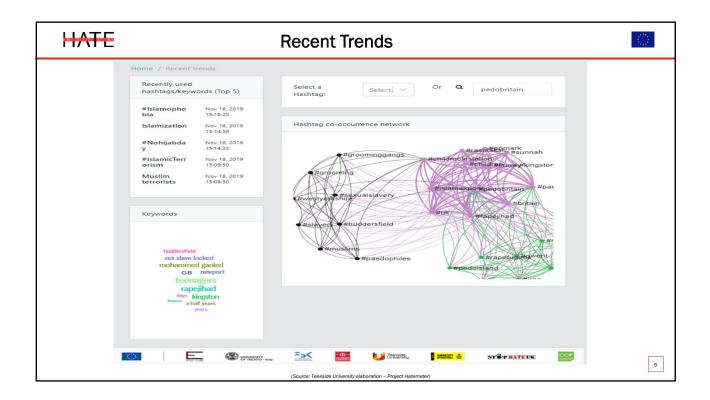


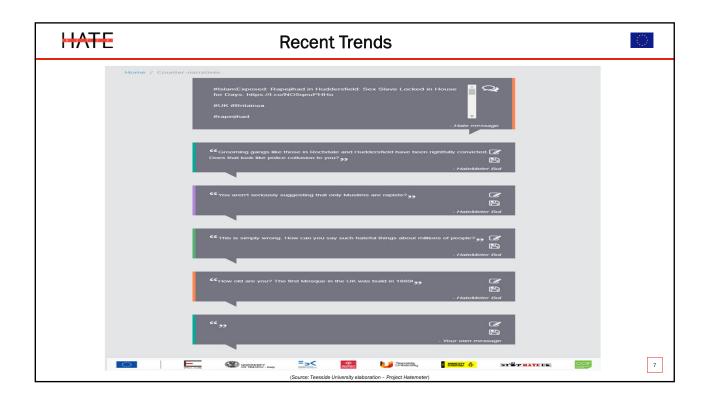


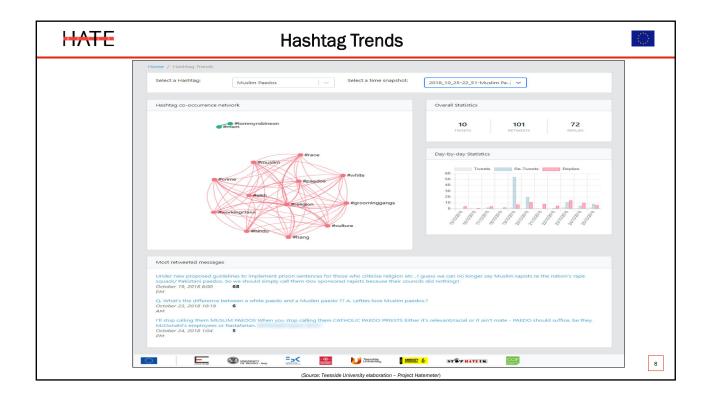


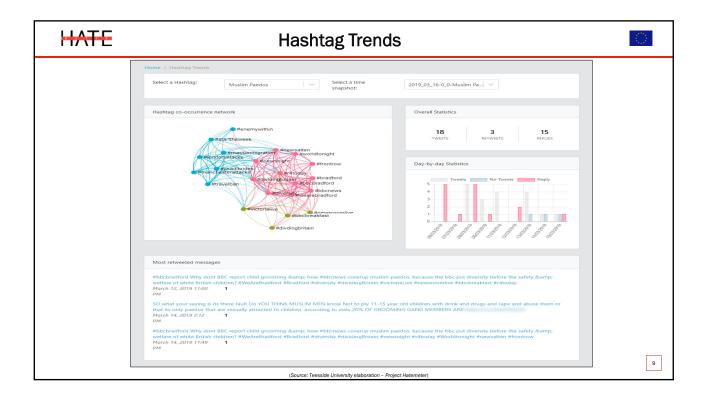


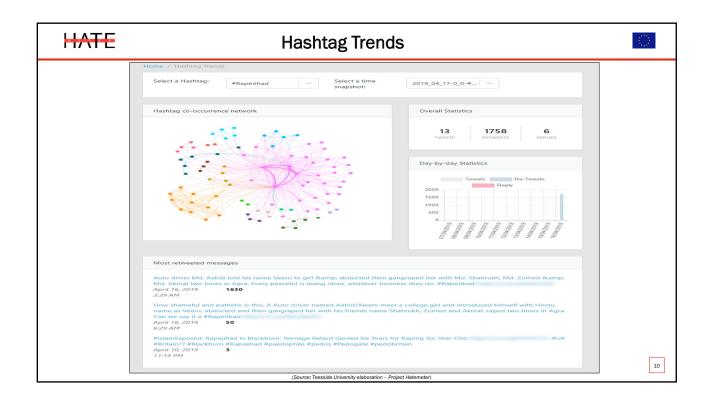


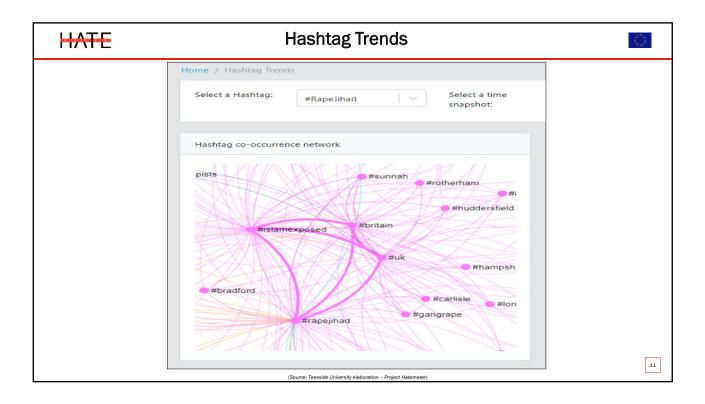


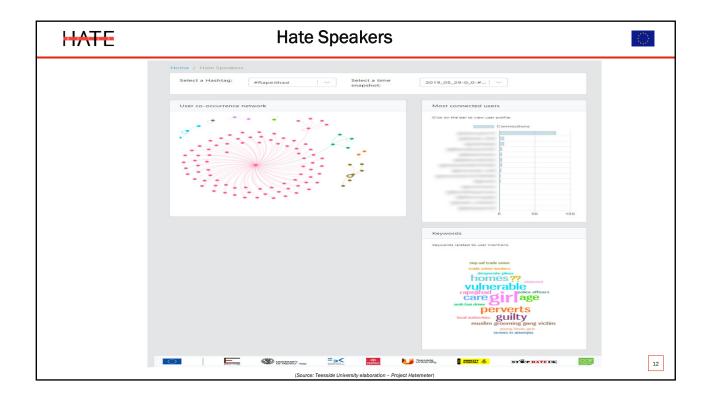


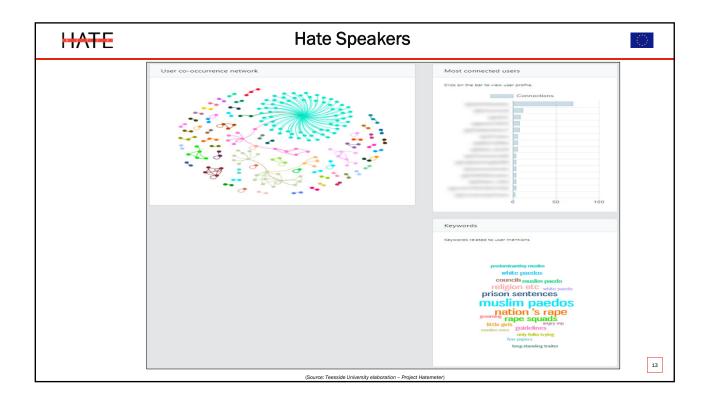


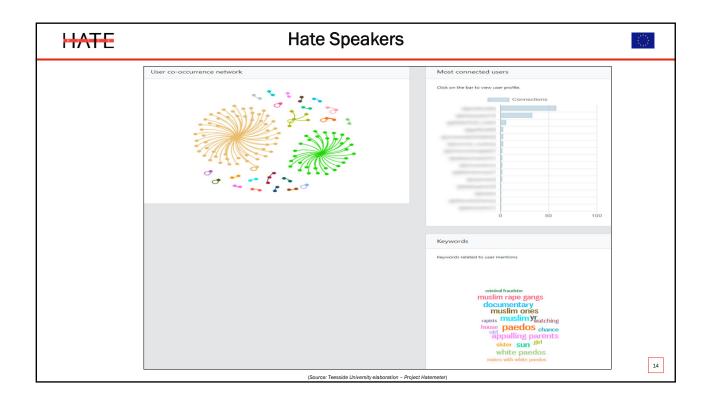


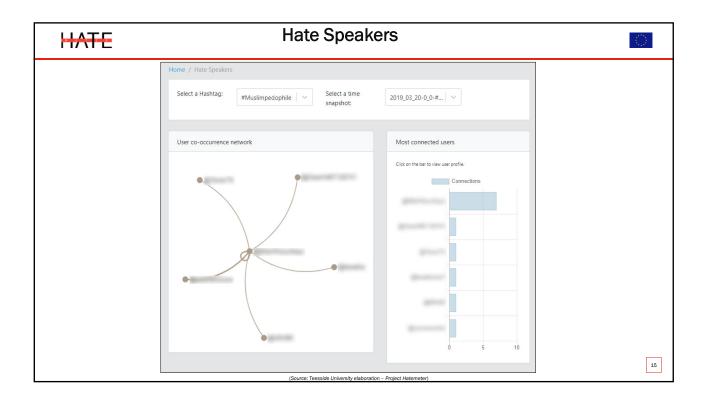






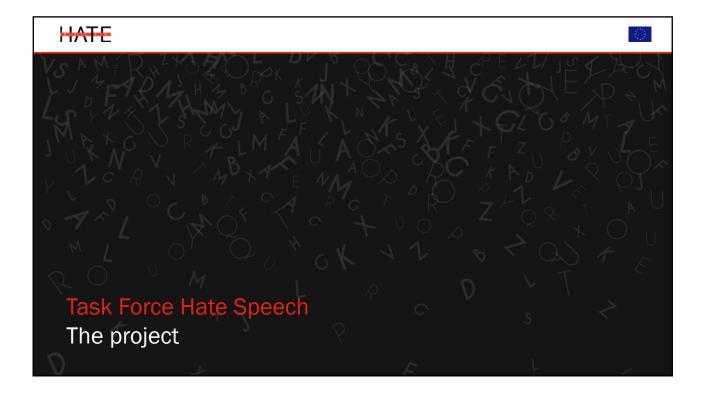








HATE	
Hatemeter Platform and Amnesty International Task Force	
Hate Speech	
The Collaboration	
Amnesty International Italy tested the Hatemeter Platform during the two Piloting Sessions thanks to the project called "Task Force Hate Speech":	
1 Pilot (from January to March 2019)	
Projects' aknowledgement     Data collection	
<ul> <li>2 Pilot (from June to September 2019)</li> <li>Training and testing</li> </ul>	
- Platform evaluation	
	F
	:



## Task Force Hate Speech

#### The Project

The Task Force Hate Speech is a project formed by motivated, skilled and trained activists.

#### Main objectives

- reduce the level of hate in the conversation
- spread impartial and objective information

#### Actions

- monitoring, contrasting and preventing the spread of online hate speech.
- acting in conversations where the level of discrimination and hatred is unacceptable.

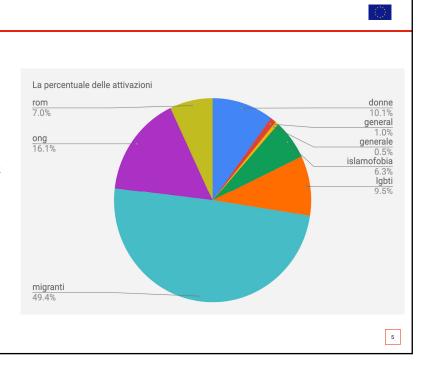


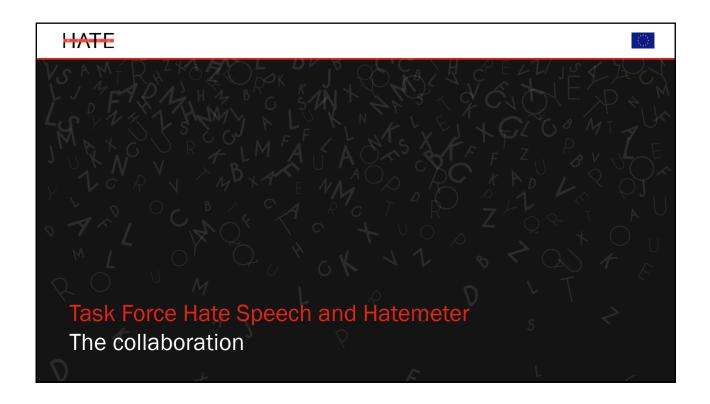
## Task Force Hate Speech

#### The Project

The main targets of online hate speech in Italy are the categories of migrants and refugees, solidarity, the LGBTI category and Islam.

Islamophobic speeches represent 6% of task force's activations,





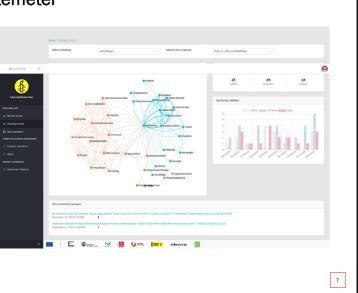
## Task Force Hate Speech and Hatemeter

#### **The Collaboration**

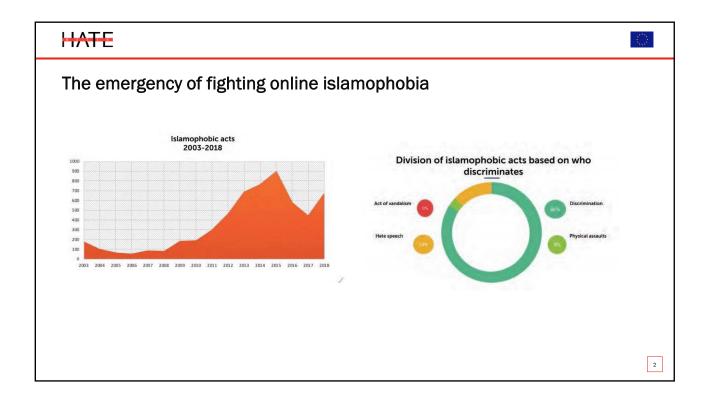
Amnesty International Task Force Hate Speech started working jointly with the Project Hatemeter focusing on the contrast of the online Islamophobia.

The collaboration started in the training sessions and continue during the piloting of the Hatemeter Platform, as well as in the development of the awareness raising campaign.

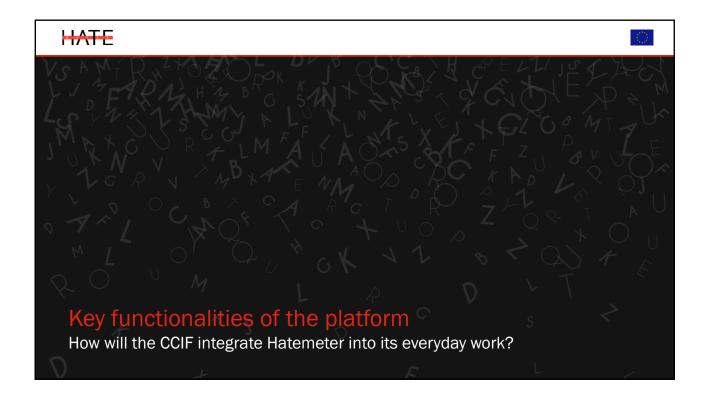
The Task Force gave support to the creation of the Hatemeter Platform by providing examples of hate speech and Islamophobic discourse followed by counter narrative examples.







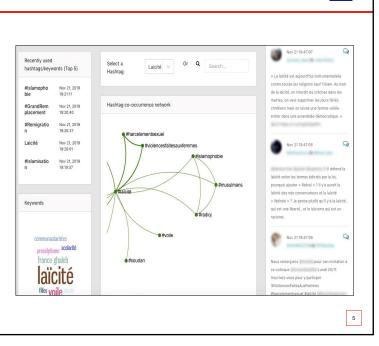
# <section-header><image><image><section-header><section-header><text><text><text><text>

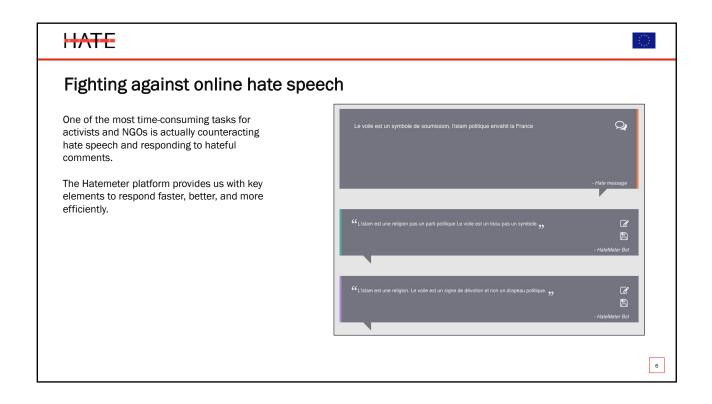


### Real time data acquisition

The Hatemeter platform provides concrete evidence and proof of the concepts used by Islamophobia, by showing the use of these concepts in islamophobic tweets. Ex: laïcité

If an activist needs to analyse how a concept is being used, he or she will have real time access to such information by simply searching its corresponding hashtag in the toolbar.







HATE
UK Pilots in Context:
Points to note:
<ul> <li>Stop Hate UK is an established 'Hate Crime' organisation, providing 24/7 'third party' (independent) reporting facilities in 'commissioned areas' across the UK.</li> </ul>
<ul> <li>We routinely receive reports of online Islamophobia from members of the public, both as 'victims' and 'bystanders', which are also used to identify potential 'targets' for both reporting and counter messaging by staff and volunteers involved in our 'No HateSpeech' team.</li> </ul>
• In testing the Hatemeter Platform, we are operating within a <b>project established in 2016</b> , with pre-existing practices and methodologies. Inevitably, we have not relied solely upon usage of the Platform as our primary mechanism for identifying and responding to instances of online Islamophobia.
2

## **UK Pilots**

Positive Experience of deployment:

- As an exceptional training tool
- Enabling volunteers & staff members to quickly and accurately identify sources of Islamophobic hate speech on Twitter.
- The ability to view various user networks, including central users and 'seed' accounts.
- Confirmation and representation of the international /transnational aspects and interrelationships of online Islamophobia.
- Counter message generation as a effective catalyst (for further refinement of a 'bespoke' response).