Rights, Equality and Citizenship (REC) Programme of the European Commission (2014-2020)



Monitoring and Detecting Online Hate Speech

D3.2: Reporting Portal[†]

Abstract: This document presents the Reporting Portal of the MANDOLA project. The deliverable is structured in two parts. In the first part of the deliverable, we focus on the different content sections which have been published related to FAQs and the existing Legal Framework regarding online hate speech. In the second part of the deliverable, we provide useful links that cover the majority of the European countries for reporting online hate speech.

Contractual Date of Delivery
Actual Date of Delivery
Deliverable Security Class
Editor
Quality Assurance

September 2016 October 2016 Public Meltini Christodoulaki Marios Dikaiakos

 $^{^{\}scriptscriptstyle \dagger}$ This project is funded by the Rights, Equality and Citizenship (REC) Programme of the European Commission.

The MANDOLA consortium consists of:

FORTH	Coordinator	Greece
ACONITE	Principal Contractor	Ireland
ICITA	Principal Contractor	Bulgaria
INTHEMIS	Principal Contractor	France
UAM	Principal Contractor	Spain
UCY	Principal Contractor	Cyprus
UM1	Principal Contractor	France

Document Revisions & Quality Assurance

Internal Reviewers

1. Marios Dikaiakos (UCY)

Revisions

Version	Date	Ву	Overview
v.3.2.2	11/10/2016	Meltini Christodoulaki	Final version
v.3.2.1	10/10/2016	Marios Dikaiakos	Review
v.3.2.0	26/9/2016	Meltini Christodoulaki	First draft.

Table of Contents

D	OCUM	/IENT REVISIONS & QUALITY ASSURANCE	3
T/	ABLE (OF CONTENTS	5
1	IN	NTRODUCTION	7
2	R	EPORTING PORTAL LAYOUT	8
	2.1	Home Section	
	2.2	REPORT HATE SPEECH SECTION	9
	2.3	LEGAL FRAMEWORK SECTION	10
	2.4	FAQ SECTION	13
3	R	EPORTING PORTAL DESIGN AND HOSTING	15
	3.1	GRID BASED DESIGN	15
	3.2	BROWSER COMPATIBILITY AND WEB STANDARDS COMPLIANCE	15
	3.3	HARDWARE AND HOSTING	15
4	E.	THICS APPROVAL	17
	4.1	7	
	4.2	NOTIFICATION OF THE HELLENIC DATA PROTECTION AUTHORITY	20
5	C	ONCHISION	21

List of Figures

Figure 1 - Home page of the reporting portal	8
Figure 2 - Report Hate Speech section	
Figure 3 - Legal Framework section	
Figure 4 - Legal Framework section that provide information for specific count	ry and specific
illegal behavior and provisions	12
Figure 5 - FAQ section	14
Figure 6 - SafeLine online form for reporting online hate speech	
Figure 7 - Ethics Approval from the Ethics Committee of FORTH	19

1 Introduction

This document presents the creation of the public MANDOLA reporting portal which is the main objective of the Task 3.2. The reporting portal of MANDOLA can be accessed from the following URL: http://mandola-project.eu/portal/

The creation of the reporting portal is in the Technical Infrastructure Workstream (WS3) and its objective is to develop a reporting portal that will allow citizens to submit possible online sources of illegal hate-related speech for further investigation by the police. More specifically, the reporting portal aims to enable people to connect to the portal and report any potentially illegal hate-related speech material and criminal activities that they have noticed on the Web and to inform them about this delicate issue. For developing the portal we have used open-source software.

2 Reporting Portal Layout

In this chapter we will present the sections of the reporting portal as they are at the time of the delivery of this document. However, the reporting portal will be "live" throughout the duration of the project and will be enriched with future informative material. At the bottom of each section there are links to the official website of MANDOLA and to our links to the online social media (Twitter and Facebook).

2.1 Home Section

In this section, internet users, can find basic information about the MANDOLA project through the "Who we are" and "What we do" sections. With the first sight citizens can see the organizations that constitute the MANDOLA consortium and the 5 core objectives of the project. The Home section (see Figure 1) as it is the default landing page of the portal, it shows the 3 other main areas of the portal and facilitates the user to easily access to the desirable section.

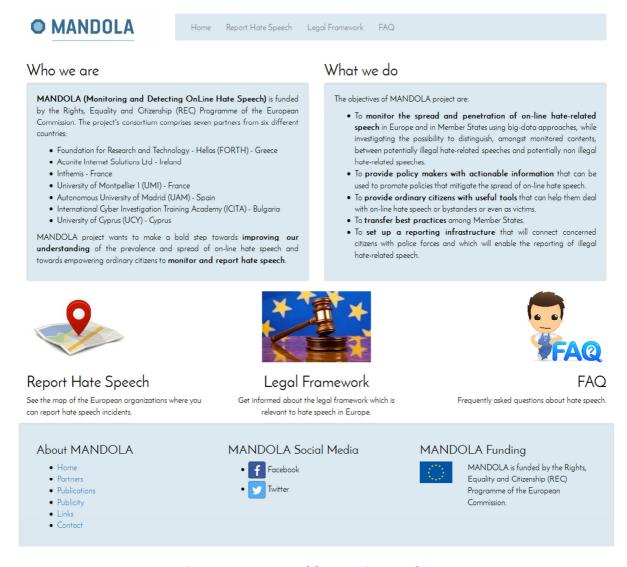


Figure 1 - Home page of the reporting portal

2.2 Report Hate Speech Section

By choosing this section, the map of Europe appears on the screen. We have found European Organization in 14 different countries where citizens can report online hate speech incidents. Users can report URLs that contain hate speech even anonymously by keeping secret their personal details. When a user clicks on a country, the organization that accepts reports for online hate speech appears (see Figure 2). Those European organizations are either Police departments or hotlines, or other non-profit organizations that as soon as they receive the report for potential hate speech, they immediately forward it to the national police.



Figure 2 - Report Hate Speech section

The list of the countries and the organization that can be accessed through this section are the following:

- Belgium
 - o http://unia.be/en/report-discrimination

- Bulgaria:
 - o http://www.safenet.bg/en/
- Check Republic:
 - o http://www.internet-hotline.cz/
- France:
 - o http://www.pointdecontact.net/click_report?language=en
- Germany:
 - o https://www.jugendschutz.net/
- Greece:
 - o http://www.safeline.gr
- Ireland:
 - o https://www.hotline.ie/
- Latvia:
 - o http://www.drossinternets.lv/page/32
 - o http://cilvektiesibas.org.lv/en/database/report-discrimination/
- Lithuania:
 - o http://www.draugiskasinternetas.lt/en
- Netherlands:
 - o http://www.meldpunt.nl/site/page.php?lang=1&pageID=23
 - o https://www.mindnederland.nl/melden/
- Poland:
 - o https://dyzurnet.pl/en/
- Portugal:
 - o http://linhaalerta.internetsegura.pt/index.php?lang=en
- Romania:
 - o http://www.safernet.ro/
- Slovenia:
 - o http://safe.si/en/spletno-oko/hotline-spletno-oko
- United Kingdom:
 - o https://secure.met.police.uk/hatecrime national/index.php?rid=2903

As soon as we have more information about other additional countries or other organization where ordinary citizens can report hate speech we will update the above list and the reporting map of this section.

2.3 Legal Framework Section

In this section, internet users can find useful information about the existing legal framework related to hate speech and hate crime issues. In this section we tried to provide in an easy and friendly manner the information that has been produced under the D2.1 (Definition of illegal hatred and implication) deliverable of our project. By clicking on the button of the Legal Framework section a second map appears on the screen (see Figure 3). At the moment, the portal can provide information on the legal framework only for the countries which are marked with the yellow color, which are Ireland, Spain, France, Belgium, Netherlands, Germany, Romania, Bulgaria, Cyprus and Greece.



Figure 3 - Legal Framework section

By clicking on a yellow marked country, we can see the four different subsections that group the content for this country. Those subsections are the following:

- Texts that appear to be particularly appropriate to sanction hatred-related content, and that are or should be covered in all countries.
- Texts that might enable to combat online discrimination even though their main objective is to combat discrimination offline.
- Civil liability
- Additional texts, less or more generic or prohibited in some countries only.

Each of the above links redirect the user to further information, the different types of behavior that can be characterized as hate speech, the civil liability out of a specific illegal behavior and the legal instruments (acts and articles) that prohibit and penalize the illegal behavior (see Figure 4).



III Get informed about the legal framework which is relevant to hate speech in Europe.



Texts that appear to be particularly appropriate to sanction hatred-related contents, and that are or should be covered in all countries

- Public incitement to violence, hatred or discrimination
- Making available xenophobic or racist material which incites or promotes hatred or violence
- Establishing or participating in organisations that promote discrimination, violence or hatred
- Publicly insulting persons for illegal hatred relating reasons
- Public defamation
- Threatening a natural person, motivated by racism or xenophobia
- Illegal hatred as an aggravating circumstance
- Public denial, condoning or trivialising war crimes, crimes against humanity, genocide and crimes against peace

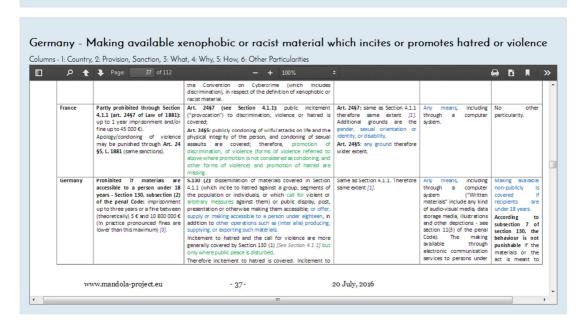


Figure 4 - Legal Framework section that provide information for specific country and specific illegal behavior and provisions

2.4 FAQ section

In this section, the visitors of the portal can find answers to the most usual questions related to hate speech. Since the crime of hate speech is not yet very clear to the majority of internet users, this section will provide a step further for understanding what hate speech is, and which types of behavior consist hate speech. With the claim of freedom of expression and freedom of speech, users may on the one hand commit the crime of hate speech on the one hand or be victims of hate speech on the other hand. The current tool will help them better realize the phenomenon of hate speech.

Out of the total number of 23 questions, the portal visitors can choose the question they wish and read the explanatory answer fully referenced (see Figure 5). The FAQ section consists of the following questions:

- 1. What is Hate Speech?
- 2. What is legally considered Hate Speech?
- 3. Who are the victims of hate speech?
- 4. I believe I have encountered hate speech. What can I do about it?
- 5. What does Freedom of Expression allow us to say?
- 6. Why words can be harmful and sometimes dangerous?
- 7. Is hate speech forbidden online?
- 8. Is Hate speech and hate crime the same thing?
- 9. I believe I have been a target of hate speech? What are my rights as a victim?
- 10. Which legal protection against Hate Speech does exist within EU?
- 11. What are providers doing about online hate speech?
- 12. Who is liable for hate speech?
- 13. Who investigates online hate speech?
- 14. Which law applies to online hate speech?
- 15. What are the penalties for those found guilty of hate speech?
- 16. Which persons are entitled to initiate legal proceedings against authors of hate speech?
- 17. Are there any restrictions in freedom of speech for audio-visual service providers?
- 18. Which rules prevent a company to hire me or fire me based on grounds of race, sex, religion or nationality or to restrict me under any internal policy rules?
- 19. What is harmful content?
- 20. Why are some harmful contents not illegal?
- 21. What is the difference between illegal content and potentially illegal?
- 22. How does the MANDOLA project support activities against online hate speech?
- 23. Where can I find more information?



Home Report Hate Speech Legal Framework FAQ

III Frequently Asked Questions

Select your question

- 1. What is Hate Speech?
- 2. What is legally considered Hate
- Who are the vicinis a.

 I believe I have encountered hate
- speech. What can I do about it? 5. What does Freedom of Expression
- sometimes dangerous?
- 7. Is hate speech forbidden online?
- 8. Is Hate speech and hate crime the same thing?

- 14. Which law applies to online hate
- speech? 15. What are the penalties for those found guilty of hate speech?
- 9. I believe I have been a target of hate 16. Which persons are entitled to initiate 20. Why are some harmful contents not speech? What are my rights as a legal proceedings against authors of
 - of speech for audio-visual service 22. How does the Mandola project
- online hate speech?

 12. Who is liable for hate speech?

 13. Who in matter that speech?

 14. Which rules prevent a company to hire me or fire matter. allow us to say?

 12. Who is liable for hate speech?

 13. Who investigates online hate speech?

 14. Who investigates online hate speech?

 15. Who investigates online hate speech?

 16. Why words can be harmful and

 17. Who investigates online hate speech?

 18. Who investigates online hate speech? to restrict me under any internal policy rules?
 - 19. What is harmful content?
- illegal?
- 3. Who are the victims of hate speech?
 4. I believe I have encountered hate

 Victim/

 Nate speech?

 10. Which legal protection against Hate
 4. I believe I have encountered hate

 Victim/

 A to there any restrictions in freedom

 Content and potentially illegal?
 - support activities against online hate speech?

14. Which law applies to online hate speech?

Due to the Internet being an egalitarian way of communication, a unilateral national content legislation does not always apply when prosecuting offences committed in such a globalised and decentralised landscape. Even though there have been efforts to agree on virtual borders in relation to cyberspace so as to regulate this type of offences, these efforts were unsuccessful since they impacted on jurisdictional and sovereignty issues. An international and multilateral legislative approach to investigate and fight against hate speech must and has progressively been adopted, even if it still faces some challenges, most of them having to do with different commitments to free speech and regulation in different countries. The Convention of Cybercrime (Budapest, 23/11/2001) was the first international treaty on crimes committed via the Internet and other computer networks, and its aim was to set out a common criminal policy pursuing the protection of society against cybercr by fostering international co-operation and developing appropriate legislation³. Its State Parties are represented by the Cybercrime Convention Committee (T-CY). There was an additional protocol to the Convention of Budapest, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems

On the basis of this agreement and on other agreements relating to penal law (such as the International Convention on the Elimination of All Forms of Racial Discrimination) or to cooperation in the penal area, bound States (see question 10) are supposed to have in their legislation a common denominator of infringements, and basic rules enabling the territorial competence of their countries in certain cases (ensuring that way that at least one penal law applies in each case), and enabling cooperation in case of assistance request coming from a foreign LEA.

For example, art. 22 of the Convention on cybercrime states that "each Party must adopt such legislative and other measures as may be necessary to establish jurisdiction over most of the offences established in accordance with the provisions of the Convention, when the offence is committed (a) in its territory; or (b) on board a ship flying the flag of that Party; or (c) on board an aircraft registered under the laws of that Party; or (d) by one of its nationals, if the offence is punishable under criminal law where it was committed or if the offence is committed outside the territorial jurisdiction of any State* (exceptions may be brought to b, c and d). Art 22's last paragraph also states that "when more than one Party claims jurisdiction over an alleged offence established in accordance with this Convention, the Parties involved shall, where appropriate, consult with a view to determining the most appropriate jurisdiction for prosecution".

Regarding the notion of "committed in its territory", in relation with materials that are available on the Internet, it may differ according to the provision of the penal legislation of each country. However, many countries consider that the offence has been committed in their own territory where one of the legal elements that constitute the offence has been committed on this territory. Some countries further consider that the commission on the territory of one of the facts that constitute together the offence is enough (such as in France). As a consequence, several legislations are likely to apply in case of illegal online content

- 1. France Vs. Yahoo! Inc.
- 2. The right to punish is the most important expression of States sovereignty See for example Huet (A), "L'impact du droit communautaire sur le droit penal", in Simon (D.) (sous la dir. de), Le droit communautaire et les metamorphoses du droit, Strasbourg PUS 2003, p. 13 s., quoted by Denys SIMON - Cahiers du Conseil constitutionnel n26 (Dossier : La Constitution et le droit penal), Constitution, "souverainete penale, droit communautaire", August 2009, available at munautaire.51489.html (last acceded on 12 July 2016).
- 3. See for example http://
- 4. See for example Estelle De Marco, Le Droit Penal Applicable sur Internet (The penal law applicable on the Internet), Master II thesis, University of Montpellier 1.1998

About MANDOLA

- Home
- Partners
- Publication Publicity

MANDOLA Social Media

Facebook

• Twitter

MANDOLA Funding

MANDOLA is funded by the Rights. Equality and Citizenship (REC) Programme of the European Commission

Figure 5 - FAQ section

- 14 -

3 Reporting Portal Design and Hosting

The reporting portal was designed by FORTH and a draft version was presented to the consortium. After we addressed most of the received comments, the portal was finalized with its section and its content in the English language and delivered at the end of August.

3.1 Grid based design

We wanted the portal to have a visual layout which is clean-cut while at the same time is easy to change in order to accommodate future needs. For this, we chose to have it designed and built using <u>Twitter Bootstrap</u>. Twitter Bootstrap is a CSS framework that allows the rapid prototyping of *grid based website designs* while working equally well when integrated into a production system.



In grid based designs, the visual blocks that comprise the website (e.g., menus, text boxes, information boxes, ads etc.) are not placed on arbitrary positions. Instead they are laid out on predefined rigid positions on a grid. This may sound restrictive but in practice the resulting design is much more efficient in communicating its contents to the visitor. This is because placing the visual blocks of the website on a grid results in *clear visual paths* and visual *structure and balance* on the design. Additionally, a grid based design also ensures consistency between the website pages and are much easier to update in order to accommodate additional content.

3.2 Browser Compatibility and Web Standards Compliance

The Reporting Portal pages have been tested to comply with the *HTML5* standard, using the *W3C Markup Validator*. We have also chosen to use *CSS3* for the portal because it greatly simplifies the implementation of aesthetic elements such as rounded element corners, element shadows etc. Without CSS3, these elements have to be pre-rendered as bitmap images and then included in the page, which degrades the semantic integrity of the produced HTML output.



3.3 Hardware and hosting

The Reporting Portal is hosted by FORTH on their premises in Heraklion. The hosting server features two Intel Xeon dual-core CPUs running at 2.66GHz and a total memory of 4GB. It is connected to the Internet through FORTH's 10 Gigabit connection to the GRNET backbone. The server has two high performance SAS disks (10k RPM) arranged as RAID-1 for fault-tolerance.

The server is protected by firewalls and is internally and externally monitored in order to minimize the risk from cyber-threats. Additionally, remote backups through the *rsync* utility are performed on a daily basis.

It is also important that the server resides in a protected physical environment. It is located in one of FORTH's data-centers. For ensuring optimal operating environment, it is equipped with industrial-strength air conditioning with more than 240.000BTUs efficiency. In power emergencies, it is supported by a UPS power supply and an external power generator which is engaged automatically on power failure. Additionally, the data-center features an automatic carbon dioxide fire-extinguishing system.

4 Ethics approval

4.1 Approval from the Ethics Committee of FORTH

In Greece, Internet users can report URLs of online hate speech to SafeLine (www.safeline.gr), which is the Greek hotline against illegal internet content. For this type of content, only the field of the URL can be reported and no any further information such as email address or other communication details or any other comments from the user. The reports can be submitted anonymously and immediately after the reception they will be forwarded to the Cyber Crime Unit of the Greek Police.

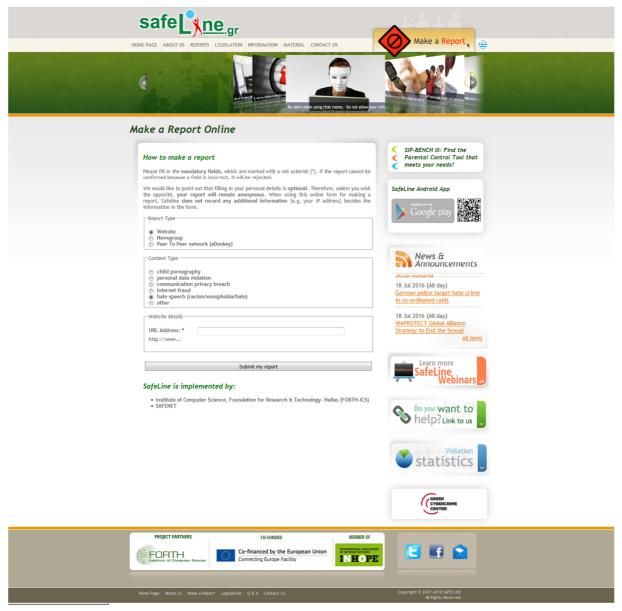
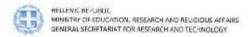


Figure 6 - SafeLine online form for reporting online hate speech

Additionally, all the URLs that will be reported as hate speech will be kept in a database where only the hotline analyst has access after the required credentials (username and password) are provided. The IPs of the users that will submit report will automatically be deleted by a software that will be running for erasing them. Since we wanted to be fully compatible with national and European law of Personal Data protection we contacted the Ethics Committee of FORTH and submitted an application for taking the permission of implementing this task.

The date of the application submission was the 1st of July and on 4th of August we received the official answer by the Committee which is shown in the following figure:





FEC APPROVAL Ref. No: 38 / Date: August 4th 2016

Applicant - FORTH personnel: Prof. Evangelos MARKATOS

Project Acronym/Title: 'MANDOLA'/'Monitoring ANd Detecting Online hAte speech'

Project Action/Deliverable: Reporting Portal Project Curation: 36 months

FEC Submission Date: 01-07-2016 FEC Issue/Approval Date: 04-08-2016 FEC Approval Expiry Date: 03-08-2017

To whom it may concern,

This is to certify that the FORTH Ethics Committee (FEC) has considered, the submittee and filled FEC Application Form Request as of .01-07-2016/PN35 (incl. all relevant and supporting documents) with respect to the FORTH Research proposal/project. [acronym: 'MANDOLA', title: Monitoring AND Detecting OnLine hAte speech] as described & submitted by the scientifically responsible person of such proposal/project Prof. Evangelos MARKATOS at FORTH-ICS for purposes of personal data protection & ethics review/approval/compliance of the aforementioned submitted research proposal/project with the most applicable national and EU legislation and certifies the following:

The FEC has reviewed that the aforementioned submitted application/research proposal/project is,

- ii) Initially and conditionally, compliant with national (GR) personal data protection laws & ethical provisions, as of today, as well as the applicable standards of good research practice; and
- (ii) potentially compliant with the recently adopted on personal data related protection matters EC Directive 95/46/EC and Regulation 2016/680/EC, which however have not been incorporated within national laws as of today, due to be until 2018; and therefore,
- (iii) obtentially and conditionally compliant with all applicable national and European laws as of today and subject to the fulfillment of the following terms and conditions:
- (iv) further documents to be re-submitted to FEC, once known and in possession of the FORTH scientifically responsible person Prof. Evangelos MARKATOS at FORTH-ICS.
- (v) further review by the applicants, interested parties and the FEC of all submitted documents and actual research work in progress in the light of upcoming legal developments with respect to personal data protection matters for purposes of compliance of the national data protection laws with the EC Directive 95/46/EC and the EC Regulation 2016/680/EC during the entire period of this project (01/10/2015-30/09/2017).

Sincerely,

Prof. N.

Coordinator FORTH Ethics Committee (FEC)

Director IMB3-FORTH

Nikolaou Plastira 100, Vassilika Vouton

GR 70013 Heraldion, Crete, GR

Tavernarakis

Phone: +30 2810 391102, Fax: +30 2810 391104

email: tavernarakis@imbb.forth.gr

Window Plastins 200 Vassilika Vastes DA 700 14 Heracidos Dorte Erescu (el. -90 2610381000 0 Fan -30 2820391005 Haust Carmati Vastesia Institut

www.forth.gc

Figure 7 - Ethics Approval from the Ethics Committee of FORTH

4.2 Notification of the Hellenic Data Protection Authority

One precondition that was set to us from the Ethics Committee of FORTH for providing the Ethics approval was to inform the Hellenic Data Protection Authority (HDPA) for the implementation of this task. For this reason, in parallel, we contacted the Hellenic Data Protection

Authority

(http://www.dpa.gr/portal/page? pageid=33,40911& dad=portal& schema=PORTAL) and we informed them about the retention of our database and the reporting form that we offer to internet users.

The discussion was made with Mr. Konstantinos Limniotis who stated that the collection of only URLs without any further personal information from the internet users does not consist serious case of collection and processing of personal data. It is extremely difficult to identify who was the user who submitted the report. As a result, our case does not come under of those where an official license from the HDPA is essential. However, he informed us that if we wish, we can simply notify the Authority about the implementation of our task.

Thus, we completed the appropriate forms and we sent them to the HDPA via email.

5 Conclusion

In this document, we presented the MANDOLA reporting portal. We provided a description of its sections along with the relevant screenshots for each section. The reporting portal will be a tool for Internet users in order to be better informed for the hate-related issues and report such content. The portal is finalized and we will make the suitable updates whenever it is required in the future for the better operation this task.