Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Foreign Affairs FDFA

ARTICLE 19

UNITED NATIONS
HUMAN RIGHTS
OFFICE OF THE HIGH COMMISSIONER

The Rabat Plan of Action in the Digital Age

Safeguarding Freedom of Expression and Freedom of Religion or Belief Online

Expert Workshop
21–22 April 2026
Glion, Switzerland

– Expert Workshop Chair's Summary –
(not an official document)

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Foreign Affairs FDFA

Background

1. From 21 to 22 April 2026, the Directorate of International Law of the Federal Department of Foreign Affairs of Switzerland, in cooperation with ARTICLE 19 and the Office of the High Commissioner for Human Rights (OHCHR), held an expert workshop titled '*The Rabat Plan of Action in the Digital Age – Safeguarding Freedom of Expression and Freedom of Religion or Belief Online*' in Glion, Switzerland. The expert workshop brought together over 50 representatives from States, the United Nations (UN), UN Special Procedures and UN Treaty Bodies, as well as civil society, technology companies and academia for two days of dynamic and interactive dialogue.

2. The Rabat Plan of Action (RPA), adopted in 2012, provides detailed and nuanced guidance on addressing advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence (hereinafter referred to as 'incitement'), as required under Article 20(2) of the International Covenant on Civil and Political Rights (ICCPR). The main objectives of the expert workshop were

a) to discuss emerging and persistent challenges related to online hatred targeting religious groups and minorities,
b) to identify good practices and practical solutions to enhance commitment and
c) to ensure the effective application and implementation of the RPA in the digital age.

3. This chair's summary aims to provide an overview of the discussions over the course of the two-day expert workshop. It does not aim to provide a comprehensive or chronological account of the discussions but rather to capture key ideas and points raised by participants. As the discussions were held under the Chatham House Rule, no comments are attributed to individual participants. The points raised in this chair's summary are not necessarily shared by the organisers nor agreed upon by the participants.

Segment I: Setting the Scene

This segment aimed to set the scene for the rest of the expert workshop. It involved a **moderated conversation between UN Special Procedures** to explore persisting and emerging challenges in addressing online hatred and how the rights to freedom of expression, religion or belief, and equality and non-discrimination are interdependent and mutually reinforcing in addressing these challenges. This conversation was followed by a **moderated panel discussion between various stakeholders** to delve deeper into the RPA, including its historical development, normative significance, and existing application and impact.

Freedom of expression, freedom of religion or belief, and equality and non-discrimination

4. The discussions in this segment underlined how the rights to freedom of expression, to freedom of religion or belief, and to equality and non-discrimination are all **mutually reinforcing and interdependent**. This is reinforced by the fact that freedom of expression is not an absolute right and can be restricted to ensure the protection of minorities, including by prohibiting incitement to discrimination, hostility or violence as required under Article 20(2) of the ICCPR. They emphasised that criminalisation must be an exceptional last resort tool for restricting the right to freedom of expression.

5. Participants looked at the **dichotomy outlined in the original RPA** – on the one hand, minorities are often under protected and become victims of incitement, while on the other hand, they are often disproportionately targeted by overbroad legislation purporting to address hate speech, thereby limiting their capacity to express themselves. They emphasised the need to assess whether restrictions on the right to freedom of expression protect religious minorities or those holding political and economic power.

6. Participants in this segment also emphasised that the right to freedom of expression applies to discourses and ideas that are not necessarily shared by the government or the general population, including **dissenting or minority views**. They highlighted how religion can be used as a pretext to suppress the expression of some minority groups.

Hate speech in the digital age

7. The discussions in this segment considered how technologies have democratised speech and allowed people, including from minority groups, to reach wider audiences. However, participants highlighted that these very technologies have also been abused **to spread hatred against religious or belief minorities** and how those seeking to spread hate have become emboldened online. Participants repeatedly stressed that the distinction between online and offline violence is at times overstated, and how they are interlinked, while both have real impacts.

8. Participants highlighted the acute need to strengthen **content moderation policies and practices** across major social media platforms, including by expanding content moderation and factchecking efforts and ensuring adequate investment in regions experiencing high levels of hate speech, regardless of whether they are considered priority markets.

9. Participants also emphasised the need to take into account aspects such as content curation, advertising or monetising practices. They stressed the need for **smart regulation of social media platforms that focused on systemic processes**, such as through reforms of the platform's business models, and the need for human rights due diligence.

10. While participants recognised that artificial intelligence tools can have advantages in moderating the sheer scale of hate speech online, they also recognised the **serious limitations of such automated tools**, especially in detecting and understanding the necessary context and nuance to assess potential instances of hate speech. In addition, the sophistication of these tools varies significantly in different languages.

Normative importance of the RPA

11. This segment also explored how the RPA, while not being legally binding, has weight and legitimacy from being created through a consensual, multistakeholder process and by outlining a **universal methodology grounded in international human rights law**.

12. Participants attached particular importance to **preventative measures outlined in the RPA** that deal with the root causes of incitement, through tools such as education, training and trust building initiatives. In this vein, they reiterated that the RPA's focus is not solely on the question of when prohibition may be justifiable, but it also offers a calibrated response with a broad range of **positive policy measures** and civil measures.

13. Participants drew attention to the **RPA's recommendation for comprehensive anti-discrimination legislation** as crucial in dealing with the root causes of incitement.

Existing application, challenges and impact of the RPA

14. This segment further looked at how the **RPA has gained traction in different contexts**, including being incorporated into over 400 UN documents, as well as being applied at the regional and national levels. Participants noted that the Tirana Framework for Confronting Intolerance, a document adopted in 2025 to address structural discrimination and ensure that freedom of expression supports inclusivity and equality across all societies, calls on States to redouble their efforts to implement the RPA and other existing frameworks, such as the Beirut Declaration on Faith for Rights.

15. Participants outlined **holistic best practices to ensure the implementation of the RPA at the national level**, including implementing awareness raising and education campaigns, strengthening the capacity of law enforcement and the judiciary, improving access to justice, and engaging in multilateral processes, all in collaboration with civil society and faith-based leaders.

16. Participants also discussed how **UN Treaty Bodies have explicitly relied on the RPA** in certain cases, while implicitly incorporating principles from the six-part test in their wider work. They discussed that despite Treaty Bodies being State-focused mechanisms, they are beginning to address the responsibilities of private actors, including the need for the latter to conduct human rights due diligence.

17. Participants emphasised that the RPA diagnoses a problem and creates a framework that responds to a problem that is increasingly complex in the digital age. They highlighted that there are many technology companies beyond just social media companies that do not have established human rights policies or sufficient oversight mechanisms in place, such as those running chatbots, and participants emphasised the **need to engage a broader range of technologies companies** when implementing this framework.

18. Participants highlighted how the six-part test can become a useful tool for social media companies when decisions have been escalated and there is a need for more nuanced judgement calls. However, they expressed the **challenge of social media companies in applying the six-part test in a comprehensive manner** given the scale of content online and need for pragmatism. Similarly, participants questioned to what extent outsourcing the responsibilities under the six-part threshold test to these companies is appropriate.

19. Participants emphasised that the **RPA remains relevant in the digital age**, but no single actor can achieve its objectives alone and how there is an ongoing need for further multistakeholder dialogue. They explored the need to intensify dialogue and bring additional actors into the conversation, such as criminologists and data scientists, as well as to break down the silos of different organisations working on hate speech. They highlighted the International Day of Countering Hate Speech, proclaimed by the United Nations General Assembly in its resolution 75/309, as an opportunity to scale up awareness and implementation of the RPA.

Segment II: Identifying Key Challenges and Issues

This segment aimed to delve deeper into identifying challenges. It involved **exchanges** on the evolving digital dynamics that are reshaping the creation, dissemination and amplification of online hate speech as well as on the impacts, drivers and intersectional dimensions of online religious hatred on affected groups, as well as barriers to redress.

Emerging challenges in the digital age

20. The second segment explored new and emerging challenges related to hate speech online. Participants focused on the growing use of **algorithms** which are accelerating the dissemination and amplification of hate speech, for example by pushing people towards radicalisation. Participants noted that advanced **artificial intelligence systems** are also being developed which could help in moderating content with greater precision, yet recognised their unpredictability and the challenges these tools have in detecting nuance, especially in diverse contexts and languages. The discussion also focused on automated tools leading to ‘false positives’, where minorities have their accounts banned or content taken down as a result of positively appropriating traditionally hateful terminology or discussing issues that affect their communities.

21. Participants discussed **chatbots and large language models in particular as the next frontier of content generation**, and how studies showed that these technologies are sycophantic and confirm existing biases of the user. They explored how content is not just recommended in personalised manner but is now being generated in personalised manner. In this context, participants recommended looking not just at the content layer but also at the infrastructure layer of these technologies, including the training of their datasets.

22. Participants highlighted that while private messages can be used as a vehicle to target individuals on the basis of their religion or belief, **the right to privacy and tools for encryption are essential** for religious and other minorities to communicate and organise safely.

Roles of technology companies, including their business models

23. This segment also delved into the role of technology companies, especially social media companies, in dealing with hate speech. Participants noted how **technology companies' community standards** can be the starting point in dealing with hateful content on their platforms, notably policies on hate speech but also on violence and incitement, bullying and harassment, misinformation, and other relevant policies.

24. Participants emphasised the need to understand and consider the **economic incentives in the creation, dissemination and amplification of hate speech online**, including focusing on the profit-driven business models of technology companies which can have a trickle-down effect on content creators monetising engaging content, as well as the role of advertisers. They emphasised that an economic analysis could help inform which stakeholders have responsibilities in this context and should be part of the conversation, as well as allowing for a tailored approach that looks less at content and more at structures, systems and processes.

25. Participants suggested that, while the RPA is a helpful guide, **social media platforms operate at scale** and cannot feasibly conduct a detailed analysis of millions of pieces of content each day and the six-part threshold test alone cannot guide real-time content moderation or victim protection. Participants also highlighted the risks of outsourcing decisions on incitement to private companies, especially through intermediary liability which creates incentives and risks for over-removal of content to avoid enforcement actions, including fines.

Impacts of hate speech on religious groups and minorities

26. The discussions looked at the impacts of **hate speech on different religious groups, including against Christian, Jewish and Muslim groups and other religious or belief minorities**. Participants highlighted the need for an intersectional approach, and how women from religious minorities often face multiple forms of discrimination on the basis of both their gender and their religion or belief.

27. Participants explored the impact of **hate speech on the basis of race, colour or descent**, including how this problem is compounded by intergenerational trauma and the accumulation of discrimination and violence through generations, including through the legacies of slavery. They also noted that Article 4 of the International Convention on the Elimination of All Forms of Racial Discrimination requires States Parties to declare an offence punishable by law all dissemination of ideas based on racial superiority or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin.

28. Participants highlighted **antisemitism and Holocaust denial** as an urgent problem. They raised how antisemitism is present in daily life and there is a culture of fear and hostility which is often shaped and magnified by the online environment.

29. Participants emphasised that **online hate speech does not exist in isolation**, but rather contributes to, precedes and takes place alongside offline discrimination and violence, and fosters a climate of intolerance that spills into schools, workspaces and other social settings. Participants highlighted the issue of cumulative patterns, and how exposure not only to a single post but rather to a series of harmful posts or comments online can lead to violence.

30. Participants highlighted the **issue of impunity** and how many victims do not report online hate speech due to their mistrust in the system or because the legal response is too often slow.

31. Participants also highlighted that many **anti-blasphemy laws** are still implemented at the national level although this violates international human rights law. They stressed that instances of blasphemy are often seen as a tension between the rights to freedom of expression and freedom of religion or belief. However, these laws often fuel division and religious intolerance by shutting down interfaith dialogue, and also risk legitimising human rights violations against religious or belief minorities.

32. Participants stressed the need to ensure a **positive environment where everyone can exercise their right to freedom of expression**, and noted the chilling effects of online hate speech that can silence minority voices. They shared how minority groups often have to think twice before posting online due to the disproportionate harassment they face online, leading to self-censorship, or avoiding positions or professions in the public eye.

33. Participants emphasised that the issue of hate speech requires action but urged the need for a careful and balanced approach when applying restrictions given that the right to **freedom of expression is fundamental for human dignity, rule of law, peace, security, sustainable development and democracy**.

34. Participants noted that the RPA remains essential as a tool in countering hate speech in line with international human rights law. However, participants highlighted some **challenges in applying the six-part test of the RPA in the digital age**, in particular as online hate speech can be coordinated, masked through anonymity or humour, or rapidly amplified to make an impact beyond the original intent of the speaker.

35. Participants discussed how, in the digital age, content that may not meet the six-part test on an individual level can nevertheless, when combined with other discriminatory content, have cumulative impacts that are equally harmful to those targeted, in particular given scale, speed, and volume. At the same time, participants emphasised that the **RPA must retain its high threshold for prohibited speech**, including the sixth threshold factor of likelihood of harm. They emphasised that chipping away at certain factors might inadvertently lead to censorship, including of the voices that initiatives on hate speech seek to protect.

Impacts of hate speech during crisis, including armed conflict

36. This segment also delved into the additional **risks in online hate speech during situations of crisis**, such as armed conflict and elections. Participants focused on situations of armed conflict and how hate speech is used to spread fear, discredit humanitarian action, and influence behaviours of civilians, including through encouraging violence. Participants noted how the use of images and videos to spread fear through artificial intelligence tools has made matters worse.

37. Participants also explored how these **narratives in armed conflict are often not generated within a given country but also created and shared across borders**, sometimes with financial motivations. Participants gave examples of where hate speech on online platforms not only reflected but shaped ethnic violence during armed conflict, including through actors in the diaspora spreading dehumanising narratives from outside of the country. They also noted how hate speech is often a tool for politicians who spread it through their networks for political gain domestically.

38. Participants discussed potential **markers and contextual classifiers to trigger moderation and identify incitement in armed conflict**, such as explicit calls for violence. Participants stressed that while it was uncontroversial that international human rights law and international humanitarian law apply concurrently in conflict situations, what this concurrent application can mean in specific contexts or with respect to specific types of content is often unsettled.

39. Participants also noted **normative gaps in operationalising Article 20(1) of the ICCPR on the prohibition of propaganda for war**, including in the digital age where it is often interrelated with incitement. They also highlighted how Article 20(1) of the ICCPR does not have a framework comparable to the RPA and the importance of developing additional guidance and clarification in this area.

Segment III: Finding Solutions

This segment aimed to find solutions to the challenges identified. It involved a session with participants **breaking out into working groups**, each tasked with examining how specific components of the six-part threshold test can be applied in the digital age, by sharing ideas, policy positions and best practices regarding the respective responsibilities of different stakeholders.

Challenges and considerations in applying the six-part test

40. The working groups fed back **challenges and considerations in operationalising the six-part test in the digital age**. Some of the discussions extended beyond the strict question of the application of the six-part test.

41. Participants discussed the challenges in applying the six-part test to **content generated by large language models, chatbots and anonymous accounts, coded language,** as well as 'likes' and reposts rather than the post itself.

42. Participants noted platforms may receive millions of reports, possibly making a case-by-case application of the six-part test impractical. They noted that since **harm can unfold in minutes,** the pace of the analysis required by the test risks leaving victims unprotected. The risk of irreversible harm necessitates precaution and prevention.

43. Participants discussed whether assessments of intent to incite might be **complicated by the monetary incentives of engaging content**, fuelled by the business models of large social media platforms, as well as the **principle of foreseeability of harm** and whether determinations of prohibited hate speech should consider whether a speaker could or should have foreseen a serious risk of discrimination, hostility or violence, including in light of the history, status and level of digital literacy.

44. Participants discussed the **transnational nature of online hate speech, including its extraterritorial effect** and the challenges posed by jurisdictional limitations. They also discussed the need to **consider the role of algorithms in amplifying hateful content**, as well as possibilities to utilise metrics to trace a piece of online content and determine the extent of its dissemination.

45. Participants discussed how the accessibility and visibility of some content that meets the threshold under the six-part test can be in the public interest, including being important for accountability purposes – for example speech of a government leader. In this context, they discussed the need to further define and **apply the principle of public interest and newsworthiness** to online content.

46. Participants considered the need to **determine whether online posts call for violence or contain dehumanising language**, especially in relation to audiovisual content which can mask incitement through imagery or humour.

47. Participants stressed the importance of **working with civil society and journalists when assessing the broader social and political context** in environments or moments prone to experience increased levels of hate speech, especially in high-risk settings. They also discussed the **possibilities of mapping where engagement and dissemination are concentrated and adding weight to speech during crisis periods** where online speech is more likely to lead to offline discrimination, hostility and violence, including during armed conflict.

48. Participants discussed the relevance of the **speaker's follower count and the nature of their online audience and content**, including a possible history of using hateful narratives to generate engagement.

49. Participants also put forward a case for a **victim-centred approach to interpreting the RPA**, which includes considering foreseeable harm, recognising patterns and cumulative impacts of posts, accounting for algorithmic amplification, and enabling interim reversible measures, such as de-amplification or temporary restrictions while preserving full analysis for final adjudication. Related discussions focused on distinguishing between the purpose of the six-part test in justifying the *ex ante* prohibition of certain content and broader human rights-based content moderation efforts undertaken in line with Article 19(3) of the ICCPR.

Roles of States

50. Participants stressed the importance of States promoting and protecting the rights to freedom of expression and freedom of religion or belief and **creating an enabling environment** to maximise opportunities to counter religious intolerance online.

51. They emphasised that public officials themselves are often the source of hate speech and disinformation online, and should **refrain from spreading such speech** which can legitimise and facilitate incitement.

52. Participants noted that some States demand social media companies to take down content that should be protected under freedom of expression standards and that States must refrain from the problematic practice to use hate speech as a justification to **impose internet shutdowns and other restrictions in violation of international law**, given that they create an information void, more likely to be filled with inciting or false content, and are inefficient in curbing disinformation or hate speech.

53. Participants discussed the **Digital Services Act (DSA) of the European Union** which does not itself define hate speech, although the revised EU Code of Conduct on Countering Illegal Hate Speech Online, integrated into the DSA framework, refers to the definition in the Framework Decision 2008/913/JHA of 28 November 2008. They noted that the DSA focuses on structural issues, including transparency about their recommender systems, notice and action mechanisms, and appeals through an internal complaint handling system.

54. Participants brought experiences from the Middle East and North Africa region where there have been different legislative initiatives and polices around tackling hate speech online. They raised the challenges of implementing legal safeguards without an **independent judiciary, free civic space, a strong media ecosystem, and the safety of journalists**.

Roles of technology companies

55. Participants emphasised that technology companies have a responsibility to address the spread of hate speech on their platforms through **human rights due diligence, reformed business models and transparency.** They discussed other policy measures, such as ensuring content moderation policies are aligned with international human rights norms and standards, investing in digital, media and information literacy, and providing accessible appeal and remedy mechanisms.

56. Participants recommended to not just focus on social media platforms but also **technology companies throughout the whole ecosystem**, such as developers of large language models, which should all be grounded in international human rights law. They suggested that through such an approach, any guidance would remain relevant throughout future technological advances and developments.

57. Participants also raised the **market power of large technology companies**, whose dominance shields them from competition and incentives to meaningfully challenge hate speech on their platforms, and suggested that recommendations should address human rights implications.

58. Participants also discussed gradated measures that platforms can adopt while content is under review, such as through **factchecking or blurring, labelling, demoting, demonetising, deprioritising, or removing content**.

59. Participants suggested that **partnerships between technology companies and civil society** may be beneficial to direct users to relevant information. An example was shared of collaborations with Meta and TikTok to provide users with and redirect them to authoritative resources on Holocaust distortion and denial.

60. Participants highlighted the importance of social media platforms to **resist requests from governments to censor content** in violation of international human rights law, while recognising possible practical difficulties, including threats to block entire platforms.

Roles of international organisations, civil society and other stakeholders

61. As in earlier segments, participants emphasised **the role of prevention and dealing with the root causes of hate speech** by all stakeholders, including through capacity building, trainings, peer-to-peer learning and interfaith dialogue.

62. Participants emphasised the importance of civil society **conducting trainings on the RPA** and its implementation in the digital age, especially as the misunderstandings surrounding international human rights law often leads to frustration.

63. Participants emphasised that the **RPA is complementary to other UN mechanisms and processes**, including Human Rights Council Resolution 16/18 and the Istanbul Process. They encouraged engagement with related initiatives, including UN B-Tech and UNESCO, and emphasised the importance of coordination with regional human rights courts.

Conclusions

64. The expert workshop reaffirmed the continued relevance and normative strength of the RPA in addressing incitement to discrimination, hostility or violence in the digital age. While the digital age has transformed the creation, dissemination, amplification and consumption of hate speech, the RPA remains essential for ensuring a calibrated and human rights-based response.

65. The discussions throughout the expert workshop highlighted that effectively tackling online hate speech requires moving beyond content-focused approaches toward systemic and preventive action, including all stakeholders. This includes addressing the business models and economic motivations of technology companies, strengthening human rights due diligence and transparency across the sector, and investing in awareness raising, education and trust-building initiatives. This also includes ensuring an intersectional approach that focus on the experiences of religious groups and minorities.

66. The expert workshop also identified significant challenges, particularly in applying the RPA's six-part threshold test at scale to online expression, navigating the role of artificial intelligence, and addressing the transnational nature of online harms. These challenges call for innovative solutions and sustained dialogue.

67. Ultimately, the workshop showed that the rights to freedom of expression, freedom of religion or belief, and equality and non-discrimination are mutually reinforcing and interdependent. While the RPA is adaptable and can be interpreted to be fully operationalised in the digital age, meaningful implementation requires coordinated action by States, technology companies, international and regional organisations, civil society, academia, faith leaders and other stakeholders.

68. The organisers of the expert workshop will prepare a set of observations and non-binding recommendations to provide initial guidance in understanding appropriate steps to combat online hate in the digital age, based on the parameters set out by international human rights law. Over time, these observations and recommendations should gain practical relevance and inform policy and practice in addressing hate speech online. To that end, efforts will be made to introduce references to them in relevant resolutions and processes, particularly those of the UN Human Rights Council and the UN General Assembly.