Report of OSCE-ODIHR Expert Meeting

Role of the Internet industry in addressing hate on the Internet

Amsterdam, 10 May 2010

Introduction

The issue of hate on the Internet and its potential to stimulate hate crimes is a matter of growing concern for a number of countries and inter-governmental organizations, including the OSCE. In recent years, there is a perception that the occurrence of hate-inciting material on the Internet has grown exponentially. In order to address this issue, the OSCE Ministerial Council mandated the ODIHR in Ministerial Council Decision 9/09 of December 2009[1] to explore this phenomenon further and to identify practical steps to be taken.

There is an expanding body of evidence confirming the relationship between online hate and hate crimes. This issue was explored at an expert meeting held on 22 March 2010 in Warsaw on *Incitement to Hatred vs. Freedom of Expression: Challenges of combating hate crimes motivated by hate on the Internet*. The high level of interest in this topic which was demonstrated at the meeting confirmed that this issue deserves further attention by the OSCE. One conclusion of the meeting was that neither individual States nor intergovernmental organizations can tackle this issue successfully without close cooperation with other actors.

Focus and outcomes

With this background, the ODIHR convened a meeting in Amsterdam on 10 May to foster discussion of these issues among representatives of States, international organizations, NGOs, and the Internet industry. The Amsterdam meeting was a direct follow up to the Warsaw meeting, where one of the main conclusions was the need to hold a meeting with the representatives of the Internet industry. In order to enable in-depth and focused discussion, the Amsterdam meeting was limited to a small group of participants.[2] The main goal of the meeting was to engage with representatives of the Internet industry in an open discussion on how to co-operate effectively to respond to manifestations of hate on the Internet, without curtailing freedom of expression.

[1] Athens Ministerial Council Decision 9/09 of 2 December 2009 on Combating Hate Crimes, http://www.osce.org/documents/cio/2009/12/41853_en.pdf.
[2] A full list of participants can be found at the end of this report.

It was considered vital to have Internet industry representatives involved in these discussions, since without their cooperation and support, implementation of any recommended measures would be very difficult, if not impossible. Meetings of this type are an important contribution to the policy-making process as they provide a platform for a variety of stakeholders, including civil society, to express their opinions and concerns and to generate recommendations that can be conveyed to governments and to the broader OSCE community.

This report presents a summary of the main topics discussed in the three sessions of the Amsterdam meeting, as well as recommendations formulated for OSCE participating States, ODIHR, civil society and the Internet industry. Views and positions presented in this report do not necessarily represent the policy or position of the OSCE or ODIHR.

Freedom of expression

Several participants stressed that freedom of expression is a fundamental right and warned that any recommended measures to address hate on the Internet should be assessed carefully against their possible impact on freedom of expression. One recommendation was that any infringement of freedom of expression on the Internet (including removal of "objectionable" content) should be mandated by a judicial authority, not based only on individual judgement.

Some participants commented that there seems to be a growing view in a number of countries that access to the Internet is an individual right, closely linked to freedom of expression; thus any restriction on such access might be seen as a violation of this right.

Legislation and its enforcement

Participants generally agreed that there is no need to adopt new legislation regulating Internet content. Instead, they considered that the first priorities should be proper enforcement of existing legislation and closer cooperation among different actors in the implementation and enforcement of such legislation. Some participants noted that while laws on Internet content exist at the national level, there are no such laws on the international level, except within the European Union (EU).

The lack of international standards results in part from the absence of an international consensus on appropriate ways to address hate speech.

Some speakers noted that enforcement of laws related to inappropriate or illegal Internet content (such as child pornography, pirated software or audio-visual material, or hate-inciting content) is very difficult and usually ineffective. A number of participants pointed out that laws are implemented effectively only when there is a political will to do so. Where political leadership is lacking, the issue of "cyberhate" is very low on the agenda of law enforcement agencies, leading to slow or inadequate official responses to such cases.

An important issue related to legislation is the need to establish the roles and responsibilities of different actors involved in dealing with Internet content. These include Internet Service Providers (ISPs), Internet companies, law enforcement agencies, NGOs, independent monitoring groups and complaints mechanisms. In this regard, participants made positive reference to the EU e-commerce directive[3] which established a "notice and take down procedure". The e-commerce directive stipulates that once an ISP or owner of a website is notified of illegal content hosted on the site/service, it is the owner's duty to remove it or face legal responsibility for hosting such content. This EU directive was seen as a possible guiding principle for any further measures.

Impact of hate-inciting content

Many participants voiced concerns about the impact of web 2.0 on various communities and the occurrence of hate-inciting material on sites with user created content and on social networking sites. In the view of some participants, the latest technological developments on the Internet – which have increased the role of individual Internet users through user created content (web 2.0) – have resulted in mainstreaming previously marginal views through the power of social networking sites.

[3] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce).
http://www.ebu.ch/CMSimages/en/leg_ref_ec_directive_e_commerce_080600_tcm6-4338.pdf.

With the increasing number of Internet sites that propagate hate, the communities targeted by such material often feel threatened and unsafe. The impact of hate-inciting online content is therefore not limited to the individual level, but affects entire communities. As a result, addressing this trend effectively requires looking at its impact on communities and involving communities in the response.

Other participants stressed that the objectionable content found on some sites merely reflects broader problems in society, with all its positive and negative aspects. Internet sites, therefore, should not be blamed for the picture they show, since they only mirror the attitudes and opinions already existing in society.

Challenges of the Internet industry

Representatives of the Internet industry stressed that the industry is aware of the problem of online hate and is committed to address it appropriately. The representatives welcomed the invitation to participate in a dialogue on this subject. However, they noted that the Internet industry is faced with three types of challenges in addressing the issue of hate on the Internet:

- Technical challenges regarding what kinds of monitoring and prevention are technically feasible, in light of the huge amount of data and content uploaded by users;
- Legal challenges, including unclear definitions of what kind of content is and is not legal; and
- Political challenges, including both the amount of attention given to this issue and conflicting views in different countries on what constitutes hate content and what is legitimate expression. Currently, there is no consensus in the industry or at the international level on what "hateful" or "objectionable" content actually means.

Industry representatives contended that due to these limitations, it should not be up to them to decide what is acceptable and what is not in terms of Internet content. The industry needs clear guidelines based on national and international law. Since the definition of objectionable or inappropriate content is vague and inconsistent, a more precise definition is needed if Internet companies are to use such terminology in their terms of service.

Some participants suggested that codes of conduct and principles of corporate responsibility could be developed and used as guidance for companies on how to establish their own policies on "cyberhate". It was stressed that if such policies are to be effective, an independent monitoring system should be set up as well.

Net neutrality

Another topic mentioned by several participants is the issue of Internet neutrality. Some industry representatives noted that global Internet companies would not agree to any form of censorship since they adhere to the principle of net neutrality. They noted that the Internet industry needs a trusted framework for discussion of contentious issues such as "cyberhate". They therefore appreciated the consultative approach based on partnership and discussion, as reflected in this meeting, and considered it far preferable to the "naming and shaming" approach sometimes adopted by critics. In this regard, the Global Network Initiative[4] was mentioned as an example of a collaborative approach towards addressing attempts to restrict freedom of speech and privacy by some governments.

One particular aspect which has to be taken into consideration in drafting recommendations for the Internet industry is the competition among different platforms and websites. Users are aware of this competition and if one online platform enacts strong restrictions on publishing "hateful" content, users will simply migrate to a different platform with fewer restrictions.

Judicial panels

As a practical solution to the problem of establishing a competent body to issue decisions to remove certain online content, it was suggested to set up judicial panels tasked with reviewing and

[4] The Global Network Initiative was set up as a reaction to increasing government pressure in many regions of the world on the information and communications technology industry (ICT) to comply with domestic laws and policies in ways that may conflict with the internationally recognized human rights of freedom of expression and privacy. The main aim of the initiative is to protect and advance freedom of expression and privacy in the ICT sector. More information is available at: http://www.globalnetworkinitiative.org/index.php.

deciding upon removal of "objectionable" Internet content. Industry representatives welcomed this idea, since having the "takedown notice" (request to remove online content) originating from an independent judicial panel would be far preferable to the current situation, when such requests can be made by a number of different individuals or bodies. This idea needs to be further elaborated though, since there are some practical issues which need to be solved, such as the jurisdiction of such panels, their capacity to deal with potentially large workloads, the possible need for pre-screening mechanisms, and the question of appeal mechanisms.

Dialogue

A number of participants pointed out that the dialogue among different actors involved in addressing the hate on the Internet should be further strengthened. Law enforcement agencies, the Internet industry and NGOs should all have a role, since only through such common dialogue can solutions be identified and consensus reached. Participants suggested using existing fora and mechanisms more effectively to ensure full participation of all stakeholders.

Terms of service as a tool for removal of hateful/illegal content

An issue highlighted by many participants was the use of ISP terms of service agreements as an effective tool for removal of hate-inciting content. Experience shows that the approach taken by some Internet companies to remove hate-inciting material on the basis of breaches of their terms of service agreements is much more effective than using the criminal justice system to identify and prosecute the authors of such material. Identifying the authors of any type of illegal online content is often difficult and time-consuming, and rarely leads to removal of the content in question if it is hosted outside of national jurisdiction. Moreover, due to the high number of potential cases, criminal justice systems would be overloaded if they attempted to process all cases.

Some participants proposed exploring the idea of harmonizing the terms of service used by different Internet companies in order to have a global and unified response to hate-inciting content on the Internet. According to some NGO representatives, 80 per cent of the terms of service

agreements used by Internet industry are almost identical, so harmonization should be possible, once there is political will to do so.

NGO representatives described the effectiveness of one of their current methodologies, which is based primarily on the use of terms of service agreements and dialogue with the Industry. Under this approach, NGOs have had a high ratio of success in having hate inciting Internet content removed once the hosting company, website owner or ISP is notified of its existence by a complaints bureau. The ratio of content removal after notification by complaints bureaus is 90 per cent in the Netherlands and 80 per cent in Germany.

On the other hand, it was also noted that there are some major Internet companies which react only if cases of "cyberhate" hosted on their sites are reported in the media, since their complaints mechanisms do not work properly or can not handle the volume of complaints they receive. It was the general view of participants that the Internet industry can successfully tackle hate on the Internet only by combining comprehensive terms of service agreements with a robust, visible, accessible and transparent complaints handling mechanism, which ensures swift and timely response. Some participants argued that owners and administrators of websites – including social networking sites – have a responsibility to set up complaints mechanisms which fulfil these criteria.

Hyves – an example of a self-regulatory complaints mechanism

The representative of Hyves, which is the largest social networking site in the Netherlands described in detail its complaints mechanism as an illustration of a robust and efficient self-regulatory mechanism supported by a clear policy on inappropriate content based on national legislation. Hyves uses two methods for notification of objectionable content on their social networking website. The first method requires at least ten notifications (originating from ten different Internet Protocol (IP) addresses) to flag content as inappropriate by clicking on a specific icon visible on every page of the Hyves website. The second option is to send an email to the Hyves help desk containing the link to the objectionable content. Types of content which could be flagged as "inappropriate" include profile, group, picture or public postings, but not a message in a discussion. After the complaint is received, the material is temporary disabled, reviewed by the

Hyves staff and either deleted or made visible again if it does not violate the terms of service. Both the author of the complaint and the owner/creator of the content in question are notified of the outcome within 48 hours after the complaint was filed. Any content which is deleted cannot be uploaded again from the same IP address.

The Internet industry representatives acknowledged that there is a need for training of complaints handling staff to recognize and understand the meaning of hate symbols and codes used by various organized hate groups in order to assess the nature of online content. It was suggested that existing online resources (such as the ODIHR database of hate symbols[5] or the Anti-Defamation League database of hate symbols[6]) should be used more broadly and made available to the complaints handling staff of ISPs for this purpose.

Law enforcement agencies

Several speakers noted that the current level of cooperation among law enforcement agencies on the international level is not sufficient and needs to be enhanced. At the same time, they stressed, the capacity and skills of national law enforcement structures to handle cases related to "cyberhate" needs to be further strengthened. In order to enhance and institutionalize cooperation between law enforcement agencies and the Internet industry, points of contact for Internet related issues and cybercrime should be established in participating States' national law enforcement agencies. Creation of specialized police units tasked with combating cybercrime, including hate on the Internet, would also help to enhance and institutionalise co-operation between police and ISPs.

[5] This database is not publicly available and will be launched in the second half of the year.
[6] The Anti-Defamation League database is available at http://www.adl.org/hate_symbols/default.asp.

Recommendations:

The following provides a summary of recommendations that were discussed at the expert meeting. These recommendations were not adopted by the participants and they do not necessarily reflect the consensus of the meeting participants. They are directed to participating States, OSCE institutions, civil society and the Internet industry.

Recommendations to governments and international organizations

Freedom of expression

- Ensure that the Internet continues to be an open and public forum for freedom of expression and free media. The right to receive and disseminate information – including through the Internet – is a basic human right as promulgated by the Universal Declaration of Human Rights and numerous human rights treaties and documents;
- Ensure that laws prohibiting bias-motivated speech are not being enforced in a discriminatory or selective manner to impede or silence dissent, political criticism or alternative opinions.

Studies and analysis

- Conduct studies of the possible relationship between racist, xenophobic, and anti-Semitic speech on the Internet and the commission of bias-motivated crimes;
- Investigate the link between the existence of "hit lists" on the Internet and the commission of hate crimes;
- Commission a study on potential hate crime offenders, looking at different stages involved in planning and carrying out of hate crimes and the psychological aspects involved;
- Evaluate existing empirical research, to identify the extent of hate material on the Internet, the impact of exposure to such hate material on young people and the material's direct link to hate crimes.

Education/prevention

- Develop educational programmes and training materials for young people about bias-motivated expression on the Internet;
- Promote and support media literacy programmes, including technical and textual Internet literacy;
- Develop, test, analyze and broadly implement educational concepts aimed at combating online hate speech;

- Increase parental awareness of widely available filtering software.

Legislation and law enforcement

- Internet Service Providers shall not be held generally liable for content without them having been made aware of the illegality of the information transmitted and given sufficient time to act appropriately;
- Explore possibilities for improved implementation of current legislation prior to recommending adoption of new laws;
- Refrain from introducing vague and over-broad definitions related to hate speech while adopting new legislation;
- Strengthen enforcement of existing legislation addressing "cyberhate", including by investigating and fully prosecuting criminal incitement to violence on the Internet;
- Train investigators and prosecutors on how to address bias-motivated crimes incited by material posted on the Internet;
- Share information on successful training programmes as part of an exchange of best practices;
- Support specialisation of law enforcement and prosecutors dealing with cybercrime and hate crimes;
- Strengthen and institutionalise the dialogue between law enforcement agencies, the Internet industry and civil society on issues related to "cyberhate";
- Enhance international law enforcement co-operation on "cyberhate" and cybercrime-related issues;
- Create official points of contact in the law enforcement agencies of participating States for Internet related issues, in order to streamline cooperation and exchange of information;
- Explore the idea of setting up a trusted authority (such as an independent judicial panel) tasked with taking decisions on "objectionable" online content and issuing "take down notices" to Internet Service Providers (ISPs).

Monitoring and complaints mechanisms

- Support self-regulatory and independent monitoring mechanisms that collect and share data and statistics on hate on the Internet.

Cooperation with other actors

- Involve ISPs and Internet companies in any future discussions on the topic of hate on the Internet;
- Develop cross-border exchanges of information on best practices adopted by different countries and organizations;

- Build partnerships among national agencies, NGOs and the Internet industry to monitor instances of hate on the Internet;
- Facilitate States' agreement to find universally acceptable responses to reduce the harm caused by online hate material through:
    - The identification and dissemination of voluntary agreements between Internet providers and users that balance freedom of expression against the need to reduce harm;
    - Work with civil society to counter the negative narrative contained in hate inciting material;
    - Support for NGO efforts to address "cyberhate".

Recommendations to the Internet Industry:

- Employ clear and comprehensive terms of service as a basis to take appropriate action against sites inciting hate;
- Encourage ISPs to inform parents about ways in which they can exercise greater supervision of their children and protect them from viewing objectionable material on the Internet;
- Use and promote industry codes of conduct, ethical guidelines and principles as a tool for addressing online hateful content;
- Develop and implement accessible, visible and transparent online complaints mechanisms supported by a robust system ensuring timely handling of complaints;
- Enhance mechanisms available to users of Internet sites to flag inappropriate content;
- Set up mechanisms which would enable users to moderate the content of online communities they participate in;
- Explore the idea of harmonizing statutes dealing with objectionable or inappropriate content in the terms of service used by international Internet companies;
- Analyze the potential of using existing ethical policies to steer the response of the Internet industry to "cyberhate".

Recommendations for NGOs

- Increase efforts to monitor the Internet for hate-inciting content and publicize the findings;
- Actively challenge hate material on the Internet;
- Lobby ISPs to implement terms of service agreements including a clause on hate-inciting material;
- Promote consumer awareness of which ISPs host hate speech and which do not in order to allow consumers to make informed decisions.

8:45-9:00	**Registration of participants, coffee**
9:00 – 9:30	**Welcome and Opening Remarks** Speakers: **Larry Olomofe,** Deputy Head of Tolerance and Non-Discrimination Department **Daniel Milo**, Adviser on combating Racism and Xenophobia **Ženet Mujić,** Office of OSCE Representative on Freedom of the Media

Agenda

9:30 - 12:30	**Session 1:** **Challenges of addressing "cyberhate": Perspective of the Internet Industry** Speakers: **Thaima Samman**, former legal director of Microsoft France and associate general counsel in Microsoft Europe, Middle East and Africa concerning public policy questions **Suzanneke Niemeijer**, Senior Community Manager of HYVES, the largest social network site in the Netherlands
12:30 - 14:00	**Lunch**
14:00 - 15:30	**Session 2: Identification of Recommendations**
15:30-15:45	**Coffee break**
15:45 - 17:00	**Session 3: Recommendations continued**
17:00 - 17:30	**Closing Session: Summary and conclusions**

List of Participants:

1. Larry Olomofe, Deputy Head of Tolerance and Non-Discrimination Department, ODIHR
2. Daniel Milo, Adviser on combating Racism and Xenophobia, ODIHR
3. Thaima Samman, former legal director of Microsoft France and associate general counsel in Microsoft Europe Middle East and Africa
4. Suzanneke Niemeijer, Senior Community Manager, HYVES social networking site, The Netherlands
5. Lisette Abercrombie, HYVES social networking site, The Netherlands
6. Suzette Bronkhorst, Secretary General, INACH, The Netherlands
7. Ronald Eissens, Board Member, INACH, The Netherlands
8. Ženet Mujić, OSCE Office of Representative on Freedom of the Media
9. Paul Giannasi, UK Ministry of Justice
10. Craig Barnes, UK Ministry of Justice
11. Stefan Glaser, Jugendschutz.net, Germany
12. Alexander Verkhovsky, SOVA Center for Information and Analysis, Russia
13. Peter Robbins, Internet Watch Foundation, UK
14. Sille Jansen, Complaints Bureau on Discrimination on the Internet (MDI), The Netherlands