

Alexander Verkhovsky

Director of SOVA Center for Information and Analysis, Russia

Data-collection and fact-finding

The importance of data collection can hardly be questioned as far as hate crimes are concerned, i.e. violent crimes motivated by racist and similar prejudice. The need to collect data on hate crimes is now almost universally recognized, and certain progress is evident in this sphere. However, it is not the case with hate speech (we use this term here as synonymous to "incitement to hatred", since terminological debates on the matter are not so relevant to the purpose of this report). Significant inconsistency of domestic legal norms across countries and substantial differences in the interpretation of international standards, compounded by differences in the legal language expressing these norms¹ – all hinder data collection on hate speech.

It is noteworthy that the objective of monitoring hate speech (as opposed to hate crimes and discrimination) is not set forth in the Outcome Document of the Durban Review Conference, published in 2009. However, the Durban Declaration and Program of Action adopted in 2001 did set forth fairly broad objectives of such monitoring: *"to collect, compile, analyze, disseminate and publish reliable statistical data at the national and local levels and undertake all other related measures which are necessary to assess regularly the situation of individuals and groups of individuals who are victims of racism, racial discrimination, xenophobia and related intolerance"* (para. 92).² This and subsequent paragraphs were undoubtedly focused on violence and discrimination in particular, where a victim or victims can be easily identified. In contrast, with hate speech it is not always easy to identify a victim or group of victims even where incitement to violence is involved. Yet, resolutions of the Durban Conference implied that monitoring should include "xenophobia and related intolerance", i.e. intolerance manifested in the form of statements as well. The wording of these resolutions, however, must have been too vague in this respect. Apparently, it is the reason why the Durban Review Conference did not revisit this objective.

In Europe, at the Council of Europe level we can observe an evolving awareness of the need to monitor incitement to hatred, but progress has been slow so far. General Policy Recommendations of The European Commission against Racism and Intolerance (ECRI) do not contain any recommendations concerning data collection on racist statements and hate speech in general. In particular, such monitoring has not been included in General Policy Recommendation # 2 as a task of Specialised bodies to combat racism, xenophobia, antisemitism and intolerance at national level. We believe, though, that this task would be very appropriate for such Specialised bodies.

At the end of the ECRI Seminar on Combating racism while respecting freedom of expression in November 2006, the following conclusion was made: *"There still is a lack of regular and systematic data collection regarding racist expression. We need to find mechanisms enabling better collection of data and information on racist expression. On the basis of the data and information*

¹ Tarlach Mcgonagle. International and European legal standards for combating racist expression: selected current conundrums // Expert seminar "Combating racism while respecting freedom of expression". Proceedings. ECRI web-site. 2007 (http://www.coe.int/t/dghl/monitoring/ecri/activities/22-Freedom_of_expression_Seminar_2006/NSBR2006_proceedings_en.pdf).

² Durban Declaration and Program of Action // UN web-site (http://www.un.org/durbanreview2009/pdf/DDPA_full_text.pdf).

collected, it should then be possible to develop targeted policies which bring practical solutions to the problems identified."³

At the same seminar, Ms Beate Winkler, Director of the European Monitoring Centre on Racism and Xenophobia (EUMC), said that relevant data are being collected only in some European countries, and most of the data come from the law enforcement,⁴ which is obviously just a fraction of all hate speech, even if we take into account only its most aggressive forms. At the said Seminar, the task of collecting data on racist statements was mentioned quite in connection with Specialised bodies to combat racism, xenophobia, antisemitism and intolerance at national level.⁵ But again, it was all about collecting data on how laws are enforced, rather than to monitoring of racist statements per se.

We should ask ourselves what kind of monitoring we need. To do so, we first have to articulate how we want to use the findings.

The first answer is evident from the practice of NGOs and law enforcement agencies. Such monitoring would enable effective law enforcement.

The second answer has been repeatedly voiced by international organizations. It is necessary to collect data on law enforcement by country and to devise a measure for evaluating its efficiency. But in order to evaluate the performance, collecting law enforcement data alone – which is usually what most international organizations, including ECRI, request from the states – is not enough. It is equally important to know about cases missed by the law enforcement. To achieve this, a broader monitoring of hate speech is required. Accordingly, a comparison of legal norms and actual situations in different countries is needed.⁶ All this requires universally applicable criteria. Certainly, domestic norms cannot serve as such criteria. However, the wording of treaties such as ICCPR and ICERD is not sufficiently transparent.⁷

For some reason, the third answer to the question is rarely raised in discussions. Legislation against hate speech is a very convenient tool enabling authorities to impose excessive restrictions on the freedom of expression. It is not just about finding the delicate balance – however important (and attractive to the theoretically-minded) this topic may be – but also about being mindful of the fact that national and local authorities in underdeveloped democracies intentionally abuse anti-hate speech laws and mask politically and otherwise motivated repression of liberties behind a rhetoric of tolerance and civil dialogue. Moreover, once the authorities realize how convenient this tool is for them, they may deliberately modify the national legislation to enable such abuse. Monitoring should expose any abuse masked as efforts to counter hate speech with the same energy and focus as it documents progress or shortcomings in this area.

³ Michael Head. Main findings and conclusions // Expert seminar "Combating racism while respecting freedom of expression". Proceedings.

⁴ Beate Winkler. The recording and monitoring of racist expression: the challenges ahead // Ibid.

⁵ Michael Head. Op.cit.

⁶ In principle, the UN framework may be used to encourage convergence of relevant national legislations. However, this idea still causes too many objections of various kinds and from different parties.

⁷ Here I also refer to Agnes Callamard. Conference room paper // Expert seminar on the links between articles 19 and 20 of the International Covenant on Civil and Political Rights. "Freedom of expression and advocacy of religious hatred that constitutes incitement to discrimination, hostility or violence". 2008. October 2-3. Pp. 7-31 (http://www2.ohchr.org/english/issues/opinion/articles1920_iccpr/docs/experts_papers/Callamard.doc).

Probably the biggest challenge in data collection is to create the classification of pronouncements we should target.

Taken together, the second and third paragraphs above reveal a potential conflict. National legislations and the legal language of international norms do not create a universally acceptable framework for identifying statements which could be categorized as manifestations of xenophobia and incitement to hatred, discrimination or even violence. (The mentioned para. 92 of Durban Final Document suggested: "*Such statistical data should be disaggregated in accordance with national legislation.*" But it is not easy to do even at the national, let alone international level.) However, we cannot disregard the national and international legal language, since we are going to assess the effect of norms formulated in these languages. Therefore we face a challenging task of developing a classification of statements to include elements of existing legal systems without depending on them entirely. While we cannot foresee a solution, we can assume that in order to be effective, such classification should be as simple as possible.

I would suggest that whatever the classification, its key test should be the degree of public danger a statement poses. Once we agree to apply this test, we can then identify at least some types of dangerous statements, such as appeals to violence and incitement to discriminatory practices. It is very important to distinguish between literal appeals and those dependent on a specific context. The latter are not necessarily criminalized, but it does not mean that they should be overlooked in monitoring.

On the other hand, when context-specific appeals (to whatever) are involved, we should bear in mind that where an expert engaged in monitoring will see a veiled appeal, an ordinary reader might not notice anything. Incidentally, similar situations occur in courts: even though a judge may request an expert opinion of a particular text, for a public appeal to be illegal it must be addressed to, and understood by lay public. Therefore we should make a distinction between texts addressing a broad audience and those addressed to a specific group using a specific language (e.g. followers of a religious doctrine or a youth subculture).

There is an important distinction between appeals to violence and appeals to discrimination, and this distinction is very important for data collection purposes. While racist violence is condemned in virtually all countries, therefore it is assumed that any call to violence must be subject to monitoring, there is no consensus regarding discrimination. Such consensus has yet to be reached through a broad-based public debate. Such a debate, however, should address whatever views existing in society that approve of various types of discrimination. This means that the monitoring tool (and the law, in my opinion) should distinguish between radical calls for discrimination clearly disapproved of by an overwhelming majority and those that are not yet universally rejected. The boundary between them is understandably vague and varies from country to country, but it should not be a fundamental difficulty. Rather than a detailed classification of statements, the important thing to us is to draw a distinction between things that are radical and unacceptable in a society and things that are controversial, but not yet universally disapproved of (to reiterate, the terms "unacceptable" and "controversial" refer to de facto social norms in a particular country, rather than our own or someone's personal opinion).

While appeals to discrimination are usually easy to identify, statements often referred to as "incitement to hatred" are more evasive, since both the law enforcement and monitoring may face the nuances of meaning inherent in the word "hatred" or its synonyms used in various jurisdictions and instruments. It would be more appropriate for monitoring purposes to use a short list of identifiable types of statements that fall under the term "incitement to hatred." Our organizations'

several years of experience with media monitoring⁸ show me that building such a classification is not an easy task, but can result in a workable monitoring tool as long as you can avoid excessively detailed categorization and tolerate vague boundaries between different types of statements. Admittedly, it would be much more difficult to develop such a classification to be used on a European (or even global) scale, but it is worth a try!

A few challenges may arise that need to be taken into account.

Firstly, any classification of statements for purposes of monitoring should steer clear of archaic explanatory models still existing in society, unless they are directly reflected in the legislation. For example, the Russian criminal law refers to incitement to hatred against individuals based on certain group characteristics they share. Public debate, however, is dominated by the long-established notion of "interethnic hostility" suggesting some kind of a symmetric conflict, rather than aggression (including verbal aggression) expressed by some people against representatives of certain groups. I believe that researchers – including those engaged in monitoring – should resist these counterproductive concepts which have nothing to do with the law, even though such concepts may be popular.

Secondly, while comparing statements based on the degree of public danger, we should be aware of the context. The speaker's status is of key importance. In particular, it applies to public officials and prominent political figures. Not accidentally, this is the focus of ECRI Declaration on the use of racist, antisemitic and xenophobic elements in political discourse, adopted on 17 March 2005.⁹ The situation where a statement was made is important, as well as publicity of the statement (see a more detailed consideration below).

Thirdly, closely related to "incitement to hatred" are statements that humiliate certain people. Laws of different countries vary broadly in how they treat such statements. It is probably impossible for any monitoring to capture the totality of such statements, because the ultimate test for how "humiliating" they are is the reaction of the targeted person, rather than the content of the statement – and such reaction may be influenced by a variety of factors. Any monitoring (and any law as well) should not address, but rather leave alone any statements that cannot be prohibited in a democratic society, even though such statements may sound offensive or demeaning to someone. For example, this obviously refers to emotionally coloured social class criticism engaged in by every left-wing politician for the past two centuries. But then it also applies to indirect criticism involving ethnicity or religion, as long as ideas or practices, rather than individuals, are targeted by critical statements, even demeaning ones. Certain traditional practices are widely condemned in Europe (e.g., clitoridectomy practiced by some cultures). Some practices are criticized by certain public circles, which tend to be quite broad (e.g., ritual slaughter of animals practiced by a number of religions). And finally, sharp criticism of religious beliefs – the so-called "defamation of religion" – and more broadly, criticism of any ideology of belief cannot be subject to such monitoring.

To sum up, it is important to remember that we are talking about monitoring potentially illegal statements which violate restrictions imposed by international law. Such monitoring is certainly no

⁸ The most recent major publication: Galina Kozhevnikova. Hate Speech and Elections: Federal and Regional Levels. M.: SOVA Center, 2008.

⁹ Available at ECRI website http://www.coe.int/t/dghl/monitoring/ecri/activities/14-Public_Presentation_Paris_2005/Presentation2005_Paris_Declaration_en.asp

See also a presentation for the 2005 Paris Conference: Jean-Yves Camus. The use of racist, antisemitic and xenophobic arguments in political discourse // ECRI web-site (http://www.coe.int/t/dghl/monitoring/ecri/activities/14-Public_Presentation_Paris_2005/Presentation2005_Paris_study_en.asp).

substitute for appropriate investigation and should not pretend to be one, therefore including a statement into a monitoring report does not mean reporting it to the police (let alone anticipating of a court verdict). But on the other hand, the scope of such monitoring should not be as broad as that of research of ethnic, religious and similar intolerance in mass media and elsewhere. Intolerance is certainly a broader concept and it should be addressed by other monitoring efforts.

Admittedly, intolerance in an extended sense is indeed a prerequisite for incitement to hatred, and the latter is part of the former. But incitement to hatred is only a fraction of various and numerous intolerant pronouncements (especially if we look at the European mainstream media). That said monitoring incitement to hatred and monitoring intolerance in a broader sense may use a similar methodology, at least for the sake of making the findings comparable. Moreover, projects with a broader scope may offer valuable lessons. International projects of the latter type have been emerging recently, e.g. the EU Agency for Fundamental Rights (FRA).¹⁰

Fourthly, we all need to agree on how we measure "publicity" of certain statements. The level of publicity clearly determines the level of public danger posed by inappropriate pronouncements. But can we say that very limited publicity should mean "zero danger" for law enforcement and monitoring purposes? This is what it usually means in practice when it comes to statements voiced in a very narrow circle of associates. Assessing the level of publicity on the Internet is much more challenging, though. Any statement becomes public unless protected by a password. Therefore, ECRI Policy Recommendation # 7 demands that any statements made on the Internet be considered "public", as well as words pronounced during private meetings.¹¹

This recommendation appears controversial to me. While any words posted on the Internet are potentially accessible, their location usually presumes a very narrow audience – often just a few people. Thus the level of public danger posed by racist pronouncements is negligible and can be compared to drunken chatter in a bar. Indeed, public criticism of such pronouncements draws undeserved and far broader attention than they otherwise could attract.

Law cannot provide quantitative criteria to determine the level of publicity, since there is no universal and effective test of publicity. Whether or not a statement is public is up to the law enforcement and courts to determine. However, for monitoring purposes it is possible to develop reasonable, if not perfect, quantitative criteria of publicity, which are also applicable to the Internet.

Data collection always raises the issue of how representative the data are. In our case the issue is particularly serious.

Collecting complaints is a natural and therefore most common method of data collection by NGOs. Complaint are usually submitted by persons who identify with a group they believe are offended by the statement in question, or by persons who associate with this group in some way (e.g. human rights activists defending representatives of the same group). Information from such sources, however, should be viewed through a critical lens just like any other statements – not because we should be suspicious of members of certain groups or NGOs, but because what is called "proximity bias" is a real problem. One good example of many would be differences in the interpretation of certain facts related to old or recent history (it certainly does not apply to some universally recognized facts, such as some cases of genocide; it is sufficient to recall some "historical disputes"

¹⁰ For details, see Pilot Media Project // FRA web-site (http://fra.europa.eu/fraWebsite/research/projects/proj_pilotmedia_en.htm).

¹¹ ECRI General Policy Recommendation N°7: National legislation to combat racism and racial discrimination // ECRI web-site. 2002 (http://www.coe.int/t/dghl/monitoring/ecri/activities/GPR/EN/Recommendation_N7/Recommendation_7_en.asp).

in the Balkans or the Caucasus). Even though complaints made by such people are important for measuring the level of public resonance caused by a particular pronouncement, these complaints must be additionally verified by those who carry out the monitoring. It is important to additionally verify not just the assessment of facts, but also the assessment of whether or not a particular statement was indeed offensive or demeaning.

Another essential source of data is the law enforcement practices in this sphere – at least to the extent that such information may be available to the organization engaged in a monitoring effort (which is not always the case). Monitoring the law enforcement practices, however, is very important for evaluating their efficiency, detecting abuse and encouraging a further discussion of domestic legal norms.

But even taken together, complaints and law enforcement records cannot serve as a source of representative data for the monitoring purposes – even in countries with a highly developed civil society where a lot of such complaints occur, and in the countries where law enforcement agencies take the problem of hate speech seriously – let alone countries where few citizens are really concerned over the problem of hate speech, while police and judges tend to ignore the problem altogether or address it in a clearly selective manner based on who is the target of hate speech.

On the other hand, total monitoring of the print media, TV and radio, not to mention the Internet, is simply impossible because of the sheer amount of data – especially since just reading (listening, watching) is not enough; monitoring also involves entering the findings into a database, marking up and rating at least a few required parameters (social status of the author, level of publicity of the statement in question, etc.).

Consequently, monitoring of mass media – complementary to sources such as complaints and law enforcement data – should be selective and cover different types of mass media in terms of their format and type of audience in particular. However, we should not forget that our data collection, rather than cover all manifestations of intolerance, should only focus on particularly offensive words that may fall under the definition of incitement to hatred. This type of statements is few in mainstream media of democratic countries, and in choosing media outlets for monitoring we should think of the potential effectiveness of our monitoring. It means that we should not choose media on the basis of their popularity as our key consideration. On the contrary, media outlets selected for monitoring will not represent the entire range of mass media in a society, but rather its most problematic segments.

It is also likely that continuous monitoring will be too costly. If this is the case, we will need to monitor at different periods of time characterized by different intensity of intolerance manifestations. It may not always be possible to identify such periods a-priori. It is known, for example, that any polemics tend to escalate before elections, but other factors can serve as catalysts as well. Therefore, rather than study the issues in real time, a retroactive approach may be a better solution.

All these considerations need to be consistent with the monitoring objectives described above. These objectives may be achieved either through immediate response to certain public statements¹² and also through a slow and deliberate quantitative analysis of publications produced in the past period, based on parameters such as subjects and objects of incitement and types of statements in the accepted taxonomy.¹³ Ideally, both approaches should be used.

¹² The Swedish Quick Response project may serve as an example of this approach. See the project website at <http://www.quickresponse.se/in-english/>

¹³ The latter approach has been used by our center.

To sum up, the type of monitoring needed to study incitement to hatred requires considerable resources. Even the development of its methodology takes substantial effort, particularly since this methodology should lend itself to comparative analysis, at least initially at the regional, if not global scale, where a region is characterized by comparable legislation and practices. In Europe, these may be several EU countries and simultaneously several post-Soviet countries. Their number may be gradually expanded to cover the entire territory of the Council of Europe.

I would assume that no NGO or research institution will be able to handle the task on their own. It is not just about a shortage of human and financial resources. An equally important concern is an authoritative consensus on methodology that needs to be achieved. Therefore, it needs to be a cooperative project involving the academia, NGOs, governments and international organizations.