D1.4: Advisory Board Meeting 1 †

Abstract: This deliverable reports the work done in creating the MANDOLA Advisory Board and preparing its first meeting, as well as the proceedings of this meeting.

Contractual Date of Delivery          September 2016
 Actual Date of Delivery              October 2016
 Deliverable Security Class           Public
 Editor                               Nikos Frydas
 Contributors                        Evangelos Markatos
                                      Melini Christodoulaki
 Quality Assurance                    Marios Dikaiakos

† This project is funded by the Rights, Equality and Citizenship (REC) Programme of the European Commission.
The *MANDOLA* consortium consists of:

<table>
<thead>
<tr>
<th>Institution</th>
<th>Role</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORTH</td>
<td>Coordinator</td>
<td>Greece</td>
</tr>
<tr>
<td>ACONITE</td>
<td>Principal Contractor</td>
<td>Ireland</td>
</tr>
<tr>
<td>ICITA</td>
<td>Principal Contractor</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>INTHEMIS</td>
<td>Principal Contractor</td>
<td>France</td>
</tr>
<tr>
<td>UAM</td>
<td>Principal Contractor</td>
<td>Spain</td>
</tr>
<tr>
<td>UCY</td>
<td>Principal Contractor</td>
<td>Cyprus</td>
</tr>
<tr>
<td>UM1</td>
<td>Principal Contractor</td>
<td>France</td>
</tr>
</tbody>
</table>
# Document Revisions & Quality Assurance

## Internal Reviewers

1. Marios Dikaiakos (UCY)

## Revisions

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>By</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.4.1</td>
<td>22/10/2016</td>
<td>Editor</td>
<td>Amendments / corrections from AB members</td>
</tr>
<tr>
<td>1.4.0</td>
<td>11/10/2016</td>
<td>Editor</td>
<td>First draft.</td>
</tr>
</tbody>
</table>
# Table of Contents

**DOCUMENT REVISIONS & QUALITY ASSURANCE** ........................................................................................................ 3

**TABLE OF CONTENTS** .............................................................................................................................................. 4

1 **INTRODUCTION** .................................................................................................................................................. 6

2 **AIMS & OBJECTIVES OF THE ADVISORY BOARD (AB)** .................................................................................. 7

   2.1 **THE OBJECTIVES OF THE MANDOLA AB** .................................................................................................. 7

   2.1.1 **AB duties in general** ............................................................................................................................ 7

   2.1.2 **AB duties in particular** ........................................................................................................................ 7

   2.2 **AB CONSTRAINTS** .......................................................................................................................................... 8

   2.3 **AB MEMBERSHIP** ......................................................................................................................................... 8

3 **METHODOLOGY USED TO POPULATE THE AB** .............................................................................................. 9

   3.1 **ATTRIBUTES OF AB CANDIDATES** ........................................................................................................... 9

   3.2 **POPULATION OF THE SUPER-LIST** .......................................................................................................... 10

   3.3 **CANDIDATE SCORE** .................................................................................................................................... 14

   3.3.1 **Primary Area of Expertise [of the AB member]** .................................................................................... 14

   3.3.2 **Type of Organization [to which the AB member belongs]** ....................................................................... 14

   3.3.3 **[MANDOLA] Member [proposing an AB candidate]** .............................................................................. 15

   3.3.4 **[MANDOLA] Member Organization [proposing an AB candidate]** ......................................................... 15

   3.3.5 **Nationality [of the AB candidate]** ........................................................................................................ 15

   3.3.6 **Area of Residence [of the AB candidate]** ............................................................................................... 15

   3.3.7 **Gender [of the AB candidate]** ................................................................................................................ 16

   3.4 **PRIORITIZING AB CANDIDATES** ............................................................................................................... 16

4 **PROCEEDINGS OF THE ABI** ............................................................................................................................ 18

   4.1 **WELCOME/INTRODUCTION/ADVISORY BOARD** ...................................................................................... 18

   4.2 **SHORT INTRODUCTION TO MANDOLA** .................................................................................................. 18

   4.3 **TECHNICAL INFRASTRUCTURE** ............................................................................................................... 19

   4.4 **DEFINITION OF HATE SPEECH & LEGAL FRAMEWORK** ...................................................................... 19

   4.5 **BRAINSTORMING PANEL I** ....................................................................................................................... 21

   4.5.1 **Panel I / Question 1** ............................................................................................................................ 21

   4.5.2 **Panel I / Question 2** ............................................................................................................................ 23

   4.5.3 **Panel I / Question 3** ............................................................................................................................ 24

   4.5.4 **Panel I / Question 4** ............................................................................................................................ 26

   4.6 **BRAINSTORMING PANEL II** ...................................................................................................................... 28

   4.6.1 **Panel II / Question 1** ........................................................................................................................... 28

   4.6.2 **Panel II / Question 2** ........................................................................................................................... 30

   4.6.3 **Panel II / Question 3** ........................................................................................................................... 31

   4.6.4 **Panel II / Question 4** ........................................................................................................................... 33

5 **CONCLUSIONS & LESSONS LEARNED** .......................................................................................................... 34

6 **APPENDIX A: AGENDA OF ABI (ADVISORY BOARD MEETING I)** ..................................................................... 35

7 **APPENDIX B: ADVISORY BOARD PRESENTATION** .......................................................................................... 36

8 **APPENDIX C: INTRODUCTION TO MANDOLA PRESENTATION** ...................................................................... 38

9 **APPENDIX D: TECHNICAL INFRASTRUCTURE PRESENTATION** ................................................................. 41

10 **APPENDIX E: DEFINITION OF HATE SPEECH & LEGAL FRAMEWORK PRESENTATION** ...................... 44

11 **APPENDIX F: BRAINSTORMING PANEL I / QUESTION 1** ............................................................................... 47

12 **APPENDIX G: BRAINSTORMING PANEL I / QUESTION 2** ............................................................................ 48

13 **APPENDIX H: BRAINSTORMING PANEL I / QUESTION 3** ............................................................................ 49

14 **APPENDIX I: BRAINSTORMING PANEL I / QUESTION 4** ............................................................................ 50

15 **APPENDIX J: BRAINSTORMING PANEL II / QUESTION 1** ............................................................................ 51
1 Introduction

This document comprises the following chapters:

1. Chapter 2 [*Aims & Objectives of the Advisory Board (AB)*]: This chapter describes the aims & objectives of the Advisory Board, as well as the practical constraints taken into consideration, when examining candidates.

2. Chapter 3 [*Methodology used to populate the AB*]: This chapter describes the methodology used to populate the Advisory Board.

3. Chapter 4 [*Proceedings of the AB1*]: This chapter gives the proceedings of the Advisory Board Meeting 1 (*AB1*).

4. Chapter 5 [*Conclusions & Lessons Learned*]: This chapter gives the conclusions and lessons learned from AB1.

5. The document includes the following appendices:
   a. Appendix A: Agenda of AB1 (Advisory Board Meeting 1)
   b. Appendix B: Advisory Board presentation
   c. Appendix C: Introduction to MANDOLA presentation
   d. Appendix D: Technical Infrastructure presentation
   e. Appendix E: Definition of Hate Speech & Legal Framework presentation
   f. Appendix F: Brainstorming Panel I / Question 1
   g. Appendix G: Brainstorming Panel I / Question 2
   h. Appendix H: Brainstorming Panel I / Question 3
   i. Appendix I: Brainstorming Panel I / Question 4
   j. Appendix J: Brainstorming Panel II / Question 1
   k. Appendix K: Brainstorming Panel II / Question 2
   l. Appendix L: Brainstorming Panel II / Question 3
   m. Appendix M: Brainstorming Panel II / Question 4
2 Aims & Objectives of the Advisory Board (AB)

This chapter describes the aims & objectives of the Advisory Board, as well as the practical constraints taken into consideration.

The aim of the task undertaken is to compose the optimum AB, under the practical constraints of the project.

The Chapter comprises the following sections:
1. The Objectives of the MANDOLA AB
2. AB Constraints
3. AB Membership

2.1 The Objectives of the MANDOLA AB

Setting up an Advisory Board “that will steer this project” is the goal of WS1.3. The delivery of the following outputs is part of the project’s contractual obligations:

1. D1.4 Advisory Board Meeting 1  Target group: ALL  
2. D1.5 Advisory Board Meeting 2  Target group: ALL

The current document constitutes deliverable D1.4.

2.1.1 AB duties in general

In general, an Advisory Board provides non-binding strategic advice. Among the reasons for creating an AB are the following:

- Seek expertise outside MANDOLA.
- Complement existing strengths.
- Counsel on issues raised by MANDOLA.
- Become a resource for MANDOLA managers.
- Provide un-biased ideas.
- Monitor project performance.

2.1.2 AB duties in particular

According to the MANDOLA project objectives, the Advisory Board should have the following characteristics:

- AB will steer the project.
- AB will help spread the project message well beyond participant Member States.
- AB will assist the promotion of the developed technologies and tools.
- AB will provide valuable feedback & market guidelines on progress & results.
- AB will further enhance impact & dissemination of MANDOLA’s ideas.
- AB will foster dialogue & debate.
- AB will serve as a source of expertise.
2.2 AB Constraints

Project constraints place an upper limit of 20 to the number of external AB members who reside outside Brussels. In addition, the AB members must be EU residents.

The meeting room made available has a capacity of 25. This implies that with a total of nine internal AB members, the external AB members should be restricted to 16.

MANDOLA project partners are grateful to the European Office of Cyprus, in Rue du Luxembourg 3, Brussels, who made their meeting room available, free of charge.

2.3 AB Membership

In general, AB members must be individuals

1. with personal qualities and
2. representing an important entity, where important is understood to mean important for the project, and
3. with knowledge of the issues the project deals with and
4. with good command of English and
5. with the ability to be present at the AB meetings in Brussels.

Given the above and the project objectives (see «The Objectives of the MANDOLA AB», above), AB members shall then be drawn from:

- Academia
- NGOs
- LEA
- Internet Industry
- Government
- other
3 Methodology used to populate the AB

This chapter describes the methodology used to populate the Advisory Board.

It was decided to follow the methodology below:

1. Create a super-list of 50-60 individuals, candidates for the AB.
2. Assess the suitability of each individual across a number of attributes.
3. Combine the marks/attribute into an overall score/individual.
4. Order the individuals according to their score.
5. Invite the top 16 individuals.
6. Once an individual accepts an invitation, the individual is moved to the top of the list.
7. Once an individual declines the invitation, the individual is moved to the bottom of the list.
8. Continue until you have 16 acceptances.

3.1 Attributes of AB candidates

An optimum AB, would be one which would satisfy the aims and objectives discussed under Chapter 2 (see p. 7). The basic qualities required, from the AB, are:

1. A focused range of expertise:
   a. Child Care
   b. Cybersecurity (Leg)
   c. Cybersecurity (Tech)
   d. Hate Speech
   e. Hotline
   f. Human Rights
   g. Linguistics
2. Balanced representation of AB members’ organizations:
   a. Academia
   b. Industry
   c. Intl ORG
   d. LEA
   e. Mass Media
   f. NGO
   g. State
3. Wide and balanced representation of nationalities.

In addition to the above, it was thought appropriate to balance MANDOLA members\(^1\) (the MEMBERS) recommendations, as well as MANDOLA member organizations’ recommendations.

Finally, given that the total AB cost depends mainly on travelling expenses, it was thought that the distribution of the areas of residence should also be balanced.

Given the above, the ATTRIBUTES for assessing AB candidates are:

---

\(^1\) The team of individuals who participate in the MANDOLA project, on behalf of the Coordinator and the Principal Contractors.
3.2 Population of the Super-List

Super-List is the list of all individuals, considered for AB participation.

The Super-List was populated via recommendations from MEMBERS, who were invited on 5/11/2015 to fill a suitable recommendation form. This process lasted for almost seven months and was interactive.

The results are depicted in Figure 1 and in Figure 2.
The Super-List composition regarding the Type of Organization the AB candidates belong to, is depicted in Figure 3.

![Type of Organization](image)

**Figure 3: Type of Organization AB candidates belong**

The Super-List composition regarding the Primary Area of Expertise of the AB candidates, is depicted in Figure 4.
The Super-List composition regarding the gender distribution of the AB candidates, is depicted in Figure 5. “X” denotes unknown gender, as it was not known which individual would represent the candidate organization.

The Super-List composition regarding the Nationality of the AB candidates, is depicted in Figure 6. “European” denotes unknown nationality, as it was not known which individual would represent the candidate organization.
Figure 6: Nationality of AB candidates.

Figure 7: Area of Residence of AB candidates.
### 3.3 Candidate Score

Each candidate is awarded a mark for each of the ATTRIBUTES introduced earlier on:

A. Primary Area of Expertise [of the AB member]
B. Type of Organization [to which the AB member belongs]
C. [MANDOLA] Member [proposing an AB candidate]
D. [MANDOLA] Member Organization [proposing an AB candidate]
E. Nationality [of the AB candidate]
F. Area of Residence [of the AB candidate]
G. Gender [of the AB candidate]

The individual marks, one per ATTRIBUTE per candidate are weighted and summed to produce the candidate score. The candidate score is thus the weighted average of the candidates’ marks per ATTRIBUTE.

#### 3.3.1 Primary Area of Expertise [of the AB member]

As discussed in §3.1, above, the following expertise is considered desirable for the AB:

*Table 1: Primary Area of Expertise of AB Members and Weighting*

<table>
<thead>
<tr>
<th>#</th>
<th>Primary Area of Expertise</th>
<th>W-Ex</th>
<th>W-Ex%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Child Care</td>
<td>5</td>
<td>11%</td>
</tr>
<tr>
<td>02</td>
<td>Cybersecurity (Leg)</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td>03</td>
<td>Cybersecurity (Tech)</td>
<td>7</td>
<td>15%</td>
</tr>
<tr>
<td>04</td>
<td>Hate Speech</td>
<td>10</td>
<td>21%</td>
</tr>
<tr>
<td>05</td>
<td>Hotline</td>
<td>9</td>
<td>19%</td>
</tr>
<tr>
<td>06</td>
<td>Human Rights</td>
<td>8</td>
<td>17%</td>
</tr>
<tr>
<td>07</td>
<td>Linguistics</td>
<td>1</td>
<td>2%</td>
</tr>
</tbody>
</table>

Each entry is given a weight from 1 to 10, according to importance (field “W-Ex”). Field “W-Ex%”, indicates the % value of the weights. For example, “Hate Speech” is the most important Expertise and is, thus, given a weight of 10, etc.

Each candidate is given a mark from 1 to 10, as following: All candidates with the same Expertise, say “Human Rights”, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

#### 3.3.2 Type of Organization [to which the AB member belongs]

As discussed in §3.1, above, the following Types of Organization are considered desirable for the AB:
Each entry is given a weight from 1 to 10, according to importance (field “W-Or”). Field “W-Or%”, indicates the % value of the weights. For example, “Industry” is considered the most important Type of Organization and is, thus, given a weight of 10, etc.

Each candidate is given a mark from 1 to 10, as following: All candidates with the same Type of Organization, say “NGO”, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.3.3 [MANDOLA] Member [proposing an AB candidate]
Entries are the MEMBERS. Each entry is given the same weight.

Each candidate is given a mark from 1 to 10, as following: All candidates proposed by the same MEMBERS, say N. Frydas, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.3.4 [MANDOLA] Member Organization [proposing an AB candidate]
Entries are the MEMBERS’ Organizations. Each entry is given the same weight.

Each candidate is given a mark from 1 to 10, as following: All candidates proposed by the same MEMBERS’ Organization, say FORTH, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.3.5 Nationality [of the AB candidate]
Entries are the AB candidate’s Nationality. Each entry is given the same weight.

Each candidate is given a mark from 1 to 10, as following: All candidates of the same Nationality, say Greek, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.3.6 Area of Residence [of the AB candidate]
Entries are the Areas of Residence of the AB candidates (e.g., Austria, Belgium, Bulgaria, Cyprus, etc.). Each entry is given different weights, according to distance from Brussels. Brussels get the top mark, 10, while Cyprus get 5, etc.
Each candidate is given a mark from 1 to 10, as following: All candidates from the same Area of Residence, say Greece, are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.3.7 Gender [of the AB candidate]

Entries are “Male”, “Female” & “X” (unknown yet). “Female” is given the top mark, 10, as there are less female candidates in the Super-List. “Male” is given 5 and “X” 7 (the average between “F” & “M”).

Each candidate is given a mark from 1 to 10, as following: All, say, male candidates are sorted in order of suitability for the AB, with the most suitable scoring 10, the next most suitable 9, etc.

3.4 Prioritizing AB Candidates

Each AB candidate obtains a certain score, as described previously. According to this score, the candidate is awarded a status, and if appropriate, an invitation, or an enquiry is sent. Status is modified according to the candidate’s response.

Status takes on the following values:

1. Accepted: The candidate has accepted the invitation. The system awards extra marks, so that the candidates who accepted move on to the top of the list. The entry is coloured **bold blue**.
2. Accepted (if): The candidate has accepted the invitation, under certain conditions. The system awards extra marks, less than for “accepted”, so that these candidates move on the list, just below the previous category. The entry is coloured **bold violet**.
3. Affiliated: The candidate has been sent an invitation, wants to be a member of the AB but is not able to be physically present at AB1. Affiliated members will be sent the material of the AB1.
4. Enquiry sent: This is an exploratory letter, just short of an invitation. The entry is coloured **violet**.
5. Invitation 2b sent: This is a candidate to whom an invitation will be sent.
6. Invitation sent: This is a candidate to whom an invitation has been sent. The system awards them extra marks, more than for the category “Enquiry sent”, but less than for the category “Accepted (if)”. The entry is coloured **blue**.
7. Invitation withdrawn: This is a candidate to whom an invitation has been sent, but then withdrawn, in writing, because there was no response after three reminders.
8. Not available: This is a candidate who has changed employment.
9. Rejected: This candidate has declined the invitation.
10. Replacement: This candidate has recommended a colleague in his/her position.
11. Research: This candidate is researched with the aim of possibly changing status to “Invitation 2b sent”.
12. TBD: This candidate is researched with the aim of possibly changing status to “Research”.

2 The system subtracts marks, so that such candidates move to the bottom of the list.
The status of the Super-List, at the end of this process, is shown in Table 3 (see p. 17).

**Table 3: AB Candidates ranked in order of decreasing score, as of Aug. 2016.**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Status</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Accepted</td>
<td>874%</td>
</tr>
<tr>
<td>2</td>
<td>Accepted</td>
<td>854%</td>
</tr>
<tr>
<td>3</td>
<td>Accepted</td>
<td>851%</td>
</tr>
<tr>
<td>4</td>
<td>Accepted</td>
<td>850%</td>
</tr>
<tr>
<td>5</td>
<td>Accepted</td>
<td>842%</td>
</tr>
<tr>
<td>6</td>
<td>Accepted</td>
<td>839%</td>
</tr>
<tr>
<td>7</td>
<td>Accepted</td>
<td>835%</td>
</tr>
<tr>
<td>8</td>
<td>Accepted</td>
<td>828%</td>
</tr>
<tr>
<td>9</td>
<td>Accepted</td>
<td>828%</td>
</tr>
<tr>
<td>10</td>
<td>Accepted</td>
<td>819%</td>
</tr>
<tr>
<td>11</td>
<td>Accepted</td>
<td>814%</td>
</tr>
<tr>
<td>12</td>
<td>Accepted</td>
<td>812%</td>
</tr>
<tr>
<td>13</td>
<td>Accepted</td>
<td>810%</td>
</tr>
<tr>
<td>14</td>
<td>Accepted</td>
<td>801%</td>
</tr>
<tr>
<td>15</td>
<td>Accepted</td>
<td>796%</td>
</tr>
<tr>
<td>16</td>
<td>Accepted</td>
<td>780%</td>
</tr>
<tr>
<td>17</td>
<td>Accepted (if)</td>
<td>707%</td>
</tr>
<tr>
<td>18</td>
<td>Accepted (if)</td>
<td>705%</td>
</tr>
<tr>
<td>19</td>
<td>Invitation sent</td>
<td>664%</td>
</tr>
<tr>
<td>20</td>
<td>Invitation sent</td>
<td>649%</td>
</tr>
<tr>
<td>21</td>
<td>Invitation sent</td>
<td>635%</td>
</tr>
<tr>
<td>22</td>
<td>Invitation sent</td>
<td>627%</td>
</tr>
<tr>
<td>23</td>
<td>Invitation sent</td>
<td>597%</td>
</tr>
<tr>
<td>24</td>
<td>Invitation sent</td>
<td>589%</td>
</tr>
<tr>
<td>25</td>
<td>Enquiry sent</td>
<td>451%</td>
</tr>
<tr>
<td>26</td>
<td>Enquiry sent</td>
<td>446%</td>
</tr>
<tr>
<td>27</td>
<td>Enquiry sent</td>
<td>409%</td>
</tr>
<tr>
<td>28</td>
<td>Research</td>
<td>336%</td>
</tr>
<tr>
<td>29</td>
<td>Research</td>
<td>143%</td>
</tr>
<tr>
<td>30</td>
<td>TBD</td>
<td>137%</td>
</tr>
<tr>
<td>31</td>
<td>TBD</td>
<td>120%</td>
</tr>
<tr>
<td>32</td>
<td>TBD</td>
<td>117%</td>
</tr>
<tr>
<td>33</td>
<td>TBD</td>
<td>113%</td>
</tr>
<tr>
<td>34</td>
<td>TBD</td>
<td>112%</td>
</tr>
<tr>
<td>35</td>
<td>TBD</td>
<td>112%</td>
</tr>
<tr>
<td>36</td>
<td>TBD</td>
<td>106%</td>
</tr>
<tr>
<td>37</td>
<td>TBD</td>
<td>101%</td>
</tr>
<tr>
<td>38</td>
<td>TBD</td>
<td>99%</td>
</tr>
<tr>
<td>39</td>
<td>Research</td>
<td>74%</td>
</tr>
<tr>
<td>40</td>
<td>TBD</td>
<td>73%</td>
</tr>
<tr>
<td>41</td>
<td>TBD</td>
<td>61%</td>
</tr>
<tr>
<td>42</td>
<td>Rejected</td>
<td>-88%</td>
</tr>
<tr>
<td>43</td>
<td>Replacement</td>
<td>-90%</td>
</tr>
<tr>
<td>44</td>
<td>Affiliated</td>
<td>-90%</td>
</tr>
<tr>
<td>45</td>
<td>Replacement</td>
<td>-91%</td>
</tr>
<tr>
<td>46</td>
<td>Affiliated</td>
<td>-93%</td>
</tr>
<tr>
<td>47</td>
<td>Affiliated</td>
<td>-93%</td>
</tr>
<tr>
<td>48</td>
<td>Rejected</td>
<td>-93%</td>
</tr>
<tr>
<td>49</td>
<td>Not available</td>
<td>-94%</td>
</tr>
<tr>
<td>50</td>
<td>Affiliated</td>
<td>-94%</td>
</tr>
<tr>
<td>51</td>
<td>Invitation with</td>
<td>-94%</td>
</tr>
<tr>
<td>52</td>
<td>Rejected</td>
<td>-95%</td>
</tr>
<tr>
<td>53</td>
<td>Rejected</td>
<td>-96%</td>
</tr>
<tr>
<td>54</td>
<td>Not available</td>
<td>-98%</td>
</tr>
<tr>
<td>55</td>
<td>Replacement</td>
<td>-98%</td>
</tr>
<tr>
<td>56</td>
<td>Affiliated</td>
<td>-98%</td>
</tr>
<tr>
<td>57</td>
<td>Rejected</td>
<td>-99%</td>
</tr>
<tr>
<td>58</td>
<td>Affiliated</td>
<td>-100%</td>
</tr>
<tr>
<td>59</td>
<td>Affiliated</td>
<td>-100%</td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As a result of the above procedure the 16 external members of the Advisory Board were finally selected.
4 Proceedings of the AB1

This chapter gives the proceedings of the Advisory Board Meeting 1 (AB1). The chapter will be partitioned into the AB1 Agenda items (see “Appendix A: Agenda of AB1 (Advisory Board Meeting 1)” Appendix A: Agenda of AB1 (Advisory Board Meeting 1).

4.1 Welcome/Introduction/Advisory Board

Nikos Frydas welcomed the AB members and went on to present briefly the procedure by which the AB external members were selected. For the presentation see Appendix B: Advisory Board presentation, in p. 36.

Following that, each AB member introduced him/her-self.

AB Internal Members:

1. Albena Spasova ICITA
2. Alvaro Ortigosa UAM
3. Christian Castane UM1
4. Cormac Callanan ACONITE
5. Evangelos Markatos FORTH
6. George Pallis UCY
7. Meltini Christodoulaki FORTH
8. Nikos Frydas FORTH
9. Ronan Hardouin INTHEMIS

NOTE: The names of the 16 external AB members are currently withheld.

4.2 Short Introduction to MANDOLA

Evangelos Markatos from FORTH, the project leader, made a short introduction to the MANDOLA project activities:

1. Monitoring the on-line hate speech in the EU. The importance of this activity is nicely determined by Lord Kelvin: “If you cannot measure it you cannot improve it”.
2. Frequently asked questions.
3. Legal:
   a. What is hate speech?
   b. Legal framework in EU member states.

For the presentation see Appendix C: Introduction to MANDOLA presentation, in p. 38.
4.3 Technical Infrastructure

George Pallis from UCY, described the technical infrastructure needed to monitor the spread and penetration of on-line hate-related speech, as well as the necessary reporting tools that will connect citizens with the police.

The presentation briefly referred to the following:

1. Monitoring Dashboard
2. Reporting Portal
3. Data Collection & Processing
4. Monitoring Dashboard Architecture
5. Data Analysis
6. Multi-lingual Corpus
7. Social Scientists
8. Smartphone app

The presentation gave rise to an interesting discussion. Issues discussed include the following:

- Importance and difficulty of measuring tweets & websites.
- Importance of the distinction of the culture and the language.
- Recent progress of artificial intelligence.
- Change of the meaning of words with times.
- Project database use for other categories of illegal content.

For the presentation see Appendix D: Technical Infrastructure presentation, in p. 41.

4.4 Definition of Hate Speech & Legal Framework

Ronan Hardouin, from INTHEMIS, described the work done so far on the Definition of Hate Speech. In particular, a great amount of work has been done on a comparative analysis of the following EU states:

1. Belgium
2. Bulgaria
3. Cyprus
4. France
5. Germany
6. Greece
7. Ireland
8. Netherlands
9. Romania
10. Spain

... for which 18 “potentially” illegal behaviours were identified.

Among the findings the following are included:

- Important disparities between legislations.
- Lack of proper transpositions of International and European legal instruments.
- Coexistence, at the domestic levels, between different provisions targeting close behaviours.

The presentation gave rise to an interesting discussion, which focused on the definition of hate speech, in general, and in particular in categorizing web content and tweets.

For the presentation see Appendix E: Definition of Hate Speech & Legal Framework presentation, in p. 44.
4.5 Brainstorming Panel I

In this session, four questions were given to the AB. For each question, the members wrote their answers on sticky notes, which were then collected, read, displayed on the wall and recorded for processing.

4.5.1 Panel I / Question 1

**Question:** “*What seems to be the most pressing category in hate speech today: LGBT?, racism?, migration?, or other?*”.

**Answers:**

1. **LGBT**: $\frac{1}{2} + \frac{1}{2} = 1$
2. **Racism**: $1 + \frac{1}{2} + 1 + \frac{1}{2} + 1 + 1 + 1 = 7$
3. **Migration**: $\frac{1}{2} + 1 + 1 + \frac{1}{2} + \frac{1}{2} + 1 = 4\frac{1}{2}$
4. **Depends on the region. CEE: LGBT [\frac{1}{2}], W. Europe: Migration [\frac{1}{2}] = 1$
5. **Anti-Islamic**: $\frac{1}{2}$
6. **Refugee / asylum seeker**: $\frac{1}{2} + \frac{1}{2} = 1$
7. **Religion**: $\frac{1}{2} + 1 + \frac{1}{2} + \frac{1}{2} = 2\frac{1}{2}$
8. **Most pressing. Need to understand freedom & responsibility**: 1
9. **Other**: 1
10. **The general l?ck of a ???????4 ~ what constitutes hate speech**: 1
11. **Sexual stereotypes**: 1
12. **Antisemitism**: $\frac{1}{2}$

The above findings may be grouped as following:

---

3 Every member has one ‘vote’. Hence, if a member gives n answers (n=1,2,...) to a question, then each of the member’s answers carries a weight of 1/n.

4 Not clear
A. Sexual: 2½
B. Racism: 7
C. Migration/refugee: 6
D. Religion: 3½
E. Other: 3
4.5.2 Panel I / Question 2

**Question:** “What do you expect to be the most important pressing issue in hate speech in 5 years from now?”

**Answers:**

1. **Migration:** $1 + 1 + \frac{1}{2} + \frac{1}{2} = 3\frac{1}{2}$
2. **Racism:** $1 + 1 + 1 = 3$
3. **Xenophobia:** $\frac{1}{2} + 1 + \frac{1}{2} = 2$
4. **Migration [½] in combination with increase of racism related to religion [½] and lack of integration:** 1
5. **Aliens:** 1
6. **Education of children who will be the future citizens to respect other’s rights:** 1
7. **Islamophobia:** $1 + \frac{1}{2} + \frac{1}{2} = 1\frac{1}{2}$
8. **I don’t know:** 1
9. **LGBT:** 1
10. **Afraid lack of knowledge to debate with responsibility:** 1
11. **Sexual stereotypes:** $\frac{1}{2}$
12. **Hoaxes:** 1
13. **No change:** 1
14. **We risk creating too many hate speech laws sliding down to extreme censorship – the pressing need is their abolition 😊:** 1
15. **Trump:** $\frac{1}{3}$

The above findings may be grouped as following:

A. **Sexual:** $1\frac{1}{2}$
B. **Racism/Xenophobia:** $6\frac{1}{3}$
C. **Migration/refugee:** 3%
D. **Religion:** $2\frac{1}{3}$
E. **Other:** 6%

Comparing the findings of this Question, with those of Question 1, the following findings emerge:

A. **Sexual:** Down from 11% to 8% (27% 🎉)
B. **Racism/Xenophobia:** No change (32%)
C. **Migration/refugee:** Down from 27% to 19% (30% 🎉)
D. **Religion:** Down from 16% to 12% (25% 🎉)
E. **Other:** Up from 14% to 30% (133% 🎊)
4.5.3 Panel I / Question 3

**Question:** “If you could pass one law about hate speech today (either national or in the EU), what would that law be about?”

**Answers:**

1. EU law equivalent to US 1st Amendment to clarify boundary of free speech.
2. A law about religion (freedom & tolerance).
3. We don’t need any more laws, we need other solutions.
4. Laws are not always the solution. More funding for research and education.
5. Humour law to avoid humour being considered as hate speech.
6. Start with building mechanisms to implement properly and fully whatever legislation we have. Enforce better cooperation with authorities upon providers (F/B, YouTube).
7. I would pass the anti-hate speech law. I’m still waiting for mandola’s definition of hate speech, though 😊.
8. Law about regulation. ISPs to have more responsibility in controlling content and for social media to have the obligation to enhance reporting mechanisms.
9. Give equal rights to LGBT.
10. Not allow social media to publish hate speech content.
11. Clear definition combined with social service as sanction for hate speech for individuals’ effective sanction. Aggravating circumstances for hate speech of organisations.
12. EU law → a clear definition on hate speech with harmonise sanction and an obligation for European countries to comply.
13. One law = Reduce + simplify.
14. Prohibition of incitement of hatred in the employment context: Harassment of workers due to certain grounds (sexism, racism, homophobia) or declarations of not hiring individuals belonging to certain groups, or discouraging them from application to certain positions.
15. I would encourage positive speech.
16. Emigration + racism.
17. To criminalize false news making, from entrance right parties / political reasons.
18. A law fully transposing the framework decision on racism, to all found in Article 2A of the EU Charter for Fundamental Rights.
19. One European law to fight against a common approach of hate repression.
20. Facebook should allow LE and HATE get all requested data.
The answers above can be categorized in more than one ways. One such way is the following:

A. Freedom of speech: 4 (20%)
B. Clear/simple definition of hate speech: 4 (20%)
C. No need for more laws: 3 (15%)
D. Internet industry responsibilities: 3 (15%)
E. Racial discrimination: 2 (10%)
F. Legislation about religion: 1 (5%)
G. Sexual discrimination: 1 (5%)
H. Workplace discriminations: 1 (5%)
I. Spreading ‘false news’: 1 (5%)
4.5.4 Panel I / Question 4

Question: “Are the reporting mechanisms of hate speech today enough? If not, how would you improve them?”

Answers:
1. No, not enough. Establish and a code of conduct for FB, YouTube, Instagram, regarding reporting.
2. Not familiar with reporting. I suppose that this means that they must be improved. Diffusion of relative info. Education.
3. NO! Establish different ?? of hate speech and not ??????????.
4. The mechanisms that exist should not overlap with any other. These mechanisms should also be limited to authorities they can act on referred incidents.
5. No improvement with AI.
6. The current reporting mechanism isn’t enough. We need more educational measures on hate and also to teach that free speech comes with responsibilities.
7. Creation of a European reporting Centre lead by Europol + one report centre by country linked to this European Centre.
8. Educate internet user (i.e. you) in school about reporting tool.
9. Basta to wasting money of the tax payers. Prevent crimes – not words! 😊
10. No! No idea how.
11. Make it easier to ANONYMOUSLY report hate speech.
12. Feedback should be given to the user who reported about the actions that were undertaken.
14. Each country should have a reporting point for hate speech. The answers to the report should be rapid to remove illegal content.
15. No, you need a specialized unit, to deal with citizens reports and to educate citizens.
16. Current reporting “HOTLINE”, concept is enough. BUT: Would need far greater resources, technology, training AND security for personnel.
17. Not enough: Joint cooperation CSOs, national Government and IT companies. Build capacities of monitors according [to] European standards so that to have comparable data. Involvement of law enforcement.
18. Reporting mechanism should be imposed with more sophisticated techniques – automatic alert systems.

5 Not clear
19. More automatization to detect hate speech. Reporting mechanism that provide feedback to users, and ....... Guide them to find support. Better coordination among organizations, companies daily w/ Reports → Identify Best Practices.

20. No. Reporting mechanisms are not sufficient. Improve them through the research of AI / Machine Learning.
4.6 Brainstorming Panel I

In this session, four questions were given to the AB. For each question, the members wrote their answers on sticky notes, which were then collected, read, displayed on the wall and recorded for processing.

4.6.1 Panel I / Question 1

**Question:** “What are the difficulties for industry to respond in this area? Complexity? Legality, Liability?, or Other?”

**Answers:**

1. **Goodwill** and Commitment.
2. **Authority**, Validity, Cost, Liability.
3. Lack of legal clarity, Misunderstanding of what constitutes Internet industry and who may be responsible, Complexity of hate speech decisions, Cross-jurisdictional nature of services.
4. **Legality**.
5. The state shifting its responsibility to industry.
6. **Freedom of speech**.
7. Lack of understanding, knowledge, especially for SMEs. Lack of strong regulation to oblige them to comply. Lack of good “incentives” to take this seriously.
8. Enforcement: Industry should withdraw from removing material unless found illegal by court or on the demand of state prosecutors, not on the basis of their codes of conduct.
10. Other priorities dealing with hate speech is not their core business, so not a priority.
11. **Complexity**: Conflict btw “community standard” and perception of people who report.
12. **Legality**: They are not judge! They can’t know what an illegal content is.
13. **Complexity, Legality.**
14. **Legality and liability.**
15. **Liability, freedom of speech, cultural feelings.**
16. **Legality.**
17. Accusation of overstepping boundaries, private policing, etc. Terrified of intro of legislation for liability.
18. **Lack of interest.**
19. **Freedom of speech.**
20. **Complexity.** Difficulty to strike a balance between competing values and rights.
22. Just an academic here: I **don't know** what outer world looks like.

The responses above can be categorized as following:

A. **Legality:** $1 + 1 + 1 + \frac{1}{2} + 1 + \frac{1}{2} + 1 = 5\%$ (27%)

B. **Freedom of speech:** $\frac{1}{5} + 1 + 1 + 1 + 1 = 5\%$ (24%)

C. **Complexity/Cost:** $1 + \frac{1}{2} + 1 + \frac{1}{2} + 1 + \frac{1}{2} + 1 = 4\%$ (20%)

D. **Liability:** $\frac{1}{2} + 1 + 1 + \frac{1}{2} = 2\%$ (13%)

E. **Other:** $\frac{1}{2} + 1 + 1 = 2\%$ (12%)

F. **Don't know:** 1 (4%)
4.6.2 Panel II / Question 2

**Question:** “Are there useful working models for this space: INHOPE?, INACH?, Other?”.

**Answers:**

1. **INHOPE**: 1+1+1+1
2. I am all ears to learn and apply back home: 1
3. **Not useful** working models: 1
4. **Not familiar** with these models: 1
5. Hate speech usually directed at specific individuals, **no new/emerging user empowerment approaches/tools**: 1
6. **Je ne sais pas**: 1
7. **OTHER! The topic is too specific**: 1
8. **EUROPOL, C3I**: 1
9. Yes **INHOPE / INACH** type models are a useful starting point, but these need considerable reworking (with appropriate legislative backing) to be workable and effective: 1
10. **INHOPE**. Not exactly familiar enough to suggest something specific.
11. **INACH**. Mandola reporting portal: 1

The responses above can be categorized as following:

A. **INHOPE**: 4 + ½ + 1 = 5½ (39%)
B. **Other**: 1 + 1 + 1 = 3 (21%)
C. **Don't know**: 1 + 1 + 1 = 3 (21%)
D. **INACH**: ½ + 1 = 1½ (11%)
E. **EUROPOL/C3I**: 1 (7%)
4.6.3 Panel II / Question 3

**Question:** “What are the Challenges for current reporting points responding to hate speech?”.

**Answers:**

1. To **assess** hate speech. To protect **free speech** while countering hate-related comments.
2. The Internet – **Trans border**. Speech & **Cultural differences**.
3. Absence of **clarity** about what hated speech is, and it’s a good thing too.
4. To be able to **action the report**. To have real impact. If non-profit, HR & **financial** resources. Sustainability. **Coordinate** their efforts with other stakeholders in the field.
5. **Legal** complexity. **Resources**, inability to take decisive action to have content **removed**. Especially **across jurisdictions**.
6. No **follow up** procedures in place. So even if it is reported no measures are taken.
7. **Funding, Legal** uncertainty. **International cooperation** versus cultural contexts.
8. Verification of information. Lack of service support for those who are targeted → consequences → **under-reporting**.
9. Lack of **quick response** from the part of enforcement authorities. Lack of instruments for systematic **monitoring**.
10. **Analysis & Legality**.
11. **Legality** or illegality of ??? ???. Responsibility what **follow up** to give?
12. The **legality** regarding the removal of illegal content.
13. People do not **report** hate speech.
14. **Legal, Financial**.
16. Privacy. Data protection. Copyright **law**. Hate speech definition (sorry, joke).
17. Doing what people are **aware** when they have hate speech.
18. **Funding**.
19. Determine **legal/illegal**. **Analyse**. Report to political level. PPP.

The responses above can be categorized as following:

**A. Legal issues:** \(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} = 5(1/12)\)

**B. Reporting/Analysis:** \(\frac{1}{2} + 1 + 1 + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + 1 + \frac{1}{12} = 5\)

**C. Effectiveness:** \(\frac{1}{2} + \frac{1}{4} + 1 + \frac{1}{7} + \frac{1}{8} + 1 = 3(7/12)\)
D. Funding: \( \frac{1}{3} + 1 + \frac{1}{4} + \frac{1}{4} = 2\left(\frac{5}{12}\right) \)

E. Coordination: \( \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1\left(\frac{5}{12}\right) \)

F. Hate speech definition: 1

G. Free speech: \( \frac{1}{2} \)
4.6.4 Panel II / Question 4

Question: “When does Hate Speech lead to Hate Crime? What conditions to you need for hate crime to occur after hate speech?”

Answers:

1. Research has shown that there is a relation between hate-speech and hate-crime however, this question seems to be an open problem.
2. Lack of response to stop it at an early stage leads to a chain of hate speech message that encourage to go further.
3. Conspiracy theory.
4. Hate crime has been decided before the hate speech. Hate speech is the warning.
5. Fear & Ignorance.
6. When it heard and believed.
7. Transitional economically poor societies with a domineering religious and political propaganda.
8. Not necessarily but cultivates the social acceptance (not reacting) to hate crime. Legitimates crime.
9. Radicalization.
11. When it is sustainable.
12. In cases when no one reacts to hate speech and informs the responsible bodies.
13. Fundamentalism.
5 Conclusions & Lessons Learned

This chapter gives the conclusions and lessons learned from AB1.

The size of the AB (16 external and 9 internal members) appears to be working very well, as the time available (6.5 hours gross time) allowed each member to be able to contribute more than one times. If space allows, AB2 may grow to 20+9 members.

Some topics, if presented adequately, may benefit from an AB debate, but the time available would allow only one (perhaps two) such cases.

The ‘sticky-notes’ brainstorming sessions are very productive and allow for the collection of hard evidence from each member. In AB1 it was possible to conduct eight such sessions. The results are available in the current report (§4.5 in p. 21 and §4.6 in p. 28).
6 Appendix A: Agenda of AB1 (Advisory Board Meeting 1)

First MANDOLA Advisory Board Meeting - AGENDA
October 5, 2016
European Office of Cyprus, Rue du Luxembourg 3, 2nd floor B-1000 Brussels

10:00-10:20 Welcome/Introductions/Advisory Board Nikos Frydas
10:20-10:40 Short Introduction to MANDOLA Vangelis Markatos
10:40-11:00 Technical Infrastructure George Pallis
11:00-11:20 Definition of hate speech and Legal Framework ¹ Ronan Hardouin
11:20-11:40 Coffee Break
  • Questions:
    o If any, what seems to be the most pressing category in hate speech today: LGBT? racism? migration? other?
    o What do you expect to be the most important pressing issue in hate speech in 5 years from now?
    o If you could pass one law about hate speech today (either national or in the EU), what would that law be about? In particular, what do you think about punishing hate speech whatever the motivating victim's characteristics (physical, genetic, psychological, philosophical, or behavioral)?
    o Are the reporting mechanisms of hate speech today enough? If not, how would you improve them?

13:00-14:00 Lunch Break
14:00-15:00 Responding to on-line hate speech - FAQs ² Cormac Callanan,
Best Practice Guide Vangelis Markatos & Nikos Frydas
  • Discussion:
    o What are considered the best-of-breed responses to hate speech online?
    o What are the difficulties for industry to respond in this area? Complexity? Legality, Liability? Other?
    o Are there useful working models for this space: INHOPE? INACH? Other?
    o Challenges for current reporting points responding to hate speech?

15:00-15:20 Coffee Break
15:20-16:15 Discussion (continued)
16:15-16:30 Closing session

¹ See D2.1: Intermediate Report - Definition of Illegal Hatred ..., in http://mandola-project.eu/publications/.
² See D4.1: FAQ on Responding to on-line hate speech, in http://mandola-project.eu/publications/.
7 Appendix B: Advisory Board presentation
Welcome!

- Let us now move on with our Agenda.
- Before that, may we introduce each other?

Thank you

Nikos Frydas
FORTH (ext. consultant)
8 Appendix C: Introduction to MANDOLA presentation

MANDOLA: Monitoring and Detecting on-line Hate Speech

Evangelos Markatos
FORTH and U of Crete

Our Background
- Partners of the project are active in
  - Cyber security
  - Cybercrime research and education
  - Illegal Internet Content Reporting

Our Background – Cyber security
- SysSec: European Network of Excellence in Cybersecurity
  - Founding coordinator
  - Editor of Red Book in Cybersecurity
- NIS (Network and Information Security) Platform
- Strategic Research and Innovation Agenda
- ECSO: European Cybersecurity Organization
  - Founding members (Spain and Greece)
  - (all) member of the Partnership Board

Our Background – Cybercrime Research and Education
- National Centers of Excellence in cybercrime
  - Founding coordinators in
    - Spain, Greece, and Bulgaria
- Safeline: Greek Hotline to report illegal Internet content
  - Founder and first coordinator
- SENTER: European Network of National Centers of Excellence in cybercrime
  - Founding member and WP leader

What do we want to do in MANDOLA?
- Monitor
  - the spread and penetration of
  - on-line hate speech in EU
- Call for proposals (from 2014) reads:
  "Monitoring and reporting on hate crime and on-line hate speech (HATE)"

Why?
- If you can not measure it you can not improve it
  Lord Kelvin
- Measurements are at the core
  - of solid conclusions
  - and sound actions
How to measure it?

- Twitter-based approach
  1. Collect tweets
  2. Apply a hate-detection filter
  3. Apply sentiment analysis
    - To find hate speech (http://www.risk.org/)

Various views I

- Category
  
<table>
<thead>
<tr>
<th>Category</th>
<th>Hate Rate per Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Religion</td>
<td>20%</td>
</tr>
<tr>
<td>Gender</td>
<td>15%</td>
</tr>
<tr>
<td>Nationality</td>
<td>10%</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>5%</td>
</tr>
<tr>
<td>Age</td>
<td>3%</td>
</tr>
<tr>
<td>Gender</td>
<td>2%</td>
</tr>
</tbody>
</table>

Various views II

- Timeline

Various views III

- Per country Category

Other activities?
Frequently Asked Questions

- FAQ Book
- What to do?
  - I believe I have encountered hate speech. What can I do about it?
  - Why words can be harmful and sometimes dangerous?
  - I believe I have been a target of hate speech. What are my rights as a victim?
- More at the talk at 14:00

Other activities? Legal

- Definition
  - What is hate speech
- Legal Framework
  - In Member States
- More at the talk at 11:00
Who?
- FORTH (coordinator), GR
- U of Cyprus, CY
- AIS, IE
- ICITA, BG
  - Bulgarian Center of Excellence in cybercrime
- UAM, SP
- Spanish Centers of Excellence in cybercrime
- UMO, FR
- Member of the French CoE in cybercrime
- INTHEMIS, FR

Why are we here today?
- We want your advice
  - Advisory Board
- You know a lot about this area
  - Can you share some of your knowledge?
  - Can you share some of your experience?
  - Some of your wisdom?

How are we going to get this advice?
- Interactive Brain storming sessions
- Post it notes
- Share your advice
  - Even half-baked ideas
- All ideas!
  - In research all ideas are welcome

Brain storming sessions
- We are going to ask you questions:
  - e.g. "If you could pass one law in hate speech, what would it be?"
  - Think "out of the box"
  - Share your knowledge
  - Share your ideas

A final note
- If you want to make a short presentation
  - Of your activities
  - Let me know
  - We have some slots
  - < 5 minutes long...

Thank you
Evangelos Markatos
FORTH and U of Crete
9 Appendix D: Technical Infrastructure presentation

- MANDOLA: Monitoring and Detecting on-line Hate Speech
  - Technical Infrastructure
  - George Pallis
  - University of Cyprus

Objectives

- To provide the technical infrastructure needed to monitor the spread and penetration of online hate-related speech using automated tools based on big-data analysis.
- To provide the necessary reporting tools that will connect citizens with the police enabling them to report online hate-related speech (such as hate-related web sites) and hate-related criminal activities.

Monitoring Dashboard

- Heat Map
- Hot Spot

Monitoring Dashboard

- Statistics
- Heat map:
  - Geo-clustering
  - Date range bar
  - Filtering Categories
  - Hot-spot density map
  - Hate-rate metric
  - Date range filtering
  - Categories filtering
  - Country drill-down
- Statistics:
  - Time-line chart
  - Zoom functionality
  - Data zoom aggregation
  - Language usage chart
  - Hate rate per category chart
  - Hate rate per country chart
  - Hate rate per city chart
  - Countries per category chart
  - Cities per category chart
  - Time-line per category chart
  - Hate strength gauge
Data collection and processing
- The Data collection consists of two sub-modules that are responsible for collecting data from Twitter and Google API.
- Twitter data collection framework (developed by UCY) is based on the use of multiple DSN accounts, which are engaged in a distributed collection process without violating the terms of use. Retrievals about 11 million tweets per day.
- Google data collection framework (developed by UAI) is based on a meta-search engine that provides the content and information for in-depth analysis of possible hate-related speech.
- No sensitive data is stored during the processing. The only information stored is a) the hate processing output, b) the date it was published or updated, c) the language and d) the location. The location is converted in geo-hash with accuracy reduced to the level of city.
- The processing and storing is in line with article 7(1)(c) of the Personal Data Protection (Protection of the Human) Laws of 2001 to 2012 in Cyprus (N. 138(0)/2001), which was submitted to the Cyprus Data Protection Commissioner on 19th December 2015.

Data Analysis
- The data streams are handled through a distributed publish-subscribe messaging system named Apache Kafka.
- Kafka feeds the hate-speech data analysis module and also collects a data sample set in order to create the multi-lingual corpus.
- The hate-speech data analysis module utilizes sentiment analysis tools via the NLTK platform in order to classify content whether it is hate-related speech or not.
- Classification process is carried out via Apache Spark Streaming stream-source framework. Apache Spark framework includes a suite of machine learning algorithms, called libraries.
- When the processing is done, the output is stored in the hate speech database (MongoDB). The Dashboard is connected with the MongoDB via an API that is used to retrieve data required for the various types of data visualization supported by the Dashboard.

Multi-lingual Corpus
- The multi-lingual corpus is used to train the classification model that exist in the hate-speech data analysis module.
- Social scientists classify the hate-related content in the given Twitter and Google sample set, based on its strength and its categories.
- The hate filtering module is used to conduct automatically an initial filtering of the sample set so that social scientists receive for review more relevant content (i.e., hate-related).
- The corpus has been initially built from hate databases such as the counter-hate database iHateWeb, and from hate-related sentiment lexicons containing seed words such as the APHNK lexicon.
Social Scientists

Smartphone app
- A smartphone app will be developed that will allow anonymous reporting of hate-related speech material that has been noticed in the Web and the social media.
- The smartphone app will enable people to report hate-speech including racism-related and hate-related material.
- The app will be compatible for iOS, Android and Windows mobile devices.

Thank you!

Questions?
10 Appendix E: Definition of Hate Speech & Legal Framework presentation
T2.1 – Definition of hate speech

18. In short – illegal in all or almost studied States:
- Publicly inciting hatred or violence or discrimination directed against a group of persons or a member of such a group, based on any ground or certain characteristics of the victim
- Making available to the public xenophobic or racist material which incites hatred or violence or discrimination, which promotes hatred, discrimination, or violence, through a computer system
- Publicly insulting a person or a group of persons based on any ground or certain characteristics of the victim;
- Public defamation, based on any ground or certain characteristics of the victim;
- Direct or indirect discrimination, including harassment, in certain specified areas, by reason of some characteristics of the victim

T2.1 – First outcomes

- Important disparities between legislations:
  - Most of transpositions of international and European instruments have not been done the same way, in addition to the fact that most countries provide for additional prohibitions
  - Example (one of the most common offence): incitement to hatred or violence, directed against a group of persons / a person determined on the basis of their race, national or ethnic origin, and religion, (eventually) if the incitement is either carried out in a manner likely to disturb public order or is threatening, abusive or insulting (CFD 2008/913/2HA)

T2.1 – First outcomes – Legislations disparities

- Example “Incitement to hatred / violence” (follow up)
  - 10 countries criminalise the incitement to hatred, but only 8 of them criminalise the incitement to violence;
  - 8 countries additionally criminalise the incitement to discrimination (not mentioned in the CFD, but in the Convention on cybercrime);
  - Only 3 countries impose an additional condition, and therefore only prohibit the public incitement to hatred if it is either carried out in a manner likely to disturb public order (2 countries), or public peace (1 country), or if it is threatening, abusive or insulting (1 country, alternatively to the disturbing of public order).
T2.1 – Definition of hate speech

- T2.1 – First outcomes
  - Lack of proper transpositions of International and European legal instruments:
    - Disparities that are noticed are often, firstly, the result of a lack of proper transposition of International and European instruments.
    - For ex., only 1 country (Cyprus) fully criminalises the public condoning, denying or grossly trivialising crimes against peace, genocide, against humanity and war crimes directed against a group / a person defined by reference to race, colour, religion, descent or national / ethnic origin (2008/913/JHA).
    - The other countries punish parts of it (ex:1: Holocaust deny; ex:2: to justify, deny or grossly palliate a crime committed against peace and humanity)

- T2.1 – First conclusions
  - Difficulty to provide a list of behaviours that are prohibited in all the studied member States (unless reducing this list to a very limited number of illegal acts)... therefore difficulty to provide a simple and short definition;
  - Non-equal treatment between victims, even legal insecurity, since one action might be punished or not depending on the characteristics of the victim that based the action, in one context or the other, even in one single country, and depending on the country that will be competent to judge the case.

- T2.1 – Final deliverable
  - Will include an attempt to a simplified definition of illegal hate speech (ongoing);
  - Will include a deeper analysis of the right to freedom of expression;
  - Should include recommendations to policy makers, based on final outcomes;
  - Any comment or wish for a particular focus of the analysis are welcome!
11 Appendix F: Brainstorming Panel I / Question 1

Brainstorming Panel I

- What seems to be the most pressing category in hate speech today:
  - LGBT?
  - racism?
  - migration?
  - other?
12 Appendix G: Brainstorming Panel I / Question 2

Brainstorming Panel I

- What do you expect to be
- the most important pressing issue in hate speech
- in 5 years from now?

Migration

Still racism

Aliens

LGBT

Racism

Migration

Acceptance

I don’t know

5 years

HOAXES

Ko course!
Appendix H: Brainstorming Panel I / Question 3
14 Appendix I: Brainstorming Panel I / Question 4
15 Appendix J: Brainstorming Panel II / Question 1

Brain Storming Panel II

- What are the difficulties for industry to respond in this area?
  - Complexity?
  - Legality, Liability?
  - Other?
Appendix K: Brainstorming Panel II / Question 2

- Are there useful working models for this space:
  - INHOPE?
  - INACH?
  - Other?
17 Appendix L: Brainstorming Panel II / Question 3
Brainstorming Panel II / Question 4

- When does Hate Speech lead to Hate Crime?
- What conditions do you need for hate crime to occur after hate speech?