联 合 国 A/HRC/59/28

大 会

Distr.: General
22 April 2025
Chinese
Original: English

人权理事会
第五十九届会议
2025 年 6 月 16 日至 7 月 11 日
议程项目 2 和 3
联合国人权事务高级专员的年度报告以及
高级专员办事处的报告和秘书长的报告

促进和保护所有人权——公民权利、政治权利、
经济、社会及文化权利，包括发展权

纪念《防止及惩治灭绝种族罪公约》七十五周年闭会期间会议

联合国人权事务高级专员办事处的报告

概要

本报告根据人权理事会第 49/9 号决议提交，简要介绍了 2023 年 12 月 4 日举行的纪念《防止及惩治灭绝种族罪公约》七十五周年闭会期间会议的议事情况。小组讨论主要探讨了社交媒体平台的作用以及试图传播仇恨的人员以这些平台为工具在现实世界制造歧视和暴力的问题。

请回收

一. 导言

1. 人权理事会根据第 49/9 号决议于 2023 年 12 月 4 日举行了一次纪念《防止及惩治灭绝种族罪公约》七十五周年闭会期间会议。

2. 本次会议分三场举行，由人权理事会主席瓦茨拉夫·巴莱克主持。联合国人权事务高级专员福尔克尔·蒂尔克致开幕辞。随后，亚美尼亚外交部长阿拉拉特·米尔佐扬发表了开幕致辞。最后，秘书长防止灭绝种族问题特别顾问艾丽斯·瓦伊里穆·恩德里图也发表了开幕致辞。

3. 第一场会议的嘉宾包括：联合国人权事务高级专员办事处专题活动、特别程序和发展权司司长 Peggy Hicks；消除种族歧视委员会委员 Mehrdad Payandeh；秘书长防止灭绝种族问题特别顾问办公室高级政治事务干事 Simona Cruciani。

4. 第二场会议的嘉宾包括：寻求真相、正义、赔偿和保证不再发生问题特别报告员 Fabián Salvioli；国际刑事法院前检察官 Luis Moreno Ocampo；联合国人权事务高级专员办事处(人权高专办)人权条约处处长 Ibrahim Salama；罗兴亚 Maìyafuìnor 协作网络执行主任 Yasmin Ullah。

5. 第三场会议的嘉宾包括：数字权利基金会创始人兼 Meta 监督委员会成员 Nighat Dad；联合国人权事务高级专员办事处高级人权干事 Scott Campbell；秘书长防止灭绝种族问题特别顾问办公室人权干事 Maria Westergren；“立即普及”组织执行董事 Brett Solomon。

二. 开幕致辞和发言

6. 人权理事会主席在介绍性发言中回顾，小组讨论会系根据理事会第 49/9 号决议举行。在该决议中，理事会要求在第五十六届会议之前举行一次闭会期间会议，讨论社交媒体平台的作用以及试图传播仇恨的人员以这些平台为工具在现实世界制造歧视和暴力的问题。

7. 联合国人权事务高级专员在开幕辞中回顾说，《防止及惩治灭绝种族罪公约》是在《世界人权宣言》通过前夕通过的首项联合国人权条约。75 年后，这两份文书仍然相互交织共振，是促进人权的基本要素。柬埔寨、卢旺达、前南斯拉夫和其他国家的经验教训充分证明，防止种族灭绝和将肇事者绳之以法至关重要。

8. 高级专员强调，发生种族灭绝行为之前必定会有基于种族、族裔、宗教或其他不可改变的特征的可识别和有针对性的歧视制度。这种有系统的针对性攻击往往是由受拥戴或有权威的人物实施的，他们对特定社群进行非人化。这些人物的言论通过社交媒体宣传运动得以放大，往往会使针对这些社区的暴力、歧视和敌意常态化，最终或将酿成种族灭绝的恶果。因此，预防措施至关重要，但惩戒手段也不可或缺。对灭绝种族罪的问责始于有效、迅速、彻底和公正的调查以及后续起诉。问责还可能包括非司法措施，如揭露真相、承认、纪念、教育和赔偿。

9. 亚美尼亚外交部长阿拉拉特·米尔佐扬在开幕致辞中强调，人权理事会关于灭绝种族罪的决议享有广泛的跨区域支持和共识。这反映出国际社会一致认为，必须通过保护人权和族裔、宗教、种族和民族群体来防止种族灭绝。对许多族裔

和宗教群体而言，这个世界是一个十分危险的地方，因为全球范围内都表现出了消灭它们的意图。社交媒体平台上的错误信息有可能传播偏见、仇外心理、不容忍、负面刻板印象和污名化，目的是对这些群体进行非人化。

10. 米尔佐扬先生指出，若能识别早期预警信号，许多大规模暴行本可避免。在纳戈尔诺－卡拉巴赫，逾 10 万名难民被赶出他们世世代代的祖居地，现于亚美尼亚寻求庇护。人类的未来中，不容忍、种族主义、仇恨和仇外心理没有容身之地。

11. 秘书长防止灭绝种族问题特别顾问艾丽斯·瓦伊里穆·恩德里图在开幕致辞中指出，新技术在传播可导致灭绝种族的仇恨言论和煽动暴力行为中扮演重要角色。将仇恨言论、歧视和“他者叙事”工具化并非新概念。此类手段曾出现在二战期间对犹太人的大屠杀、卢旺达对图西人的种族灭绝、波斯尼亚和黑塞哥维那斯雷布雷尼察的种族灭绝以及其他许多可被视为种族灭绝的严重侵犯人权事件之前。社交媒体平台和新技术目前以前所未有的速度助长了仇恨言论的扩散和放大。

12. 恩德里图女士列举了联合国为防止种族灭绝所做的努力，首先回顾了缅甸问题国际独立实况调查团如何记录了社交媒体和其他在线平台上发布的仇恨言论最终造成缅甸境内的严重侵犯人权行为。秘书长 2019 年推出《联合国消除仇恨言论战略和行动计划》，其中他优先重视解决和打击仇恨言论的根源和影响。通过该战略和行动计划后，为决策者和从业者编制了关于如何处理网上仇恨言论的政策文件和指南。

三. 利用社交媒体平台放大可能导致种族灭绝的仇恨言论

A. 嘉宾的发言

13. Hicks 女士在发言中指出，社交媒体最初曾被盛赞为一个能够通过连接全球民众、揭露偏远地区侵犯人权行为而彻底改变宣传倡导工作的工具。然而，目前，政治领导人和公众人物也利用它来破坏民主，散布虚假信息和煽动仇恨。目前社交媒体充斥着大量攻击性内容。国际社会有责任发现那些对于传播可能引发种族灭绝的仇恨言论负责的在线平台，理解这一问题并找到应对方法。

14. Hicks 女士重点介绍了欧洲联盟基本权利署 2023 年发表的一份报告，其中该署全面考察了 4 个国家的 4 个社交媒体平台，并认定这些平台上存在大量厌女和种族主义内容。尽管社交媒体只是现实世界的折射，但其受到的关注实属必然，原因有二：覆盖范围和传播速度。覆盖范围指社交媒体甚至可以让地方性内容产生全球影响。就速度而言，社交媒体的“新陈代谢”可使内容在几分钟内获得数千次点击，每次点击都可能引发再发布和传播。

15. Hicks 女士指出，联合国已制定重要工具来应对网上仇恨言论。世界各国政府和企业普遍采用《关于禁止构成煽动歧视、敌意或暴力的鼓吹民族、种族或宗教仇恨言论的拉巴特行动计划》，包括其中所载的六项门槛测试，用以确定煽动歧视、敌意和暴力的行为。

16. Hicks 女士强调，在处理网上仇恨言论时，不要陷入虚假的取舍。虽然有些人可能会争辩说，保护表达自由与有效应对仇恨言论不可兼得，但这两个目标实

则相辅相成。保护表达自由是营造可信、安全的网络环境的最有效途径。必须避免简单的要么删除、要么保留内容的二元论方法。有证据表明，加大力度限制和压制言论往往会加剧对机构的不信任感，进而扩大滥用和两极分化的范围。因此，必须探索和使用更广泛的敏感内容处理方案，如降低有害内容的可见度或添加警示标签。

17. Hicks 女士强调，社交媒体平台上的有害内容，社交媒体公司虽然难辞其咎，但政府和终端用户同样责无旁贷。部分公司表现得更负责任，同意接受透明度评估并发布透明度报告，而另一些则明确表示不会按国际标准审核内容。必须鼓励那些接受过内容审核培训的公司做得更多，而不是更少。对积极采取措施的公司应予以表彰，对消极抵制者则要点名批评。研究人员需要获取应用程序的编程接口权限，以便分析平台动态并制定最佳解决方案。不幸的是，很多公司正在退出这种合作。

18. Hicks 女士指出，透明度报告虽有帮助，但远远不够，因为需要实时掌握挑战和对策的有关信息。若违规行为发生数月后才提交报告，就无法实现实时响应。某大型社交媒体公司委托开展的一项研究表明，该公司在审核涉及巴勒斯坦的内容时，未能平等对待阿拉伯语和希伯来语内容。这证明，公司必须改进针对具体语境和具体语言的对策。语境和语言方面的挑战对人类而言已是难题，自动化系统在处理这些挑战方面效力更低。

19. Hicks 女士指出，各国政府应负起主要责任，为社交媒体公司的行为设置有效护栏，包括要求其采取措施提高透明度和加强问责制。但若监管仓促或不当，反而可能强化不民主和歧视性的做法，限制言论自由，压制异见并损害其他多项权利。

20. Payandeh 先生首先强调了种族主义仇恨言论与可能引发包括种族灭绝在内的大规模暴行的煽动种族仇恨及暴力行为之间的关联。社交媒体平台可能放大仇恨言论，加剧民族、族裔、种族或宗教对立。与传统媒体相比，社交媒体的特性导致责任归属可能被摊薄。传统媒体传播仇恨言论时，责任明确涉及言论发布者、传播上述言论的媒体公司及监管失职或未能作出妥当反应的国家当局。而社交媒体因常通过虚假账号散布仇恨言论，使得责任认定更为复杂。

21. Payandeh 先生提到，算法可能固化现有种族偏见，强化结构性的不平等和歧视。但不能因聚焦算法而误认为监管社交媒体上的仇恨言论主要是一项技术方面的挑战。这本质上仍是人为的决定，因此内容方面的责任最终仍应由传播仇恨言论的个人和机构承担。算法亦可发挥积极作用，因其能有效对抗社交媒体上的负面刻板印象和误导性叙事。

22. Cruciani 女士首先强调了仇恨言论与灭绝种族行为之间的关联，这一关联在波斯尼亚和黑塞哥维那、卢旺达等多地事件中均有体现。虽然仇恨言论不尽相同，但根据《公民及政治权利国际公约》第二十条，凡是导致煽动暴力、敌意或歧视的言论都应予以禁止或限制。即使是受表达自由权保护的仇恨言论，由于其可能演变为更致命的形式并引发线下暴力乃至种族灭绝，仍然具有危险性。

23. Cruciani 女士指出，否认种族灭绝的言论被视为未来可能发生种族灭绝事件的一项指标。否认种族灭绝的行为将肇事者和战犯变为英雄，同时播下恐惧、猜忌与仇恨的种子。这种形式的言论造成社群和国家撕裂，阻碍愈合进程与同理心

的建立。它还质疑已经作出的司法裁决，从而破坏正义。这种做法直接违背了曾经历种族灭绝和暴行罪的社会的最大利益。

24. Cruciani 女士指出，需要加大投入制定积极主动的战略，并就灭绝种族及相关暴行罪以及网上否认和歪曲历史行为的危害加强宣传教育。否认和歪曲灭绝种族行为可能导致受害者再次受到伤害，使历史上的歧视模式和负面刻板印象循环往复，并可能煽动对曾经历种族灭绝社群的敌意和暴力。社交媒体公司和网络平台采用联合国对仇恨言论的定义至关重要。它们还必须制定社区准则，以全面识别和删除否认或歪曲历史上的种族灭绝及相关暴行罪的网上内容。在审核内容时，社交媒体公司还须咨询各种利益攸关方，如民间社会组织和经历过暴行罪的群体代表。

B. 讨论情况

25. 在小组讨论的互动部分，下列代表团作了发言：巴基斯坦(代表属于伊斯兰合作组织成员的联合国会员国)、哥斯达黎加(同时代表保护责任之友小组)、伊拉克、美利坚合众国、以色列、巴勒斯坦国、大不列颠及北爱尔兰联合王国、卡塔尔、丹麦、乌克兰、土耳其、伊朗伊斯兰共和国、亚美尼亚、中国、巴西、委内瑞拉玻利瓦尔共和国。

26. 下列非政府组织作了发言：独立人权委员会、联合国观察组织、米赞人权中心、世界犹太人大会、全球无杀戮研究中心、国际人权服务社、人权与移民庇护协会。

27. 来自各种不同法系、传统、文化和宗教背景的多国代表一致认为，社交媒体平台已成为助长仇恨言论和非人化的新型增强渠道。代表们认同，很明显，仇恨言论攻势、种族优越论调和目标社群的边缘化先于种族灭绝行为，这些模式是种族灭绝的前兆。代表们一致认为，每个国家都有责任保护其人民免遭种族灭绝。

28. 两位代表强调，当前亟需加强对仇恨言论和煽动暴力行为的监测，并更有效地运用预防工具，以禁止暴行罪，减少网络言论的滥用。代表们表示关切的是，社交媒体平台上仇恨言论日益泛滥。

29. 大多数发言的代表认为，需要在网络言论审核方面加强国际合作。他们强调促进线上线下平等在全球都具有重要意义，并一致认同促进真相、正义和问责对预防种族灭绝的发生至关重要。他们同意需要加大全球努力，确保社交媒体成为促进人权的关键工具。

30. 许多代表对社交媒体在宣传和传播有关冲突各方的错误信息和仇恨内容方面扮演的角色表示失望。他们一致认为，必须坚定不移地致力于预防种族灭绝和其他暴行，因为如果仇恨不受约束，暴力定会随之而来。代表们认同，历史上所有的大规模屠杀都始于歧视、仇恨言论和对特定群体的非人化，例如卢旺达的图西人被比作蟑螂，德国的犹太人被比作啮齿动物和寄生虫，缅甸的罗辛亚人被比作跳蚤、荆棘和入侵物种。一些代表赞扬某些在线平台已经采取行动处理仇恨言论和类似关切。

31. 许多发言的代表表示关切的是，以色列利用社交媒体传播仇恨、煽动、歧视、虚假信息、深度伪造、审查、骚扰和隐形封禁，这助长了针对加沙的巴勒斯

坦人的暴力、敌意和种族灭绝。另一些代表则对社交媒体被用于宣扬反犹主义和消灭犹太人和以色列的行为表示关切。

32. 几位代表承认，监测仇恨言论的努力不应不当限制表达自由权。

33. Cruciani 女士回应说，联合国力图通过《联合国消除仇恨言论战略和行动计划》打击仇恨言论，该计划包含打击仇恨言论及煽动政治或公民暴力行为的指令。这一工具为各利益攸关方提供了指导方针，说明了何种言论构成仇恨言论以及仇恨言论在何时构成煽动行为。联合国还通过由 17 个联合国机构组成的“仇恨言论问题工作组”开展相关工作，该工作组负责审查仇恨言论的根本成因和社会影响，并就如何与信息技术公司和社交媒体公司接触制定相关建议。

34. Payandeh 先生在回答中重点指出了消除种族歧视委员会关于打击种族主义仇恨言论的第 35 号一般性建议(2013 年)以及该委员会关于防止和打击执法人员种族定性行为的第 36 号一般性建议(2020 年)，其中包括对算法和人工智能的讨论。在这两份一般性建议中，该委员会强调了应采取的措施，包括透明度、监督、评估算法对人权的影响以及各国在区域和国际范围内对社交媒体的监管。委员会还强调，必须确保法律不会被用来针对特定边缘化群体以压制其意见自由。

35. Hicks 女士在回应时指出，要求开展人权尽责的趋势日益增强，对公司防范网上仇恨言论的义务产生了重大影响。必须密切监测政治领导人的声音，以确保他们不会放大有害内容。在进行内容审核时，必须纳入那些最直接受到有害内容影响的人的观点，如妇女、少数群体成员、种族歧视受害者和边缘化群体。还应加强联合国与利益攸关方之间的合作，同时利用人权框架和《拉巴特行动计划》应对社交媒体上的仇恨言论。

四. 国家在防止数字空间内直接和公开煽动种族灭绝方面的义务

A. 嘉宾的发言

36. Salvioli 先生首先重申国家有义务防止煽动种族灭绝，并重申必须建立进程和机制，确保问责和促进线上线下和平。他指出，必须遵守寻求真相、正义、赔偿和保证不再发生问题特别报告员 2023 年向人权理事会提交的报告[1] 中概述的关于真相、司法、赔偿和保证不再发生的国际标准。

37. Salvioli 先生强调，必须调查和记录种族灭绝行为，并将肇事者绳之以法。各国必须力求惩治犯下灭绝种族罪行者，量刑应与其罪行的严重程度相称。还须向受害者及其家属公开道歉，并提供康复手段和赔偿。防止种族灭绝的体制措施与纪念活动和保证不再发生之间存在直接、密切的关联。

38. Salvioli 先生着重指出，若不对灭绝种族罪予以适当惩治，将传递出这种极端罪行可被容忍的信号，从而导致正义缺失和有罪不罚现象。应修订体制和法律框架，并落实适当的审查程序和立法，以确保国内法符合国际义务。

39. Salvioli 先生表示，不应将新技术用来传播仇恨言论，或假借表达自由之名掩盖仇恨言论之实。历史表明，在没有歧视的环境中，危害人类罪绝无滋生可

[1] 见 A/HRC/54/24。

能。尽早将人权教育作为一门课纳入学校课程，有助于培养人权文化和社会意识，使仇恨言论和歧视被社会边缘化并遭到唾弃。

40. Ocampo 先生在发言中称，防止数字空间内直接和公开煽动种族灭绝，既是国家的责任，也是社交媒体公司的责任。各国需加强监督网络仇恨言论及其与种族灭绝之间的关联。即使已经提请其注意，各国往往还是会拒绝承认正在发生的种族灭绝行为或其风险，1994 年安全理事会对卢旺达暴行的轻描淡写即为明证。

41. 联合国坚持认为，必须由国际法院或国际刑事法院作出裁决，方可将某一局势定性为种族灭绝，Ocampo 先生建议联合国修改这一立场。这一要求对惩治而言确有必要，但忽视了预防的必要性。因此，由于法院迟迟无法作出裁决，实施预防措施的能力受到影响，即使存在明显风险，工作人员也无法采取明确行动防止种族灭绝。

42. Ocampo 先生强调，国际法院在 2007 年《防止及惩治灭绝种族罪公约》的适用(波斯尼亚和黑塞哥维那诉塞尔维亚和黑山)一案的判决中宣布，当一国知悉或理应知悉存在发生灭绝种族的严重风险时，即产生预防和行动义务。这意味着预防义务并不仅仅在实际发生灭绝种族行为时产生，而是在根据现有信息足以合理认为将发生灭绝种族行为时即已产生。一旦达到这一证据标准，就应启动刑事调查作为预防手段。

43. Ocampo 先生坚持认为，社交媒体上的预警信号可能达到必要的证据标准，表明一国已充分知悉灭绝种族的严重风险。这反过来足以触发国家采取预防或威慑行动。各国必须根据国际法院 2007 年对波斯尼亚和黑塞哥维那诉塞尔维亚和黑山一案的判决实施预防措施。

44. Salama 先生首先强调了在落实国家义务方面的三个主要挑战。首要问题是如何界定煽动暴力行为，同时最大限度保障表达自由。其次是如何管理非国家行为体，对其施加影响，使其对自身行为、传播内容和公众参与承担法律责任。再次是如何在极具技术复杂性的数字空间落实国家义务，社交媒体在数字空间可能迅速激化并升级事态。

45. Salama 先生重点强调了《拉巴特行动计划》中的若干补救措施，包括区分煽动暴力与正当表达自由的门槛测试。某些国家的视听监管机构、欧洲人权法院以及 Meta 监督委员会均采用该门槛来区分表达自由与内容审查。

46. Ullah 女士在发言中援引联合国难民事务高级专员公署印度尼西亚办事处一则关于罗兴亚人种族灭绝问题可持续解决方案的社交媒体帖文。该帖文获得约 80 个点赞支持，却收到近 1,500 条据称是印度尼西亚国民的人发表的反对罗兴亚人出现在印度尼西亚的评论。据观察，部分账号是在评论前数小时才创建的，似乎旨在匿名散布针对罗兴亚人所谓“难以管束”、“问题重重”及“缺乏教养”的极端观点。此类相同论调曾在缅甸境内广泛传播，表明种族灭绝行为与种族灭绝文化之间存在不可分割的关联。

47. Ullah 女士回顾说，缅甸敏昂莱大将曾通过 Facebook 煽动对罗兴亚人的种族灭绝，其帖文将罗兴亚人和孟加拉裔在缅长期困境归咎于前政府，同时宣称新政府将着力解决这一问题。大赦国际强调，Facebook 早在 2011 年就通过多项内部研究证实算法对罗兴亚人造成的伤害；然而，这种伤害对人权构成的风险没有得

到及时处理。相反，其算法机制强化了仇视罗兴亚人的内容。由于冲突中人权保护缺位，缅甸政府得以将社交媒体作为武器来延续暴力。

48. Ullah 女士表示，鉴于种族灭绝文化极易被他国效仿，必须确保各国持续参与并合作执行和落实防止和处理大规模暴行的措施。在孟加拉国、印度尼西亚、马来西亚和泰国，部分由于仇恨宣传，罗兴亚人已被视为国家安全威胁。各国应参照人权理事会第 16/18 号决议，采取行动防止基于宗教或信仰的不容忍、负面刻板印象、污名化、歧视、煽动暴力及暴力行为。

B. 讨论情况

49. 在小组讨论的互动部分，下列代表团作了发言：利比亚(同时代表阿拉伯集团)、乌克兰、希腊、俄罗斯联邦、亚美尼亚、伊拉克、巴勒斯坦国、丹麦、阿尔及利亚、阿塞拜疆、约旦、突尼斯、伊朗伊斯兰共和国。

50. 下列非政府组织也作了发言：国际人权服务社、人权与移民庇护协会、法律援助会──法律服务于人、全球无杀戮研究中心。

51. 许多代表深表关切的是，巴以冲突中滥用社交媒体平台的行为，主要是以色列的这种行为，加剧了仇恨言论，助长了两极分化。许多代表一致认为，各国肩负着确保国际人权规范和标准在数字空间内得到落实以防有害思想传播的首要责任。许多代表表示失望的是，有人利用虚假信息对巴勒斯坦人进行非人化，并为犯下暴行进行开脱。

52. 几位代表指出，以色列政府通过数字压制手段使巴勒斯坦人在网上噤声，包括传播虚假信息、内容审查、网络骚扰、人肉搜索和隐形封禁。几位代表呼吁停火并采取司法措施，防止暴行罪，并保证不再发生。

53. 一些代表指出，可使用数字技术生成、传播并放大错误信息。他们敦促各国不要在国内或国际上为政治或其他目的开展或资助虚假信息宣传运动。一位代表强调指出，俄罗斯联邦不断利用现代技术和虚假信息作为针对乌克兰的致命武器，公开策划侵略战争，煽动种族灭绝，否认乌克兰作为一个民族国家存在的权利。一位代表呼吁防止灭绝种族罪行和保护责任办公室在修订《暴行罪分析框架》时处理预防的数字层面。

54. 一些代表承认，新技术有可能成为传播仇恨的放大器，包括否认对犹太人的大屠杀等种族灭绝行为，引起对数字时代未来的严重关切。人工智能与合成媒体的作用及其潜在风险与进步仍存在不确定性。各国有集体责任果断采取行动，在尊重言论和表达自由权的前提下，明确他们在防止数字空间内煽动种族灭绝行为方面的义务。

55. 一位代表援引 2022 年 12 月在第四届惩治灭绝种族罪全球论坛上通过的《埃里温联合行动宣言》，强调有效预防灭绝种族罪取决于加强预警机制、保存历史记忆、了解真相的权利和确保问责。有必要确立新的人权，特别是网上的权利。

56. 几位代表指出，需要有更多的国家通过和批准《灭绝种族罪公约》。他们还敦促那些已经批准《公约》的国家履行其根据《公约》承担的义务，努力防止任何形式的煽动种族灭绝行为。几位代表还恳请各国采取必要措施，打击线上线下一切可能导致种族灭绝的行为。

57. Salvioli 先生在回应时强调指出，国际社会必须将《灭绝种族罪公约》和《国际刑事法院罗马规约》纳入其国内法。各国必须在防止和惩治种族灭绝和其他暴行的行动中消除任何双重标准，而不论是谁犯下了违法行为。

五. 社交媒体平台在应对煽动暴力和歧视的内容方面的作用

A. 嘉宾的发言

58. Dad 女士在发言中强调指出，践踏人权行为和暴行仍在全球各地持续发生，特别是针对土著人民和少数群体的此种行为。在遭受系统性镇压和种族灭绝的地区，社交媒体的效果因其日益增长的全球影响力而被放大。2019 年受威胁民族指数证明了社交媒体在煽动暴力方面的普遍作用。通过人工智能和深度伪造产生的媒体以前所未有的速度强化了仇恨。

59. Dad 女士指出，社交媒体公司没有采取足够的行动，未能充分认识到它们在动荡和易发生冲突地区助长暴力和分裂方面扮演的角色。然而，正如利比亚内战期间 Facebook 群组所证明的那样，这些公司在传播有价值的信息、提供公共服务以及监控军事和叛乱运动方面也发挥了积极作用。

60. Dad 女士指出，社交媒体公司在审核内容时面临着复杂的任务。它们必须在遵守国家法律和维护国际人权标准之间取得平衡，同时恪守对用户表达自由的承诺。更多社交媒体平台应设立监督委员会，并采用语境化方法处理存在争议的内容。在审核内容时，语言因素也同样重要，因为俚语或习语在某一特定语境中可能有害，而在另一种语境中则无关紧要。社交媒体公司还必须监测内容政策的执行过严或过松问题，尤其当执行政策可能对以特定语言或在特定地区发布的内容存在偏见时。

61. Campbell 先生在发言中强调，社交媒体在社会中发挥着积极与消极的双重作用。社交媒体公司对应对煽动暴力与歧视的内容负有重大责任。它们必须审查自身政策，使其符合《工商企业与人权指导原则》。工商企业应当尊重人权，避免通过自身活动造成或助长不利人权影响，并解决涉及其自身的任何负面影响。

62. Campbell 先生指出，社交媒体公司应努力深入了解“煽动行为”的门槛，并熟悉《拉巴特行动计划》和其中概述的六部分测试。社交媒体组织必须采取一切适当措施，确保其政策、业务、通知程序和补救措施为各方所理解，并在全球各地得到一致执行。这包括就其内容管理方法、使用人工智能情况以及与各国和外部机构的数据共享做法保持透明。

63. Westergren 女士在发言中首先指出，2023 年秘书长防止灭绝种族问题特别顾问办公室与埃塞克斯大学“人权、大数据与技术”项目联合发布了题为《打击和应对网上仇恨言论：决策者和从业者指南》的政策文件。该文件提出多项重要建议，包括必须确保科技公司和社交媒体公司的商业模式、内容、审核及政策遵循人权标准。该文件的另一项建议重点指出，社交媒体公司需加大投资，以打击其平台支持的所有语言中的网上仇恨言论。投资应植根于对当地情况的全面了解，并应更多地关注仇恨言论目标群体的观点和声音。此外，投资还应包括实施政策和工具，处理煽动暴力或种族灭绝方面令人高度关切的局势。此类措施包括降频展示、用户警告、推广积极替代叙事以及开展数字素养和其他提高认识方案。这

些措施应透明并符合人权法，纳入种族灭绝风险的具体指标及其网络表现形式，包括在现实生活中引发暴力的可能性。

64. Solomon 先生在发言时首先强调，社交媒体公司在防止和减轻使用其平台所致人权侵害方面负有重大责任。鉴于社交媒体在追求超额利润的同时，助长了煽动种族灭绝、歧视与暴力内容的全球传播，因此科技公司和社交媒体公司已被敦促改善人权实践。社交媒体公司应利用其利润，向埃塞俄比亚、印度、缅甸、苏丹、巴勒斯坦国及其他地区的弱势用户分配资源，推动创建冲突或种族灭绝发生前的紧急响应机制，从而避免在危机期间仓促应对。

65. Solomon 先生补充说，鉴于仇恨言论和网上冲突常以多数社交媒体公司不熟悉的语言出现，建立具备过硬语言技能及当地和地区专业知识的内部团队也至关重要。这将更好地支持种族灭绝和其他暴行罪的预防工作，相比缺乏人工监督、脱离语境、造成灾难性后果的自动化工具和权宜之计是一大进步。

66. Solomon 先生敦促社交媒体组织理解冲突和仇恨的生命周期。“立即普及”组织和其他组织共同发布了一份危机时期内容和平台治理原则宣言。该宣言除其他外建议为即将到来的冲突监测早期预警系统，确保制定有关删除内容的规范流程及申诉机制，以评估内容是否遭到误删。社交媒体亦可用于建设和平和政治参与目的，从而缓解冲突并优先保障社群。必须认识到，国家往往是仇恨言论常态化的推手；各国在对网络实施政治管控的同时不应限制合法言论。

B. 讨论情况

67. 在小组讨论的互动部分，下列代表团作了发言：德国、法国、乌克兰、亚美尼亚、瑞士、土耳其、阿塞拜疆。

68. 一个区域性政府间组织——欧盟作了发言。下列非政府组织也作了发言：世界犹太人大会、法律援助会——法律服务于人、国际人权服务社、米赞人权中心。

69. 一位代表 27 个国家发言的代表说，各国有义务防止和惩治灭绝种族罪，并且在出现预警信号时必须作出坚决反应。科技公司也必须加强在该领域的努力，因为它们在建立基于人权的数字空间方面发挥着关键作用。打击仇恨言论的努力不能成为过度限制表达自由权的借口。许多代表对这一说法表示赞同。

70. 一位代表表示，鉴于网络空间已成为新的全球公共领域，因此需要建立一种能够将利益攸关方、各国、学术界、非政府组织以及受社交媒体和网络平台影响最大的各方聚集在一起的全球数字秩序。

71. 几位代表指出，在社交媒体平台拥有巨大权力和影响力的数字空间，实施积极主动的措施势在必行。另有一些代表则强调，必须加强透明度、问责制以及国家与社交媒体公司之间的合作，以有效防止和处理威胁人权的内容的传播。

72. 一位非政府组织代表表示，该组织与社交媒体公司、联合国系统及民间社会组织合作，共同打击否认和歪曲犹太人大屠杀的行为。该组织与联合国教育、科学及文化组织合作，推出了 www.aboutholocaust.org 网站，并将其纳入 Meta，以确保搜索犹太人大屠杀相关术语的用户被引导至可靠的信息来源。在该网站上，

还可以进行由人工智能驱动的互动体验，通过大屠杀幸存者的叙述，让全球儿童了解大屠杀成因并深化对预防概念的理解。

73. 许多发言的代表表示关切的是，倡导巴勒斯坦人权利的内容一贯遭到严重且不成比例的审查，包括删除内容、限制、下架及隐藏标签。这些代表强调，以色列政府继续对社交媒体公司施压，要求其审查巴勒斯坦人的言论，压制批评以色列政策的声音。对民间社会和人权维护者的限制严重威胁到表达自由、信息获取、政治参与及集会自由。代表们呼吁社交媒体公司解决这些问题。

74. Campbell 先生回应指出，社交媒体公司对世界上某些利润较低但冲突频发的市场投入资源不足，而这些公司本应加大力度将利润投资于人权尽责和针对具体政治形势的专项研究。这包括确保获得适当的语言技能，聘用对当地情况有深刻了解并能有效应对的国家一级分析人员。各国应采取更多措施，打击利用社交媒体作为工具宣传暴力的行为，同时要非常注意避免越权。

六. 结语

75. 恩德里图女士重申，仇恨的表达既可能是种族灭绝的前兆，也可能是种族灭绝的催化剂。历史上每起种族灭绝事件之前都出现过仇恨言论，且是这一罪行的主要诱因。各国和社交媒体公司必须熟悉《暴行罪分析框架》，用以分析灭绝种族罪、战争罪、危害人类罪及其他暴行罪。