Octopus Project

Implementing the First Protocol to the Convention on Cybercrime on Xenophobia and Racism: Good practice study

Strasbourg, 1 December 2023 (provisional)

www.coe.int/cybercrime

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

Acknowledgement

This study was prepared by the Octopus Project (Jan Kralik and Alexander Seger (Cybercrime Division, Council of Europe)) and co-authored by Alexander Brown (University of East Anglia, United Kingdom).

The Octopus Project is grateful to all those who contributed to this study, including representatives of States who provided replies to a questionnaire and comments on drafts of the study, and to members of the Cybercrime Convention Committee and other government officials as well as civil society organisations and service providers who attended two webinars to discuss preliminary findings in December 2022 and February 2023, as well as the Conference on xenophobia and racism committed through computer systems in January 2023. That conference also marked the 20th anniversary of the opening for signature of the First Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems (ETS No. 189)

Contact

Cybercrime Division
Council of Europe
Email cybercrime@coe.int

Contents

Abbreviations

ACHPR	African Charter on Human and Peoples' Rights (African Union)
ACHR	American Convention on Human Rights (Organization of American States)
CC	Criminal Code
CETS	Council of Europe Treaty Series (numbering of Council of Europe Treaties since 2005)
COVID-19	Coronavirus disease of 2019
CPC	Criminal Procedure Code
C-PROC	Council of Europe Cybercrime Programme Office
DSA	Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)
ECHR	European Convention of Human Rights
ECRI	European Commission against Racism and Intolerance (Council of Europe)
ECtHR	European Court of Human Rights
ETS	European Treaty Series (numbering of Council of Europe treaties until 2004)
HELP	Human Rights Education for Legal Professionals (Council of Europe)
HUDOC	Human Rights Documentation (database of the European Court of Human Rights)
ICCPR	International Covenant on Civil and Political Rights (United Nations)
LGBTQI+	Lesbian, gay, bisexual, transgender, queer or questioning, intersex, asexual and more
OCLCTIC	Central Office for Combating Information and Communication Technology Crime (France)
PACE	Parliamentary Assembly of the Council of Europe

1 Introduction

1.1 Purpose of this study

The Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems, ETS 189 (hereinafter referred to as the "Protocol" or the "First Protocol") was opened for signature in Strasbourg, France, on 28 January 2003.[1] In the light of the 20th anniversary of this Protocol since its opening for signature in January 2023, it was decided to prepare the present good practice study on the implementation of this treaty.[2]

The overall objective of this study is to facilitate implementation of and increased membership in this Protocol by:

- underlining the relevance of this Protocol given increasing hate speech and hate crime online;
- documenting good practices by Parties to this Protocol and other States;
- explaining key issues and concepts;
- promoting synergies between this Protocol and related instruments, tools and initiatives.

1.2 Methodology

The primary scope of this study is good practices in either implementing or reflecting the Protocol found in State Parties to the Protocol as well as in States that have been invited to accede to the Convention on Cybercrime and thus to this Protocol.[3]

Where relevant, the study also refers to examples of other States.

France, Germany, Norway, Slovakia and Spain provided responses to a questionnaire circulated for the purposes of this study. Additionally, Brazil and Serbia also contributed to it. Good practices presented in this study are thus primarily related to these "participating States". However, publicly available examples from other States have also been included.

As confirmed with participating States at the outset, while their contributions may be reflected in the study, this exercise does not constitute an assessment of their implementation of the Protocol, and the study does not provide recommendations for follow by individual States.

The study used several further methods to obtain examples of good practice. A meta-survey was undertaken of existing country reports of the European Commission against Racism and Intolerance (ECRI), looking at a randomised sample of these reports for examples of good practice related to the Protocol.[4] These references range from recommending ratification;[5] to

[1] https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=189

[2] The study and related activities were carried out under the Octopus Project. The CyberSouth project also supported these.

[3] Information on signatories, ratifications, entries into force, and States invited to accede to the Protocol is available at: www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=189

[4] The ECRI country reports databased is available at: https://www.coe.int/en/web/european-commission-against-racism-and-intolerance/country-monitoring

[5] See, for example, ECRI, *Sixth Report on Austria*, 2 June 2020, paragraph 56. Available at: https://rm.coe.int/report-on-austria-6th-monitoring-cycle-/16809e826f. See also ECRI, *Sixth Report on*

recommending the withdrawal of reservations;[6] to noting that stronger enforcement was needed, such as special cybercrime units stepping up their monitoring of the Internet;[7] to noting some of the challenges faced by States in achieving implementation, including jurisdictional challenges;[8] to recommending particular forms of institutional capacity building such as supporting and intensifying the work of special cybercrime units, including through better hiring and training of police officers at scale, so that enforcement can respond to the demand created by increasing amounts of online hate speech and hate crime.[9]

Use was also made of references to the Protocol and examples of good practice in its implementation, via key term searches within the HUDOC database of case law of the European Court of Human Rights (ECtHR),[10] and of relevant academic literature on hate speech laws and hate crime laws.

The text of this study is based on an initial draft prepared by Alexander Brown[11] which was then further reviewed by staff of the Octopus Project and by experts of participating States.

The study was further complemented by inputs obtained during two webinars in December 2022 and February 2023 and the Conference on the occasion of the 20th anniversary of the Protocol held in Strasbourg in January 2023 and attended by representatives of governments, civil society organisations, human rights associations, technology companies and other interested stakeholders.

2 About the Protocol on Xenophobia and Racism

2.1 Racism and xenophobia: the relevance of the First Protocol

The First Protocol focuses on a threat that today is more relevant than it was when it was opened for signature twenty years ago, namely acts of a racist or xenophobic nature committed through computer systems. As the Explanatory Report to the First Protocol points out:

> "the emergence of international communication networks like the Internet provides certain persons with modern and powerful means to support racism and xenophobia and enables them to disseminate easily and widely expressions containing such ideas."[12]

Belgium, 18 March 2020, paragraph 54. Available at: https://rm.coe.int/ecri-sixth-report-on-belgium-/16809ce9f0

[6] See, for example, ECRI, *Fifth Report on Romania*, 5 June 2019, paragraph 42. Available at: https://rm.coe.int/fifth-report-on-romania/168094c9e5

[7] See, for example, ECRI, *Fourth Report on Portugal*, 9 July 2017, paragraph 77. Available at: https://rm.coe.int/fourth-report-on-portugal/16808b59c3

[8] See, for example, ECRI, *Fifth Report on the Czech Republic*, 13 October 2015, paragraph 51. Available at: https://rm.coe.int/fifth-report-on-the-czech-republic/16808b5664

[9] See, for example, ECRI, *Second Report on Montenegro*, 19 September 2017, paragraph 33. Available at: https://rm.coe.int/second-report-on-montenegro/16808b5942. See also ECRI, *Fourth Report on Finland*, 9 July 2013, paragraph 101. Available at: https://rm.coe.int/fourth-report-on-finland/16808b5714

[10] The HUDOC database is available at: https://hudoc.echr.coe.int/

[11] Expert consultant on hate speech law based at the University of East Anglia (UEA9: The opinions expressed are solely his own and do not express the opinions of UEA.

[12] Explanatory Report to the First Protocol, paragraph 3.

It is not only the quantity, scale, and reach of digital communication that makes online hate speech and hate crime different to offline forms and potentially more harmful. What also matters is the combination of the size of the audience combined with other features: the anonymity of the speaker; the instantaneousness of communicative exchanges; the facility for vast numbers of people to participate in "piling on" against targets; the permanence of the content; the way digital communication is capable of entering even the most private spaces; and the extent to which people have become captive audiences to online hate speech given that abstaining from social media and switching off email is simply not a viable option.[13] Added to this is the excruciating realisation that the moment of humiliation may be permanent and endlessly replayable.

Digital communication technologies are both forms and sources of further globalization. Technologies can enable migration and the movement of people, but in turn these processes "can lead to exclusion and increased inequality, very often along racial and ethnic lines".[14] Wars and conflicts in Afghanistan, Syria, Yemen, Ukraine and elsewhere have created a rapid increase in flows of migrants and refugees; their integration is rarely smooth. Moreover, current wars and conflicts are accompanied by waves of mis- and disinformation.

The Russian aggression against Ukraine includes propaganda alleging, for example, that Ukrainians are actually Russians, that "Ukraine is not a real country", that Ukrainian'ness is a "specific disorder of the mind", that defenders of Ukraine are guilty of "genocide", that the Russian invading forces are "peacekeepers", that Ukraine is governed by "Nazis", and that the invasion is not a "war" but a "liberation". The Russian State and its proxies have uploaded propaganda, disinformation, and hate speech against Ukraine and Ukrainians onto social media platforms like Telegram in order to facilitate the sharing of such content on mainstream platforms such as X (formerly known as Twitter) and Facebook, the aim being to make it harder for the latter to trace the original source of the content.[15] The Russian propaganda machine has created large numbers of "news websites" that mimic the style and appearance of Western sites in order to create fake stories about the war that could be promoted on Twitter and Facebook.[16] Moreover, the Russian aggression has been preceded and is accompanied by eliminationist rhetoric that some consider to amount to evidence of genocidal intent towards Ukrainian people.[17]

Furthermore, the COVID-19 pandemic has increased the use of digital communication technologies throughout societies. This increase in interconnectivity has created fertile grounds for the creation, transmission and amplification of conspiracy theories concerning the origin and sources of spread of the virus, which in turn fed into undercurrents of racism and xenophobia. UN organisations[18] and academic researchers[19] alike have observed that instances of vilifying, stigmatizing and discriminatory language targeting not merely people of Asian descent but also migrants more

---

[13] See Citron (2014); Delgado and Stefancic (2014); Cohen-Almagor (2015); Brown (2017d); Brown (2018); Brown (2020: 42–44); Vidgen et al. (2021); and Hassan et al. (2022).

[14] Explanatory Report to the First Protocol, paragraph 2.

[15] See Turner (2022).

[16] See Scott (2022).

[17] Russia's Eliminationist Rhetoric Against Ukraine: A Collection (justsecurity.org)

[18] See the UN International Organization for Migration (IOM), *COVID-19 Analytical Snapshot #6: Stigmatization & discrimination*. Available at: www.iom.int/sites/g/files/tmzbdl486/files/our_work/ICP/MPR/covid-19_analytical_snapshot_6_-_stigmatization_and_discrimination.pdf. See also the UN Department of Global Communications (DGC), *COVID-19: UN Counters Pandemic-related Hate and Xenophobia*. Available at: www.un.org/en/coronavirus/covid-19-un-counters-pandemic-related-hate-and-xenophobia

[19] See Roberto et al. (2020).

generally increased as a result of COVID-19, and in all regions of the world. The vilification of out-groups defined by perceived threat of carrying viruses has also exacerbated wider patterns of racism, xenophobia and hate speech in general, with resentment, mistrust, hatred, and contempt spilling over into attacks against other historically victimised groups including Jews, Roma Communities, and LGBTQI+ people. The Internet with its social media platforms has become the go to vehicle for the propagation of such ideas in many parts of the world.

These are several of the reasons why the First Protocol has never been more relevant than today, as pointed out also in the key messages of the Conference on xenophobia and racism committed through computer systems in January 2023.[20]

While as of October 2023, 68 States are Parties to the Convention on Cybercrime, 35 States are Parties to the First Protocol, most of them members of the Council of Europe but also Morocco, Paraguay and Senegal. Canada and South Africa have signed but not yet ratified it.[21] Any Party to the Convention can become a Party to the First Protocol. Also the States invited to accede to the Convention – there were 21 of those by October 2023[22] – may become a Party to both, the Convention and its Protocols. Senegal deposited the instruments of accession to the Budapest Convention on Cybercrime and its Additional Protocol on Xenophobia and Racism at the same time[23], and so did Morocco[24] and Paraguay[25].

2.2 The place of the First Protocol within the framework of the Convention on Cybercrime

In 2001, the Convention on Cybercrime (the Budapest Convention) was opened for signature.[26] This was the first international treaty specifically focused on cybercrime and electronic evidence, and it remains the most relevant legally binding treaty today. It provides States with substantive criminal provisions, procedural powers with safeguards applicable to any type of offence where the evidence is on computer system and international cooperation provisions applicable to any type of crime where the evidence is on computer system.

The committee drafting the Convention between 1997 and 2001 had discussed the possibility of including additional content-related offences[27], such as the distribution of racist propaganda through computer systems. However, the committee did not reach consensus on the criminalisation of such conduct. Noting the complexity of the issue, it was decided that that type of conduct be covered in a separate protocol.

---

[20] https://rm.coe.int/2542-111-key-messages-xr-conference-2023-v5-eng/1680aa2379

[21] Information on signatories, ratifications, entries into force, and States invited to accede to the Protocol is available at: https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=189

[22] Information on signatories, ratifications, entries into force of the Convention on Cybercrime is available at: https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=185

[23] https://www.coe.int/en/web/cybercrime/-/accession-by-senegal-to-the-budapest-convention-on-cybercrime-and-its-protocol-on-xenophobia-and-racism

[24] https://www.coe.int/en/web/cybercrime/-/accession-by-senegal-to-the-budapest-convention-on-cybercrime-and-its-protocol-on-xenophobia-and-racism.

[25] https://www.coe.int/en/web/cybercrime/-/paraguay-ratifies-the-convention-on-cybercrime-and-the-additional-protocol-on-xenophobia-and-racism

[26] Council of Europe, Budapest, 23 November 2001, European Treaty Series No. 185.

[27] Text of the Budapest Convention contains provisions concerning computer-related offences and one-content related offence (child pornography), while the First Protocol contains provisions concerning the content-related offences of racism and xenophobia on the Internet. See Esposito (2002: 2-3).

In January 2003, the "Additional Protocol to the Convention on Cybercrime concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems" was then opened for signature.[28]

This Protocol focuses on substantive criminal law by requiring Parties to criminalise several acts of a xenophobic and racist nature. It can be joined by Parties to the Budapest Convention.[29]

While Articles 3 to 7 of the First Protocol establish obligations on the part of Parties regarding the appropriate regulation of conduct,[30] Article 8 deals with the relationship between the Convention and the First Protocol. Paragraph 1 stipulates that some of the provisions of the Convention apply, mutatis mutandis, to the First Protocol. Through its paragraph 2, States may use the procedural powers and international cooperation provisions of the Convention also in relation to offences provided in the First Protocol.[31]

Furthermore, in May 2022, the Second Additional Protocol to the Convention on Cybercrime on enhanced co-operation and disclosure of electronic evidence (CETS 224) was opened for signature.[32] The Second Protocol addresses the need for more effective tools to obtain evidence on computer systems in foreign, unknown, or multiple jurisdictions in order to protect individuals and their rights against crime. It provides for measures for enhanced cooperation such as direct disclosure of subscriber information[33] or the expedited disclosure of stored computer data in an emergency[34]. The Second Protocol supplements both the Convention and the First Protocol[35], and as indicated in its Article 2, its measures may be applied to specific criminal investigations and proceedings concerning the offences established pursuant to the Convention as well as the First Protocol.[36] The Convention on Cybercrime and its two additional protocols thus work in harmony. The more countries that join these three instruments, the more comprehensive the international response to crime online will be.

2.3 Structure and scope of the Protocol

Opened for signature in January 2003, the First Protocol embodies a growing normative recognition among States both that "acts of a racist and xenophobic nature constitute a violation of human rights and a threat to the rule of law and democratic stability" and that "national and international law need to provide adequate legal responses to propaganda of a racist and xenophobic nature committed through computer systems".[37]

---

[28] Council of Europe, Strasbourg, 28 January 2003, European Treaty Series No. 189.

[29] See Council of Europe, Strasbourg, 28 January 2003, European Treaty Series No. 189, Article 9-10.

[30] See Brown (2020: 30–32).

[31] This was reconfirmed in a Guidance Note adopted on 27 June 2023, in which the Cybercrime Convention Committee (T-CY) underlined that the procedural powers and the tools for international and cross-border cooperation of the Budapest Convention and its Second Protocol apply to evidence in electronic form of any criminal offence and not only to offences against and by means of computer systems or data. See T-CY(2023)6 Guidance note Scope of powers. Available at: https://rm.coe.int/t-cy-2023-6-guidancenote-scope-of-powers-v9adopted/1680abc76a.

[32] Council of Europe, Strasbourg, 12 May 2022, European Treaty Series No. 224.

[33] Ibid, Article 7.

[34] Ibid, Article 9.

[35] Ibid, Article 1.

[36] Ibid, Article 2.

[37] The Protocol, Preamble.

The Protocol primarily addresses digital or online content of a racist and xenophobic nature, which Article 2 defines as

> "… any written material, any image or any other representation of ideas or theories, which advocates, promotes or incites hatred, discrimination or violence, against any individual or group of individuals, based on race, colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors."

Chapter II of the Protocol (Articles 3 to 7) establishes obligations on the part of Parties regarding "legislative and other measures as may be necessary to establish as criminal offences" the following conduct:

- dissemination of racist and xenophobic material through computer systems (Article 3);
- racist and xenophobic motivated threat (Article 4);
- racist and xenophobic motivated insult (Article 5);
- denial, gross minimisation, approval or justification of genocide or crimes against humanity (Article 6).

Chapter III through Article 8 addresses relations between the Convention and the First Protocol, and Chapter IV contains final provisions (Articles 9 to 16).

The Protocol aims to provide a unifying framework for how States address through criminal law the issue of acts of a racist and xenophobic nature committed through a computer system. The Preamble to the Protocol describes Parties as being "convinced of the need to harmonise substantive law provisions concerning the fight against racist and xenophobic propaganda".[38] In this regard, the Protocol prepared the ground for subsequent efforts at harmonisation in Europe.

For example, paragraph 12 of the Council of the European Union Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law (2008) reads:[39] "Approximation of criminal law should lead to combating racist and xenophobic offences more effectively, by promoting a full and effective judicial co-operation between Member States." Indeed, the European Union Guidance Note on the Practical Application of Council Framework Decision (2018) stipulates that "Due regard is given to other relevant EU standards, in particular on the rights of victims of crime, as well as international standards enshrined in relevant international and regional instruments" and cites the Protocol in the accompanying footnote.[40]

Importantly, the Protocol seeks to provide a unifying framework not simply between member States of the Council of Europe but also between member and non-member States that are Parties, thus reflecting the broad base of Parties to the Convention on Cybercrime. The First Protocol may play a role in solidifying and enshrining global anti-hate norms, such as are embodied in other key international anti-hate instruments. These internationalist aspirations of the Protocol are expressed in various places within the Preamble including for example the United Nations International Convention on the Elimination of All Forms of Racial Discrimination of 21 December 1965.

One of the challenges to broad adherence to the Protocol is divergence in both free speech norms and anti-hate norms across the world. Therefore, the Protocol is written and intended to be

[38] The Protocol, Preamble.

[39] Council of the European Union, 28 November 2008, 2008/913/JHA.

[40] EU High Level Group on combating racism, xenophobia and other forms of intolerance, *Guidance Note on the Practical Application of Council Framework Decision* (Brussels: European Commission, 2018), p. 2.

interpreted in ways that accommodate the particular historical experiences, social make-up, constitutional essentials and domestic legal systems, and predominant ideologies of Parties. This contextualism is articulated in the Preamble to the Protocol:

> "Mindful of the need to ensure a proper balance between freedom of expression and an effective fight against acts of a racist and xenophobic nature"; "Recognising that this Protocol is not intended to affect established principles relating to freedom of expression in national legal systems".[41]

It follows from this that the First Protocol gives flexibility to Parties to address certain types of conduct, as they deem appropriate, either through criminal law or other means.[42]

The contextualist approach is also manifest in both procedural and substantive aspects of the Protocol. Two examples illustrate this:

- Article 12 gives Parties the right to enter various reservations and declarations upon signing, or when depositing their instruments of ratification (acceptance, approval or accession), both carrying over existing reservations and declarations to the Convention on Cybercrime and new reservations as explicitly provided for in the Protocol. More specifically:
  - Article 3(2) gives Parties the right not to establish as criminal offences the distributing, or otherwise making available, racist and xenophobic material to the public through a computer system, when this advocates, promotes or incites discrimination that is not associated with hatred or violence, provided that other "effective remedies" are available;
  - Article 3(3) gives Parties the right not to establish criminal offences of distributing, or otherwise making available, racist and xenophobic material to the public through a computer system, due to established principles in its national legal system concerning freedom of expression; Article 5(2)(b) gives Parties the right not to establish criminal offences of aggravated public insult; and Article 6(2)(b) gives Parties the right not to establish criminal offences of genocide denial.
- The First Protocol gives Parties the right to formulate their own domestic laws on acts of a racist and xenophobic nature committed through a computer system in ways that reflect their own context. It permits Parties to add additional qualifying conditions. For example:
  - Article 5(2)(a) gives Parties the right to establish criminal offences of aggravated public insult that include as an element of the conduct that victims are exposed to hatred, contempt or ridicule;
  - Article 6(2)(a) gives Parties the right to establish criminal offences of denying, grossly minimising, approving or justifying acts of genocide and other atrocity crimes that include as an element of the conduct the intention to incite hatred, discrimination or violence.

[41] The Protocol, Preamble.

[42] See for example, Key messages to the Conference on xenophobia and racism committed through computer systems (Strasbourg, 30-31 January 2023). Available at: https://rm.coe.int/2542-111-key-messages-xr-conference-2023-v5-eng/1680aa2379.

Together these features enable States to join the First Protocol while not being bound by all of its obligations. They also permit countries to counter acts of a racist and xenophobic nature committed through computer systems by non-criminal means, taking into consideration established principles of their national law. In other words, a State could criminalise some of the acts identified in the Protocol but not others, based on that State's context and fundamental principles of domestic law.

The Cybercrime Convention Committee (T-CY) will remain an important forum for sharing experiences in the implementation of the First Protocol by its Parties and addressing issues and challenges involved.

2.4 Hate speech and hate crime

Among the most important and thorny concepts are those of "hate speech" and "hate crime".

The term "hate speech" has no globally agreed definition in human rights instruments. However, Recommendation CM/REC(2022)16 of the Council of Europe Committee of Ministers on combating hate speech defines it as:

> "(…) all types of expression that incite, promote, spread or justify violence, hatred or discrimination against a person or group of persons, or that denigrates them, by reason of their real or attributed personal characteristics or status such as "race", colour, language, religion, nationality, national or ethnic origin, age, disability, sex, gender identity and sexual orientation."[43]

The key thing to note here is that the scope of protected characteristics is much wider than xenophobic or racist behaviour (see also Chapter 3.1.1.7. of this study for further discussion).

Further, Recommendation CM/REC(2022)16 distinguishes between hate speech prohibited under criminal law, hate speech subject to civil or administrative law, and offensive or harmful types of expression which are not sufficiently severe to be legitimately restricted under the European Convention on Human Rights, but nevertheless call for alternative responses, , including but not limited to victim support, education, and counter-speech.[44] While both hate crime and hate speech may be countered by non-criminal means, the main difference between them is that hate crime, by definition, always meets the threshold of criminalisation. Hate speech generally does not meet this threshold except in certain circumstances defined by law (which varies widely by country).

When it comes to hate speech prohibited under criminal law; according to the First Protocol, it is the expression itself against persons for the reason that they belong to a protected group[45] that is criminalised. No criminal offence is constituted when such expression is not directed against persons because they belong to a protected group. The hate speech Recommendation further

[43] Recommendation CM/Rec(2022)16 of the Committee of Ministers to member States on combating hate speech (Adopted by the Committee of Ministers on 20 May 2022 at the 132nd Session of the Committee of Ministers), paragraph 1(2).

[44] Recommendation CM/REC(2022)16, see paragraphs 45 and 55).

[45] Pursuant to the provisions of the First Protocol hatred, discrimination or violence, have to be directed against any individual or group of individuals, for the reason that they belong to a group distinguished by "race, colour, descent or national or ethnic origin, as well as religion, if used as a pretext for any of these factors". See for example, Explanatory Report to the First Protocol,paragraph 17.

specifies that personal characteristics could be real or attributed.[46] Victims thus do not need to prove their belonging to a minority.[47]

A clear legal framework should, therefore, be established to prevent and combat hate speech, while criminal law should only be applied as a last resort and for the most serious expressions of hatred.[48]

"Hate crimes always comprise two elements: a criminal offence committed with a bias motive"[49] (also known as the animus model)[50].

In other words, hate crimes refer to aggravated crimes that do have corresponding or parallel basic or base offences, whereas hate speech does not have such a parallel.[51] This means that the conduct itself – for example, threats with the commission of a serious criminal offence – is criminalised, even if not directed against persons for the reason that they belong to a protected group. The Explanatory Report to the First Protocol recognises that most legislations provide for the criminalisation of threat in general.[52] Based on this logic, Article 4 of the First Protocol, which criminalises racist and xenophobic threat, can be considered a hate crime, as a threat itself (base crime) is usually criminal. Hate crimes may thus attract greater penalties if directed against persons for the reason that they belong to a certain group.

Although freedom of expression concerns may still arise in relation to online hate, they are less serious when it comes to hate crimes than hate speech. This is because hate crimes criminalise actual behaviour that is already criminal even when not directed against persons for the reasons that they belong to a protected group while as regards hate speech, it is the expression itself directed against persons for the reasons that they belong to a protected group that is criminalised, while conduct lacking the element of "being directed against protected group" does not constitute a criminal offence. This is also why Article 4 of the First Protocol that may be considered a hate crime, does not allow for the possibility to make a reservation, as threats are criminalised generally by most domestic laws and the freedom of expression concerns are less serious. Other offences under the Protocol allow for possibilities to enter various reservations or to add additional qualifying conditions. This is further evidence that the First Protocol is not intended to affect established principles relating to freedom of expression in domestic legal systems.

---

[46] Recommendation CM/REC(2022)16, see paragraph 1 (2).

[47] See Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraph 21.

[48] Recommendation CM/REC(2022)16, see paragraphs 7-9.

[49] OSCE, *Hate Crime Laws*, p. 16.

[50] However, while certain organisations and member States use the term "bias motivation" as an operational framework for hate crime, the term hate element under hate crime is broader and encompasses not only the animus model which uses motivation as the legal test, but also the discriminatory selection model. The discriminatory selection (or "group selection") model requires that the offender intentionally selected his or her victim from the protected group, but unlike the animus model, proof of prejudice, bias, hostility, or hatred is not necessary to formally establish for liability to ensue. Thus, hate element in hate crime ensures that hate crime legislation based on the "animus" model as well as legislation based on the discriminatory selection model are incorporated into the definition.

[51] See Brown and Sinclair (2023: ch. 6). See also OSCE, Hate Crime Laws: A Practical Guide, 9 March 2009, pp. 16, 25–26. Available at: https://www.osce.org/odihr/36426.

[52] Explanatory Report to the First Protocol, paragraph 33.

Traditionally, hate crime offences can take the form of sentence-enhancement provisions or separate offences, in either case based on the aggravated circumstances.[53] For example, Article 4 of the EU Council Framework Decision creates an obligation on Parties to "take the necessary measures to ensure that racist and xenophobic motivation is considered an aggravating circumstance, or, alternatively that such motivation may be taken into consideration by the courts in the determination of the penalties." Furthermore, paragraph 21 of the ECRI GPR No. 7 recommends that States establish in criminal law that racist hostility is an "aggravating circumstance"[54]. The First Protocol is silent on this matter, leaving it up to Parties to take into account their domestic context and legal principles.

Although hate speech and hate crime are separate legal concepts, they are not unrelated. Hate speech can contribute to a culture of violence and intolerance which may escalate into violence and abuse. In some cases, hate speech can thus pave the way for hate crimes, sometimes even genocide, crimes against humanity or war crimes[55]. Moreover, there are instances in which conduct described as hate speech may not constitute hate speech prohibited under criminal law and does not constitute a hate crime as such. This evidence of hate speech may be submitted to courts as evidence of the commission of key elements of a hate crime, such as when the defendant's use of a racist slur during the commission of a crime (for example, assault) is used as evidence of the crime being aggravated by bias or hostility based on race.[56]

As indicated at the beginning of this section, the Protocol does not contain an explicit reference either to hate speech or hate crime. It thus gives flexibility to Parties to decide and develop different approaches to implementing offences under the Protocol. States may determine whether to criminalise acts of a racist and xenophobic nature committed through computer systems under their domestic hate speech or hate crime laws taking into consideration their national context. Furthermore, in case of existing parallel basic or base hate crime provisions, Parties are free to decide whether to criminalise acts of a racist or xenophobic nature, for example, through a sentence-enhancement provision under base or basic crime provision or through a separate offence (both based on the aggravated circumstances). The First Protocol also allows Parties not to attach criminal liability to the conduct under Article 3(1) provided that other effective remedies (for example, civil or administrative) are available (see Article 3(2) of the First Protocol), add other qualifying conditions or reserve the right not to apply, in whole or in part, certain provisions of the First Protocol, as already indicated above.

2.5 Hate speech versus free speech

2.5.1 Freedom of expression: rights and restrictions

Given that freedom of expression concerns are greater when it comes to hate speech than hate crimes, consideration should be given how to reconcile hate speech and free speech.

Freedom of expression is protected by various international instruments of a global or regional character such as the International Covenant on Civil and Political Rights (ICCPR), Article 19;

[53] Brown and Sinclair (2023: ch. 6).

[54] European Commission Against Racism and Intolerance (ECRI) of the Council of Europe: ECRI General Policy Recommendation No. 7 on National Legislation to Combat Racism and Racial Discrimination, 13 December 2002, revised 7 December 2017

[55] See Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraph 18.

[56] For further discussion of these overlaps, see Brown and Sinclair (2023: ch. 6).

European Convention of Human Rights (ECHR), Article 10[57]; American Convention on Human Rights (ACHR), Article 13,[58] or the African Charter on Human and Peoples' Rights (ACHPR), Article 9.[59] All of these provisions allow for certain restrictions as the freedom of expression is thus not an absolute right.

Although there are several regional or national approaches to restricting the freedom of expression, the so-called "three-part test" to assess whether a restriction to the right has been legitimate is more or less applied by most international judicial institutions or national constitutional courts. The following describes how the test is applied by the European Court of Human Rights (ECtHR).[60]

According to Article 10(1) of the European Convention of Human Rights (ECHR):

> "everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers".[61]

As stated in paragraph 11 of the Explanatory Report to the First Protocol, "Article 10 of ECHR recognises the right to freedom of expression, including the freedom to hold opinions and to receive and impart information and ideas".[62]

It is applicable not only to "information" or "ideas" that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population.[63] Although such ideas may be perceived negatively, they may still fall within the scope of free speech. Free speech is thus a crucial component of a democratic society, allowing persons to express their ideas freely and without fear, and the protection of Article 10 is extended equally to communication on the Internet.[64]

However, the ECtHR also held that a State's actions to restrict the right to freedom of expression are justified according to paragraph 2 of Article 10 of the ECHR when certain conditions are met, including when such ideas or expressions violate the rights of others. Article 10(2) of the ECHR reads as follows:

> "The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of

---

[57] https://www.echr.coe.int/european-convention-on-human-rights.

[58] https://treaties.un.org/doc/publication/unts/volume%201144/volume-1144-i-17955-english.pdf

[59] https://au.int/sites/default/files/treaties/36390-treaty-0011_-_african_charter_on_human_and_peoples_rights_e.pdf
For details on Article 9 of the African Charter, including on requirements for restricting this right, see the Declaration of Principles of Freedom of Expression and Access to Information in Africa adopted in Banjul in 2019.

[60] See also Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraphs 8-9.

[61] ECHR, Article 10 paragraph 1.

[62] See Explanatory Report to the First Protocol, paragraph 11.

[63] Application no. 5493/72, Handyside v the United Kingdom, 7 December 1976, paragraph 49.

[64] Application No. 36769/08, Ashby Donald and Others v. France, 10 January 2013, paragraph 34.

information received in confidence, or for maintaining the authority and impartiality of the judiciary."

Formulations similar to Article 10(2) of the ECHER can be found in the ICCPR, ACHR or ACHPR. The First "Protocol, on the basis of national and international instruments, establishes the extent to which the dissemination of racist and xenophobic expressions and ideas violates the rights of others".[65]

When cases come before the ECtHR, the Court examines the question of free speech versus hate speech on a case-by-case basis. The ECtHR's case-law in this respect relates primarily to Article 10 (freedom of expression) of the ECHR. Other provisions may be at stake as well, namely Articles 8 (right to respect for private and family life), 14 (prohibition of discrimination) and 17 (prohibition of abuse of rights). The ECtHR distinguishes between, on the one hand, cases dealing with "the gravest forms of 'hate speech'", which the ECtHR considers not only with respect to Article 10 (freedom of expression), but also Article 17 (Prohibition of abuse of rights) of the ECHR; and, on the other hand, cases dealing with "[the] 'less grave' forms of 'hate speech'", which are connected only to Article 10.

It is important to note here that Article 10(2) of the ECHR expressly permits content-based restrictions on speech that are "necessary in a democratic society" based on fundamental interests including "public safety" and "the protection of the reputation or rights of others".[66] Article 17 adds the further dimension that some speech constitutes an abuse of the right to freedom of expression such as by exploiting this right in order to justify, promote or perform acts that, for example, would infringe the rights of others – incitement to hatred, discrimination or violence being a case in point.[67]

As mentioned above, the ECtHR uses a step-by-step analysis (the so-called "three-part test") when examining whether a restriction according to Article 10(2) of the ECHR is legitimate. It should be noted that for such restriction to be legitimate, all three parts of the test must be met.

The ECtHR examines whether the interference was:

- "prescribed by law";
- "pursued one of the legitimate aims" within the meaning of Article 10(2);
- "necessary in a democratic society".

First, the condition that any interference must be prescribed by law must meet the requirements of accessibility and foreseeability. A person must be able to foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail. These consequences do not need to be foreseeable with absolute certainty, as experience showed that to be unattainable.[68] The ECtHR pointed out that the scope of foreseeability depends on the context in which the restrictive measures in question are used.[69] In assessing the foreseeability of a law, the ECtHR also verifies the quality of the law, with regard to its clarity and precision. It should also be noted that the condition "prescribed by law" does not necessarily call for criminalisation, and countries are not obliged to criminalise all forbidden expressions. As emphasised in the previous section

[65] See Explanatory Report to the First Protocol, paragraph 11.

[66] Ibid, at paragraphs. 34–5.

[67] See ECtHR, *Guide on Article 17 of the European Convention on Human Rights: Prohibition of Abuse of Rights*, 31 August, 2022. Available at: www.echr.coe.int/Documents/Guide_Art_17_ENG.pdf.

[68] Application No. 27510/08, Perinçek v. Switzerland, 15 October, 2015, paragraph 131.

[69] Application No. 201/17, Magyar Kétfarkú Kutya Pártv.Hungary, 20 January 2020, paragraph 99.

some expressions that constitute hate speech may be addressed through civil or administrative remedies. This is also recognised by the First Protocol.[70]

Second, if the ECtHR finds that the interference was prescribed by law, it examines whether the interference with the freedom of expression serves to advance one of the legitimate aims set forth in Article 10(2). No other aims as those set therein are considered to be legitimate grounds restricting freedom of expression.

Third, after the Court establishes that the interference pursued a legitimate aim, it continues with examining whether the interference in question was necessary in a democratic society. The Court has developed in its case-law the autonomous concept of whether an interference is "proportionate to the legitimate aim pursued". This is determined by having regard to all the circumstances of the case using criteria established in the ECtHR's case-law and with the assistance of various principles and interpretation tools.[71]

When finally determining whether the restriction under Article 10 was legitimate, it considers the words used and the context in which they were published, as well as to their potential impact. The ECtHR takes into consideration factors such as content of the expression, intent of the speaker, likelihood of harm, political and social context at the time the expression was made, the significance and credible nature of the potential harm, the degree to which the person is well known in the society, how the expression is disseminated, or the target of the statement. The Court has also emphasised the importance of the interplay between these factors, rather than any one of them taken in isolation, in determining the outcome of the case.[72]

2.5.2 Exclusion from the protection of the Convention

Some expressions of racist and xenophobic hate speech may engage Article 17 of the ECHR. Article 17, regarded as one of the most far-reaching provision of the ECHR, applies however only in exceptional circumstances. It is examined by the ECtHR if it is immediately clear that the applicant attempted to rely on a right under the ECHR to engage in an activity or to perform acts that are clearly contrary to the values of the ECHR and aimed at the destruction of the rights and freedoms laid down in it.[73] Article 17 prevents applicants from relying on the ECHR in order to perform or justify, for example, expressions of hatred, violence, xenophobia and racism, or anti-Semitism, in cases where it is evident that the applicant relied on his right to freedom of expression, and through this right intended to violate other rights under the ECHR. If it is not immediately clear that Article 17 is applicable, the ECtHR continues to follow its analysis under Article 10(2).

2.5.3 Balancing rights

Furthermore, when the right to freedom of expression interferes with other rights protected under the ECHR, the ECtHR examines whether a proper balance was struck by national authorities between the protection of the right to freedom of expression and other rights or values guaranteed by the ECHR.[74] If the balance is disproportionate and the protection of other rights prevails, free

[70] See for example, Explanatory Report to the First Protocol, paragraph 32.

[71] European Court of Human Rights, Guide on Article 10 of the European Convention on Human Rights. Freedom of Expression, updated on 31 August 2022, paragraph 88.

[72] Application no. 27510/08, Perinçek v. Switzerland, 15 October 2015, paragraph 208. See also Recommendation CM/REC(2022)16, paragraph 4 and Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraphs 32-35.

[73] Application no. 27510/08, Perinçek v. Switzerland, 15 October 2015, paragraph 114-115.

[74] Application no. 27510/08, Perinçek v. Switzerland, 15 October 2015, paragraph 274.

speech protection is not granted. The right to respect for private life (Article 8 of the ECHR) is the article that comes into question most frequently. However, in order for Article 8 to come into play, an attack on a person's reputation must attain a certain level of seriousness and be made in a manner causing prejudice to personal enjoyment of the right to respect for private life.[75] It should be emphasised that the specific features of the Internet may be taken into account in assessing the level of seriousness in order for an attack on personal reputation to fall within the scope of Article 8.[76]

Relevant factors that the Court may consider are:

- the characteristics of the group (size, its degree of homogeneity, its particular vulnerability or history of stigmatisation, and its position vis-à-vis society as a whole);
- the precise content of the negative statements regarding the group; and
- the form and context in which the statements were made, their reach, the position and status of their author, and the extent to which they could be considered to have affected a core aspect of the group's identity and dignity.[77]

In its balancing exercise the Court may also consider principles formulated in its case-law related to Article 10. Namely, whether the statements were made against a tense political or social background; whether those statements, fairly construed and examined in their immediate or wider context, could be seen as a direct or indirect call to violence or as a justification for violence, hatred or intolerance; it will also assess the manner in which the impugned statements were made and their capacity – direct or indirect – to lead to harmful consequences.

The ECtHR's approach to these types of cases can thus be described as highly context specific.[78]

2.6 Related standards, tools and initiatives[79]

The following are examples of relevant Council of Europe standards, tools and initiatives:

- Recommendation CM/Rec(2023)2 of the Committee of Ministers to member States on rights, services and support for victims of crime.[80] It calls upon member States to actively seek out and remove any barriers in the access to justice for victims of crime. It can be accessed here.

[75] Applications no. 1759/08, 50766/10 and 50782/10, Kaboğlu and Oran v. Turkey, 2018, paragraph 65.

[76] Application no. 58781/13, Arnarson v. Iceland, 2017, paragraph 37.

[77] ECtHR, Key theme Articles 8, 13 and 14 Protection against hate speech, 19 September 2022, p. 3. See also Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraph 32-34.

[78] Application no. 10851/13, Király and Dömötör v. Hungary, 2017, paragraph 73-74.
See also the Guide on Article 10 of the European Convention of Human Rights with a detailed list of cases in the appendix.

[79] This section contains only some examples. For a more comprehensive list please consult for example, UN Office of the High Commissioner for Human Rights, Compendium of international and regional standards against racism, racial discrimination, xenophobia and related intolerance : note / by the Office of the High Commission [that is, Commissioner] for Human Rights, 2004. Available at: https://digitallibrary.un.org/record/512629?ln=en

[80] Committee of Ministers of the Council of Europe, 15 March 2023, CM/Rec(2023)2.

- Recommendation CM/Rec(2022)16 of the Committee of Ministers to member States on combating hate speech[81]: It contains a set of guidelines aimed at preventing and combating hate speech, both online and offline. It calls on governments to develop comprehensive strategies to prevent and fight hate speech, including the adoption of an effective legal framework and implementing adequately calibrated and proportionate measures. The recommendation can be accessed here.

- Recommendation CM/Rec (2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries[82]: It provides for guidelines for States on actions to be taken vis-à-vis internet intermediaries with due regard to their roles and responsibilities. They focus on the rights to non-discrimination, freedom of expression and information, freedom of association, privacy, education, protection and safety, as well as access to remedies. It can be accessed here.

- ECRI General Policy Recommendation N°5 (revised) on preventing and combating anti-Muslim racism and discrimination (2021)[83]: adopted in 2000 and revised in 2021, it acknowledges that the phenomenon of anti-Muslim hatred and prejudice is multi-layered and intersectional, and includes stigmatisation and discrimination in various areas of life as well as anti-Muslim hate speech and hate crime, including online. This instrument provides comprehensive guidance to governments on addressing anti-Muslim racism and discrimination in four specific areas: policies and institutional coordination, prevention, protection, prosecution / law enforcement. It can be accessed here.

- ECRI General Policy Recommendation N°6 combating the dissemination of racist, xenophobic and antisemitic materiel via the internet (2000)[84]: It requests Governments to take the necessary measures, at national and international levels, to act effectively against the use of Internet for racist, xenophobic and antisemitic aims. It can be accessed here.

- ECRI General Policy Recommendation N°7 (revised) on national legislation to combat racism and racial discrimination (2017)[85]: Adopted in 2002 and revised in 2017 it sets out the key elements which should feature in a comprehensive national legislation to effectively combat racism and racial discrimination. The scope of the Recommendation is very wide and covers all branches of the law, that is, constitutional, criminal, civil and administrative. It can be accessed here.

- ECRI General Policy Recommendation N°9 (revised) on preventing and combating Antisemitism (2021)[86]: adopted in 2004 and revised in 2021, it sets out a comprehensive set of legal and policy measures to help governments prevent and combat antisemitism, including online (denial and distortion of Holocaust, the continued spread of hatred against Jews by a wide range of individuals and groups). This instrument provides

[81] Committee of Ministers of the Council of Europe, 20 May 2022, CM/Rec (2022)16.

[82] Committee of Ministers of the Council of Europe, 7 March 2018, CM/Rec(2018)2.

[83] ECRI General Policy Recommendation N°5 (revised) on preventing and combating anti-Muslim racism and discrimination, adopted on 16 March 2000 and revised on 8 December 2021.

[84] ECRI General Policy Recommendation No. 6 on Combating the Dissemination of Racist, Xenophobic, and Antisemitic Material via the Internet, 15 December 2000.

[85] ECRI General Policy Recommendation No. 7 on National Legislation to Combat Racism and Racial Discrimination, 13 December 2002, revised 7 December 2017.,.

[86] ECRI General Policy Recommendation No. 9 on Preventing and Combating Antisemitism, 25 June 2004, revised 1 July 2021.

guidance in four main fields of action against antisemitism, including contemporary forms: policies and institutional coordination, prevention / education, protection, and prosecution / law enforcement. It can be accessed here.

- ECRI General Policy Recommendation N°15 (2015)[87]: this instrument calls for speedy reactions by public figures to hate speech; promotion of self-regulation of media; raising awareness of the dangerous consequences of hate speech; withdrawing financial and other support from political parties that actively use hate speech; and criminalising its most extreme manifestations, while respecting freedom of expression. It can be accessed here.

- ECRI Statement on preventing and combating ultra-nationalistic and racist hate speech and violence in relation to confrontations and unresolved conflicts in Europe (2021)[88]: alarmed by the use of inflammatory rhetoric and the wide dissemination of hateful and dehumanising content, notably on the internet, in the context of confrontations and unresolved conflicts in Europe, ECRI calls upon all stakeholders, including at the highest political level, to take preventive action, ensure accountability and engage in confidence-building measures. It can be accessed here.

- No Hate Speech Movement Youth Campaign: It was launched by the Council of Europe in May 2013, extended until the end of 2017, by which time it had been launched in 45 countries including both member and non-member States of the Council of Europe. Driven by the need to counter on-line hate speech in all its forms, including those that most affect young people, such as cyber-bullying and cyber-hate, racism and other forms of discrimination. More information available here.

- Committee of Ministers Action Plan on the fight against violent extremism and radicalisation leading to terrorism (2015)[89] which had two objectives: 1. to reinforce the legal framework against terrorism and violent extremism; and 2. to prevent and fight violent radicalisation through concrete measures in the public sector, in particular in schools and prisons, and on the Internet. It can be accessed here. The final report of the SG on its implementation of 2018 provides an overview of the implementation of the Action Plan.

- Parliamentary Assembly Recommendation 1805 (2007) – Blasphemy, religious insults and hate speech against persons on grounds of their religion[90]: In this PACE Recommendation the Parliamentary Assembly of the Council of Europe considers that blasphemy, as an insult to a religion, should not be deemed a criminal offence. This is in line with the First Protocol, which stipulates that hatred, discrimination or violence directed against any individual or group for the reason of religion shall be criminalised only if used as a pretext for hatred, discrimination or violence based on based on race, colour, descent or national or ethnic origin. The recommendation can be accessed here.

---

[87] ECRI General Policy Recommendation No. 15 on Combating Hate Speech, 8 December 2015.

[88] ECRI Statement on preventing and combating ultra-nationalistic and racist hate speech and violence in relation to confrontations and unresolved conflicts in Europe, adopted at ECRI's 85th plenary meeting (30-31 March 2021).

[89] Committee of Ministers, Action Plan on the fight against violent extremism and radicalisation leading to terrorism, Brussels, 19 May 2015, CM(2015)74.

[90] Text adopted by the Assembly on 29 June 2007 (27th Sitting).

- Parliamentary Assembly Resolution 2275 (2019) - The role and responsibilities of political leaders in combating hate speech and intolerance[91]: It considered that a range of measures is necessary to counter hate speech, ranging from self-regulation, particularly by political movements and parties, and in the statutes and rules of procedure of national and local elected bodies, to civil, administrative and criminal legislation prohibiting and sanctioning its use, which should be considered as a last resort. It can be accessed here.

- Parliamentary Assembly Resolution 2317(2020) - Threats to media freedom and journalists' security in Europe[92]: Although it does not contain explicit reference to online racism or xenophobia, it aims at strengthening media freedom in all its aspects, including the safeguarding of editorial independence and of the ability to investigate, criticise and contribute to public debate without fear of pressure or interference. It can be accessed here.

- No Hate Parliamentary Alliance[93]: A platform for parliamentary activities to tackle intolerance, hate speech and all forms of racism including Afrophobia, anti-Gypsyism, antisemitism, Islamophobia and LGBTI-phobia. It has organised several national and international activities in co-operation with national parliaments. More information available here.

- Future instruments of the Council of Europe framework: Work is underway on developing a complementary comprehensive Recommendation on combating hate crime[94], which is focused more broadly on addressing the criminal law aspects of hate, including where committed through, or facilitated by the internet.

- Cybercrime Programme Office of the Council of Europe: C-PROC is responsible for assisting countries worldwide in strengthening their legal systems capacity to respond to the challenges posed by cybercrime and electronic evidence on the basis of the standards of the Budapest Convention on Cybercrime and its Protocols, including the First Protocol.[95]

- Parliamentary toolkit on hate speech (2023)[96]: The aim of the document is to provide parliamentarians with a toolkit for effective protection of victims of hate speech. Particular focus is also paid to combating online hate speech. The document outlines how to ensure that existing or proposed hate speech laws are compatible with human rights, while protecting the victims of hate speech. It can be accessed here.

- HELP[97] Tutored Course on hate speech and hate crime: Aimed in assisting legal professionals throughout Europe to understand hate crime and hate speech and deal with them in their daily work. It can be accessed here.

---

[91] Text adopted by the Assembly on 10 April 2019 (15th Sitting).

[92] Text adopted by the Assembly on 28 January 2020 (4th Sitting).

[93] See No Hate Parliamentary Alliance.

[94] See PC/ADI-CH Committee of Experts on Hate Crime – Committee of experts on hate crime (PC/ADI-CH) (coe.int)

[95] Council of Europe Office on Cybercrime in Bucharest: C-PROC activity report for the period October 2021 – December 2022, SG/Inf(2023)1, 9 January 2023, p. 15-16.

[96] Document prepared by Chara Bakalis, Reader in Law, Oxford Brookes University, in cooperation with the Secretariat of the Committee on Equality and Non-Discrimination of the Parliamentary Assembly of the Council of Europe, February 2023.

[97] Council of Europe: Human Rights Education For Legal Professionals (HELP).

- Models of Governance of Online Hate Speech[98]: The study seeks to map and explain but also to critically evaluate governance tools for online hate speech across Europe. It can be accessed here.

3 Implementing the Protocol on Xenophobia and Racism: Good practices

3.1 Criminal justice responses

3.1.1 Legislation

This subsection explores examples of good practice in implementing the Protocol in terms of domestic legislation. Many examples are available of the criminalisation, or other forms of prohibition, of the conduct of persons who use computer systems to commit acts of a racist or xenophobic nature according to Articles 3 to 7. The text of domestic provisions mentioned in this section can be found in the Appendix to this study.

3.1.1.1 Use of terms and specific elements of the offences

Racist and xenophobic material is defined in Article 2(1) of the Protocol as:

> "any written material, any image or any other representation of ideas or theories, which advocates, promotes or incites hatred, discrimination or violence, against any individual or group of individuals, based on race, colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors."

This definition refers to conduct to which the content of the material may lead, rather than to the expression of feelings/belief/aversion as contained in the material. The definition builds upon existing national and international definitions and documents. The Explanatory Report to the First Protocol further explains the specific terms used in the definition of "racist and xenophobic material" to provide greater clarity and guidance to the Parties. [99]

It should be noted that criminal law measures – including those of the First Protocol – are an important part of the response but should be used as the last resort.[100] And the offences prescribed in the First Protocol – like those of the Convention – must be committed "intentionally" for criminal liability to apply.

Furthermore, it is important to highlight the element "without right". According to the Explanatory Report to the First Protocol, this element:

> "reflects the insight that the conduct described is not always punishable per se, but may be legal or justified not only in cases where classical legal defences are applicable, like consent, self-defence or necessity, but where other principles or interests lead to the exclusion of

[98] Brown, A. (2020) Models of Governance of Online Hate Speech. Strasbourg: Council of Europe.

[99] See Explanatory Report to the First Protocol, paragraphs 10-22.

[100] See for example, Key messages of the Conference on xenophobia and racism committed through computer systems (Strasbourg, 30-31 January 2023). Available at: https://rm.coe.int/2542-111-key-messages-xr-conference-2023-v5-eng/1680aa2379

> criminal liability (for example, for law enforcement purposes, for academic or research purposes). The expression "without right" derives its meaning from the context in which it is used. Thus, without restricting how Parties may implement the concept in their domestic law, it may refer to conduct undertaken without authority (whether legislative, executive, administrative, judicial, contractual or consensual) or conduct that is otherwise not covered by established legal defences, excuses, justifications or relevant principles under domestic law."[101]

For example, **Slovakia** does not criminalise conduct relating to such a material that is demonstrably produced, distributed, put into circulation, made publicly accessible or kept in possession for the purpose of educational, collection or research activities.[102] Such possession is, therefore, deemed to be with right.

Another illustration of the application of the concept "without right" – although not in relation to a criminal offence – is found in **Australia's** s. 18C of the Racial Discrimination Act 1975 that makes it unlawful for a person to perform a public act that is "reasonably likely, in all the circumstances, to offend, insult, humiliate or intimidate another person or a group of people", where "the act is done because of the race, colour or national or ethnic origin of the other person or of some or all of the people in the group". In addition to this, s. 18D sets out the following multi-pronged exemption: "Section 18C does not render unlawful anything said or done reasonably and in good faith: (a) in the performance, exhibition or distribution of an artistic work; or (b) in the course of any statement, publication, discussion or debate made or held for any genuine academic, artistic or scientific purpose or any other genuine purpose in the public interest; or (c) in making or publishing: (i) a fair and accurate report of any event or matter of public interest; or (ii) a fair comment on any event or matter of public interest if the comment is an expression of a genuine belief held by the person making the comment."

3.1.1.2 Article 3 – Dissemination of racist and xenophobic material through computer systems

Article 3 of the First Protocol reads:

> Article 3 – Dissemination of racist and xenophobic material through computer systems
>
> 1 Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, the following conduct:
>
> distributing, or otherwise making available, racist and xenophobic material to the public through a computer system.
>
> 2 A Party may reserve the right not to attach criminal liability to conduct as defined by paragraph 1 of this article, where the material, as defined in Article 2, paragraph 1, advocates, promotes or incites discrimination that is not associated with hatred or violence, provided that other effective remedies are available.
>
> 3 Notwithstanding paragraph 2 of this article, a Party may reserve the right not to apply paragraph 1 to those cases of discrimination for which, due to established principles in

[101] Explanatory Report to the First Protocol, paragraph 24.

[102] Article 130(8) of the Criminal Code.

its national legal system concerning freedom of expression, it cannot provide for effective remedies as referred to in the said paragraph 2.

Arguably, the obligation under Article 3 has been legislatively fulfilled, albeit with greater or lesser degrees of compliance in terms of the exact wording and the precise scope of the legislation, by several Parties, as well as by some other States. Examples of relevant domestic laws can be found in the criminal codes of Armenia,[103] Finland,[104] Germany,[105] Lithuania,[106] North Macedonia,[107] the Netherlands,[108] Norway,[109] Serbia,[110] Slovakia,[111] and Spain[112].

Some examples for illustration:[113]

- **Armenia** – Section 226 of the Criminal Code:

  Incitement of national, racial, or religious hostility

  1. Actions targeted at incitement of national, racial, or religious hostility, at racial superiority or humiliation of national dignity — shall be punished by a fine in the amount of two-hundred-fold to five-hundred-fold of the minimum salary or by imprisonment for a term of two to four years.
  2. The acts provided for in part 1 of this Article, which have been committed — (1) publicly or by use of mass media; (2) by use of violence or threat thereof; (3) by use of official position; (4) by an organised group — shall be punished by imprisonment for a term of three to six years.

- **Finland** – Chapter 11, Section 10 of the Criminal Code:

  Agitation against a population group

  A person who makes available to the public or otherwise disseminates among the public or keeps available to the public information, an opinion or another message where a certain group is threatened, defamed or insulted on the basis of its race, colour, birth, national or ethnic origin, religion or belief, sexual orientation or disability or on another comparable basis shall be sentenced for agitation against a population group to a fine or to imprisonment for at most two years.

- **Serbia** – Article 387(4) of the Criminal Code (introduced in 2009 when ratifying the Protocol):

  (4) Whoever spreads or otherwise makes publicly available texts, images or any other representation of ideas or theories advocating or encouraging hatred, discrimination or

[103] s. 226 of the Criminal Code.
[104] Chapter 11, s. 10 of the Criminal Code.
[105] s. 130(2) of the Criminal Code ("Volksverhetzung").
[106] Article 170 of the Criminal Code.
[107] Article 319 of the Criminal Code.
[108] Article 137(d) and Article 422b of the penal code.
[109] Article 185 of the penal code.
[110] Article 387 of the criminal law.
[111] Article 130(7)(c) of the Criminal Code.
[112] Article 510(1) of the Criminal Code.
[113] See appendix for additional and more detailed examples.

> violence against any person or group of persons based on race, colour, religious affiliation, nationality, ethnic origin or other personal property, shall be punished with imprisonment of three months to three years.

Some examples in relation to Article 3 that reflect domestic context include the prohibition of disseminating material that incites hatred against people on grounds of protected characteristics by means of administrative, civil or human rights law, including forms of anti-discrimination laws. Examples of relevant domestic laws can be found in **Mexico**[114] or **New Zealand**[115]. It is worth reiterating that Article 3(2) explicitly allows for States not to attach criminal liability to conduct that "advocates, promotes or incites discrimination that is not associated with hatred or violence, provided that other effective remedies are available". The Explanatory Report to the First Protocol clarifies as follows: "For instance, those remedies may be civil or administrative".[116]

3.1.1.3 Article 4 – Racist and xenophobic motivated threat

Article 4 of the First Protocol reads:

> Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, the following conduct:
>
> threatening, through a computer system, with the commission of a serious criminal offence as defined under its domestic law, (i) persons for the reason that they belong to a group, distinguished by race, colour, descent or national or ethnic origin, as well as religion, if used as a pretext for any of these factors, or (ii) a group of persons which is distinguished by any of these characteristics.

Germany has fulfilled the requirements of Article 4 by extending the applicability of existing criminal provisions on threats in general through a further hate crime provision. Threats are covered by s. 126 (Disturbing public peace by threatening the commission of offences) and s. 241 of the Criminal Code (Threatening commission of a serious criminal offence). Under s. 241 of the Criminal Code, offences committed publicly, for instance on the internet, incur a higher penalty. More importantly, the Criminal Code includes a further hate crime provision explicitly recognising "the offender's motives and objectives, in particular including racist, xenophobic, antisemitic or other motives evidencing contempt for humanity" as an aggravating factor or circumstance in the commission of criminal acts such as threats.[117] The adoption of such provision was recommended by ECRI in its Fifth Report on Germany in 2014,[118] and the subsequent adoption of the provision was welcomed by ECRI in its Sixth Report.[119]

---

[114] Article 9(27), in conjunction with Article 4, of the Federal Law for the Prevention and Elimination of Discrimination.

[115] Article 61 of the Human Rights Act 1993.

[116] Explanatory Report to the First Protocol, paragraph 32.

[117] s. 46(2) of the Criminal Code.

[118] ECRI, *Fifth Report on Germany*, 25 February 2014, paragraph 10. Available at: https://rm.coe.int/fifth-report-on-germany/16808b5683.

[119] ECRI, *Sixth Report on Germany*, 17 March 2020, paragraph 69. Available at: https://rm.coe.int/ecri-report-on-germany-sixth-monitoring-cycle-/16809ce4be.

Section 46 Criminal Code – General principles

(1) The offender's guilt provides the basis on which the penalty is fixed. The effects which the penalty can be expected to have on the offender's future life in society are to be taken into account.

(2) When fixing the penalty, the court weighs the circumstances which speak in favour of and those which speak against the offender. The following, in particular, may be taken into consideration: the offender's motives and objectives, in particular including racist, xenophobic, antisemitic, gender-specific, anti-sexual orientation or other motives evidencing contempt for humanity (…).

Section 126 Criminal Code – Disturbing public peace by threatening to commit offences

(1) Whoever, in a manner which is suitable for causing a disturbance of the public peace, threatens to commit

1. breach of the peace as designated in section 125a sentence 2 nos. 1 to 4,
2. murder under specific aggravating circumstances (section 211), murder (section 212) or genocide (section 6 of the Code of Crimes against International Law) or a crime against humanity (section 7 of the Code of Crimes against International Law) or a war crime (section 8, 9, 10, 11 or 12 of the Code of Crimes against International Law),
3. grievous bodily harm (section 226),
4. an offence against personal liberty under section 232 (3) sentence 2, section 232a (3), (4) or (5), section 232b (3) or (4), section 233a (3) or (4), each to the extent that it represents a serious criminal offence, section 234, 234a, 239a or 239b,
5. robbery or extortion with use of force or threat of force (sections 249 to 251 or section 255), 6. a serious criminal offence constituting a public danger under sections 306 to 306c or section 307 (1) to (3), section 308 (1) to (3), section 309 (1) to (4), section 313, section 314 or section 315 (3), section 315b (3), section 316a (1) or (3), section 316c (1) or (3) or section 318 (3) or (4) or
7. a less serious criminal offence constituting a public danger under section 309 (6), section 311 (1), section 316b (1), section 317 (1) or section 318 (1) incurs a penalty of imprisonment for a term not exceeding three years or a fine. (2) Whoever, despite knowing better and in a manner which is suitable for causing a disturbance of the public peace, pretends that the commission of one of the unlawful acts referred to in subsection (1) is imminent incurs the same penalty.

Section 241 Criminal Code – Threatening commission of serious criminal offence

(1) Whoever threatens a person with the commission of a serious criminal offence against that person or a person close to him or her incurs a penalty of imprisonment for a term not exceeding one year or a fine.

(2) Whoever, despite knowing better, pretends to another person that the commission of a serious criminal offence against that person or a person close to him or her is imminent incurs the same penalty.

Incorporation of Article 4 of the First Additional Protocol into the domestic law of **Serbia** is reflected in paragraph 6 of Article 387 of the Criminal Code:

(6) Whoever publicly threatens to commit a criminal offence punishable with imprisonment of four and more years against a person or group of persons because of a particular race, colour, religion, nationality, ethnic origin or because of other personal property, shall be punished with imprisonment of three months to three years.

In **Spain**, conduct under Article 4 of the Protocol may be covered by the offence of threats punishable under Article 170 of the Spanish Criminal Code, or – if the elements required by that offence are not met – the generic aggravating circumstances of Article 22.4 may be applied:

Article 170 Criminal Code

1. Should the intimidation be of a harm which constitutes a criminal offence is intended to cause fear among the inhabitants of a location, ethnic, cultural or religious group, or a social or professional group, or any other group of persons, and if serious enough for such harm to be inflicted, the respective higher degree of penalties than those foreseen in the preceding Article shall be imposed.
2. A sentence of imprisonment from six months to two years shall be applied to those who, for the same purpose and severity, publicly call for violent deeds to be committed by armed gangs, organisations or terrorist groups.

Article 22 Criminal Code

The following are aggravating circumstances:
(…)
4. Committing the criminal offence for racist or anti-Semitic reasons, or another kind of discrimination related to ideology, religion or belief of the victim, ethnicity, race or nation to which he belongs, his gender, sexual orientation or identity, reasons related to gender, illness suffered or disability.

3.1.1.4 Article 5 – Racist and xenophobic motivated insult

Article 5 of the First Protocol reads:

Article 5 – Racist and xenophobic motivated insult

1 Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, the following conduct: insulting publicly, through a computer system, (i) persons for the reason that they belong to a group distinguished by race, colour, descent or national or ethnic origin, as well as religion, if used as a pretext for any of these factors; or (ii) a group of persons which is distinguished by any of these characteristics.

2 A Party may either: a require that the offence referred to in paragraph 1 of this article has the effect that the person or group of persons referred to in paragraph 1 is exposed to hatred, contempt or ridicule; or b reserve the right not to apply, in whole or in part, paragraph 1 of this article.

France, for example, meets this requirement through Article 33 of the Press Law of 1881 (as amended[120]). It establishes that the basic or base offence of insult – as defined in Article 29 ("Any insulting expression, terms of contempt or invective that do not contain the imputation of any fact is an insult") and when done by means of communication to the public as defined in Article 23 (including "any means of communication to the public by electronic means") – when committed "against a person or group of persons because of their origin or their membership or non-membership or non-membership of a particular ethnic group, nation, race or religion", "shall be punishable by one year's imprisonment and a fine of 45,000 euros." Moreover, Article 33 also extends the offence to cover insults against people based on other protected characteristics: "The penalties set out in the previous paragraph will apply to any insult committed under the same conditions against a person or group of persons on the grounds of their sex, sexual orientation, gender identity or disability."[121]

It is furthermore worth noting here that when hate is not an aggravating element, the sanction for the offence of insult under French law is only a fine and not imprisonment. This demonstrates how in France insult when aggravated by hate is treated as a more serious offence.

> Article 33 Press law of 1881 (version in force since 26 August 2021)
>
> (…)
>
> Sera punie d'un an d'emprisonnement et de 45 000 euros d'amende l'injure commise par les mêmes moyens envers une personne ou un groupe de personnes à raison de leur origine ou de leur appartenance ou de leur non-appartenance à une ethnie, une nation, une race ou une religion déterminée.
>
> Sera punie des peines prévues à l'alinéa précédent l'injure commise dans les mêmes conditions envers une personne ou un groupe de personnes à raison de leur sexe, de leur orientation sexuelle ou identité de genre ou de leur handicap.
>
> Lorsque les faits mentionnés aux troisième et quatrième alinéas du présent article sont commis par une personne dépositaire de l'autorité publique ou chargée d'une mission de service public dans l'exercice ou à l'occasion de l'exercice de ses fonctions ou de sa mission, les peines sont portées à trois ans d'emprisonnement et à 75 000 euros d'amende.
>
> En cas de condamnation pour l'un des faits prévus par les troisième et quatrième alinéas, le tribunal pourra en outre ordonner :
>
> 1° L'affichage ou la diffusion de la décision prononcée dans les conditions prévues par l'article 131-35 du code pénal ;

Germany implements Article 5 through several provisions of its Criminal Code. First, speakers who use a computer system to intentionally and publicly insult persons or groups of persons based on protected characteristics, or, indeed, any persons, can be charged with the criminal offence of insult.[122] As mentioned above, where insult has a racist or xenophobic motive it may be considered

[120] As last amended by Law n°2021-1109 of 24 August 2021
https://www.legifrance.gouv.fr/loda/id/LEGITEXT000006070722

[121] « Sera punie des peines prévues à l'alinéa précédent l'injure commise dans les mêmes conditions envers une personne ou un groupe de personnes à raison de leur sexe, de leur orientation sexuelle ou identité de genre ou de leur handicap. »

[122] s. 185 of the Criminal Code.

an aggravating factor.[123] Second, persons who use a computer system to intentionally and publicly insult persons or groups of persons based on protected characteristics in a manner that constitutes an attack on their dignity, can be charged with a criminal offence of hate speech.[124] Application of these provisions to material disseminated using a computer system is covered by the Criminal Code, according to which "'content' means that which is contained in writings, on audio or visual media, on data carriers, in images or other materialised content or which is also transmitted independently of any storage using information or communication technologies."[125]

Serbia considers the conduct covered by Article 5 to be a criminal offence against honour and reputation (insult), and after the ratification of the First Additional Protocol the legislator decided to prescribe this criminal offence in Chapter XVII of the Criminal Code: "Ruining the Reputation for Racial, Religious, Ethnic or other Affiliation". The substance of Article 174 CC reads:

> Whoever publicly ridicules a person or a group because of particular race, colour, religion, nationality, ethnic origin or other personal characteristic, shall be punished with a fine or imprisonment up to one year.

In **Australia**, examples of good practice in relation to Article 5 that reflect domestic context, include the prohibition of intentionally and publicly insulting persons for reasons of protected characteristics by means of domestic administrative, civil or human rights law. For an example in the federal legislation of Australia, see s. 18C of the Racial Discrimination Act 1975. In fact, this Australian provision is closer to a hate speech law than a hate crime law because it does not depend on a parallel or basic offence of insult. Nevertheless, what is key is that it is a civil rather than criminal law provision. As Gelber and McNamara explain:

> The civil laws require a person who believes an incident of hate speech has occurred to lodge a complaint in writing with a human rights authority (for example, the Anti-Discrimination Board in New South Wales, or the Australian Human Rights Commission under federal law). The authority investigates the complaint to ascertain whether vilification has occurred, and seeks to conciliate a confidential settlement between the complainant and respondent. The kinds of remedies that can be provided include an agreement to desist, apologise, or publish a retraction, or to conduct an educational campaign in a workplace. A complainant may terminate a complaint and commence civil proceedings in a tribunal (in a State or territory) or the Federal Court (under federal law).[126]

This Australian provision might be relevant for Parties that consider making a reservation pursuant to Article 5(2)b of the Protocol because they intend not to criminalise the relevant conduct, but who still wish to become parties to the Protocol by introducing non-criminal provisions to address this conduct.

[123] s. 46(2) of the Criminal Code.

[124] s. 130(2)(1)(c) of the Criminal Code.

[125] s. 11 (3) of the Criminal Code.

[126] See Gelber and McNamara (2015: 637).

3.1.1.5 Article 6 – Denial, gross minimisation, approval or justification of genocide or crimes against humanity

Article 6 of the First Protocol reads:

> Article 6 – Denial, gross minimisation, approval or justification of genocide or crimes against humanity
>
> 1 Each Party shall adopt such legislative measures as may be necessary to establish the following conduct as criminal offences under its domestic law, when committed intentionally and without right:
>
> distributing or otherwise making available, through a computer system to the public, material which denies, grossly minimises, approves or justifies acts constituting genocide or crimes against humanity, as defined by international law and recognised as such by final and binding decisions of the International Military Tribunal, established by the London Agreement of 8 August 1945, or of any other international court established by relevant international instruments and whose jurisdiction is recognised by that Party.
>
> 2 A Party may either
>
> a require that the denial or the gross minimisation referred to in paragraph 1 of this article is committed with the intent to incite hatred, discrimination or violence against any individual or group of individuals, based on race, colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors, or otherwise
>
> b reserve the right not to apply, in whole or in part, paragraph 1 of this article.

Article 6 explicitly relates not merely to Holocaust denial but also the denial, minimisation or justification of any genocide or crime against humanity recognised as such by the decisions of international courts established by relevant international instruments. As such, it also applies to the denial of the Rwandan genocide, for instance.[127] Notably, **Rwanda** is among the few non-European countries with genocide denial laws in place.[128] Its denialism law is an example of good practice reflecting Article 6, even though Rwanda has joined neither the Convention on Cybercrime nor the Protocol.

The obligation under Article 6 has been legislatively met, with greater or lesser degrees of compliance in wording and scope, by several Parties, as well as by some States who have been invited to accede such as by the Czech Republic,[129] France,[130] Germany,[131] Israel,[132] Lithuania,[133]

[127] See also the Explanatory Report to the First Protocol, paragraph 40.

[128] See Article 4 of Law No. 33bis/2003 of 2003 Repressing the Crime of Genocide, Crimes Against Humanity and War Crimes.

[129] Article 261 of the Criminal Code (as amended in 2001), replaced by Article 405 of the Criminal Code (as amended in 2010).

[130] Article 24-bis of Law on the Freedom of the Press of 1881 (as amended by the Gayssot Act of 1990, No. 90-615).

[131] s. 130 of the Criminal Code.

[132] Denial of Holocaust (Prohibition) Law (No. 5746-1986).

[133] Article 170-2 of the Criminal Code.

Luxembourg,[134] Poland,[135] Romania,[136] Serbia,[137] Slovakia,[138] Slovenia,[139] or Switzerland.[140] With the exception of Slovenia and Lithuania, all of them already had the relevant legislation in place prior to becoming Parties to the First Protocol.

Some examples for illustration:[141]

- **Czech Republic** – Section 405 Criminal Code: Denial, Impugnation, Approval and Justification of Genocide

  Whoever publicly denies, impugns, approves, or attempts to justify Nazi, Communist or any other genocide, or other crimes of the Nazis and Communists against humanity, shall be sentenced to imprisonment for six months to three years.

- **Germany** – Section 130 (3) Criminal Code

  (3) Whoever publicly or in a meeting approves of, denies or downplays an act committed under the rule of National Socialism of the kind indicated in section 6 (1) of the Code of Crimes against International Law in a manner suited to causing a disturbance of the public peace incurs a penalty of imprisonment for a term not exceeding five years or a fine.

- **Rwanda** – Article 4 Law No. 33n bis/2003 of 2003 Repressing the Crime of Genocide, Crimes Against Humanity and War Crimes

  Shall be sentenced to an imprisonment of ten (10) to twenty (20) years, any person who will have publicly shown, by his or her words, writings, images, or by any other means, that he or she has negated the genocide committed, rudely minimised it or attempted to justify or approve its grounds, or any person who will have hidden or destroyed its evidence. Where the crimes mentioned in the preceding paragraph are committed by an association or a political party, its dissolution shall be pronounced.

- **Switzerland** – Article 261bis (4) (6) Criminal Code

  4. any person who publicly denigrates or discriminates against another or a group of persons on the grounds of their race, ethnic origin, religion or sexual orientation in a manner that violates human dignity, whether verbally, in writing or pictorially, by using gestures, through acts of aggression or by other means, (…),
  6. shall be liable to a custodial sentence not exceeding three years or to a monetary penalty.

---

[134] Article 457(3) of the Criminal Code.

[135] Article 55 of the Act of 18 December 1998 on the Institute of National Remembrance Commission for the Prosecution of Crimes against the Polish Nation.

[136] Article 6 of Emergency Ordinance No. 31/2002 of 13 March 2002 (approved with amendments and additions by Law No. 107/2006 of 27 April 2006, and amended and supplemented by Law No. 217 of 23 July 2015).

[137] Article 387(5) of the Criminal Law.

[138] Arts. 130(7)(d) and 422d of the Criminal code.

[139] Article 300(1) of the Criminal Code (as amended in 2004), replaced with Article 297(2) of the Criminal Code (as amended in 2008).

[140] Article 261-bis of the Criminal Code (as amended by Article 1 of the Federal Act of 18 June 1993 and entered into force on 1 January 1995 following referendum FF 1993 II 868-69).

[141] See appendix for additional and more detailed examples.

3.1.1.6 Article 7 - Aiding and abetting

Article 7 of the First Protocol reads:

> Article 7 - Aiding and abetting
>
> Each Party shall adopt such legislative and other measures as may be necessary to establish as criminal offences under its domestic law, when committed intentionally and without right, aiding or abetting the commission of any of the offences established in accordance with this Protocol, with intent that such offence be committed. Article 7 creates a Party's obligation to establish criminal offences of abetting and aiding the commission of any of the other criminal offences envisaged under the Protocol under Articles 3–6, if such abetting and aiding is done intentionally and without right. For example, although the transmission of racist and xenophobic material through the Internet requires the assistance of service providers as a conduit, a service provider that does not have the criminal intent cannot incur liability under this section. Thus, there is no duty on a service provider to actively monitor content to avoid criminal liability under this provision.

The Explanatory Report to the First Protocol clarifies that the intentionality clause built into Articles 3 to 6 places a high bar on legitimate criminalisation of service providers by stating: "Persons cannot be held criminally liable for any of the offences in this Protocol, if they have not the required intent. It is not sufficient, for example, for a service provider to be held criminally liable under this provision, that such a service provider served as a conduit for, or hosted a website or newsroom containing such material, without the required intent under domestic law in the particular case. Moreover, a service provider is not required to monitor conduct to avoid criminal liability."[142] This is without prejudice to domestic laws that criminalise conduct of persons who fail to provide assistance based on court orders. In **Brazil**, for example, based on existing laws relating to contempt of court, senior managers of social media platforms and other Internet services providers may be held in contempt of court for failing to comply with court orders requiring them to take down illegal content, including illegal hate speech.

In **Germany**, Article 7 of the Protocol is implemented through s. 26 (Abetting) and s. 27 (Aiding) of the Criminal Code.

3.1.1.7 Protected characteristics

One notable aspect of the Protocol is the relatively narrow list of protected characteristics. The Protocol addresses conduct of a racist or xenophobic nature, which Article 2 defines in terms of the protected characteristics of race, colour, descent or national or ethnic origin, as well as religion but only when it is used as a pretext (or proxy) for these other characteristics, meaning it is not considered as a standalone or independent characteristic in its own right under the Protocol. Whilst it is good practice under the Protocol to criminalise conduct that is based on race, ethnicity, nationality, and religion, this does not mean that States may not take a different path. On the contrary, it is also in keeping with the spirit of the Protocol to criminalise conduct that is based on other protected characteristics in addition to those explicitly listed in the Protocol. The aforementioned 2022 Recommendation of the Committee of Ministers on combating hate speech proffers the following longer list of protected characteristics associated with the term "hate speech": "personal characteristics or status such as "race", colour, language, religion, nationality, national or ethnic origin, age, disability, sex, gender identity and sexual orientation."[143] According

[142] Explanatory Report to the First Protocol, paragraph 25.

[143] Ibid.

to the Explanatory Memorandum to the Recommendation, "The list of grounds is purposefully open-ended."[144] This longer list also chimes with the wider ecosystem of national and international hate speech laws.[145] Several years earlier, in 2016, ECRI offered a definition of hate speech as part of its General Policy Recommendation No. 15 on combating hate speech, which provides a similar list of protected characteristics that ends with the open-ended clause "and other personal characteristics or status".[146] In its General Policy Recommendation No. 17 on preventing and combating intolerance and discrimination against LGBTI persons, ECRI recommends to ensure that a comprehensive legal framework is developed to effectively prevent hate speech, including online hate speech, and prosecute anti-LGBTI hate speech of a criminal nature and other hate crimes (i.e. on the basis of based on actual or perceived sexual orientation, gender identity and sex characteristics).[147]

The dilemma of how many protected characteristics to include can be illustrated by incitement to hatred laws. Incitement to hatred laws are subject to significant jurisdictional variation not only in terms of how acts that count as inciting hatred are specified but also in terms of which protected characteristics are included.[148]

In **England** and **Wales**, for example, the offence of incitement to racial hatred concerns the use of threatening, abusive or insulting words or behaviour with the intent or likelihood of stirring up racial hatred, whereas the offences of incitement to religious hatred and incitement to sexual orientation hatred are defined far more narrowly in terms of using threatening words or behaviour with intent to stir up religious or sexual orientation hatred (Parts 3 and 3A of the Public Order Act 1986). But currently there is no offence of incitement to gender identity hatred, for example.

By contrast, in some other countries, such as the **Netherlands**,[149] hate speech laws cover not only the characteristics of race, ethnicity, nationality, and religion but also other protected characteristics, including "sex" but not "gender identity". In some jurisdictions, such as in **Queensland, Australia**, hate speech laws cover, amongst other things, the grounds not of "sex" but instead "gender identity".[150]

In **Malta**,[151] hate speech laws cover, amongst other things, "gender" and "gender identity". In **Norway**,[152] hate speech laws cover, amongst other things, "gender identity or gender expression". In **Uruguay**,[153] the relevant legislation covers, amongst other things, "sexual identity". In

[144] Explanatory Memorandum to the Recommendation of the Committee of Ministers on combating hate speech, paragraph 19.

[145] For a discussion of the full spectrum of protected characteristics that either are already or could be associated with hate speech laws, see Brown (2016) and (2017c).

[146] ECRI, GPR No. 15, Strasbourg, 21 March, 2016. Available at: https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech/16808b5b01. As regards references to "race", see ECRI's opinion on the concept of "racialisation" (adopted on 8 December 2022 at ECRI's 87th plenary meeting), available at: https://rm.coe.int/ecri-opinion-on-the-concept-of-racialisation/1680a4dcc2.

[147] ECRI, GPR No. 17, Strasbourg, published on 28 September, 2023. Available at: https://rm.coe.int/general-policy-recommendation-no-17-on-preventing-and-combating-intole/1680acb66f.

[148] See Brown (2015: 26–28).

[149] Article 137d of the penal code.

[150] See paragraphs 124A(1) and 131A(1) of the Anti-Discrimination Act 1991.

[151] Article 82A of the Criminal Code.

[152] Article 185 of the Criminal Code.

[153] Article 149 of the Criminal Code.

Canada[154] and **France**,[155] the relevant legislation covers, amongst other things, both "sex" and "gender identity". Finally, in some jurisdictions hate speech laws cover, amongst other things, subcategories or sub-identities of sex and gender identity, such as "gender identity or intersex variations of sex characteristics", in **Tasmania, Australia**,[156] and "transgender persons", in **New South Wales, Australia**.[157] Since the Protocol should be interpreted by States based on their own context, it can be argued that all of these States are demonstrating good practice despite having different configurations of protected characteristics.

Note, however, extending lists of protected characteristics that are relevant to domestic hate speech and hate crime laws could also be open to abuse by countries, such as if they expand the scope of protected characteristics in a way that facilitates the authoritarian suppression of social and political dissent, precisely the sort of suppression that the human right to freedom of expression is supposed to mitigate against.

In **Russia**[158], for example, the police and military soldiers are recognised as a "social group" for the purposes of interpreting and applying Article 282 of the Criminal Code: "Actions aimed at inciting hatred or enmity and humiliating the dignity of an individual or a group of individuals on the grounds of gender, race, ethnic origin, language, background, religious beliefs or membership of a social group".

However, this is where judicial interpretation comes back in. It is important for international human rights courts to be able to distinguish between, on the one hand, reasonable extensions of the list of protected characteristics based on country context in keeping with the spirit of international instruments like the Protocol, and, on the other hand, the abuse of such instruments as cover for censorship of dissent.

As another example, consider how the ECtHR has handled cases involving Russian hate speech laws. Article 29 of the Constitution of **Russia** protects the right to freedom of speech but also qualifies this right by permitting the prohibition of "Propaganda or agitation inciting social, racial, national or religious hatred and enmity". Along these lines, the Criminal Code of Russia prohibits "Actions aimed at inciting hatred or enmity and humiliating the dignity of an individual or a group of individuals on the grounds of gender, race, ethnic origin, language, background, religious beliefs or membership of a social group, committed publicly or through the mass media".[159] The idea of social hatred, as in inciting hatred against a social group, is not merely an extremely broad protected characteristic but quasi open-ended in the sense that almost any group could potentially be classified as a social group. Notably, the Russian State has previously used this hate speech law to prosecute individuals for publishing content that allegedly amounted to inciting hatred against Russian soldiers defined as a social group, including content allegedly inciting hatred against Russian soldiers involved in the war in Chechnya. In some of these cases, the convicted individuals have made complaints to the European Court of Human Rights (ECtHR). In some cases, the ECtHR has concluded that certain statements did not in fact amount to incitement to hatred against Russian soldiers whilst other statements did.[160] Importantly, the ECtHR has not rejected per se the inclusion of "social group" as a protected characteristic under national hate speech law.

---

[154] Article 319 of the Criminal Code.

[155] Article R. 625-7 of the Criminal Code.

[156] See paragraph 19 of the Anti-Discrimination Act 1998.

[157] See paragraphs 38R-38S of the Anti-Discrimination Act 1977.

[158] Neither a Party to the Convention nor to the First Protocol.

[159] Article 282(1) of the Criminal Code.

[160] See *Stomakhin v. Russia*, ECtHR, Strasbourg, 9 May, 2018, No. 52273/07, at paragraphs 107 and 114.

This reflects the court's broader commitment to the margin of appreciation doctrine: the doctrine that as an international human rights court it should afford to States and national courts a degree of discretion, latitude, and diversity based on country context in how they interpret human rights obligations.[161] However, the ECtHR has sometimes blocked unreasonable interpretations of social hate speech. In *Savva Terentyev v. Russia*,[162] for example, the ECtHR held that the prosecution, conviction and handing down of a suspended sentence against an individual for sharing online a provocative comment criticising police abuses constituted a violation of the right to freedom of expression under Article 10(1) of the European Convention and was not "necessary in a democratic society" under designated limits to that right established in Article 10(2) and ECtHR case law, even with the margin of appreciation.

3.1.1.8 Reservations and declarations

Article 12 of the First Protocol gives the right to Parties at the time of signature or when depositing its instrument of ratification, acceptance, approval, or accession to enter various reservations and declarations, both extending existing reservations and declarations to the Convention on Cybercrime and making new reservations[163] and declarations[164] under the First Protocol. So far, from 35 Parties to the First Protocol, 16 have made reservations or declarations under the Protocol.

In addition, a Party to the First Protocol may make reservations that are available in the Convention under Article 22(2) and Article 41(1), even if it did not make these when joining the Convention. However, to date, no Party has made use of this possibility.

3.1.1.8.1 Article 3 – Dissemination of racist and xenophobic material through computer systems

Paragraphs 2 and 3 permit a reservation in very limited circumstances (they should be read in conjunction and in sequence)[165]:

- First, a Party has the possibility not to attach criminal liability to the conduct contained in this article where the material advocates, promotes or incites discrimination that is not associated with hatred or violence, provided that other effective remedies are available. For instance, those remedies may be civil or administrative. **Croatia** made such a reservation.
- Second, a Party may reserve the right not to apply paragraph 1 to those cases of discrimination for which, due to established principles in its national legal system concerning freedom of expression, it cannot provide for effective remedies as referred to in paragraph 2. The scope of the reservation may be further restricted by the Party by requiring that the discrimination is, for instance, insulting, degrading, or threatening a group of persons.[166] **Denmark**, **Finland**, **Iceland**, **Norway** and **Poland** made such reservation.

Croatia is the only Party so far that made a declaration under paragraph 2 to not attach criminal liability to cases where the material, as defined in Article 2(1), advocates, promotes or incites discrimination that is not associated with hatred or violence. It will thus rely on other effective remedies in those cases.

---

[161] See Legg (2012).

[162] ECtHR, Strasbourg, 28 August 2018, No. 10692/09.

[163] Arts 3, 5 and 6.

[164] Article 5, paragraph 2a and Article 6 paragraph 2a allow for possibility to require additional elements

[165] See Explanatory Report to the First Protocol.

[166] Explanatory Report to the First Protocol, paragraph 32.

Of the Parties that made the reservation, **Denmark** reserved the right to fully or to partially refrain from criminalising acts covered by Article 3(1).

Finland, **Iceland** and **Norway** largely relied on the wording provided in paragraph 3, and reserved the right not to apply Article 3(1) to those cases of discrimination for which, due to established principles in its national legal system concerning freedom of expression, it cannot provide for effective remedies as referred to in Article 3(2).

Poland specified in its reservation that the necessary condition to consider conduct referred to in Article 3(1) a criminal offence, is that the discrimination is associated with violence or hatred, as referred to in Article 3(2).

3.1.1.8.2 Article 5 – Racist and xenophobic motivated insult

Article 5 allows for the possibility either to:

- require pursuant to paragraph 2(a) through a declaration, that the conduct must also have the effect that the person or group of persons are, not only potentially, but actually exposed to hatred, contempt or ridicule[167] (**Greece** made such a declaration); or
- go even further and pursuant to paragraph 2(b) reserve the right to not apply in whole or in part (**Denmark**, **Finland**, **Iceland**, **Norway** and **Romania** made such a reservation) with respect to Article 5(1)[168].

Greece is the only Party that entered a declaration under paragraph 2(a), requiring that the offence in Article 5(1) has the effect that the person or group of persons referred to in paragraph 1 is exposed to hatred, contempt or ridicule.

Of Parties that entered a reservation under paragraph 2(b) of this article, only **Romania** specifically indicated that it will not (fully) apply paragraph 1.

Finland specified that it reserves the right not to apply, in whole or in part, Article 5(1) to cases where the national provisions on defamation or ethnic agitation are not applicable.

Iceland, due to established principles of its national legal system concerning freedom of expression, reserved the right not to apply, in whole or in part, Article 5(1).

Norway reserved the right not to apply Article 5(1) except for offences of hatred.

Denmark relied on the wording of Article 5(2)(b) and reserved the right to fully or to partially refrain from criminalising acts covered by Article 5(1).

3.1.1.8.3 Article 6 – Denial, gross minimisation, approval or justification of genocide or crimes against humanity

As is the case of the previous article, with respect to Article 6, Parties are allowed either to:

- require pursuant to paragraph 2(a), through a declaration, that the denial or the gross minimisation referred to in Article 6(1), is committed with the intent to incite hatred, discrimination or violence against any individual or group of individuals, based on race,

[167] Explanatory Report to the First Protocol, paragraph 37.

[168] Explanatory Report to the First Protocol, paragraph 38.

colour, descent or national or ethnic origin, as well as religion if used as a pretext for any of these factors (**Greece**, **Lithuania**, **Poland**, **Sweden** and **Ukraine** made such a declaration/reservation); or

- make use of a reservation pursuant to paragraph 2(b), allowing a Party not to apply this provision in whole or in part (**Denmark**, **Finland**, **Iceland**, **Monaco**, **Montenegro**, **Netherlands** and **Norway** made such a reservation).[169]

Greece and **Lithuania** relied in their declaration made under paragraph 2(a) to a large extent to the text as stipulated in the said provision. This is also the case for **Ukraine**, although the different terms provided for in its declaration may be due to use of different terms in their national legislation through which the First Protocol is implemented.

Although not specifically provided for in the First Protocol, **Poland** and **Sweden** made a reservation under paragraph 2(a) that the conduct referred to in paragraph 1 (in the case of **Poland**) or the denial or the gross minimisation (in the case of **Sweden**) must be committed with intent as specified further in paragraph 2 of the said Article

Of Parties that made reservations to this article, **Denmark** and **Monaco** reserved the right to fully or to partially refrain from criminalising acts covered by Article 6(1).

Finland, due to established principles in its national legal system concerning freedom of expression, reserved the right not to apply, in whole or in part, Article 6(1) to cases where the national provisions on ethnic agitation are not applicable.

Iceland, due to established principles in its national legal system concerning freedom of expression, reserved the right not to apply, in whole or in part, Article 6(1).

Montenegro in its reservation indicated that it requires that the denial or the gross minimization, approval or justification of acts constituting genocide or crimes against humanity, be committed with the intent to incite hatred, discrimination, or violence against an individual or group of individuals based on race, colour, descent or national or ethnic origin, as well as religion if used as pretext for any of these factors, or otherwise.

Netherlands specified that it would comply with the obligation to criminalise the denial, gross minimisation, approval or justification of genocide or crimes against humanity laid down in Article 6(1) of the Protocol where such conduct incites hatred, discrimination or violence on the grounds of race or religion.

Norway reserved the right not to apply paragraph 1 of this Article, except for offences of hatred.

3.1.2 Reporting mechanisms

Even though not explicitly provided for in the Protocol, it is in keeping with the basic ethos or spirit of the Protocol for States to assist victims of unlawful online hate speech and hate crime offences. One way of doing this is via reporting mechanisms. Thus, it is good practice for governmental institutions to either directly put in place or indirectly support the creation within civil society of dedicated reporting mechanisms (for example, online platforms, telephone hotlines) that enable victims and third parties to report relevant instances. Reporting mechanism that contribute to countering hate offences more effectively may vary from country to country depending on context and institutional capacity. However, in general such mechanisms may enable reports to be sent

[169] Explanatory Report to the First Protocol, paragraph 43.

to and received by (a) law enforcements authorities (for example, special police units, special public prosecutors, regulators), or (b) civil society organisations (for example, trusted flaggers, institutions of regulated self-regulation) who assess and filter reports before passing them on to relevant law enforcement authorities and/or to the relevant social media platforms and other services providers.

An example of a reporting mechanism of type (a) is briefly outlined in ECRI's Fourth Report on **Finland**: "In March 2010, the police launched an online service for reporting, for example, racist or xenophobic material on the Internet."[170]

Likewise, in **France** the Central Office for Combating Information and Communication Technology Crime (OCLCTIC) provides Internet users with the PHAROS platform (Plateforme d'Harmonisation, d'Analyse, de Recoupement et d'Orientation des Signalements), which enables them to report illegal content and behaviour on the Internet.[171] The offences that can be reported are: pedophilia and child pornography; expression of racism, anti-Semitism and xenophobia; incitement to racial, ethnic and religious hatred; and terrorism and apology for terrorism.

When users come across illegal content, they can report it on the platform by entering the date on which they viewed the content, the site on which they viewed it, the type of infringement and the link to the content in question.

The reports are then analysed and evaluated by a team of police officers and gendarmes assigned to Pharos to determine whether or not they are infringements. Once the information has been cross-checked, relevant services, such as the national police, the gendarmerie, customs, the General Directorate for Competition, Consumer Affairs and Fraud Control (DGCCRF) in France, or INTERPOL, are alerted. An investigation is then opened under the authority of the Public Prosecutor.

In **Germany**, members of the public can file a complaint of suspected unlawful hate speech or hate crime offences that occur on the Internet by directly contacting any local police station.[172] Moreover, law enforcement authorities in many federal States (*Länder*), including Brandenburg, for example, operate public-facing websites that enable individuals to file criminal complaints using a special online reporting mechanism.[173] The Ministry of Justice of Bavaria in 2022 established the "Respect!" portal permitting the public to report online hate.[174] This portal is also used in Baden-Württemberg and other Länder.

In fact, **Germany** has an integrated reporting landscape, meaning that in addition to having in place reporting mechanisms of type (a) there is also an emphasis on governmental agencies taking proactive steps to fund, support and signpost to members of the public reporting mechanisms of type (b). At the heart of this system are initiatives and programmes at the federal and state levels which provide information and advice about preventing, dealing with and reporting unlawful hate speech and hate crime offences that occur on the Internet. Examples include a regularly organised Germany-wide Action Day on Combating Online Hate Posts. Its aim is to raise public awareness.

---

[170] ECRI, Fourth Report on Finland, 9 July 2013, paragraph 100. Available at: https://rm.coe.int/fourth-report-on-finland/16808b5714.

[171] https://www.internet-signalement.gouv.fr/PharosS1/

[172] Information available at: https://www.polizei-beratung.de/opferinformationen/.

[173] Information available at: https://polizei.brandenburg.de/page/how-to-file-charges-online/63001.

[174] https://meldestelle-respect.de/

In addition, under the "Live Democracy!" programme, the Federal Ministry for Family Affairs, Senior Citizens, Women and Youth funds the work of one "Democracy Centre" in each of Germany's "Länder". These Democracy Centres coordinate a wide range of advisory services by civil society actors, including mobile advisory services for preventive work and advisory services for the victims and survivors of right-wing, racist and xenophobic violence, in the field of extremism and group-focused enmity. They also refer victims to reporting mechanisms, whether of law enforcement agencies or civil society organisations. For example, a civil society app called "MeldeHelden", operated by HateAid, has a form through which victims and third parties can report unlawful hate speech and hate crime offences that occur on the Internet. HateAid is funded under the "Live Democracy!" federal programme as part of the Competence Network Against Internet Hate Crime.

Similarly, the Federal Association of Departments for Research and Information on Antisemitism (Bundesverband RIAS e.V.) is a federal government-funded civil society project with a nationwide reporting mechanism (that is, online portal) facilitating the reporting and recording of antisemitic incidents in Germany.[175] RIAS is also part of the Competence Network against Antisemitism funded as part of the federal programme "Live Democracy!".

In addition to this, the public-facing website of Germany's federal law enforcement authorities[176] contains links to a national civil society reporting mechanisms called "Respect!". This anti-hate online reporting mechanism is run by Youth Foundation Baden-Württemberg at the Democracy Centre. Its staff review the reports and, where necessary, takes further action, including passing reports on to relevant law enforcement authorities, social media platforms or services providers.

A further example of public awareness raising activities can be found in **Slovakia**, as the Police of the Slovak Republic periodically informs on its social media accounts (Facebook, Instagram) about possibilities for the public to report cases related to xenophobia and racism.

An example of entrusting monitoring of the Internet to a non-governmental entity is the conclusion of a memorandum of understanding between the National Criminal Agency of the Presidium of the Police of **Slovakia** with the non-governmental organisation DigiQ.[177] Based on this memorandum, DigiQ identifies and reports racist and xenophobic content on the Internet.

Countries might furthermore consider developing a "one-stop" reporting system whereby victims or third parties can use (i) a single official reporting system and (ii) a reporting system that allows them to report multiple hate speech content and/or hate crime incidents simultaneously, and which in turn shares the information with a range of appropriate law enforcement agencies. This has the benefit of reducing the burden on victims of having to report the same content or incident via several different systems.

Although not itself a reporting mechanism and also not explicitly provided for in the Protocol, a related example of good practice in assisting victims of hate speech and hate crime offences that occur on the Internet is the provision of victim support and guidance. In **Spain**, for example, the National Office for Combating Hate Crimes, which is under the responsibility of the Secretariat of State for Security, a division of the Ministry of the Interior, has a public-facing website that offers

[175] Information available at: https://www.antisemitismusbeauftragter.de/Webs/BAS/EN/fight-against-antisemitism/initiatives/rias-department/rias-department-node.html

[176] See www.polizeifuerdich.de (the '*Hilfsangebote*' (Help) section). See also www.polizei-beratung.de.

[177] DigiQ is also a member of the INACH International Network against Cyber Hate. https://www.inach.net/digiq/#

a list of resources for victim assistance differentiated by provinces[178] with contact details for a Victim Assistance Service, including a telephone number and email address.

Similarly, in **Germany** the Federal Government in conjunction with the 16 federal States (*Länder*) provides a public-facing website that contains a list of resources for victim assistance, differentiated by kinds of crime, including cybercrime and hate crime ("politically motivated crime").[179] A map search can be used to locate and contact local victim advisory and counselling services across Germany, many of which are run by police forces.[180]

A challenge with respect to such mechanisms is that of malicious reporting,[181] and the question of what measures, if any, should be taken to counter the problem of persons who report content that is not manifestly unlawful or for the purpose of bringing another person into legal jeopardy. People who file reports to the police or flag content to social media platforms may have various motivations and contextual triggers for doing so, many of which are appropriate and reasonable.[182] However, in the small minority of instances where malicious reporting does occur, it may pose a credible threat to the free speech of users or subscribers who post or share what is in fact legal content. To counter this problem, Article 1 of the Avia Law ("Loi Avia") in **France** amends existing legal provisions to create, amongst other things, an offence of malicious reporting, punishable by one year imprisonment and 15,000 euros fine.[183] However, on 18 June 2020 the French Constitutional Court (Conseil Constitutionnel) struck down as unconstitutional most of the key provisions of the Avia Law, finding that the substantial restrictions on freedom of expression and commercial freedoms embodied in the law were neither necessary, proportionate nor suited to the legislative goal of combating unlawful content online.[184] This decision puts into doubt the future of the law in general and, therefore, indirectly the new offence of malicious reporting. Perhaps there are other, less restrictive but equally effective, measures available to address the problem of malicious reporting. Alternative measures might include training sessions and best practice guidelines designed to ensure that social media platforms, civil society organisations (for example, trusted flaggers), and law enforcement authorities treat all reports as "unproven unless proven".

3.1.3 Statistics

Another practice not explicitly provided for in the Protocol but nevertheless in keeping with the spirit of the Protocol, is the collection, management and transparency of official statistics on instances of unlawful hate speech and hate crime offences that occur online. Such practice may be beneficial for various reasons, including to give recognition to victims and to show that their experiences are taken seriously by the State; to provide an empirical basis for statements that hate speech and hate crime are real problems that need to be addressed; to provide impetus for new legislative, law enforcement and other initiatives or better implementation of existing measures; to justify budget allocation and to determine where to allocate more resources; to assess whether existing measures have been effective in countering hate speech and hate crime; or to provide a comparison between the number of cases that are investigated, prosecuted and lead to successful convictions respectively.

---

[178] Available at: https://encuestadelitosdeodio.ses.mir.es/publico/encuestas/mapaRecursos.html

[179] Information available at: https://www.polizei-beratung.de/opferinformationen/

[180] Information available at: https://www.polizei-beratung.de/opferinformationen/beratungsstellensuche/

[181] See Brown (2020: 121–2).

[182] See, for example, Naab et al. (2018).

[183] Available at: https://www.assemblee-nationale.fr/dyn/15/textes/l15t0419_texte-adopte-provisoire.pdf

[184] Décision No. 2020-801 DC, du 18 juin 2020. Available at: https://www.conseil-constitutionnel.fr/sites/default/files/as/root/bank_mm/decisions/2020801dc/2020801dc.pdf

In **Germany**, for example, the statistical offices of the federal States receive individual data sets from criminal prosecution authorities, the administrative offices of the public prosecutors at local courts (*Amtsanwaltschaften*), public prosecution offices (*Staatsanwaltschaften*), and court registries. Following a plausibility check, the individual data sets are processed by the statistical offices of the federal States both for use in their own publications and in order that they can be sent to the Federal Statistical Office (*Statistisches Bundesamt*), which then compiles the data into federal statistics. In terms of the sorts of crimes highlighted in the Protocol, the Federal Criminal Police Office reporting service (*Bundeskriminalamt*) publishes annual statistics on "politically motivated crime".[185] Importantly, they also include statistics on the category "*Hasspostings*" (hate posts), if such cases have been reported to the police and can be categorised as politically motivated crime (that said, direct messages are not classed as "hate posts"). At the level of federal States, the Bavarian Ministry of Justice has published annual reports – "Hate-Speech-Bilanz" of the Bavarian justice system – for 2021 and 2022 with data on criminal proceedings concerning hate and hate speech online.[186]

In addition, Germany's "RED Database" ("Rechtsextremismusdatei RED")[187] comprises records of all preliminary investigations in Germany by the Federal Criminal Police Office (BKA), the Federal Police (Bundespolizei), the Federal Constitutional Protection Service (Bundesverfassungsschutz), the Military Security Service (Militärischer Abschirmdienst) as well as the constitutional protection services and criminal police offices of the 16 Länder. The survey methodology was significantly revised for the reporting year 2013, creating a reliable, nationally-standardised database. The "RED Database" covers offences pursuant to sections 86, 86a, 125, 125a, 130, 131, 211, 212, 223 to 231, 340, 306 to 306f of the German Criminal Code, as well as other offences where, following a legal assessment of the facts and circumstances, there are indications the offender had a right-wing extremist, xenophobic, or antisemitic motivation.

In terms of partnerships between governmental agencies and civil society organisations in the collection of data and production of statistics, the "Monitoring and Transferplatform MOTRA: Internet Monitoring" project is a joint civil security research project involving a number of partner organisations from the realms of science and practice in Germany but which is headed by the Federal Criminal Police Office. Its "Internet Monitoring" subproject monitors and analyses digital spaces, interactions and discourses in order to better understand radicalisation processes and extremist activities. The primary objective is to get a more precise overview of trends, topics and actors in the digital realm, including trends related to producers and targets of racist and xenophobic content on the Internet.[188]

Good practices can also be found in **Spain**. Every year statistics on hate crimes reported in Spain are compiled by the Ministry of the Interior and include information on profiles of perpetrators and victims, and whether crimes were committed via the Internet. In addition, the Attorney General's Office covers in its annual report data on hate crime proceedings, including statistics on convictions and acquittals. In addition, the Computer Crime Unit and the Hate Prosecutor's Office (divisions of

[185] See report for 2022 at: https://www.bka.de/SharedDocs/Downloads/DE/UnsereAufgaben/Deliktsbereiche/PMK/2022PMKFallzahlen.pdf?__blob=publicationFile&v=3

[186] Hate-Speech-Bilanz 2021 der bayerischen Justiz / 2.317 Verfahren wegen Hass und Hetze im Internet / Bayerns Justizminister Eisenreich: „Hass und Hetze im Netz haben sich zu einer echten Gefahr für die Demokratie entwickelt. Der Kampf gegen … – Bayerisches Landesportal https://www.rechtundpolitik.com/justiz/justizministerium-bayern/hate-speech-bilanz-2022-der-bayerischen-justiz/

[187] https://www.bfdi.bund.de/DE/Buerger/Inhalte/Polizei-Strafjustiz/National/RED.html

[188] Information available at: https://www.motra.info/radikalisierungsmonitoring/internetmonitoring/ .

the State Attorney General's Office) provide granular statistics on hate crimes committed using a computer system and on different forms of hate crime.[189]

The Guide for Criminal Justice Statistics on Cybercrime and Electronic Evidence (2020), developed jointly by the Council of Europe and INTERPOL, serves as a reference for enhancing specialized cybercrime capabilities of law enforcement and criminal justice systems in various national contexts. It presents measures these authorities can adopt to collect, process and maintain statistics on cybercrime and electronic evidence.

3.1.4 Specialised authorities

The Protocol creates obligations for Parties to criminalise certain forms of hate speech and hate crime, that is, acts of a racist and xenophobic manner committed through computers systems. It matters for the implementation of the Protocol not merely whether the relevant laws are in place but also whether they are properly enforced.

A key instrument of enforcement in several countries (for example, France, Finland, Germany, Slovakia and Spain) is that of specialised authorities. These are dedicated criminal justice authorities (operating at national, regional or local levels) with both expertise and special responsibilities in dealing with precisely the sort of unlawful content highlighted in Articles 3–7 of the Protocol. They may take the form, for instance, of special community police officers, special police cybercrime units or special public prosecutors for cybercrime.

For example:

- ECRI's fourth country report on **Finland** states: "The authorities have also informed ECRI that virtual community police officers are operating on the Internet and facilitate contact with the police."[190]

- In **Spain**, the Public Prosecutor's Office coordinates an "area of specialisation in cybercrime" which includes *inter alia* responsibility for combating unlawful hate speech and hate crime offences committed in the digital environment. This area of specialisation comprises a network of 150 prosecutors specialising in cybercrime.

- In **France** the national centre for combating online hate is embedded in the Paris public prosecutor's office, created by the law of 24 June 2020. The national centre centralises the handling of the most significant and complex cases. It is also the main contact for the PHAROS platform for harmonisation, analysis, cross-checking and orientation of reports. Since its creation, the centre has dealt with more than 140 cases. The Paris public prosecutor's office thus provides its expertise to all jurisdictions and can take on any case relating to the fight against online hate. It is composed of two magistrates, two legal assistants, and a specialist assistant.

Although the activities of specialised authorities vary between countries, taken together as a body of examples of good practice, relevant activities may include one or more of the following:

- monitoring the Internet for unlawful content;

[189] Information available at: https://www.fiscal.es/memorias/memoria2022/FISCALIA_SITE/index.html.

[190] ECRI, *Fourth Report on Finland*, 9 July 2013, paragraph 100. Available at: https://rm.coe.int/fourth-report-on-finland/16808b5714.

- developing algorithms to detect unlawful content online;

- specially trained police officers and/or community police officers working undercover online to gather intelligence and evidence on individuals suspected of committing criminal offences;

- assessing and, if appropriate, investigating reports of unlawful content, including reports sent by other law enforcement authorities, members of the public, civil society organisations, and social media platforms or other Internet services providers;

- entering into dialogue or liaising with services providers, as well as with other law enforcement authorities, governmental agencies, and independent institutions that have been accredited to play a role in the regulation of services providers, to establish agreed procedures or processes of substantive co-operation;

- organising and/or participating in special meetings, events or conferences bringing together a range of specialised authorities and other stakeholders to intensively discuss and promote improved co-ordination;

- determining local jurisdiction in cases of unlawful content;

- requesting information from Internet services providers about the identities of users or subscribers suspected of creating or sharing unlawful content and/or applying to courts to issue production orders for such information;

- sending administrative notice and take-down requests to Internet services providers concerning content deemed to be unlawful by the special authorities and/or applying to courts to issue judicial notice and take-down orders for unlawful content, as well creating special priority channels so that the relevant requests and judicial orders can be sent directly to the relevant personnel or divisions within Internet services providers to facilitate timely removal of unlawful content;

- launching criminal investigations against people who create or share unlawful content and, if appropriate, passing on the results of the investigations to other relevant public prosecutors;

- participating in periodic "action days" where multiple law enforcement authorities come together to simultaneously undertake mass or large-scale enforcement actions, including crackdowns on suspected creators and distributors of unlawful content and informing Internet services providers of possible unlawful content.

3.1.5 Public/private cooperation for criminal justice purposes

3.1.5.1 Guidelines on law enforcement/service provider cooperation

Co-operation frameworks may include *domestic co-operation* among different law enforcement authorities, other governmental agencies, civil society organisations and social media platforms and other services providers *within* a country but also *international co-operation between* Parties.

Regarding cooperation between criminal justice authorities and service providers, the Octopus Conference held by the Council of Europe in 2008 adopted a set of "Guidelines for cooperation

between law enforcement and Internet service providers".[191] Among other things, they promoted partnerships between law enforcement authorities and service providers, standard requests formats, the appointments of single points of contact, and so on. In December 2008, the European Court of Human Rights made reference to these guidelines in the decision K.U. v. Finland.[192] Although the guidelines did not result from a formal process of negotiations, they helped shape cooperation between law enforcement and service providers, including with major multi-national service providers, ever since.

3.1.5.2 Tools of the Convention on Cybercrime and its Second Protocol

While such guidelines for cooperation may be extremely useful, the disclosure of computer data for authorities to identify perpetrators and obtain evidence for use in criminal proceedings is often an interference with the rights of individuals, and, therefore, requires a clear legal basis and must be subject to conditions and safeguards.

Tools that meet these requirements are available in Articles 16 to 21 of the Convention on Cybercrime. These provisions are also applicable to investigate and obtain evidence related to the offences of xenophobia and racism criminalised under the First Protocol.

A particularly useful provision for cooperation between criminal justice authorities and service providers is that of Article 18 on production orders. Through such orders, law enforcement may order a "person" – including a service provider – to produce specified data needed in a specific criminal investigation. This means that the law enforcement is not required to seize or search servers; service providers only produce the data needed. A sub-provision of Article 18, namely 18(1)(b) provides a legal basis for ordering a service provider that is not with in the territory of the law enforcement authority but is offering its services in that territory, to produce specified subscriber information.[193]

"Subscriber information" is most often needed in the early stages of a criminal investigation to determine, for example, the owner of a social media or email account or the user of an Internet Protocol (IP) address involved in unlawful hate speech or online hate crimes.

The Second Additional Protocol to the Convention on Cybercrime, opened for signature in May 2022, provides additional tools for cooperation with service providers. Building on Article 18 of the Convention, Article 7 of the Second Protocol permits that production orders for subscriber orders are sent directly to a service provider in another Party; that service provider is then required to produce that information.

In fact, the procedural powers and provisions for international cooperation of the Convention and tools of the Second Additional Protocol are not only applicable to the offences of the Convention and its First Protocol, but may also be applied to obtain electronic evidence related to offences – including hate speech and hate crimes – criminalised under other international agreements.[194]

---

[191] Council of Europe/Project on Cybercrime (2008): Guidelines for cooperation between law enforcement and Internet service providers. Available at: https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016802fa3ba

[192] https://hudoc.echr.coe.int/fre?i=001-89964

[193] See T-CY Guidance Note on the production of subscriber information. Available at: https://rm.coe.int/doc/09000016806f943e

[194] See the T-CY Guidance Note on the scope of procedural powers and cooperation provisions. Available at: https://rm.coe.int/t-cy-2023-6-guidancenote-scope-of-powers-v9adopted/1680abc76a

3.1.6 International cooperation

The First Protocol to the Convention on Cybercrime underlines the importance of international cooperation on xenophobia and racism.[195]

A major obstacle preventing international cooperation on xenophobia and racism, and on hate speech and hate crime more broadly, is the "dual criminality" requirement, meaning that the conduct constituting an offence in the requesting State also needs to be an offence in the requested State. This may be especially challenging in cases where content is created and viewed in a Party to the First Protocol, but in possession or control by a person in a territory of a country that is not a Party, for example, where strict constitutional protections of freedom of expression have often led to hate speech laws being ruled unconstitutional by courts. Between Parties to the First Protocol this requirement should normally be met, although it is not always straightforward given the possibility of reservations and declarations.

The tools of the Convention on Cybercrime with respect to international cooperation are also applicable to the offences criminalised pursuant to the First Protocol, including:

- the general provisions and principles of Articles 23 to 28 of the Convention, such as on extradition, spontaneous information and others;
- the specific provisions of Articles 29 to 35 of the Convention, ranging from requests for expedited preservation of stored computer data to the network of 24/7 points of contact.

The same applies to the tools of the Second Additional Protocol already referred to above, such as those pertaining to:

- flexible approach to language requirements for sending and receiving requests for cooperation (Article 4);
- direct cooperation with private sector entities in other Parties, including requests for domain name registration information (Article 6), orders for the disclosure of subscriber information (Article 7);
- expedited means of government-to-government cooperation, including giving effect to production orders (Article 8), cooperation in emergencies (Articles 9 and 10), making use of video conferencing (Article 11) or joint investigation teams and joint investigations (Article 12).

3.1.6.1 Mutual assistance

As an example of good practice with regards to mutual legal assistance, in **Slovakia** requests for mutual legal assistance are handled by prosecutors of the International Department of the General Prosecutor's Office that have relevant experience with international co-operation. This good practice is related, in particular, to United States First Amendment issues. As indicated by Slovak authorities, despite the U.S. not being a Party to the Protocol, it provides cooperation on the basis of the Budapest Convention concerning some of the offences.

[195] See Preamble and Explanatory Report to the First Protocol, paragraph 3.

Although the Convention already provides for appropriate legal basis for international cooperation in relation to electronic evidence for any type of crime, the added value of the First Protocol is in establishment of harmonising substantive criminal law in the fight against racism and xenophobia online. As a consequence, international cooperation is facilitated, for example regarding requirements of double criminality.[196]

Similarly, in **France** international cooperation is carried out through the work of the International Mutual Assistance in Criminal Matters Office, which is part of the Director of the Directorate of Criminal Affairs and Pardons (DACG) of the French Ministry of Justice. The Office assists the courts in drafting and transmitting requests for international mutual assistance in criminal matters. These requests include requests for access to subscription, traffic or content data hosted outside the territory of an EU Member State. Having a dedicating team of experts with special responsibility for mutual legal assistance may be especially important when it comes to dealing with jurisdictional challenges, for example, which are commonplace in respect of illegal hate speech and hate crime offences committed through computer systems.

Examples of coordinated enforcement actions may include joint action days when multiple law enforcement authorities from different States conduct simultaneous large-scale crackdowns, including raiding locations, making arrests, shutting down websites or accounts, interrogating suspects, etc. Consider the Europe-wide **EUROPOL** coordinated Joint Action Days against acts of a racist and xenophobic nature. The first Joint Action Day was held in November 2020 on Germany's initiative;[197] the second was held in April 2022 on France's initiative.[198]

3.1.7 Capacity building and capacity management

Another important element of States working effectively to investigate and prosecute acts of a racist and xenophobic nature committed through computer systems is institutional capacity building and capacity management.

3.1.7.1 Capacity building

Capacity building includes training. In **Germany**, for example, the Federal Criminal Police Office provides a wide range of police training including regular training courses dealing with various legal and technical aspects tackling unlawful hate speech and hate crime offences online. In addition, the German Judicial Academy, a training institution funded jointly by the Federal Government and the federal States, provides a wide range of supra-regional training courses for judges and public prosecutors, including courses covering criminal offences related to acts of a xenophobic and racist nature. In 2022, for instance, training sessions on "Racism—A Challenge for the Judiciary" and "Combating Right-Wing Extremism and Right-Wing Terrorism" were available. These sessions are generally in high demand, which is an that judges and public prosecutors regard continuous professional training in relation to these offences to be very relevant.

Similar good practice in training can be found in **Spain**. Here the National Office for Combating Hate Crimes is in charge, among other things, of promoting and coordinating the enforcement of

[196] See e.g. Explanatory report to the First Protocol, paragraph 3.

[197] Information available at: www.europol.europa.eu/media-press/newsroom/news/stopping-hate-speech-online-europol-coordinates-first-europe-wide-action-day

[198] Information available at: https://www.europol.europa.eu/media-press/newsroom/news/tackling-hate-crime-across-europe-second-joint-action-day-targets-over-170-individuals

Spanish hate crime and hate speech laws.[199] As part of its activities, the Office coordinates the training of law enforcement authorities in this field, including the provision of specific training to the National Police Corps, the Civil Guard, and local police officers. The Office also coordinates a annual national seminar for training local police officers in matters relating to both offline and online unlawful hate speech and hate crime offences.[200] Along similar lines, each year the Public Prosecutor's Office of Spain coordinates several training courses on hate speech and hate crime for prosecutors working in this area, including special courses on unlawful hate speech and hate crime and training days on computer crime in general. In addition, the Computer Crime Unit of the State Attorney General's Office provides training sessions on unlawful hate speech and hate crime offences on the Internet within the framework of initial and continuous professional training for legal professionals. Furthermore, in 2018, more than ten ministries of the Spanish Government signed an institutional co-operation agreement on the fight against racism, xenophobia and other forms of intolerance. Part of the agreement was about a framework for the training of officials of law enforcement authorities and other relevant ministries. The agreement also covered the capacity building for awareness raising. At present, several ministries of the Spanish Government are also working with the Spanish Television Academy to develop capacities for awareness raising involving television broadcasts.

In **Slovakia**, regular training for police officers is provided by the National Criminal Agency of the Presidium of the Police. Trainings for judges, prosecutors and other relevant judicial personnel are provided by the Judicial Academy, an independent educational institution with nationwide coverage. It holds regular annual trainings focusing on various topics, including the concept of freedom of expression in the virtual space, or hate speech and hate crimes.

3.1.7.2 Capacity management

Turning to institutional capacity management, one obstacle to implementing the Protocol can be difficulties faced by the criminal justice system as a whole in dealing with significant increases in the numbers of criminal cases and in the numbers of convicted criminals – both resulting from enacting new legislation and/or building greater law enforcement and judicial capacity in the area of unlawful hate speech and hate crime offences committed through computer systems. Therefore:

- Within the court system, an example of good practice in capacity management could be the use of temporary injunctions against suspected illegal content pending a full criminal case. This would facilitate the timely removal of content and mitigate against cases being on hold for months or years.

- Similarly, countries could opt to classify illegal hate speech and hate crime offences if committed through a computer system as falling below a threshold required to trigger a jury trial. Dispensing with a jury trial for cases involving online perpetrators could facilitate the handling of a larger number of cases each year, especially when presided over by specially trained judges or magistrates.

Another example of good practice in capacity management may be using special data gathering systems to prevent duplication of prosecution cases. In **Brazil**, for example, there are several points of entry for notifications, including government and non-government channels. However, a

[199] Information available at: www.interior.gob.es.

[200] See, for example, Secretariat of State for Security, *Second Action Plan to Combat Hate Crimes* [English version] (Madrid: Ministry of the Interior, 2022). Available at: https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/Delitos-de-odio/descargas/II-PLAN-DE-ACCION-DE-LUCHA-CONTRA-LOS-DELITOS-DE-ODIO-english-version-copy-00001.pdf.

process has been developed to centralise all notifications coming in via different channels in one database to collect the data and make them available directly to state or federal prosecution offices in order to avoid duplicate investigations. Whenever a piece of content or incident generates several online notifications, via several channels, a non-governmental organization will gather all of them onto one database, collect the evidence, and make it available to the competent prosecution service.

A further example of good practice in capacity management may be minimising the use of custodial sentences as sanctions for criminal offences and instead relying more heavily on fines and suspended sentences, where feasible. Reducing the number of offenders being given custodial sentences may help take the pressure off criminal justice systems, which are often operating at near capacity due to budgetary and infrastructure limitations. There are different ways of achieving this sort of capacity management:

- First, sentencing guidelines could be promulgated that establish clear standards on imposing lesser sentences for criminal offences perpetrated online.[201]

- Second, relevant legislation could be amended to remove custodial sentences for example for first-time offenders.

Parties are free to pursue different sanction regimes. Specifically, Article 13 of the Convention on Cybercrime creates obligations to "adopt such legislative and other measures as may be necessary to ensure that the criminal offences established […] are punishable by effective, proportionate and dissuasive sanctions, which include deprivation of liberty." Article 8(1) of the Protocol makes it clear that Article 13 of the Convention on Cybercrime shall apply mutatis mutandis to the Protocol. Importantly, the Explanatory Report to the Convention on Cybercrime states the following:

> "The article leaves open the possibility of other sanctions or measures reflecting the seriousness of the offences, for example, measures could include injunction or forfeiture. It leaves to the Parties the discretionary power to create a system of criminal offences and sanctions that is compatible with their existing national legal systems."[202]

3.2 Service providers

3.2.1 Role of service providers in addressing racism and xenophobia online

Typically, service providers may directly address unlawful hate speech and hate crime offences on their services or platforms though the work of their legal compliance teams. Legal compliance teams monitor and remove unlawful content with greater or lesser degrees of proactivity depending on the platform or provider. In addition, legal compliance teams receive and respond to report forms, notices, or referrals relating to content that is suspected of being illegal based on local laws. Such reports can be received via intra-site reporting mechanisms, as well as directly from law enforcement authorities and from civil society organisations (for example, trusted flaggers).

As an example of good practice, most social media platforms and other service providers adopt "terms of service" which specify that users may not post or share "unlawful" or "illegal" content, including unlawful or illegal hate speech content. It is important to recognise that "[l]egal

[201] See (Brown 2020: 123–4).

[202] Explanatory Report to the Convention on Cybercrime, Budapest, 23 November, 2001, paragraph 130. Available at: https://rm.coe.int/16800cce5b.

compliance is not the same as moderation because legal compliance concerns local laws and the Internet platform's terms of service on illegal content in general […], whereas moderation concerns the platform's community standards or content policies including but not limited to standards or policies on hate speech."[203] That said, in practice it may be unfeasible to draw a sharp distinction between the work of legal compliance teams and the work of moderation or content policy teams (and, indeed, oversight boards) insofar as a significant proportion of reported or flagged hate speech content, for example, may be, as a matter fact, both unlawful and in violation of the relevant content policies.[204]

Once service providers are seen as having an important role to play in combating unlawful hate speech and hate crime offences committed through computer systems, however, this can create special co-operational challenges. Consider this passage taken from the 2020 Council of Europe study on models of governance of online hate speech:

> "One main challenge lies in navigating the different sorts of notifications that the police and/or public prosecutors can give to Internet platforms about suspected unlawful or illegal hate speech content. On the one hand, notifications made by the police and/or public prosecutors to Internet platforms could be non-binding or advisory only (administrative notifications). One potential benefit is that the police and/or public prosecutors are able to advise Internet platforms swiftly and without undertaking full investigations and seeking court orders. This may create opportunities for expedited interventions. These sorts of notifications also leave the Internet platforms (legal compliance teams, for example) with control over the final decision about whether to take down content based on the notifications but also their own assessments. Then again, because administrative notifications do not reflect full investigations and legal hearings, they are not based on the highest standards of due process. Moreover, they might give the impression of the police and/or public prosecutors exerting undue influence over Internet platforms. On the other hand, notifications made by police or public prosecutors to Internet platforms could be legally binding insofar as they are based on court rulings obtained by the police or public prosecutors (judicial notifications). This may ensure higher levels of due process, and remove any uncertainty or ambiguity about whether Internet platforms have an obligation to remove the content upon notification because these would be judicial "notice and take down" orders. Then again, this could be a slower process and may impede swift interventions. It also removes control from Internet platforms, which, depending on one's perspective, may not be a good thing."[205]

Reflecting on this dilemma, good practices may involve obtaining judicial notifications wherever feasible, as a sort of gold standard, and where necessary using administrative notifications, but at the same time having an ongoing, transparent and fair dialogue between law enforcement authorities and Internet platforms, so that processes and expectations are clear to both sides.

3.2.2 Regulatory frameworks for service providers

Although not expressly provided for in the Protocol, some States may take the initiative in creating regulatory frameworks as part of the overall package of governance of acts of a racist and xenophobic manner perpetrated online. Regulatory frameworks can be used by governmental agencies (for example, ministries of government, law enforcement authorities, regulators) to monitor, incentivise and where necessary sanction social media platforms and other services providers in relation to the reporting and removal of unlawful hate speech and hate crime offences

[203] See Brown (2020: 86).

[204] Ibid (52–58).

[205] Ibid, (134).

that occur on their platforms. Regulatory frameworks may be seen as non-criminal means and alternative measures to the use of criminal laws and more traditional law enforcement tools envisaged in the Protocol. For example, they can involve administrative laws requiring social media platforms to promptly remove illegal content upon a complaint. Council of Europe Recommendation CM/Rec(2022)16 on combating hate speech (May 2022)) and ECRI's relevant general policy recommendations provide a catalogue of such measures.[206]

In **Germany**, for example, social media platforms and other Internet services providers are obliged, under the NetzDG regulatory framework which entered into force in 2017, to provide a transparent and effective mechanism for users to report unlawful content (see s. 3(1) of NetzDG). NetzDG also defines statutory response times. Accordingly, providers are required by law to take note of and examine reports without delay and to remove or block manifestly unlawful content within 24 hours and other unlawful content within a maximum of seven days (see s. 3(2) of NetzDG). These response times may be exceeded in an individual case where an assessment is dependent on the truthfulness/falsehood of a claim or the decision is passed on to an accredited independent institution under the system of "institutions of regulated self-regulation". Non-compliance with these requirements represents a regulatory offence which may incur an administrative fine of up to 5 million euros or up to 50 million euros in the case of legal persons. The amount of the fine is dependent, among other things, on the number of users registered with the social media platform in question in Germany and the severity of the violation. The Federal Office of Justice is responsible for monitoring compliance with these requirements. In July 2019, for example, Germany's Federal Office of Justice fined Facebook 2 million Euros, among other things, because its NetzDG reporting form was too difficult to find.[207]

In the words of the 2020 Council of Europe study on models of governance of online hate speech:

> "The NetzDG Act is in one sense Germany taking out an insurance policy against online hate speech: namely, not simply leaving it to trust that Internet platforms will undertake thorough, effective and timely removal of unlawful hate speech content—whether directly via Internet platforms' legal compliance teams or indirectly through their content moderation practices—but making it a legal requirement, enforced or backed up with fines, for Internet platforms to remove unlawful hate speech."[208] Furthermore, there is evidence to suggest that the insurance policy is achieving some good results. For example, in its sixth country report on Germany, ECRI makes the following observation about NetzDG's positive effects: "During the country visit, ECRI was informed about the positive effects of [NetzDG]: the large social network providers have invested considerable resources in applying the law in an efficient manner. Many stakeholders confirmed that the most serious and open forms of hate speech have disappeared from the large platforms and thus do not any more reach the big number of their users."[209]

An instrument shaping the architecture of Internet governance in Europe is the **European Union's** Digital Services Act (DSA)[210] of October 2022. The Act creates obligations that are slightly less

206 See e. g. Committee of Ministers of the Council of Europe, 20 May 2022, CM/Rec (2022)16. See also ECRI's General Policy Recommendations aimed at preventing and combating hate speech (published in 2016), antisemitism (published in 2021) and anti-Muslim racism and discrimination (published in 2022).

207 ECRI, Sixth Report on Germany, 17 March 2020, paragraph 76. Available at: https://rm.coe.int/ecri-report-on-germany-sixth-monitoring-cycle-/16809ce4be

208 Brown (2020: 9).

209 ECRI, Sixth Report on Germany, paragraph 52.

210 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2022.277.01.0001.01.ENG&toc=OJ%3AL%3A2022%3A277%3ATOC

rigid than NetzDG. Insofar as authorities ("national judicial or administrative authorities") decide to issue orders for social media platforms to "act against" (for example, remove) specific illegal content (including but not limited to illegal hate speech), then these authorities have obligations to include reasons why the content is considered to be illegal content; and in turn platforms have obligations to respond to the orders, "without undue delay, specifying the action taken and the moment when the action was taken" (Article 8 of the DSA). In addition, platforms have obligations to "put mechanisms in place to allow any individual or entity to notify them of the presence on their service of specific items of information that the individual or entity considers to be illegal content" (Article 14). Furthermore, platforms have obligations to "process any notices that they receive under the mechanisms […], and take their decisions in respect of the information to which the notices relate, in a timely, diligent and objective manner" (Article 14(6) of the DSA). Furthermore, Article 18 of the DSA creates an obligation for providers to inform the law enforcement or judicial authorities about any information giving rise to a suspicion that a criminal offence involving a threat to the life or safety of a person or persons has taken place, is taking place or is likely to take place.

4 Lessons learnt and recommendations

4.1 Challenges and opportunities

The "First Protocol has never been more relevant than today", as noted by the Conference on Xenophobia and Racism via Computer Systems held in January 2023 on the occasion of the 20th anniversary of the Protocol:[211]

- "With the increasing use of digital technologies, online xenophobia and racism have been spreading considerably. Online hate is more prevalent and may be more harmful than hate offline, as perpetrators often act more spontaneously and anonymously, with a wider reach and lasting impact on victims.
- The Russian aggression against Ukraine is accompanied by hate speech and propaganda celebrating strikes and justifying attacks, including on civilian infrastructure. [212]
- Increased flows of refugees and migrants resulting from armed conflict provide a further fertile ground for online hate."

Addressing offences that occur online and are related to content, raises a number of complex challenges. For example:

- Criminalising hate speech and hate crime while protecting the freedom of expression and while maintaining a free, open and global internet.
- Different standards and approaches to hate speech and hate crime across States, including among Parties to the Convention on Cybercrime.[213] This in turn hinders international cooperation in criminal matters in this field.

[211] https://coe.int/en/web/cybercrime/international-conference-on-xenophobia-and-racism-committed-through-computer-systems

[212] See e.g. Faloppa, F., Gambacorta, A., Odekerken, R., van der Noordaa, R. (2023) Study on Preventing and Combating Hate Speech in Times of Crisis. Strasbourg (Council of Europe).

[213] For this reason, the offences of the First Protocol were not included in the Convention on Cybercrime but in a additional protocol. By end-2023, 35 out of 68 Parties to the Convention were also Parties to the First Protocol.

- The role and responsibilities of service providers operating in multiple jurisdictions with different legal requirements.
- The question of when it is necessary and proportionate to apply criminal law measures as opposed to non-criminal measures.

As shown in the present study, governments do implement this Protocol as part of a broader policy or of a range of measures to address the increasing problem of online hate speech and hate crime. Being a Party to the First Protocol has a number of benefits that may help directly or indirectly address some of these challenges:

- More consistent legal framework:
  - Implementing the provisions of the First Protocol in domestic law specifies the conduct that constitutes a criminal offence. This also provides greater clarity for service providers offering their services in a Party to the Protocol.
  - With regard to international cooperation with other Parties to the Protocol, it helps meet the dual criminality requirement.

- The procedural and international cooperation tools of the Convention on Cybercrime and its Second Protocol are available for investigating and securing evidence related to the offences of the First Protocol. This is particularly important given the cross-border nature of many of these offenses.

- Implementing the First Protocol contributes to increased protections of victims of online racism and xenophobia. Victims will have a greater expectation to obtain justice.

- The Protocol – in particular as part of a broader set of measures – contributes to greater awareness and education about the harms of xenophobia and racism committed through computer systems, and encourages the development of measures to prevent these forms of hate.

Good practices in the implementation of the Protocol are available as shown in the present study. They range from the enactment of criminal legislation to reporting mechanisms, statistics, specialised authorities, public/private and international cooperation and to guidelines or regulations for service providers.

4.2 Recommendations

The following recommendations can be drawn from this study, the contributions received and the workshops and conference held between December 2022 and February 2023:

- A broad range of measures and cooperation across sectors, organisations and stakeholders is needed to effectively counter online xenophobia and racism. Criminal law measures – including those under the First Protocol – are an important part of the response but should be used as the last resort. Non-criminal means and alternative measures should be pursued, including through regulations requiring social media platforms to moderate content and to promptly remove illegal content upon a complaint. Council of Europe Committee of Ministers' Recommendation CM/Rec(2022)16 on combating hate speech (May 2022) and relevant general policy recommendations of the European Commission against Racism and Intolerance (ECRI) provide a catalogue of such measures.

- States should ensure an appropriate legal framework – in line with the Protocol – for criminalising acts of online xenophobia and racism and ensure, in particular, that the offences are precisely defined in their domestic law. When doing so, Parties should ensure a proper balance between the freedom of expression and effective action against acts of a racist and xenophobic nature. While the Protocol requires Parties to criminalise online xenophobia and racism, it permits them to take into consideration established principles of domestic law relating to the freedom of expression, giving flexibility to Parties to address certain types conduct, as they deem appropriate, either through criminal law or other means.

- State shall seek to prevent any misuse of provisions criminalising acts of online xenophobia and racism which may impact on media freedom, and ensure that sanctions are not applied in a discriminatory or arbitrary way against journalists.

- States may extend the list of protected characteristics associated with offences of online xenophobia and racism but should not use this extension as a pretext for suppressing legitimate dissent and should always consider the right to freedom of expression.

- States may either directly put in place or indirectly support the creation of dedicated reporting mechanisms (for example, online platforms, telephone hotlines) that enable victims and third parties to report suspected unlawful hate offences that occur on the Internet.

- The Convention on Cybercrime and its First and Second Protocols complement each other. While the First Protocol focuses on substantive criminal law by requiring Parties to criminalise a number of acts of a xenophobic and racist nature, the tools of the Budapest Convention on Cybercrime, in particular the procedurals powers to investigate cybercrime and collect electronic evidence in relation to any crime as well as the provisions on international cooperation, are available to enforce the provisions on xenophobia and racism under the First Protocol. The same is true for the Second Additional Protocol to the Convention on Cybercrime on enhanced cooperation and disclosure of electronic evidence that was opened for signature in May 2022. States are therefore encouraged to become parties to all three instruments.

- The Cybercrime Programme Office of the Council of Europe (C-PROC) should support reforms of legislation, training and specialisation (including specialised authorities) and other criminal justice measures helping States address the challenges of online xenophobia and racism.

- Within the Council of Europe further synergies should be sought between the Budapest Convention and its Protocols, the Lanzarote Convention, the Istanbul Convention, the Convention on the Prevention of Terrorism, as well as the European Commission on Racism and Intolerance (ECRI). Soft law recommendations, resolutions and guidelines of the Committee of Minsters and other bodies should be made use of.

- Additional information on xenophobia and racism should be made available at the Council of Europe's online resource on cyberviolence.

5 Appendix

5.1 Appendix 1: Examples of domestic law

5.1.1 Article 3 – Dissemination of racist and xenophobic material through computer systems

5.1.1.1 Argentina

Law No. 23.593 of 9 Sept. 1988

Article 3
Those who participate in an organisation or carry out propaganda based on ideas or theories of superiority of a race or a group of persons of a certain religion, ethnic origin or colour, with the aim of justifying or promoting racial or religious discrimination in any form, shall be sentenced to between one month and three years' imprisonment.

The same penalty shall be incurred by those who by any means encourage or incite persecution or hatred against a person or groups of persons on account of their race, religion, nationality or political ideas.

5.1.1.2 Armenia

Criminal Code – Section 226

Incitement of national, racial, or religious hostility

1. Actions targeted at incitement of national, racial, or religious hostility, at racial superiority or humiliation of national dignity — shall be punished by a fine in the amount of two-hundred-fold to five-hundred-fold of the minimum salary or by imprisonment for a term of two to four years.

2. The acts provided for in part 1 of this Article, which have been committed — (1) publicly or by use of mass media; (2) by use of violence or threat thereof; (3) by use of official position; (4) by an organised group — shall be punished by imprisonment for a term of three to six years.

5.1.1.3 Finland

Criminal Code, Chapter 11

Section 10

Agitation against a population group
A person who makes available to the public or otherwise disseminates among the public or keeps available to the public information, an opinion or another message where a certain group is threatened, defamed or insulted on the basis of its race, colour, birth, national or ethnic origin, religion or belief, sexual orientation or disability or on another comparable basis shall be sentenced for agitation against a population group to a fine or to imprisonment for at most two years.

5.1.1.4 Germany

Section 130 of the Criminal Code (*Strafgesetzbuch*, StGB) (Incitement of masses)

Section 130 Incitement of masses

(1) Whoever, in a manner which is suitable for causing a disturbance of the public peace, 1. incites hatred against a national, racial, religious group or a group defined by their ethnic origin, against sections of the population or individuals on account of their belonging to one of the aforementioned groups or sections of the population, or calls for violent or arbitrary measures against them or 2. violates the human dignity of others by insulting, maliciously maligning or defaming one of the aforementioned groups, sections of the population or individuals on account of their belonging to one of the aforementioned groups or sections of the population incurs a penalty of imprisonment for a term of between three months and five years.

(2) Whoever

1. disseminates material (section 11 (3)) or makes it available to the public, or offers, supplies or makes available to a person under 18 years of age material (section 11 (3)) which a) incites hatred against one of the groups referred to in subsection (1) no. 1, sections of the population or individuals on account of their belonging to one of the groups referred to in subsection (1) no. 1, or sections of the population, b) calls for violent or arbitrary measures against one of the persons or bodies of persons referred to in letter (a) or c) attacks the human dignity of one of the persons or bodies of persons referred to in letter (a) by insulting, maliciously maligning or defaming them, 2. makes content referred to in no. 1 (a) to (c) available to a person under 18 years of age or to the public through broadcasting or telemedia services or 3. produces, purchases, supplies, stocks, offers, advertises or undertakes to import or export material (section 11 (3)) of such content referred to in no. 1 (a) to (c) in order to use it or parts obtained from it within the meaning of no. 1 or 2 or to facilitate such use by another incurs a penalty of imprisonment for a term not exceeding three years or a fine.
(3) Whoever publicly or in a meeting approves of, denies or downplays an act committed under the rule of National Socialism of the kind indicated in section 6 (1) of the Code of Crimes against International Law in a manner which is suitable for causing a disturbance of the public peace incurs a penalty of imprisonment for a term not exceeding five years or a fine.
(4) Whoever publicly or in a meeting disturbs the public peace in a manner which violates the dignity of the victims by approving of, glorifying or justifying National Socialist tyranny and arbitrary rule incurs a penalty of imprisonment for a term not exceeding three years or a fine.
(5) Subsection (2) no. 1 and no. 3 also applies to material (section 11 (3)) of such content referred to in subsections (3) and (4). Whoever makes content referred to in subsections (3) and (4) available to a person under 18 years of age or available to the public through broadcasting or telemedia services incurs the same penalty specified in subsection (2) no. 2.
(6) In the cases under subsection (2) nos. 1 and 2, also in conjunction with subsection (5), the attempt is punishable.
(7) In the cases under subsection (2), also in conjunction with subsection (5), and in the cases under subsections (3) and (4), section 86 (3) applies accordingly.

5.1.1.5 Lithuania

Article 170. Incitement against Any National, Racial, Ethnic, Religious or Other Group of Persons
1. A person who, for the purposes of distribution, produces, acquires, sends, transports or stores the items ridiculing, expressing contempt for, urging hatred of or inciting discrimination against a group of persons or a person belonging thereto on grounds of sex, sexual orientation, race, nationality, language, descent, social status, religion, convictions or views or inciting violence, a physical violent treatment of such a group of persons or the person belonging thereto or distributes them

shall be punished by a fine or by restriction of liberty or by arrest or by imprisonment for a term of up to one year.
2. A person who publicly ridicules, expresses contempt for, urges hatred of or incites discrimination against a group of persons or a person belonging thereto on grounds of sex, sexual orientation, race, nationality, language, descent, social status, religion, convictions or views
shall be punished by a fine or by restriction of liberty or by arrest or by imprisonment for a term of up to two years.
3. A person who publicly incites violence or a physical violent treatment of a group of persons or a person belonging thereto on grounds of sex, sexual orientation, race, nationality, language, descent, social status, religion, convictions or views or finances or otherwise supports such activities
shall be punished by a fine or by restriction of liberty or by arrest or by imprisonment for a term of up to three years.
4. A legal entity shall also be held liable for the acts provided for in this Article.

5.1.1.6 Mexico

Article 9.- (The first paragraph is repealed first and the others are traversed in their order)

Based on the provisions of the first constitutional article and the second paragraph of Article 1, fraction III of this Law are considered to be discrimination, other:
(…)
XXVII. Incite hate, violence, rejection, mockery, injury, persecution or exclusion;

Article 1.- The provisions of this Law are of public order and of social interest. The object of the same is to prevent and eliminate all forms of discrimination that are exercised against any person under the terms of Article 1 of the Political Constitution of the United Mexican States, as well as to promote equality of opportunities and treatment.

III. Discrimination: For the purposes of this law, discrimination shall be construed as discrimination, exclusion, restriction or preference which, by way of action or omission, with or without intention, is not objective, rational or proportional and has as its object or result in hindering, restricting, preventing, undermining or annulling the recognition, enjoyment or exercise of human rights and freedoms, where it is based on one or more of the following reasons: ethnic or national origin, skin colour, culture, sex, gender, age, disabilities, social, economic, health or legal, religious, physical appearance, genetic characteristics, migratory status, pregnancy, language, opinions, sexual preferences, identity or political affiliation, marital status, family situation, family responsibilities, language, criminal history or any other reason;

There will also be an expression of discrimination against homophobia, misogyny, any manifestation of xenophobia, racial segregation, anti-Semitism, as well as racial discrimination and other related forms of intolerance;

Article 4.- Any discriminatory practice which has as its object or effect the prevention or cancellation of the recognition or exercise of rights and the exercise of rights is prohibited. real equality of opportunity in terms of Article 1. constitutional and Article 1, second paragraph, fraction III of this Law

5.1.1.7 Netherlands

Criminal Code
Section 137d

1. Any person who publicly, either verbally or in writing or through images, incites hatred of or discrimination against persons or violence against their person or property because of their race, religion or beliefs, their sex, their hetero- or homosexual orientation or their physical, mental or intellectual disability, shall be liable to a term of imprisonment not exceeding one year or a fine of the third category.

2. If the offence is committed by a person who makes a profession or habit of it or by two or more persons in concert, a term of imprisonment not exceeding two years or a fine of the fourth category shall be imposed

5.1.1.8 New Zealand

Human Rights Act 1993 - Article 61 Racial disharmony

(1)
It shall be unlawful for any person—
(a) to publish or distribute written matter which is threatening, abusive, or insulting, or to broadcast by means of radio or television or other electronic communication words which are threatening, abusive, or insulting; or
(b) to use in any public place as defined in section 2(1) of the Summary Offences Act 1981, or within the hearing of persons in any such public place, or at any meeting to which the public are invited or have access, words which are threatening, abusive, or insulting; or
(c) to use in any place words which are threatening, abusive, or insulting if the person using the words knew or ought to have known that the words were reasonably likely to be published in a newspaper, magazine, or periodical or broadcast by means of radio or television,—
being matter or words likely to excite hostility against or bring into contempt any group of persons in or who may be coming to New Zealand on the ground of the colour, race, or ethnic or national origins of that group of persons.
(2)
It shall not be a breach of subsection (1) to publish in a newspaper, magazine, or periodical or broadcast by means of radio or television or other electronic communication a report relating to the publication or distribution of matter by any person or the broadcast or use of words by any person, if the report of the matter or words accurately conveys the intention of the person who published or distributed the matter or broadcast or used the words.
(3)
For the purposes of this section,—
newspaper means a paper containing public news or observations on public news, or consisting wholly or mainly of advertisements, being a newspaper that is published periodically at intervals not exceeding 3 months
publishes or distributes means publishes or distributes to the public at large or to any member or members of the public
written matter includes any writing, sign, visible representation, or sound recording.

5.1.1.9 North Macedonia

Causing hatred, discord or intolerance on national, racial, religious or any other discriminatory ground
Article 319
(1) Whosoever by force, maltreatment, endangering the security, mocking of the national, ethnic, religious and other symbols, by burning, destroying or in any other manner damaging the flag of the Republic of Macedonia or flags of other States, by damaging other people's objects, by desecration of monuments, graves, or in any other discriminatory manner, directly or indirectly, causes or excites hatred, discord or intolerance on grounds of gender, race, color of the skin, membership in marginalized group, ethnic membership, language, nationality, social background, religious belief, other beliefs, education, political affiliation, personal or social status, mental or physical impairment, age, family or marital status, property status, health condition, or in any other ground foreseen by law on ratified international agreement, shall be sentenced to imprisonment of one to five years.
(2) Whosoever commits the crime referred to in paragraph (1) of this Article by abusing his position or authorization, or if because of these crimes, riots and violence were caused against the people, or property damage to a great extent was caused, shall be sentenced to imprisonment of one to ten years.

5.1.1.10 Norway

Penal Code
Section 185. Hate speech
A penalty of a fine or imprisonment for a term not exceeding three years shall be applied to any person who with intent or gross negligence publicly makes a discriminatory or hateful statement. «Statement» includes the use of symbols. Any person who in the presence of others, with intent or gross negligence, makes such a statement to a person affected by it, see the second paragraph, is liable to a penalty of a fine or imprisonment for a term not exceeding one year.

«Discriminatory or hateful statement» means threatening or insulting a person or promoting hate of, persecution of or contempt for another person based on his or her

a. skin colour or national or ethnic origin,
b. religion or life stance,
c. homosexual orientation, or
d. reduced functional capacity.

5.1.1.11 Serbia

Criminal Code
Article 387 paragraph 4

(4) Whoever spreads or otherwise makes publicly available texts, images or any other representation of ideas or theories advocating or encouraging hatred, discrimination or violence against any person or group of persons based on race, colour, religious affiliation, nationality, ethnic origin or other personal property, shall be punished with imprisonment of three months to three years.

5.1.1.12 Slovakia

Criminal Code

Section 422b – Spreading extremist material

(1) Whoever reproduces, transports, supplies, makes available, puts into circulation, imports, exports, offers, sells, sends or distributes extremist material, shall be punished by imprisonment for one to five years.

(2) The offender shall be punished by imprisonment for three to eight years if he commits the act referred to in paragraph 1

a) a more serious way of acting,
b) publicly, or
c) as a member of an extremist group.

Section 130
(7) For the purposes of this Act, extremist material shall mean written, graphic, video, audio or audio-video works

a) of texts and declarations, flags, badges, passwords, or symbols, groups and movements that lead or led in the past to the suppression of fundamental human rights and freedoms,
b) of programmes or ideologies of groups and movements that lead or led in the past to the suppression of fundamental human rights and freedoms,
c) advocating, promoting or inciting hatred, violence or unreasonable differential treatment of groups of persons or an individual because of their belonging to one race, nation, nationality, skin colour, ethnicity, origin, or their religion, if it is an excuse for the above reasons, or
d) justifying, approving, denying or seriously derogating genocide, crimes against peace, crimes against humanity or military crimes, if the offender or an accessory to such an act was convicted by a final judgment of an international court established under international public law, the authority of which is recognised by the Slovak Republic, or by a final judgment of a court of the Slovak Republic.

(8) A material referred to in Subsection 7 shall not be deemed to be extremist material if it is demonstrably produced, distributed, put into circulation, made publicly accessible or kept in possession for the purpose of educational, collection or research activities.

5.1.1.13 Spain

Criminal Code
Article 510

1. A prison sentence of one to four years and a fine of six to twelve months shall be imposed on:
a) Those who, directly or indirectly, foster, promote or incite hatred, hostility, discrimination or violence against a group, or part thereof, or against a certain person for belonging to such a group, for reasons of racism, anti-Semitism or for other reasons related to ideology, religion or beliefs, family circumstances, the fact that the members belong to an ethnicity, race or nation, national origin, gender, sexual orientation or identity, or due to gender, illness or disability;

b) Those who produce, prepare, possess with the purpose of distributing, provide third parties access to, distribute, publish or sell documents or any other type of material or medium that, due

to the content thereof, are liable to directly or indirectly foster, promote or incite hatred, hostility, discrimination or violence against a group, or part thereof, or against a certain person for belonging to such a group, for reasons of racism, anti-Semitism or for other reasons related to ideology, religion or beliefs, family circumstances, the fact that the members belong to an ethnicity, race or nation, national origin, gender, sexual orientation or identity, or due to gender, illness or disability;

5.1.2 Article 4 – Racist and xenophobic motivated threat

5.1.2.1 Germany

Criminal Code

Section 126 Disturbing public peace by threatening to commit offences

(1) Whoever, in a manner which is suitable for causing a disturbance of the public peace, threatens to commit

1. breach of the peace as designated in section 125a sentence 2 nos. 1 to 4,
2. murder under specific aggravating circumstances (section 211), murder (section 212) or genocide (section 6 of the Code of Crimes against International Law) or a crime against humanity (section 7 of the Code of Crimes against International Law) or a war crime (section 8, 9, 10, 11 or 12 of the Code of Crimes against International Law),
3. grievous bodily harm (section 226),
4. an offence against personal liberty under section 232 (3) sentence 2, section 232a (3), (4) or (5), section 232b (3) or (4), section 233a (3) or (4), each to the extent that it represents a serious criminal offence, section 234, 234a, 239a or 239b,
5. robbery or extortion with use of force or threat of force (sections 249 to 251 or section 255),
6. a serious criminal offence constituting a public danger under sections 306 to 306c or section 307 (1) to (3), section 308 (1) to (3), section 309 (1) to (4), section 313, section 314 or section 315 (3), section 315b (3), section 316a (1) or (3), section 316c (1) or (3) or section 318 (3) or (4) or
7. a less serious criminal offence constituting a public danger under section 309 (6), section 311 (1), section 316b (1), section 317 (1) or section 318 (1) incurs a penalty of imprisonment for a term not exceeding three years or a fine. (2) Whoever, despite knowing better and in a manner which is suitable for causing a disturbance of the public peace, pretends that the commission of one of the unlawful acts referred to in subsection (1) is imminent incurs the same penalty.

Section 241 Threatening commission of serious criminal offence

(1) Whoever threatens a person with the commission of a serious criminal offence against that person or a person close to him or her incurs a penalty of imprisonment for a term not exceeding one year or a fine.

(2) Whoever, despite knowing better, pretends to another person that the commission of a serious criminal offence against that person or a person close to him or her is imminent incurs the same penalty.

Section 46 General principles

(1) The offender's guilt provides the basis on which the penalty is fixed. The effects which the penalty can be expected to have on the offender's future life in society are to be taken into account.

(2) When fixing the penalty the court weighs the circumstances which speak in favour of and those which speak against the offender. The following, in particular, may be taken into consideration: the offender's motives and objectives, in particular including racist, xenophobic, antisemitic, gender-specific, anti-sexual orientation or other motives evidencing contempt for humanity (…).

5.1.2.2 Serbia

Criminal Code
Article 387 (6)

Whoever publicly threatens to commit a criminal offence punishable with imprisonment of four and more years against a person or group of persons because of a particular race, colour, religion, nationality, ethnic origin or because of other personal property, shall be punished with imprisonment of three months to three years.

5.1.2.3 Spain

Criminal Code
Article 170

1. Should the intimidation be of a harm which constitutes a criminal offence is intended to cause fear among the inhabitants of a location, ethnic, cultural or religious group, or a social or professional group, or any other group of persons, and if serious enough for such harm to be inflicted, the respective higher degree of penalties than those foreseen in the preceding Article shall be imposed.
2. A sentence of imprisonment from six months to two years shall be applied to those who, for the same purpose and severity, publicly call for violent deeds to be committed by armed gangs, organisations or terrorist groups.

Article 22
The following are aggravating circumstances:
(…)
4. Committing the criminal offence for racist or anti-Semitic reasons, or another kind of discrimination related to ideology, religion or belief of the victim, ethnicity, race or nation to which he belongs, his gender, sexual orientation or identity, reasons related to gender, illness suffered or disability.

5.1.3 Article 5 – Racist and xenophobic motivated insult

5.1.3.1 France

Article 33 Press law of 1881 (version in force since 26 August 2021)

(…)

Sera punie d'un an d'emprisonnement et de 45 000 euros d'amende l'injure commise par les mêmes moyens envers une personne ou un groupe de personnes à raison de leur origine ou de leur appartenance ou de leur non-appartenance à une ethnie, une nation, une race ou une religion déterminée.

Sera punie des peines prévues à l'alinéa précédent l'injure commise dans les mêmes conditions envers une personne ou un groupe de personnes à raison de leur sexe, de leur orientation sexuelle ou identité de genre ou de leur handicap.

Lorsque les faits mentionnés aux troisième et quatrième alinéas du présent article sont commis par une personne dépositaire de l'autorité publique ou chargée d'une mission de service public dans l'exercice ou à l'occasion de l'exercice de ses fonctions ou de sa mission, les peines sont portées à trois ans d'emprisonnement et à 75 000 euros d'amende.

En cas de condamnation pour l'un des faits prévus par les troisième et quatrième alinéas, le tribunal pourra en outre ordonner :

1° L'affichage ou la diffusion de la décision prononcée dans les conditions prévues par l'article 131-35 du code pénal ;

Article 29

1. Any allegation or imputation of a fact which is prejudicial to the honour or consideration of the person or body to which the fact is imputed is a defamation. The direct publication or reproduction of such an allegation or imputation is punishable, even if it is made in dubious form or if it refers to a person or body not expressly named, but whose identification is made possible by the terms of the incriminating speeches, shouts, threats, writings or printed matter, placards or posters.

2. Any insulting expression, term of contempt or invective which does not contain the imputation of any fact is an insult.

Article 23

1. Those who, either by speeches, shouts or threats made in public places or meetings, or by writings, printed matter, drawings, engravings, paintings, emblems, images or any other written, spoken or pictorial material sold or distributed, or by placards or posters displayed for public view in public places or meetings, shall be punished as accomplices to an action classified as a crime or offence, sold or exhibited in public places or meetings, or by placards or posters exposed to public view, or by any means of communication to the public by electronic means, shall have directly provoked the perpetrator or perpetrators to commit the said action, if the provocation was followed by its effect.

2. This provision shall also apply when the provocation has been followed only by an attempt to commit a crime as provided for in Article 2 of the Criminal Code.

5.1.3.2 Germany

Criminal Code

Section 130 Incitement of masses

(1) Whoever, in a manner suited to causing a disturbance of the public peace,
1. incites hatred against a national, racial, religious group or a group defined by their ethnic origin, against sections of the population or individuals on account of their belonging to one of the aforementioned groups or sections of the population, or calls for violent or arbitrary measures against them or

2. violates the human dignity of others by insulting, maliciously maligning or defaming one of the aforementioned groups, sections of the population or individuals on account of their belonging to one of the aforementioned groups or sections of the population

incurs a penalty of imprisonment for a term of between three months and five years.

(2) Whoever
1. disseminates content (section 11 (3)) or makes it available to the public, or offers, supplies or makes available to a person under 18 years of age content (section 11 (3)) which a) incites hatred against one of the groups referred to in subsection (1) no. 1, sections of the population or individuals on account of their belonging to one of the groups referred to in subsection (1) no. 1, or sections of the population, b) calls for violent or arbitrary measures against one of the persons or bodies of persons referred to in letter (a) or c) attacks the human dignity of one of the persons or bodies of persons referred to in letter (a) by insulting, maliciously maligning or defaming them, or
2. produces, purchases, supplies, stocks, offers, advertises or undertakes to import or export content (section 11 (3)) as referred to in no. 1 (a) to (c) in order to use it within the meaning of no. 1 or to facilitate such use by another

incurs a penalty of imprisonment for a term not exceeding three years or a fine.

Section 185 Insult

The penalty for insult is imprisonment for a term not exceeding one year or a fine and, if the insult is committed publicly, in a meeting, by disseminating content (section 11 (3)) or by means of an assault, imprisonment for a term not exceeding two years or a fine.

Section 186 Malicious gossip (üble Nachrede)

Whoever asserts or disseminates a fact about another person which is suited to degrading that person or negatively affecting public opinion about that person, unless this fact can be proved to be true, incurs a penalty of imprisonment for a term not exceeding one year or a fine and, if the offence was committed publicly, in a meeting or by disseminating content (section 11 (3)), a penalty of imprisonment for a term not exceeding two years or a fine. table of contents Section

187 Defamation

Whoever, despite knowing better, asserts or disseminates an untrue fact about another person which is suited to degrading that person or negatively affecting public opinion about that person or endangering said person’s creditworthiness incurs a penalty of imprisonment for a term not exceeding two years or a fine, and, if the act was committed publicly, in a meeting or by disseminating content (section 11 (3)), a penalty of imprisonment for a term not exceeding five years or a fine.

5.1.3.3 Serbia

Criminal Code
Article 174
Whoever publicly ridicules a person or a group because of particular race, colour, religion, nationality, ethnic origin or other personal characteristic, shall be punished with a fine or imprisonment up to one year.

5.1.4 Article 6 – Denial, gross minimisation, approval or justification of genocide or crimes against humanity

5.1.4.1 Czech Republic

Criminal Code
Section 405 Denial, Impugnation, Approval and Justification of Genocide
Whoever publicly denies, impugns, approves, or attempts to justify Nazi, Communist or any other genocide, or other crimes of the Nazis and Communists against humanity, shall be sentenced to imprisonment for six months to three years.

5.1.4.2 France

French Press Law 1881
Article 24 bis

Those who deny, by one of the means set out in Article 23, the existence of one or more crimes against humanity as defined by Article 6 of the Statute of the International Military Tribunal annexed to the London Agreement of 8 August 1945 and which were committed either by members of an organisation declared criminal pursuant to Article 9 of the said Statute, or by a person convicted of such crimes by a French or international court, shall be punished by one year's imprisonment and a fine of €45,000.

The same penalties shall be imposed on those who deny, belittle or trivialise, by any of the means set out in Article 23, the existence of a crime of genocide other than those mentioned in the first paragraph of this article, of another crime against humanity, a crime of enslavement or exploitation of a person reduced to slavery or a war crime as defined in Articles 6, 7 and 8 of the Statute of the International Criminal Court signed in Rome on 18 July 1998 and in Articles 211-1 to 212-3, 224-1 A to 224-1 C and 461-1 to 461-31 of the Criminal Code, when :

1° This crime has led to a conviction by a French or international court;

When the acts mentioned in this article are committed by a person holding public authority or entrusted with a public service mission in the exercise or on the occasion of the exercise of his or her functions or mission, the penalties are increased to three years' imprisonment and a fine of 75,000 euros.

[Provisions declared unconstitutional by the Constitutional Council's decision no. 2016-745 DC of 26 January 2017].

The court may also order:

1° The posting or dissemination of the decision pronounced under the conditions provided for by Article 131-35 of the Criminal Code.

5.1.4.3 Germany

Criminal Code
Section 130 (3)
(3) Whoever publicly or in a meeting approves of, denies or downplays an act committed under the rule of National Socialism of the kind indicated in section 6 (1) of the Code of Crimes against International Law in a manner suited to causing a disturbance of the public peace incurs a penalty of imprisonment for a term not exceeding five years or a fine.

5.1.4.4 Israel

Israel Denial of Holocaust (Prohibition) Law 5746-1986
Definitions
1. In this Law, "crime against the Jewish people" and "crime against humanity" have the same respective meanings as in the "Nazis and Nazi Collaborators Law, 5710-1950[1]. Prohibition of Denial of Holocaust
2. A person who, in writing or by word of mouth, publishes any statement denying or diminishing the proportions of acts committed in the period of the Nazi regime, which are crimes against the Jewish people or crimes against humanity, with intent to defend the perpetrators of those acts or to express sympathy or identification with them, shall be liable to imprisonment for a term of five years. Prohibition of publication of expression for sympathy for Nazi crimes
3. A person who, in writing or by word of mouth, publishes any statement expressing praise or sympathy for or identification with acts done in the period of the Nazi regime, which are crimes against the Jewish people or crimes against humanity, shall be liable to imprisonment for a term of five years. Permitted publication
4. The publication of a correct and fair report of a publication prohibited by this Law shall not be regarded as an offence thereunder so long as it is not made with intent to express sympathy or identification with the perpetrators of crimes against the Jewish people or against humanity. Filing of charge
5. An indictment for offences under this Law shall only be filed by or with the consent of the Attorney-General.

5.1.4.5 Lithuania

Criminal Code
Article 170^2. Public Condonation of International Crimes, Crimes Committed by the USSR or Nazi Germany against the Republic of Lithuania or Inhabitants Thereof, Denial or Gross Trivialisation of the Crimes

1. A person who publicly condones the crimes of genocide or other crimes against humanity or war crimes recognised under legal acts of the Republic of Lithuania or the European Union or effective judgements passed by courts of the Republic of Lithuania or international courts, denies or grossly trivialises them, where this is accomplished in a manner which is threatening, abusive or insulting or which disturbs the public order, also a person who publicly condones the aggression perpetrated by the USSR or Nazi Germany against the Republic of Lithuania, the crimes of genocide or other crimes against humanity or war crimes committed by the USSR or Nazi Germany in the territory of the Republic of Lithuania or against the inhabitants of the Republic of Lithuania or other grave or serious crimes committed during 1990-1991 against the Republic of Lithuania by the persons perpetrating or participating in perpetration of the aggression against the Republic of Lithuania or grave crimes against the inhabitants of the Republic of Lithuania, denies or grossly trivialises them, where this is accomplished in a manner which is threatening, abusive or insulting or which disturbs the public order,

shall be punished by a fine or by restriction of liberty or by arrest or by a custodial sentence for a term of up to two years.
2. A legal entity shall also be held liable for the acts provided for in this Article.

5.1.4.6 Luxembourg

Criminal Code
Article 457- 3
(1) Any person who, either by means of false or misleading 25,000 or one of these penalties only, whoever, either through speeches, shouts or threats made in public places or 25,000 or one of these penalties only, whoever, either by speeches, shouts or threats made in public places or meetings, or by writings, printed matter drawings, engravings, paintings, emblems, images or any other written, spoken or pictorial material sold or distributed, offered for sale or displayed in public places or meetings, or by posters or billboards displayed in public view. or posters displayed for public view, or by any means of audiovisual communication, has or by any means of audiovisual communication, disputed, minimised, justified or denied the existence of one or more crimes against humanity or war crimes as defined in Article 6 of the Statute of the International Military Tribunal annexed to the London Agreement of 8 August 1945 and which were committed either by members of an organisation declared criminal under Article 9 of that Statute, or by a person convicted of such crimes by a Luxembourg, foreign or international court.

(2) Anyone who, by one of the means set out in the preceding paragraph, has contested, minimised, justified or denied the existence of one or more genocides as defined by Article 136bis of the Criminal Code, as well as of crimes against humanity and crimes against humanity and war crimes, as defined in Articles 136ter to 136quinquies of the Criminal Code of the Criminal Code and recognised by a Luxembourg or international court.

5.1.4.7 Poland

The Act of 18 December 1998 on the Institute of National Remembrance Commission for the Prosecution of Crimes against the Polish Nation
Article 1
The act regulates:
1) the recording, collecting, storing, processing, securing, making available and publishing of the documents of the state security authorities, produced and accumulated from July 22, 1944 until July 31, 1990, as well as the documents of the security authorities of the Third Reich and the Soviet Union relating to:
a) - the Nazi crimes, - the communist crimes, - other crimes against peace, humanity or war crimes, perpetrated on persons of Polish nationality or Polish citizens of other nationalities between September 1, 1939 until July 31, 1990,
b) other politically motivated reprisals, instigated by the officers of the Polish law enforcement agencies or the judiciary or persons acting on their order which were disclosed in the contents of the rulings made on the strength of the Act, dated February 23, 1991, on considering as invalid the rulings made in the cases of persons oppressed for their activities for the cause of an independent Polish State (Journal of Laws No. 34, section 149, with later amendments),
c) the actions of the state security authorities described in Article 5;
2) the procedure for the prosecution of the crimes specified in point 1, letter a;
3) the protection of the personal data of the people referred to in the documents collected in the archive of the Institute of National Remembrance.
4) performing activities in the field of public education

Article 55

Anyone who publicly and contrary to the facts denies crimes referred to in Article 1, point 1 shall be subject to a fine or the penalty of imprisonment of up to 3 years. The sentence shall be made public.

5.1.4.8 Rwanda

Law No. 33n bis/2003 of 2003 Repressing the Crime of Genocide, Crimes Against Humanity and War Crimes
Article 4

Shall be sentenced to an imprisonment of ten (10) to twenty (20) years, any person who will have publicly shown, by his or her words, writings, images, or by any other means, that he or she has negated the genocide committed, rudely minimised it or attempted to justify or approve its grounds, or any person who will have hidden or destroyed its evidence. Where the crimes mentioned in the preceding paragraph are committed by an association or a political party, its dissolution shall be pronounced.

5.1.4.9 Serbia

Criminal Code
Article 387 (5)

Whoever publicly approves of, denies the existence or significantly impairs the gravity of genocide, crimes against humanity and war crimes committed against a group of persons or a member of a group designated on the grounds of their race, colour of skin, religion, origin, State, national or ethnic affiliation, in the manner that may lead to violence or inciting hatred towards such a group of persons or a member of such a group, where such criminal offences are determined by the final judgement of a court in Serbia or of the International Criminal Court, shall be punished with imprisonment of six months to five years.

5.1.4.10 Slovakia

Criminal Code
Section 422d Denial and approval of the Holocaust, crimes of political regimes and crimes against humanity
(1) Whoever publicly denies, questions, approves or tries to justify the Holocaust, crimes of a regime based on fascist ideology, crimes of a regime based on communist ideology or crimes of another similar movement that by violence, the threat of violence or the threat of other serious harm aims to suppress the fundamental rights and freedoms of persons, shall be punished by imprisonment for six months to three years.
(2) As in paragraph 1, whoever publicly denies, approves, questions, grossly belittles or tries to justify genocide, crimes against peace, crimes against humanity or war crimes in a way that can incite violence or hatred against a group of persons or its member, if the perpetrator or participant of this act has been convicted by a valid judgment of an international court established on the basis of public international law, the jurisdiction of which is recognized by the Slovak Republic, or by a valid judgment of a court of the Slovak Republic.

5.1.4.11 Slovenia

Criminal Code
Public Incitement to Hatred, Violence or Intolerance
Article 297
(1) Whoever publicly provokes or stirs up ethnic, racial, religious or other hatred, strife or intolerance, or provokes any other inequality on the basis of physical or mental deficiencies or sexual orientation, shall be punished by imprisonment of up to two years.
(2) The same sentence shall be imposed on a person who publicly disseminates ideas on the supremacy of one race over another, or provides aid in any manner for racist activity or denies, diminishes the significance of, approves, disregards, makes fun of, or advocates genocide, holocaust, crimes against humanity, war crime, aggression, or other criminal offences against humanity.

5.1.4.12 Switzerland

Criminal Code
Article 261bis (4) (6)
4. any person who publicly denigrates or discriminates against another or a group of persons on the grounds of their race, ethnic origin, religion or sexual orientation in a manner that violates human dignity, whether verbally, in writing or pictorially, by using gestures, through acts of aggression or by other means,(…),
6. shall be liable to a custodial sentence not exceeding three years or to a monetary penalty.

5.1.5 Article 7 - Aiding and abetting

5.1.5.1 Germany

Criminal Code

Section 26 Abetting

Whoever intentionally induces another to intentionally commit an unlawful act (abettor) incurs the same penalty as an offender.

Section 27 Aiding

(1) Whoever intentionally assists another in the intentional commission of an unlawful act incurs a penalty as an aider.
(2) The penalty for the aider is determined in accordance with the penalty threatened for the offender. It must be mitigated pursuant to section 49 (1).

5.2 Appendix 2: References

Alkiviadou, N. (2018) 'The Legal Regulation of Hate Speech: The International and European Frameworks'. In: Croatian Political Science Review 55: 203–229.

Alkiviadou, N. (2020) 'The Legal Regulation of Hate Speech: The United Nations Framework as the Common Denominator for Europe and Asia'. In: European-Asian Journal of Law and Governance 10: 22–41.

Au, K. (1984) 'Freedom from Fear', Lincoln Law Review 15: 45–52.

Bakalis, C. (2015) Cyberhate: An Issue of Continued Concern for the Council of Europe's Anti-Racism Commission. Strasbourg: Council of Europe. Available at: https://edoc.coe.int/en/cybercrime/6883-cyberhate-an-issue-of-continued-concern-for-the-council-of-europe-s-anti-racismcommission.html.

Blain, M. (1995) 'Group Defamation and the Holocaust', in M. Freedman and E. Freedman (eds.) Group Defamation and Freedom of Speech: The Relationship between Language and Violence. Westport, CT: Greenwood Press.

Borgesius, F. Z. (2018) Discrimination, Artificial Intelligence, And Algorithmic Decision-Making. Strasbourg: Council of Europe.

Borovoy, A.A. (1988) When Freedoms Collide: The Case for Our Civil Liberties. Toronto: Lester and Orpen Dennys.

Brenncke, M. (2018) Judicial Law-Making in English and German Courts: Techniques and Limits of Statutory Interpretation. Cambridge: Cambridge University Press.

Brown, A. (2015) Hate Speech Law: A Philosophical Examination. Abingdon: Routledge.

Brown, A. (2016) 'The "Who?" Question in the Hate Speech Debate: Part 1: Consistency, Practical, and Formal Approaches'. In: Canadian Journal of Law & Jurisprudence 29: 275–320.

Brown, A. (2017a) 'What is Hate Speech? Part 1: The Myth of Hate'. In: Law and Philosophy 36: 419–468.

Brown, A. (2017b) 'What is Hate Speech? Part 2: Family Resemblances'. In: Law and Philosophy 36: 561–613.

Brown, A. (2017c) 'The "Who?" Question in the Hate Speech Debate: Part 2: Functional and Democratic Approaches'. In: Canadian Journal of Law & Jurisprudence 30: 23–55.

Brown, A. (2017d) 'Averting Your Eyes in the Information Age: Hate Speech, the Internet, and the Captive Audience Doctrine'. In: Charleston Law Review 12: 1–54.

Brown, A. (2018) 'What is so Special About Online (as Compared to Offline) Hate Speech? Internet Companies, Community Standards and the Extragovernmental Regulation of Cyberhate'. In: Ethnicities 18: 297–326.

Brown, A. (2020) Models of Governance of Online Hate Speech. Strasbourg: Council of Europe. Available at: https://rm.coe.int/models-of-governance-of-online-hate-speech/16809e671d.

Brown, A. (2022) An Ethics of Political Communication. Abingdon: Routledge.

Brown, A. and Sinclair, A. (2020) The Politics of Hate Speech Laws. Abingdon: Routledge.

Brown, A. and Sinclair, A. (2023) Hate Speech Frontiers: Exploring the Limits of the Ordinary and Legal Concepts. Cambridge: Cambridge University Press.

Butler, J. (1997) Excitable Speech: A Politics of the Performative. New York: Routledge.

Citron, D. K. (2014) Hate Crimes in Cyberspace. Harvard, MA: Harvard University Press.

Cohen-Almagor R. (2015) Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway. Cambridge: Cambridge University Press.

Cortese, A. (2006) Opposing Hate Speech. Westport: Praeger Publishers.

Delgado, R. (1982) 'Words That Wound: A Tort Action for Racial Insults, Epithets, and Name-Calling'. In: Harvard Civil Rights-Civil Liberties Law Review 17: 133–181.

Delgado, R. and Stefancic, J. (2014) 'Hate Speech in Cyberspace'. In: Wake Forest Law Review 49: 319–343.

Esposito, G. (2002) Workshop on racist and xenophobic content on the Internet – problems and solutions. In: International Journal of Communications Law and Policy, Issue 7, Winter 2002/2003.

Faloppa, F., Gambacorta, A., Odekerken, R., van der Noordaa, R. (2023) Study on Preventing and Combating Hate Speech in Times of Crisis. Strasbourg: Council of Europe.

Farrior, S. (1996) 'Molding the Matrix: The Historical and Theoretical Foundations of International Law concerning Hate Speech'. In: Berkeley Journal of International Law 14: 1–98.

Gelber, K. and McNamara, L. (2015) 'The Effects of Civil Hate Speech Laws: Lessons from Australia'. In: Law and Society Review 3: 631–664.

Goldberg, A. (2015) 'Hate Speech and Identity Politics in Germany, 1848–1914'. In: Central European History 48: 480–497.

Greenawalt, K. (1989) Speech, Crime, and the Uses of Language. New York: Oxford University Press.

Greenawalt, K. (1995) Fighting Words. Princeton: Princeton University Press.

Hall, P. (2019) 'Dialogues and Diversity in Korea, Japan, and France: The Contribution of International Law to Hate Speech Legislation in National and Transnational Contexts'. In: M. Kang et al. (eds.) Hate Speech in Asia and Europe: Beyond Hate and Fear. Abingdon: Routledge.

Hassan, G. et al. (2022) 'Hate online and in traditional media: A systematic review of the evidence for associations or impacts on individuals, audiences, and communities'. In: Campbell Systematic Reviews 18 (2022): e1245.

Heinze, E. (2009) 'Cumulative Jurisprudence and Hate Speech: Sexual Orientation and Analogies to Disability, Age, and Obesity'. In: I. Hare and J. Weinstein (eds.) Extreme Speech and Democracy. Oxford: Oxford University Press.

Heinze, E. (2016) Hate Speech and Democratic Citizenship. Oxford: Oxford University Press.

Keller, P. (2011) European and International Media Law: Liberal Democracy, Trade, and the New Media. Oxford: Oxford University Press.

Kors, A. C. (1991) 'Harassment Policies in the University'. In: Society 28: 22–30.

Legg, A. The Margin of Appreciation in International Human Rights Law: Deference and Proportionality. Oxford: Oxford University Press.

Matsuda, M. (1989) 'Public Response to Racist Speech: Considering the Victim's Story'. In: Michigan Law Review 87: 2320–2381.

Mello, M. (2006) 'Hagan v. Australia: A Sign of the Emerging Notion of Hate Speech in Customary International Law'. In: Loyola of Los Angeles International and Comparative Law Review 28: 365–378.

Mendel, T. (2012) 'Does International Law Provide for Consistent Rules on Hate Speech?', In; M. Herz and P. Molnar (eds.) The Content and Context of Hate Speech: Rethinking Regulation and Responses. Cambridge: Cambridge University Press.

Musolff, A. (2015) 'Dehumanizing Metaphors in UK Immigrant Debates in Press and Online Media'. In: Journal of Language Aggression and Conflict 3: 41–56.

Naab, T. K. et al. (2018) 'Flagging Uncivil User Comments: Effects of Intervention Information, Type of Victim, and Response Comments on Bystander Behavior'. In: New Media and Society 20: 777–795.

Nadelmann, E. A. (1990) 'Global Prohibition Regimes: The Evolution of Norms in International Society'. In: International Organization 44: 479–526.

Neller, J. (2018) 'The Need for New Tools to Break the Silos: Identity Categories in Hate Speech Legislation'. In: International Journal for Crime, Justice and Social Democracy 7: 75–90.

Petty, R. and Cacioppo, J. (1984) 'The Effects of Involvement on Responses to Argument Quantity and Quality'. In: Journal of Personality and Social Psychology 46: 69–81.

Richardson-Self, L. (2018) 'Offending White Men: Racial Vilification, Misrecognition, and Epistemic Injustice'. In: Feminist Philosophy Quarterly 4: 1–24.

Riesman, D. (1942) 'Democracy and Defamation: Control of Group Libel'. In: Columbia Law Review 42: 727–780.

Roberto, K. et al. (2020) 'Stigmatization and Prejudice During the COVID-19 Pandemic'. In: Administrative Theory & Praxis 42: 364–378.

Schweppe, J. (2021) 'What is a Hate Crime? In: Cogent Social Sciences 7: 1–14.

Scott, Mark. (2022) '"Grotesque" Russian Disinfo Campaign Mimics Western News Websites to Sow Dissent'. In Politico, 27 September. Available at: www.politico.eu/article/russia-influence-ukraine-fake-news/.

Timmermann, W. K. (2005) 'The Relationship between Hate Propaganda and Incitement to Genocide: A New Trend in International Law Towards Criminalization of Hate Propaganda?' In: Leiden Journal of International Law 18: 257–282.

Turner, Ben. (2022) 'Ukraine War: How Russian Propaganda Has Found a Way of 'Avoiding Detection' Online"'. Euronews, 7 October. Available at: www.euronews.com/2022/10/07/ukraine-war-how-russian-propaganda-has-found-a-way-of-avoiding-detection-online.

Vidgen, B. et al. (2021) Understanding Online Hate: VSP Regulation and the Broader Context. London: The Alan Turing Institute.

von der Leyen, U. (2019) A Union that Strives for More: My Agenda for Europe. Available at: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-nextcommission_en.pdf .

Wainer, H. (1986) Drawing Inferences from Self-selected Samples. New York: Springer-Verlag.

Weber, A. (2009) Manual on Hate Speech. Strasbourg: Council of Europe. Available at: http://www.coe.int/t/dghl/standardsetting/hrpolicy/Publications/Hate_Speech_EN.pdf.

Weinstein, J. (2009) 'An Overview of American Free Speech Doctrine and its Application to Extreme Speech': In: I. Hare and J. Weinstein (eds.) Extreme Speech and Democracy. Oxford: Oxford University Press.