



第七十七届会议

临时议程* 项目 69(b)

促进和保护人权：人权问题，包括增进人权和基本自由切实享受的各种途径

武装冲突期间的虚假信息信息与意见和表达自由**

秘书长的说明

秘书长谨向大会转递促进和保护意见和表达自由权特别报告员伊雷内·汗根据人权理事会第 43/4 号决议提交的报告。

* [A/77/150](#).

** 本报告在截止日期之后提交以反映最新情况。



促进和保护意见和表达自由权特别报告员的报告

摘要

在本报告中，促进和保护意见和表达自由权特别报告员伊雷内·汗审查了在武装冲突期间操纵信息对意见和表达自由构成的挑战。她在报告中指出，数字时代的信息环境已成为一个危险的战场，其中国家和非国家行为体在数字技术和社交媒体的帮助下，将信息武器化，以制造混乱、助长仇恨、煽动暴力和延长冲突。

特别报告员强调信息权作为一项“生存权”至关重要，人们的生命、健康和安​​全都有赖于此，她建议在武装冲突期间，人权标准应与国际人道法一道得到加强。她敦促各国重申致力于维护意见和表达自由，并确保打击虚假信息、宣传和煽动行为的行动以人权作为坚实基础。她建议社交媒体公司使其政策和做法与人权标准保持一致，并在世界各地连贯一致地适用这些标准。她最后重申，需要建设社会抵御虚假信息的能力，并推动采取多方利益攸关方办法，使民间社会以及各国、公司和国际组织参与其中。

目录

	页次
一. 导言	4
二. 概念、受害者和媒介	5
A. 概念方面的挑战	5
B. 面临风险的人员和问题	6
C. 操纵信息的媒介	9
三. 法律现状摸底	10
A. 同时适用国际人权法和人道法	10
B. 国际人权法下的信息操纵	11
C. 国际人道法下的信息操纵	13
D. 保护记者	13
E. 对人权的域外责任	14
四. 国家应对措施：关切、挑战和良好做法.....	14
A. 促进获取信息	15
B. 国家支持的虚假信息和宣传	15
C. 对记者和人权维护者的攻击	16
D. 社交媒体监管	17
E. 干扰互联网和电信	18
五. 社交媒体公司：角色和职责.....	19
A. 冲突环境中的社交媒体	19
B. 冲突期间的公司法律标准	20
C. 公司政策	21
D. 公司做法	22
六. 结论和建议	25
A. 对各国的建议	26
B. 对各公司的建议	27

一. 引言

1. 在武装冲突期间，民众处于最脆弱的状态，最需要准确、可靠的信息，以确保自身的安全和福祉。然而，恰恰是在这些情况下，他们意见和表达自由，包括“寻求、接受及传播各种消息及思想之自由”，¹ 受到战争状况以及冲突各方和其他行为体出于政治、军事和战略目的操纵和限制信息的行动最严重的制约。

2. 国家和武装团体操纵信息及信息环境长期以来一直是战争的一个特点。操纵的形式多种多样，从试图欺骗和打击敌军士气的“战争诈术”到旨在影响公众的“信息行动”和旨在煽动针对少数民族的暴力的“仇恨言论”。新的、令人严重关切的问题是，数字技术正在轻而易举、大规模和迅速地创造、传播和扩大虚假或误导性的有害信息。

3. 社交媒体平台在现代冲突中扮演着双重角色。一方面，它们使人们能够与家人、朋友和外部世界保持联系，并获得各种关键救生信息。另一方面，它们又成为传播虚假消息、宣传和仇恨言论的媒介。

4. 许多国家为了对付虚假信息，或作为其自身操纵信息的一部分举措，通过有关国家安全、反恐或“虚假新闻”的法律、攻击独立记者和人权维护者、关闭独立媒体、切断互联网和监管数字平台的方式，限制获取信息的机会，这样的方式会损害人权，加剧他们本来意欲解决的问题。

5. 意见和表达自由不是问题的一部分，而是打击虚假信息的手段，本身也是一种价值。获得各种可核查的信息来源是一项基本人权。这是受冲突影响的社会中民众的基本需要。实际上，这是一种“生存权”。它对于解决冲突、揭露侵犯人权行为以及寻求正义和追究责任也至关重要。

6. 虚假信息和其他形式的信息操纵对平民，特别是边缘化和弱势群体，以及对人权、人道主义行动和和平进程都构成加剧的风险，这突出表明，迫切需要重申各国和公司维护意见和表达自由的义务。只有通过尊重人权和人道主义原则，才能在武装冲突期间保持信息的真实性和民众的安全。

7. 特别报告员在本报告中，以其关于在促进和保护意见和表达自由的同时打击虚假信息的报告² 为基础，重点阐述国际人道法所界定的武装冲突背景下的虚假信息、宣传和构成煽动歧视、敌意或暴力的鼓吹仇恨言论(“仇恨言论”)。

8. 特别报告员在报告中审查了操纵信息的性质和影响，法律框架，以及国家和社交媒体公司的作用。她指出了挑战和威胁，包括合规情况薄弱或法律不确定的领域，以及良好做法，并主要向各国和社交媒体公司提出了建议。本报告不是一项全面的研究，也没有审查武装团体或社交媒体以外的公司的作用或网络作战。

¹ 《公民及政治权利国际公约》第十九条第二款。

² [A/HRC/47/25](#)。

9. 特别报告员承认这一问题具有复杂性和敏感性，需要进一步研究、分析和协商，并认为本报告是与有关利益攸关方接触的初步步骤。

10. 本报告受益于各国、记者、人权维护者、学者和民间社会组织的协商和提交的书面材料。³

二. 概念、受害者和媒介

A. 概念方面的挑战

11. 尽管操纵信息问题得到越来越多的关注，而且联合国已通过关于虚假信息的决议，⁴ 但关于各种概念及其相互关系仍存在很大程度的混乱。国际上没有就定义达成一致意见，突显出这些概念的复杂性以及围绕这些概念存在的争论。

12. 虚假信息、宣传和鼓吹煽动歧视、暴力和敌意的言论有一些共同特点：缺乏明确、商定的定义；在动乱和冲突环境中出现频率高；通过数字技术得到加强；更加侧重于平民而不是军事人员；对人权、民主与和平进程具有有害影响。这些概念有时有重叠。例如，某些形式的虚假信息和宣传可等同于鼓吹煽动暴力、敌意、歧视和战争罪的言论。

13. 所有这三个概念都含有某种程度的操纵、欺骗和歪曲信息的行为，目的是制造混乱，包括关于其本身含义的混乱。真实信息被颠倒黑白地称为“假新闻”或虚假信息。观点、看法和不确定的知识被歪曲，目的是抹黑来源。虚假信息被用作造成伤害的工具。来自国际机构的可核实数据，包括联合国独立专家的报告，被当作虚假信息忽视，而宣传则被当作事实宣扬。

14. 虚假信息已经存在了几千年，但在数字时代又有了新的流行性。虽然没有一个统一的虚假信息定义，但该术语越来越多地用于表示操纵虚假或误导性信息以故意欺骗和造成公众伤害。⁵ 应将其与错误信息区别开来，后者是在无意造成伤害的情况下传播的不实信息。

15. 国际法中提到了宣传，但没有对宣传进行定义。⁶ 虚假信息的目的是混淆和破坏，而宣传的目的是推进某个特定的议程或政党。这个词来源于“传播”或散布信息和观点的概念，它有一种贬义，即传播的信息可能或真或假，但具有偏见、片面性、误导性且令人情绪激动。⁷ 宣传和虚假信息可能作为“信息行动”的一

³ 提交的材料可查阅 <https://www.ohchr.org/en/special-procedures/sr-freedom-of-opinion-and-expression>。

⁴ 大会第 76/227 号决议和人权理事会 A/HRC/49/L.31/Rev.1 号决议草案；另见 A/HRC/47/25。

⁵ 见欧洲联盟委员会，《反虚假信息行为守则》加强版，可查阅 <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>。

⁶ 《公民及政治权利国际公约》第二十条第一款。

⁷ 见 Manfred Nowak, *U.N. Covenant on Civil and Political Rights CCPR Commentary* (Kehl am Rhein, Germany, N.P. Engel, 1993)。

部分而有所重叠，这种行动通常被理解为国家或政治行为体为实现政治和军事目标而影响对手或公众的观点、态度和行为的运动。

16. 国际法禁止煽动或者直接或间接怂恿或鼓动犯罪的行为。本报告中使用的“煽动”一词是指构成煽动歧视、敌意、暴力、战争罪、危害人类罪和灭绝种族罪的鼓吹民族、种族或宗教仇恨言论(通常称为“仇恨言论”)。

B. 面临风险的人员和问题

17. 越来越多的证据表明，有人操纵信息，以引发、加剧和长期维持暴力，用相互矛盾和虚假的消息增加战争的迷惑性，并助长不信任的气氛。武装冲突和虚假信息与其他不满情绪复杂地相互作用，加剧人类痛苦，助长仇恨，并以弱势群体为目标。

18. **少数和边缘化群体。**从 30 年前的卢旺达⁸ 到最近的缅甸⁹ 和埃塞俄比亚，¹⁰ 冲突各方利用大众传播平台煽动民众的仇恨，使对方丧失人性，煽动严重侵犯人权行为、战争罪、危害人类罪和灭绝种族罪。¹¹ 在某些情况下，政治领导人使用不容忍、分裂和危险的言论，否认既定事实，加剧紧张局势，并将民族、族裔和宗教团体作为替罪羊。¹² 难民、境内流离失所者和移民常常被描绘成对国家安全或社会凝聚力的威胁，以激起对他们的仇恨。¹³

19. 近年来，中非共和国的网上仇恨言论和煽动行为助长了基督徒和穆斯林之间的暴行循环。¹⁴ 在埃塞俄比亚，在线“活动人士”利用他们的脸书账户散布仇恨，煽动攻击、杀戮和驱逐其他部落。¹⁵ 在伊拉克，激进组织利用 Telegram、脸书和

⁸ 见 Montreal Institute for Genocide and Human Rights Studies, “Rwanda radio transcripts”, 可查阅 <https://www.concordia.ca/research/migs/resources/rwanda-radio-transcripts.html>。

⁹ A/HRC/39/CRP.2; 另见缅甸表达自由组织提交的资料。

¹⁰ 见 Jasper Jackson and others, “Facebook accused by survivors of letting activists incite ethnic massacres and misinformation in Ethiopia”, Bureau of Investigative Journalism, 20 February 2022, 可查阅 <https://www.thebureauinvestigates.com/stories/2022-02-20/facebook-accused-of-letting-activists-ignite-ethnic-massacres-with-hate-and-misinformation-by-survivors-in-ethiopia>。

¹¹ 见 SC/14939。

¹² 见促进和保护意见和表达自由权特别报告员及其他人，《政治家和公职人员表达自由问题联合宣言》，可查阅 https://www.ohchr.org/sites/default/files/2022-04/Joint-Declaration-2021-Politicians_EN.pdf。

¹³ 见 <https://www.un.org/en/hate-speech/impact-and-prevention/targets-of-hate>。

¹⁴ 见 Nicola Barrach-Yousefi and others, *A Lexicon of Hateful and Inflammatory Speech in the Central African Republic* (PeaceTech Lab, Washington, D.C.), 可查阅 https://static1.squarespace.com/static/54257189e4b0ac0d5fca1566/t/60edb60d3680ef421b572ff7/1626191378524/CARLexicon_English_web.pdf; 另见 <https://www.ungeneva.org/fr/news-media/news/2022/07/rca-un-expert-de-lonu-alarme-par-la-proliferation-des-fausses-informations>。

¹⁵ 见 Global Witness, “Now is the time to kill: Facebook continues to approve hate speech inciting violence and genocide during civil war in Ethiopia”, 9 June 2022, 可查阅 <https://www.globalwitness.org/en/campaigns/digital-threats/ethiopia-hate-speech/>; 另见 Tessa Knight and Beth Alexion, “Influential Ethiopian social media accounts stoke violence along ethnic lines”,

YouTube 传播仇恨和宗派分裂，迫使许多目标人群，特别是妇女，东躲西藏或逃离家园。¹⁶

20. 歌手 Hachalu Hundessa 被枪杀后的暴力和杀戮浪潮只是网上仇恨如何转化为线下暴行的一个例子。¹⁷ 在南苏丹，有人在网上使用贬损性语言，污蔑丁卡牧民部落威胁领土和族裔完整，并煽动战斗人员和武装团体寻找、攻击和杀害他们。¹⁸

21. 平民人口。关于敌对行动的地点和性质、部队或人口的流动或者安全区的存在和准入情况的不实信息可能导致人们作出错误和危险的决定。在许多情况下，平民发现自己不仅受到虚假信息的影响，而且还受到当局实行的信息封锁和切断互联网的影响。¹⁹ 当家庭无法相互沟通时，当人们无法获得关于安全局势或基本服务或人道主义援助的可靠信息时，他们就无法评估其安全保障面临的风险并作出相应的决定。埃塞俄比亚的平民称，由于缺乏可靠的信息，他们的生活“天翻地覆”。²⁰ 乌克兰马里乌波尔平民试图逃离战争的案例突显出关于人道主义走廊的可靠信息的重要性。²¹

DFRLab, 17 December 2021, 可查阅 <https://medium.com/dfrlab/influential-ethiopian-social-media-accounts-stoke-violence-along-ethnic-lines-6713a1920b02>。

¹⁶ 见 Joey Shea and Ruba al-Hassani, “Hate speech, social media and political violence in Iraq: Virtual civil society and upheaval”, The Tahrir Institute for Middle East Policy, 11 February 2021, 可查阅 <https://timep.org/commentary/analysis/hate-speech-social-media-and-political-violence-in-iraq-virtual-civil-society-and-upheaval/>; 另见 Pshtiwan Faraj and Emilie Wilson, “Deeply ingrained prejudice fuels hate speech in Iraq”, Institute of Development Studies, 9 January 2021, 可查阅 <https://www.ids.ac.uk/opinions/deeply-ingrained-prejudice-fuels-hate-speech-in-iraq/>。

¹⁷ 见联合国新闻稿, “联合国专家呼吁埃塞俄比亚允许和平抗议, 欢迎部分恢复互联网”, 2020 年 7 月 21 日, 可查阅 <https://www.ohchr.org/en/press-releases/2020/07/un-experts-call-ethiopia-allow-peaceful-protests-welcome-partial-restoration?LangID=E&NewsID=26115>。

¹⁸ 见 Dangerous Speech Project, “Dinka called ‘MTN’ in South Sudan”, 2 January 2022, 可查阅 <https://dangerousspeech.org/dinka-called-mtn-in-south-sudan/>; 另见 Community Power for Progress Organization, “Social networks ignite the war that puts the country on the brink of genocide”, 15 May 2017, 可查阅 <http://cepo-southsudan.org/news/social-networks-ignite-war-puts-country-brink-genocide>。

¹⁹ 大众媒体维护中心、纪念人权维护中心、网络自由项目和 OVD-Info 联合提交的材料。

²⁰ 见《埃塞俄比亚人权委员会/联合国人权事务高级专员办事处关于埃塞俄比亚提格雷州冲突各方据称违反国际人权和人道法及难民法行为的联合调查报告》，可查阅 <https://digitallibrary.un.org/record/3947207?ln=en>。

²¹ Deutsche Welle, “Ukraine says planned Mariupol evacuations fell short – as it happened”, April 2022.

22. 研究表明，虚假信息对妇女、²² 儿童²³ 和男女同性恋、双性恋、跨性别者、性别奇异者和间性者等人士²⁴ 的影响格外严重。在冲突环境中，往往缺乏支持，他们的处境只会更加恶化。

23. 对局势的了解失真会大大增加焦虑、恐惧和压力，再加上遭受暴力和暴行造成的创伤，会对心理健康造成长期后果。²⁵ 通过不断激发愤怒和愤慨的情绪，虚假信息还可以煽动激进形式的怨恨、极端意见和诉诸暴力。

24. 人权与人道主义行为体。冲突各方或其盟友利用虚假信息和宣传抹黑人权维护者和人道主义行为体，破坏人道主义准入和援助。例如，在乌克兰、²⁶ 阿拉伯叙利亚共和国²⁷ 和巴勒斯坦国，²⁸ 精心策划的虚假信息宣传运动散布毫无根据的指控，指控这些组织偏袒一方、从事犯罪活动或与武装团体有联系。

25. 这些虚假信息会影响人们对这些组织的看法，并诱使他们拒绝与其接触或接受其服务。抹黑行动可能影响捐助方的供资，并造成安全风险，包括对女性工作人员实施性别暴力的风险，影响人道主义工作者的存在、准入和向弱势群体提供援助的能力。²⁹ 在某些情况下，例如在埃塞俄比亚，当局限制或阻止人

²² EU DisinfoLab, “Gender-based disinformation: Advancing our understanding and response”, 20 October 2021, 可查阅 <https://www.disinfo.eu/publications/gender-based-disinformation-advancing-our-understanding-and-response/>; 另见 Lucina Di Meco and Kristina Wilfore, “Gendered disinformation is a national security problem”, Brookings Institution, 8 March 2021, 可查阅 <https://www.brookings.edu/techstream/gendered-disinformation-is-a-national-security-problem/>。

²³ 见联合国儿童基金会(儿基会)全球洞察和政策办公室, “数字错误信息/虚假信息与儿童”, 2021年8月, 可查阅 <https://www.unicef.org/globalinsight/media/2096/file/UNICEF-Global-Insight-Digital-Mis-Disinformation-and-Children-2021.pdf>。

²⁴ 见 European Parliament, Directorate-General for External Policies, Policy Department, “Disinformation campaigns about LGBTI+ people in the EU and foreign influence”, July 2021, 可查阅 [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/653644/EXPO_BRI\(2021\)653644_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/653644/EXPO_BRI(2021)653644_EN.pdf)。

²⁵ 见 Eian Katz, “Liar’s war: Protecting civilians from disinformation during armed conflict” in *International Review of the Red Cross*, No. 914, December 2021, 可查阅 <https://international-review.icrc.org/articles/protecting-civilians-from-disinformation-during-armed-conflict-914>。

²⁶ 见红十字国际委员会(红十字委员会), “乌克兰: 应对关于红十字委员会活动的错误信息”, 2022年3月26日, 可查阅 <https://www.icrc.org/en/document/ukraine-addressing-misinformation-about-icrcs-activities>。

²⁷ 见 Mel Bunce, “Humanitarian communication in a post-truth world” in *Journal of Humanitarian Affairs*, Manchester openhive, Vol. 1, No. 1, 1 January 2019, 可查阅 <https://www.manchesteropenhive.com/view/journals/jha/1/1/article-p49.xml>; 另见 Louisa Loveluck, “Russian disinformation campaign targets Syria’s beleaguered rescue workers” in *The Washington Post*, 18 December 2018, 可查阅 https://www.washingtonpost.com/world/russian-disinformation-campaign-targets-syrias-beleaguered-rescue-workers/2018/12/18/113b03c4-02a9-11e9-8186-4ec26a485713_story.html。

²⁸ 慈善与安全网提交的材料(2021年征集材料时提交)。

²⁹ 见 Rachel Xu, “你无法处理真相: 错误信息与人道主义行动”, 红十字委员会人道法和政策博客, 2021年1月15日, 可查阅 <https://blogs.icrc.org/law-and-policy/2021/01/15/misinformation-humanitarian/>。

道主义组织的工作，在它们的活动或通信不完全符合政府利益时指责它们散布虚假信息。³⁰

26. **公众信任。**虚假信息造成公众对信息真实性的不信任，进而产生广泛的社会和政治影响，阻碍和平、民主、和解与重建。³¹ 当人们不能信任信息来源时，社区就不可能建立对事实的共同理解，没有这种理解，就不可能有进行建设性交流以促成调解与和解的共同基础。在纳戈尔诺-卡拉巴赫³² 或巴勒斯坦国³³ 等长期冻结的冲突局势中，冲突地区内外各方的虚假信息、宣传和歪曲的表述几十年来一直阻碍着冲突的解决与和平进程。

C. 操纵信息的媒介

27. 虚假信息、宣传和仇恨言论并非武装冲突所特有。它们也在其他时候被加以利用，以无定形的方式在武装冲突之前或之后的紧张局势和动乱各个阶段和周期中传播。冲突的根本原因，即历史冤情、系统性不平等、歧视、社区间和族裔间的竞争、政治紧张局势和治理不善，为它们提供了一个完美的滋生地。分裂、两极分化和非人化的动态是暴力和冲突的特点，这样的动态与这种对信息的操纵相互推动、共同延续。

28. 改变战争以及和平时期虚假信息状况的利器是数字技术。在埃塞俄比亚，估计有 70% 的虚假信息是通过社交媒体传播的。³⁴

29. 技术创新使人们有可能创造出“深度造假”图像、视频和文本，令人信服地歪曲现实，而数字和社交媒体平台的商业模式和技术有助于它们大规模快速传播。³⁵ 精准投放技术、算法推荐和网络效应会助长那些激发惊讶、愤怒、厌

³⁰ 见 Kaamil Ahmed, “Ethiopia suspends aid groups for ‘spreading misinformation’” in *The Guardian*, 6 August 2021, 可查阅 <https://www.theguardian.com/global-development/2021/aug/06/ethiopia-suspends-aid-groups-for-spreading-misinformation>。

³¹ 见 Mercy Corps, “Strengthening social cohesion for violence prevention: Ten lessons for policymakers and practitioners”, March 2022, 可查阅 https://www.mercycorps.org/sites/default/files/2022-06/10-Lessons-SC-Brief_V6_EU.pdf; 另见 T.M. Sagharian-Dickey, “The importance of trust in achieving positive peace” in *The Palgrave Handbook of Positive Peace*, 31 July 2021, 可查阅 https://link.springer.com/referenceworkentry/10.1007/978-981-15-3877-3_52-1#citeas。

³² 见 European Resources for Mediation Support, “Media and disinformation in the Nagorno-Karabakh conflict and their role in conflict resolution and peacebuilding”, January 2021, 可查阅 https://www2.coleurope.eu/system/tdf/uploads/news/event_report_-_media_and_disinformation_in_the_nagorno-karabakh_conflict.pdf?&file=1&type=node&id=draft&force=。

³³ 7amleh 阿拉伯社交媒体促进中心提交的材料；另见 Sheera Frankel, “Lies on social media inflame Israeli-Palestinian conflict” in *The New York Times*, 14 May 2021, 可查阅 <https://www.nytimes.com/2021/05/14/technology/israel-palestine-misinformation-lies-social-media.html>。

³⁴ 见 European Institute of Peace, “Fake news misinformation and hate speech in Ethiopia: A vulnerability assessment”, 12 April 2021, 可查阅 <https://www.eip.org/wp-content/uploads/2021/04/Fake-News-Misinformation-and-Hate-Speech-in-Ethiopia.pdf>。

³⁵ 见 Dan Boneh and others, “Preparing for the age of deepfakes and disinformation”, Stanford University Human-Centred Artificial Intelligence, November 2020, 可查阅 https://hai.stanford.edu/sites/default/files/2020-11/HAI_Deepfakes_PolicyBrief_Nov20.pdf; 另见 Ben Buchanan and

恶或愤慨等认知和情感偏见以吸引并维持用户注意力的信息形式和内容。³⁶ 它们还强化回声室效应并使受众两极分化，沿政治和宗派路线站队。

30. 尽管社交媒体的使用有所增加，但对冲突地区的大多数人来说，传统媒体仍然是最常见的新闻来源。公益新闻是打击被操纵的信息的一个重要工具，而一些媒体，特别是国家控制的媒体，是传播虚假信息、宣传和仇恨言论的媒介。³⁷ 如果独立媒体依赖政府官员作为其唯一来源，或不愿质疑官方声明，它们也有可能扩散被操纵的信息。

31. 在线和离线环境的相互作用值得注意。例如，在也门，由于传统媒体在冲突中有倾向，也门年轻人越来越多地依赖社交媒体来获取新闻，但却发现自己受到虚假新闻的影响。³⁸ 在缅甸，各平台加强审查和内容审核，导致当局利用小册子、传单、信件和当地报纸，通过线下宣传补充在线活动。³⁹ 在这两种情况下，网上仇恨的后果都表现为线下暴力。

32. 虚假信息的增长不能仅仅归因于数字技术或冲突的根源。其媒介是多方面的，包括国家和非国家行为体、政党、武装团体以及由被买通的网络水军和公关公司支持的企业。在本报告中，特别报告员主要侧重于国家和社交媒体公司的作用和责任。

三. 法律现状摸底

A. 同时适用国际人权法和人道法

33. 国际人权法和国际人道法同时适用于武装冲突，这一点已得到公认。⁴⁰ 国际法院指出，“有些权利可能完全是国际人道法问题，而有些权利可能完全是人权法问题，还有些权利可能是涉及国际法这两个分支的问题”，⁴¹ 人权事务委员会则认为，这两种制度是“互补而不是相互排斥的”。⁴²

others, “Truth, lies and automation: How language models could change disinformation”, Center for Security and Emerging Technology, May 2021, 可查阅 <https://cset.georgetown.edu/wp-content/uploads/CSET-Truth-Lies-and-Automation.pdf>。

³⁶ 见 Matthew Shaer, “What emotion goes viral the fastest?” in *Smithsonian Magazine*, April 2014, 可查阅 <https://www.smithsonianmag.com/science-nature/what-emotion-goes-viral-fastest-180950182/>。

³⁷ “第十九条”组织提交的材料。

³⁸ 见 Ark, “Fake news and disinformation in Yemen’s conflict”, 5 June 2021, 可查阅 <https://www.ark.international/ark-blog/fake-news-and-disinformation-in-yemens-conflict>。

³⁹ 见 Andrew Nachemson and Frontier Myanmar, “Military disinformation moves offline amid Internet restrictions”, 28 January 2021, 可查阅 <https://www.frontiermyanmar.net/en/military-disinformation-moves-offline-amid-internet-restrictions/>; 另见缅甸表达自由组织提交的资料。

⁴⁰ 见国际法院关于以核武器进行威胁或使用核武器的合法性的咨询意见，1996年7月8日。

⁴¹ 见国际法院关于在巴勒斯坦被占领土修建隔离墙的法律后果的咨询意见，2004年7月9日。

⁴² CCPR/C/21/Rev.1/Add.13, 第11段。

34. 国际人权法与国际人道法一道适用，对于在冲突期间有效保护意见和表达自由权至关重要。国际人道法只有在武装冲突开始时才被触发，主要涉及军事行动的实施以及在国际和非国际冲突中对某些类别人员的保护。因此，它“只是微弱和非系统地”涉及表达自由和获取信息的问题。⁴³ 人权原则和标准可以在国际人道法未作规定、缺失或不明确的情况下提供明确性和保护。

35. 这两种法律制度具有相辅相成的性质，为在数字时代面临新出现的复杂挑战时维护意见和表达自由提供了重要的可能性。例如，国际人道法只适用于武装冲突的当事方，而人权义务则适用于参与操纵信息的更广泛的行为体。此外，国际人道法所允许的某些类型的操纵行为可根据人权制度加以限制。因此，两种制度的同时适用为校准方法提供了空间。⁴⁴ 另一方面，国际人道法比国际人权法在战时针对某些具体威胁提供了更大的保护。

B. 国际人权法下的信息操纵

36. 《世界人权宣言》第十九条和《公民及政治权利国际公约》第十九条均规定了意见和表达自由权。意见自由是一项绝对权利，即使在武装冲突期间也不受限制、约束或克减。⁴⁵

37. 表达自由包括寻求、接受及传播各种信息及思想的自由，不论其是真是假、是令人愤怒还是发人深省，也不论其疆界和所选择的任何媒体。只有为保护他人权利和名誉、国家安全、公共秩序、公共卫生或公共道德，通过合法和绝对必要的措施，才能限制这种自由。⁴⁶ 合法性原则要求通过明确、准确和公开的法律作出限制，并且不会赋予官员不适当的酌处权。必要性原则要求实施的限制应当对于达到《公约》第十九条第三款的合理目的而言是相称的，而且应作狭义解释。

38. 国际人权法不能禁止虚假信息，除非这种信息等同于构成煽动敌意、暴力和歧视的鼓吹仇恨言论。只有在符合《公约》规定的合法性、必要性和合理目的的要求时，才能对其加以限制。捏造或操纵信息本身并不是限制表达自由的充分理由。在大多数情况下，消除虚假信息的最佳办法不是法律限制，而是各种可核查的信息源的自由流动，包括通过独立、自由和多元化的媒体、值得信赖的公共信息以及媒体和数字扫盲。⁴⁷

39. 表达自由像保护其他言论一样保护宣传。宣传可能在与上述虚假信息相同的条件下受到限制。但是，战争宣传必须禁止。⁴⁸ 这项禁令被理解为只适用于违反《联合国宪章》的侵略或破坏和平行为，并限于煽动战争而不是战争期间

⁴³ 见 <https://www.geneva-academy.ch/joomlatools-files/docman-files/working-papers/Protecting%20the%20Global%20inupholdinformation%20space%20in%20times%20of%20armed%20conflict.pdf>。

⁴⁴ 见 Eian Katz, “Liar’s war”。

⁴⁵ CCPR/C/GC/34。

⁴⁶ 《公民及政治权利国际公约》第十九条第三款。

⁴⁷ A/HRC/47/25 和 A/HRC/49/L.31/Rev.1。

⁴⁸ 《公民及政治权利国际公约》第二十条第一款。

的宣传。⁴⁹ 将“战争”解释为侵略，就排除了滥用这一条款镇压国内动乱的可能性，而将其限于煽动侵略，则使受到攻击的国家能够为自卫争取支持。然而，一些国家和公司对其范围感到困惑，这突出表明需要进一步澄清。⁵⁰

40. 国际法禁止构成煽动敌意、暴力和歧视的鼓吹仇恨言论，但不要求各国将其定为刑事犯罪。人权理事会核可的《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》提供了根据六个要素评估刑事定罪适当性的指导，包括社会背景、发言者的身份和意图、言论的内容和形式、受众的性质、传播的影响范围和危害的紧迫性。⁵¹ 这些准则同样适用于冲突和其他情况。

41. 《消除一切形式种族歧视国际公约》也禁止煽动歧视、敌意和暴力，只有在“严重情况”下才需要刑事定罪。⁵²

42. 国际法允许国家在发生“危及国本”的紧急情况期间，克减某些权利，包括表达自由。⁵³ 虽然克减给予国家更大的余地来限制表达，但这并不意味着可以无限中止这项权利。克减措施必须有时间限制、相称，而且仅限于“在此种危急情势绝对必要之限度内”。这些措施不应具有歧视性，不应与该国的其他国际义务相抵触，也不应违反国际法强制性规范。此外，克减不允许采取旨在“破坏”权利本身的任何行动。⁵⁴ 因此，即使在克减的情况下，表达自由权也享有一定程度的保护。

43. 鉴于操纵意见自由的严重性，国家和社交媒体公司应给予更多关注。无论在战争还是平时时期，意见自由都受到国际人权法的绝对保护。国家或非国家行为体以强迫、非自愿或未经同意的方式操控思考过程，如灌输或“洗脑”，都侵犯意见自由。借助强大的平台推荐或精准投放进行内容集展是加强虚假信息 and 加剧政治紧张局势的关键因素，是以数字化形式对用户内心深处的思考过程进行未经同意的操纵。因此，这等于侵犯意见自由权。⁵⁵

⁴⁹ 见 Andrei Richter, “The Relationship between Freedom of Expression and the Ban on Propaganda for War in *European Yearbook on Human Rights 2015*”; 另见“第十九条”组织提交的材料。

⁵⁰ 法律与民主中心和 Meta 公司提交的材料。

⁵¹ [A/HRC/22/17/Add.4](#), 附录。

⁵² 第四条。

⁵³ 《公民及政治权利国际公约》第四条。

⁵⁴ 《公民及政治权利国际公约》第五第一款。

⁵⁵ 见 Evelyn Aswad, “Losing the freedom to be human” in *Columbia Human Rights Law Review*, 29 February 2020)。另见 [A/HRC/47/25](#), 第 33-36 段。

C. 国际人道法下的信息操纵

44. 国际人道法据称对武装冲突期间的信息操纵采取了“一种非常宽大的做法”。⁵⁶ 虚假信息运动，包括为了削弱敌方抵抗意志而使用的诈术或操纵信息、托词和其他形式的欺骗和宣传，被交战方广泛使用，根据国际人道法不属于非法。

45. 国际人道法对操纵信息有一些限制。⁵⁷ 背信弃义(旨在诱使一方给予对手国际人道法的保护，使后者能够杀害、伤害或俘虏的误导行为)受到禁止。信息行动所产生的某些有害后果也受到禁止，例如威胁使用暴力或攻击从而在平民人口中散布恐怖，鼓励违反国际人道法或煽动实施战争罪，以及威胁或命令不宽恕或攻击平民。禁止对未参加敌对行动的人实施一切形式的不人道待遇、凌辱个人尊严或侮辱性或有辱人格的待遇。⁵⁸ 这些规则适用于通过任何手段，包括作为媒体或社交媒体上煽动、虚假信息和宣传运动的一部分采取的行动。

46. 一些专家质疑，鉴于利用数字技术和社交媒体针对平民而非军方的虚假信息运动的性质和影响，这些限制是否足够。⁵⁹ 这是一个值得联合国条约机构、各国和国际组织认真考虑的合理关切。

D. 保护记者

47. 新闻媒体不受审查和不受阻碍以及记者⁶⁰ 安全地工作免于恐惧的权利不仅是意见和表达自由权的组成部分，而且也是打击虚假信息，包括在冲突环境中打击虚假信息的关键。国际人权法保护自由、独立和多元化的新闻工作，以及记者的言论自由权，⁶¹ 而国际人道法对这一问题未作规定。

48. 国际人道法将记者作为平民加以保护。蓄意杀害记者是战争罪。⁶² 新闻工作者进行宣传，即使这种活动支持战争举措，也不能使针对新闻工作者或媒体设施

⁵⁶ 见 Eian Katz, “Liar’s war”。

⁵⁷ 见 Robin Geiss and Henning Lahmann, “Protecting the global information space in times of armed conflict, Geneva Academy, February 2021, 可查阅 <https://www.geneva-academy.ch/joomlatools-files/docman-files/working-papers/Protecting%20the%20Global%20information%20space%20in%20times%20of%20armed%20conflict.pdf>. 另见 Dapo Akande, “Oxford statement on international law protections in cyberspace: The regulation of information operations and activities”, Just Security, 2 June 2021, 可查阅 <https://www.justsecurity.org/76742/oxford-statement-on-international-law-protections-in-the-regulation-of-information-operations-and-activities/>。

⁵⁸ 1949年8月12日日内瓦四公约，共同第3条。

⁵⁹ 见 Geiss and Lahmann, “Protecting the global information space” 以及 Eian Katz, “Liar’s war”。

⁶⁰ “记者”一词包括专业记者、分析员、媒体工作者、博主和其他从事新闻工作的人：见 [A/HRC/50/29](#)，第15-16段。

⁶¹ [A/HRC/50/29](#)。

⁶² 《国际刑事法院罗马规约》第八条第二款第1项。

的行为合法化。⁶³ 记者或媒体机构如果直接参与敌对行动或煽动战争罪或其他国际罪行，就可能成为合法的军事目标。⁶⁴

49. 在冲突环境中记者的保护和安全方面面临的挑战不是法律漏洞，而是缺乏政治意愿，以及冲突国家和其他各方不遵守国际法。

E. 对人权的域外责任

50. 数字技术使人们能够远程操纵信息，并跨境影响人权。然而，国际人权法的域外适用是一个复杂和有争议的问题，引起对主权和安全以及人权的关切。国际人权法和国际人道法似乎都没有明确回答一个棘手的问题，即各国在其不是当事方的冲突中从境外输入、传播或支持宣传、虚假信息或煽动言论，应负有何种责任。

51. 《公民及政治权利国际公约》第二条第一款要求缔约国尊重并确保其境内和受其管辖的所有人的《公约》所赋权利。人权事务委员会对这一规定的解释是，既包括在有关国家境内的人，也包括在有关国家境外但在其有效控制下的人。⁶⁵ 另一方面，人权事务委员会在一个涉及移交人身监护权的案件中还采取了以下立场，即如果一个国家是导致侵犯人权行为发生的因果链中的一个环节，那么它就可能对域外侵犯人权行为负责。⁶⁶

52. 这一决定表明，不仅应考虑对个人或其所处的领土实行有效控制的权力，而且应考虑对其人权实行有效控制的权力。在利用数字技术远程实施侵犯人权行为的情况下，这种办法提供了一种可能的问责途径。⁶⁷

四. 国家应对措施：关切、挑战和良好做法

53. 大会和人权理事会都已申明，应对虚假信息必须以人权为基础。⁶⁸ 各国的应对措施多种多样，既有根据国际人权法促进信息自由流动的政策和做法，也有利用有损于人权的措施打击虚假信息的行动，还有国家本身涉及制造、支持和传播虚假信息、宣传和仇恨言论的一些情况。

⁶³ 见尼尔斯·梅尔策，《国际人道法中直接参加敌对行动定义的解释性指南》（红十字委员会，日内瓦，2009年5月），可查阅 <https://www.icrc.org/en/doc/assets/files/other/icrc-002-0990.pdf>。另见法律和民主中心提交的材料。

⁶⁴ 见前南斯拉夫问题国际刑事法庭，审查北约轰炸南斯拉夫联盟共和国行动委员会提交检察官的最后报告，可查阅 <https://www.icty.org/en/press/final-report-prosecutor-committee-established-review-nato-bombing-campaign-against-federal>。

⁶⁵ CCPR/C//21/Rev.1/Add.13，第10段。

⁶⁶ CCPR/96/D/1539/2006。

⁶⁷ 见 Ido Kilovaty, “An Extraterritorial Human Right to Cybersecurity” in *Notre Dame Journal of International and Comparative Law*, vol. 10, No. 1。

⁶⁸ 大会第 76/227 号决议和 A/HRC/49/L.31/Rev.1。

A. 促进获取信息

54. 各国通过采取措施改善获得各种可核查信息来源的机会，打击虚假信息，这方面有许多良好做法的例子。其中包括关于获取信息、施政透明度、真实可信的公共信息以及促进独立、自由、多元化和多样化媒体的健全法律和政策。⁶⁹

55. 一些国家政府，特别是丹麦、芬兰、挪威和瑞典政府，在数字、媒体与信息素养方面进行了大量投资，以建立社会抵御虚假信息的能力。⁷⁰

56. 在欧洲联盟一级，已启动各种立法和政策措施，支持成员国应对虚假信息的威胁，包括《欧洲民主行动计划》、《反虚假信息行为守则》加强版和《数字服务法》规定的共同监管框架。欧洲联盟还在资助各种方案，支持民间社会反对虚假信息的举措，例如媒体和事实核对项目。⁷¹

57. 支持应对危机的举措很多。其中一个例子是，为独立记者和人权维护者从受冲突影响或镇压局势中撤离作出签证安排，以便他们能够继续从国外进行报道、监测和核对事实。

B. 国家支持的虚假信息和宣传

58. 在某种程度上，所有国家，不论其政治或意识形态色彩如何，都制作和传播宣传品。如上所述，根据国际法，国家宣传本身并不违法。当宣传战争或国家使用虚假信息，致使人们无法区分事实与假象，可能造成社会伤害或侵犯人权时，就会引发关切。在某些情况下，国家支持的宣传或虚假信息可构成煽动战争罪，如在缅甸发生的情况。⁷²

59. 国家主导或支持的虚假信息对人权、法治、民主进程、国家主权和地缘政治稳定具有强大的影响，因为国家拥有资源和影响力，也因为它们有能力同时压制国内的各种独立和批评声音，从而使官方说法无法受到质疑。⁷³ 在一份联合声明中，特别报告员和区域组织的相应任务负责人表示严重关切在俄罗斯入侵乌克兰之后，俄罗斯国有媒体关于乌克兰冲突发布虚假信息，以及俄罗斯联邦境内表达自由被削弱，媒体审查制度进一步收紧，多元化信息来源被封锁，批评声音受到压制。⁷⁴

⁶⁹ A/HRC/50/29。

⁷⁰ 见 Michael Forsman, “Media literacy and the emerging media citizen in the Nordic media welfare State” in *Nordic Journal of Media Studies*, 6 June 2020, 可查阅 <https://www.sciendo.com/article/10.2478/njms-2020-0006>。

⁷¹ 欧洲联盟提交的材料。

⁷² A/HRC/39/CRP.2；另见缅甸表达自由组织提交的资料。

⁷³ 朝鲜民主主义人民共和国是“无处不在”的宣传的极端例子，国家媒体受到严格控制，没有独立或外部媒体。见 A/HRC/25/CRP.1，第 187、197-221 段。

⁷⁴ 见非洲人权和民族权委员会、美洲人权委员会和欧洲安全与合作组织，关于乌克兰局势的联合声明，2022 年 5 月 3 日，可查阅 <https://srfreedex.org/joint-statement-on-the-situation-in-ukraine/>；另见纪念人权维护中心、大众媒体维护中心、网络自由项目和 OVD-Info 联合提交的材料。

60. 数字技术使一些国家及其代理人以及非国家行为体更容易跨境干预冲突，并以难以溯源和追责的方式传播虚假信息。在利比亚、马里、阿拉伯叙利亚共和国和也门等国的冲突中，曾有关于此类“信息行动”的指控的报道，引起了激烈的争论。⁷⁵

C. 对记者和人权维护者的攻击

61. 独立、自由和多元化的媒体通过核对事实和提供多样、可核实的信息，在打击虚假信息和国家宣传方面发挥着关键作用。正因为如此，在许多受冲突影响的国家或周边国家媒体承受严重压力的情况令人担忧。采取的措施包括驱逐外国媒体、关闭当地新闻机构、以“虚假新闻”或有悖于国际法律标准的国家安全法为由进行起诉，以及攻击记者。

62. 马里以传播军队侵犯人权的“虚假指控”为由封禁了法国媒体频道。⁷⁶ 吉尔吉斯斯坦以战争宣传罪为由起诉了一家全国性独立媒体，因为该媒体传播了一篇关于与塔吉克斯坦边界冲突的文章，与官方版本不符。⁷⁷ 数字运动是揭露叙利亚战争中虐待行为的关键。2022年初，阿拉伯叙利亚共和国将“破坏国家威望或损害民族团结的……假新闻”定为犯罪，并逮捕了与境外社交媒体网站交流的人。⁷⁸ 俄罗斯联邦通过了一项法律，规定发布任何与官方版本不同的关于乌克兰战争的新闻均为严重刑事犯罪。全面的信息封锁迫使俄罗斯独立媒体暂停活动或关闭，西方媒体离开或被俄罗斯当局阻止报道。⁷⁹

63. 国家安全和反恐法常常被用来压制批评声音，包括压制记者、人权维护者和政治反对派。其中很多法律并不符合《公民及政治权利国际公约》第十九条第三款关于合法性、必要性及合理目的的三重检验标准。例如，缅甸军政府扩大了其

⁷⁵ 见 Léa Ronzaud, Ira Hubert and Ben Nimmo “Capture the flag: Iranian operators impersonate anti-Netanyahu ‘black flag’ protestors, amplify Iranian narratives”, Graphika, 6 November 2020, 可查阅 <https://graphika.com/reports/capture-the-flag>; Africa Center for Strategic Studies, “A light in Libya’s fog of disinformation”, 9 October 2020, 可查阅 <https://africacenter.org/spotlight/light-libya-fog-disinformation/>; Graphika and The Stanford Internet Observatory, “More Troll-Kombat: French and Russian influence operations go head to head targeting audiences in Africa”, 2020, 可查阅 <https://graphika.com/reports/more-troll-kombat>; Institute for Strategic Dialogue, “Digital investigation on Syria’s disinformation” in *Deadly Disinformation: How Online Conspiracies about Syria Cause Real-World Harm* (13 July 2022), 可查阅 https://www.isdglobal.org/digital_dispatches/isds-digital-investigation-on-syria-disinformation/。

⁷⁶ 人权高专办，关于“对马里独立媒体被关闭后的关切”的声明，2022年4月29日，可查阅 <https://www.ohchr.org/en/press-briefing-notes/2022/04/concerns-independent-media-mali-after-shutdowns>。

⁷⁷ 国际言论自由交流网提交的材料。

⁷⁸ 见 Mark Frary, “Syria passes draconian cybercrime laws” in *Index on Censorship*, 6 May 2022, 可查阅 <https://www.indexoncensorship.org/2022/05/syria-passes-draconian-cybercrime-laws/>。

⁷⁹ 人权高专办关于“俄罗斯：联合国专家对‘令人窒息的’信息压制感到震惊”的新闻简报，2022年3月12日，可查阅 <https://www.ohchr.org/en/press-releases/2022/03/russia-un-experts-alarmed-cho-king-information-clampdown>。

刑法，不仅包括“假新闻”，还包括新的国家安全条款，将造成“对军队和政府的仇恨、不服从或不忠诚”定为犯罪。⁸⁰

64. 封禁媒体是对表达自由的严重限制，很少有正当理由。欧洲联盟委员会封禁了几家俄罗斯国有媒体，理由是它们散布虚假信息和宣传，对公共秩序和安全构成威胁。在一个独立媒体和事实核对人员能够质疑虚假信息、本可考虑采取其他不那么激烈的措施的地区，这项禁令的必要性和相称性受到了质疑。⁸¹

65. 记者发从前线的报道在揭露虚假信息方面发挥着关键作用，但他们自己也面临着恐吓、骚扰、绑架、暴力和因工作而被杀害的高风险。安全理事会谴责对记者和媒体工作者的袭击，并呼吁所有各方停止这种做法。⁸²

66. 根据国际法，有针对性地杀害记者是一种战争罪，然而，在 10 起案件中，有 9 起案件有罪不罚，因为各国缺乏调查和起诉的政治意愿。⁸³ 最近的一个例子是，2022 年 5 月 11 日，资深巴勒斯坦裔美国记者 Shireen Abu Akleh 在报道以色列安全部队在巴勒斯坦被占领土的一次行动时被枪杀。尽管包括特别程序任务负责人、人权事务高级专员和秘书长在内的各方多次呼吁，但以色列没有启动刑事调查或支持独立调查。⁸⁴

D. 社交媒体监管

67. 一些国家颁布了关于媒介机构责任和在线内容监管的法律，对社交媒体公司规定了过于宽泛的义务，要求其监测和删除用户生成的内容，包括虚假信息。⁸⁵ 例如，中国对社交媒体实施全面的内容控制，禁止许多外国平台，禁止批评政府、共产党和被视为不受欢迎的宗教或社会问题，要求在线博主和影响者取得许可证，最近还提出，平台必须在发布所有评论之前进行审查。⁸⁶

⁸⁰ 见 Shawn W. Crispin, “Bitter reversal: Myanmar military coup wipes out press freedom gains”, Committee to Protect Journalists, 28 July 2021, 可查阅 <https://cpj.org/reports/2021/07/bitter-reversal-myanmar-journalists-jailed-imprisoned-military-crackdown/>。

⁸¹ A/HRC/50/29, 第 62 段。然而，欧洲联盟法院维持了广播禁令：见 *RT France 诉 Council* 案，可查阅 <https://curia.europa.eu/juris/documents.jsf?num=T-125/22>。

⁸² 安全理事会第 2222(2015)号决议。

⁸³ 联合国教育、科学及文化组织关于“打击针对记者的暴力威胁和犯罪以保护所有人表达自由”的概念说明，2021 年 11 月 2 日，可查阅 https://en.unesco.org/sites/default/files/concept_note_-_idei_2021_en.pdf。

⁸⁴ AL ISR 14/2022 号来文。

⁸⁵ A/HRC/38/35, 第三.A 节，以及 A/HRC/47/25, 第 56–58 段。

⁸⁶ 见 <https://freedomhouse.org/country/china/freedom-net/2021>; 另见 Zeyi Yang, “Now China wants to censor online comments” in *MIT Technology Review*, 18 June 2022, 可查阅 <https://www.technologyreview.com/2022/06/18/1054452/china-censors-social-media-comments/>。

68. 对社交媒体的严格监管被用来限制与武装冲突有关的言论。⁸⁷ 在俄罗斯联邦，国家当局利用现有和新颁布的监管措施，迅速控制有关乌克兰战争的信息。⁸⁸ Meta 公司报告说，当局已经对脸上发布的与战争有关的内容发出了删除要求，但 Meta 没有遵守。⁸⁹ 当局最终在 2022 年 3 月屏蔽了脸书和 Instagram 以及推特。⁹⁰ 此外，谷歌在俄罗斯联邦的子公司在俄罗斯银行账户被冻结后申请破产，据报道，冻结原因与其服务中被禁止的内容、数据本地化要求以及谷歌对俄罗斯媒体机构 YouTube 频道的限制有关。⁹¹

69. 各国不应要求平台执行不符合国际人权法的条例。无论是在冲突期间还是在其他情况下，特别报告员都建议对互联网媒介机构进行“明智的监管”，以确保它们遵守人权尽责调查、有效的透明度和正当程序要求，而不是基于观点或内容进行监管。⁹²

E. 干扰互联网和电信

70. 人权理事会明确谴责切断互联网的行为，并敦促各国不要这样做。⁹³ 切断这一重要的通信手段或实行限速，只会加剧而不是打击虚假信息、宣传或煽动。

71. 使用互联网至关重要，特别是在受冲突影响的情况下，因为这可能是与外部世界沟通的唯一渠道。在遭受冲突的国家，包括埃塞俄比亚、缅甸、苏丹和阿拉伯叙利亚共和国，互联网经常被切断，这是政府控制信息流动的一种手段。⁹⁴ 断网和“限流”或减速的影响可能对人们的日常生活造成破坏性影响。这些干扰还妨碍了人权维护者和记者的监测和报道。断网与军事运动、示威和政变相关的虐待之间的关联令人不安。⁹⁵

72. 根据国际人道法，媒体设施是民用物体，即使它们正在传播支持战争的宣传。除非这些设施直接用于敌对行动，否则不得将其作为攻击目标。然而，在乌克兰、也门和加沙地带等冲突期间，电信基础设施或媒体机构的房地遭到了轰炸或炮击。⁹⁶

⁸⁷ OL RUS 4/2019 号来文。

⁸⁸ 大众媒体维护中心、纪念人权维护中心、网络自由项目和 OVD-Info 联合提交的材料以及立即普及组织提交的材料。

⁸⁹ Meta 公司提交的材料。

⁹⁰ 大众媒体维护中心、纪念人权维护中心、网络自由项目和 OVD-Info 联合提交的材料。

⁹¹ 见 Interfax, “Russian subsidiary of Google files for bankruptcy”, 17 June 2022, 可查阅 <https://interfax.com/newsroom/top-stories/80331/>。

⁹² A/HRC/47/25, 第 91 段。

⁹³ A/HRC/47/L.22, 第 11 段。

⁹⁴ 见 #KeepItOn, “The return of digital authoritarianism: Internet shutdowns in 2021”, April 2022, 可查阅 <https://www.accessnow.org/cms/assets/uploads/2022/05/2021-KIO-Report-May-24-2022.pdf>。

⁹⁵ A/HRC/50/55, 第 27 段；另见立即普及组织提交的材料。

⁹⁶ 立即普及组织提交的材料。

73. 国家的制裁范围过广，公司过于热衷的遵守，可能会干扰受制裁国和实施制裁国的网络接入和信息流动。制裁可能会产生反作用，使这些国家的公众、民间社会和人权维护者更难获得各种信息来源，伊朗伊斯兰共和国、俄罗斯联邦、苏丹和阿拉伯叙利亚共和国就出现了这样的情况，制裁也使其他地区的人更难了解这些国家正在发生的事。⁹⁷

五. 社交媒体公司：角色和职责

A. 冲突环境中的社交媒体

74. 社交媒体平台非常容易受到虚假信息、宣传和煽动的影响，因为它们依赖算法向目标用户提供特定内容和推荐，部分依据是所收集的关于这些用户的数据。来自民间社会、媒体、研究人员和国际组织的大量报告指出，社交媒体在埃塞俄比亚、⁹⁸ 缅甸、⁹⁹ 乌克兰¹⁰⁰ 和也门¹⁰¹ 传播国家宣传、极端主义内容和虚假信息方面发挥了作用。

75. 在社交媒体公司中，Meta 最受关注，因为脸书在缅甸等国加剧仇恨和暴力方面的作用，人权理事会任命的独立国际实况调查团宣布脸书是“仇恨言论的主要平台”。¹⁰² 在乌克兰战争中，Meta 建立了一个由讲俄语和乌克兰语的人组成的特别行动中心，并与当地和国际合作伙伴以及事实核对人员合作，以解决信息操纵问题，但该公司也对其仇恨言论政策实行了一项独特的“自卫表达”例外，即允许乌克兰用户表达对俄罗斯军队的抵抗和愤怒。¹⁰³

76. 信息操纵在其他平台上也很普遍。一项研究发现，YouTube 已经成为跨平台传播虚假信息和极端主义内容的一个高度货币化的基础来源。¹⁰⁴ 推特因其在一些冲突中的角色而受到批评。¹⁰⁵ TikTok 上有关乌克兰战争的内容，包括虚假信

⁹⁷ 同上。

⁹⁸ 监督委员会提交的材料(Raya Kobo 的案例)；另见 Global Witness, “Now is the time to kill”。

⁹⁹ 缅甸表达自由组织提交的材料。

¹⁰⁰ 大众媒体维护中心、纪念人权维护中心、网络自由项目和 OVD-Info 联合提交的材料。另见 Carl Miller “Who’s behind #IStandWithPutin?” in *The Atlantic*, 5 April 2022, 可查阅 <https://www.theatlantic.com/ideas/archive/2022/04/russian-propaganda-zelensky-information-war/629475/>。

¹⁰¹ 见 Hannah Porter, “A conversation on fighting disinformation in Yemen”, Yemen Policy Center, March 2022, 可查阅 <https://www.yemenpolicy.org/a-conversation-on-fighting-disinformation-in-yemen/>。

¹⁰² A/HRC/42/50, 第 72 段。

¹⁰³ 见 Meta, “Meta’s ongoing efforts regarding Russia’s invasion of Ukraine”, 26 February 2022, 可查阅 <https://about.fb.com/news/2022/02/metas-ongoing-efforts-regarding-russias-invasion-of-ukraine/>。

¹⁰⁴ 见 Paul M. Barrett and Justin Hendrix, “A platform ‘weaponized’: How YouTube spreads harmful content – and what can be done about it”, Stern Center for Business and Human Rights, June 2022, 可查阅 https://static1.squarespace.com/static/5b6df958f8370af3217d4178/t/62a38fc022745a7274601da0/1654886337000/NYU+CBHR+YouTube_Final_June10.pdf。

¹⁰⁵ 见 Torinmo Salau, “How Twitter failed Africa”, Foreign Policy, 19 January 2022, 可查阅 <https://foreignpolicy.com/2022/01/19/twitter-africa-ghana-dorsey-disinformation/>。

息大幅增加。¹⁰⁶ 微博同样注意到，不准确地列示为源自乌克兰的内容有所增加，宣布将强制在有关该冲突的帖子中自动添加地理位置。¹⁰⁷ Telegram 在言论限制方面采取不干涉的做法，在俄罗斯联邦和乌克兰，当局和公众广泛使用 Telegram 来传播有关冲突的信息，包括虚假信息和宣传。¹⁰⁸

77. 如上述例子所示，在冲突中利用社交媒体放大被操纵的信息的现象十分普遍，而且还在不断增加。各公司需要加大行动力度，防止煽动暴力和其他严重侵犯人权行为，同时确保尊重意见和表达自由。

B. 冲突期间的公司法律标准

78. 国际人权法和国际人道法都适用于武装冲突局势中的公司。正如《工商企业与人权指导原则》所述，公司有责任尊重国际公认的人权，以避免造成或加剧“负面人权影响”的方式开展业务，并预防或减轻这种影响，¹⁰⁹ 采取体现其尊重人权承诺的政策，开展人权尽责调查，并提供补救它们造成或促成的负面人权影响的程序。

79. 在冲突环境中提供信息传播手段的社交媒体或电信公司可能与武装冲突有着充分的联系，从而引发对其业务活动适用国际人道法。公司人员可能因严重违反国际人道法而被追究责任，构成战争罪，¹¹⁰ 无论是基于直接行为还是公司共谋。¹¹¹ 此外，如果他们从事可被视为直接参与敌对行动的活动，他们可能失去国际人道法规定的平民保护。¹¹² 在这种情况下，公司负有双重责任：第一，继续尊重意见和表达自由，包括信息权，第二，遵守国际人道法。

80. 工商业与人权工作组呼吁在冲突局势中(包括冲突前和冲突后)经营的公司根据加剧的风险加强尽责调查。¹¹³ 联合国已将“针对特定群体或个人的煽动性言

¹⁰⁶ 见 Sara Brown, “In Russia-Ukraine war, social media stokes ingenuity, disinformation”, MIT Sloan School of Management, 6 April 2022, 可查阅 <https://mitsloan.mit.edu/ideas-made-to-matter/russia-ukraine-war-social-media-stokes-ingenuity-disinformation>。

¹⁰⁷ 见 Weilun Soon, “How China’s tech giants, from TikTok to Tencent, are reacting to Russia’s invasion of Ukraine”, Insider, 13 April 2022, 可查阅 <https://www.businessinsider.com/how-chinas-tech-giants-reacting-to-ukraine-crisis-tiktok-tencent-2022-3?r=US&IR=T>。

¹⁰⁸ 见 Sara Brown, “In Russia-Ukraine war”。

¹⁰⁹ A/HRC/17/31, 附件。

¹¹⁰ 见《工商企业与人权指导原则》，原则 23 的评注；人权高专办，《尊重人权的企业的责任：解释性说明》，2012 年，可查阅 https://www.ohchr.org/sites/default/files/Documents/Publications/HR.PUB.12.2_En.pdf；红十字委员会，国际人道法数据库，可查阅 https://ihl-databases.icrc.org/customary-ihl/eng/docs/v1_rul_rule156。

¹¹¹ A/HRC/50/40/Add.4, 第 34 段。

¹¹² 见《1949 年 8 月 12 日日内瓦四公约关于保护国际性武装冲突受难者的附加议定书》(第一议定书)，1977 年 6 月 8 日，第五十一条第(3)款，以及人权高专办，《尊重人权的企业的责任：解释性说明》。

¹¹³ A/75/212, 第 13、19–21 和 72 段；另见《工商企业与人权指导原则》，原则 7 和 17。

论或仇恨言论增加”视为一个促使公司加强尽责调查的“危险信号”。¹¹⁴ 联合国建议各公司不仅查明和评估这些言论对人权的实际或潜在不利影响，而且还查明和评估其对冲突本身的这种影响。

81. 民间社会组织强调，必须整合人权、冲突敏感性、技术伦理和人类安全框架，以便为冲突环境中的数字公司业务制定有效的做法。它们还提议，通过一个多利益攸关方进程，以这一做法为基础制定一个示范框架。¹¹⁵

C. 公司政策

82. 乌克兰和俄罗斯联邦之间的战争以及俄罗斯联邦将独立战争报道定为刑事犯罪并禁止媒体的行动，促使社交媒体公司采取了前所未有的行动。¹¹⁶

83. 几乎所有的西方公司都已离开俄罗斯联邦或被封锁。总部设在美利坚合众国的一些公司宣布了它们为维持乌克兰的互联网接入和抵制俄罗斯联邦的审查和宣传而作出的努力，这些消息得到了广泛的报道。关于这些公司在世界其他地区冲突中采取的应对措施，或者关于总部设在西欧和北美以外的公司采取的政策和做法，人们所知甚少。不出所料，人们对各公司政策在全球的连贯一致性及其对人权承诺的深度感到关切。

84. 各公司公布的关于冲突和操纵信息的政策差别很大，有些政策还根据实地的事态发展不断演变。例如，推特发布了一项官方的“危机错误信息政策”，其中规定，公司“将对使用推特服务分享虚假或误导性信息的账户采取行动，这些信息可能会给受危机影响的人群带来伤害”。¹¹⁷ Meta 公司已经制定了对待“有风险的国家”的程序，并在公开帖子中详细说明，¹¹⁸ 据报告，Meta 正在制定一项冲突政策，此前该公司以安全问题为由，取消了让其监督委员会就冲突中的内容

¹¹⁴ 见 Gerald Pachoud and Siniša Milatović, “Heightened Human Rights Due Diligence for Business in Conflict-affected Contexts: A Guide”, United Nations Development Programme, 2022, 可查阅 https://www.undp.org/sites/g/files/zskgke326/files/2022-06/UNDP_Heightened_Human_Rights_Due_Diligence_for_Business_in_Conflict-Affected_Context.pdf。

¹¹⁵ 见 Jennifer Easterday, Hana Ivanhoe and Lisa Schirch, “Comparing guidance for tech companies in fragile and conflict-affected situations”, Policy Brief No. 125, Toda Peace Institute, March 2022, 可查阅 <https://toda.org/policy-briefs-and-resources/policy-briefs/comparing-guidance-for-tech-companies-in-fragile-and-conflict-affected-situations.html>。

¹¹⁶ 电信公司也采取了令人瞩目的措施，维护乌克兰受冲突影响地区的基础设施，并为难民提供免费通信工具，云服务也让用户随时可以使用。

¹¹⁷ Crisis information policy, Help Center, May 2022, 可查阅 <https://help.twitter.com/en/rules-and-policies/crisis-misinformation>。

¹¹⁸ Meta 公司提交的材料。

审核提供建议的邀请。¹¹⁹ 其他平台，包括 Reddit、Snapchat 和 TikTok，迄今只选择发布声明来回应特定冲突局势的具体关切。¹²⁰

85. 正如前任任务负责人所指出和现任特别报告员所重申的那样，公司应将人权标准作为默认内容纳入其服务条款、内容审核政策、规则制定和执行工作。¹²¹ 然而，除了少数例外，¹²² 各公司很少参考国际人权标准作为其政策的依据。相反，它们利用自身的诸多不同政策来制定应对冲突挑战的对策。¹²³ 这种支离破碎的做法无法为平台做法提供急需的一致性和可预测性，并有可能妨碍公司遵守国际人权法和国际人道法。

D. 公司做法

86. 前任和现任特别报告员及其他利益攸关方已经就非冲突局势提出了下文所述的许多问题。在冲突环境中，由于公司面临的风险更高，用户的脆弱性更大，因此这些问题就愈发重要。虽然已经努力改进危机应对和内容审核，但主要关切，包括对于商业模式本身的关切，基本上仍未得到解决。

87. 人权尽责。虽然有关各公司开展的与冲突有关的尽责调查的细节有限，但资源分配与问题的严重性之间似乎不匹配，而且在开展强化尽责调查方面不够及时。¹²⁴ 民间社会感到关切的是，各公司是否已制定适当的程序，查明某一特定业务环境是否存在潜在的冲突风险。¹²⁵ 自 2 月乌克兰战争爆发以来，有几家公司已经建立或加强了冲突小组，并正在制定与冲突有关的政策和程序，但关于这些小组如何运作

¹¹⁹ 见 Meta, “Meta withdraws a policy advisory opinion request related to Russia’s invasion of Ukraine”, 13 July 2022, 可查阅 <https://transparency.fb.com/oversight/oversight-board-cases/ukraine-russia-pao>; 另见 Oversight Board, “Protecting freedom of expression and human rights in Ukraine and Russia”, May 2022, 可查阅 <https://oversightboard.com/news/382264103827624-protecting-freedom-of-expression-and-human-rights-in-ukraine-and-russia/>。

¹²⁰ 见 Upvoted, “Supporting Ukraine and our community”, 2 March 2022, 可查阅 <https://www.reddit.com/blog/supporting-ukraine-and-our-community>; Team Snap, “We Support Ukraine”, 1 March 2022, 可查阅 <https://newsroom.snap.com/we-support-ukraine/>; TikTok, “Bringing more context to content on TikTok”, 4 March 2022, 可查阅 <https://newsroom.tiktok.com/en-us/bringing-more-context-to-content-on-tiktok>。

¹²¹ [A/HRC/38/35](#); [A/HRC/47/25](#), 第五节。

¹²² Meta 公司已经注意到《工商企业与人权指导原则》、国际人权法和国际人道法的适用性。见 Meta 公司提交的材料。另见 Meta, “Facebook Community Standards”, 可查阅 <https://transparency.fb.com/policies/community-standards/>, and Meta, “Corporate Human Rights Policy”, 可查阅 <https://about.fb.com/wp-content/uploads/2021/03/Facebooks-Corporate-Human-Rights-Policy.pdf>。推特已经注意到国际人道法的适用性。见 Sinéad McSweeney, “Our ongoing approach to the war in Ukraine”, 16 March 2022, 可查阅 https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine。

¹²³ 见 Meta 公司提交的材料。

¹²⁴ [A/HRC/47/25](#), 第 74-76 段。

¹²⁵ JustPeace Labs 和 “第十九条” 组织提交的材料。

以及正在应用哪些资源的具体信息很少。也不清楚是否正在建立类似的结构，以便在全球应用。

88. **内容审核。**在冲突局势中，内容审核不足会助长虚假信息、宣传和煽动，并加剧暴力。平台用户、民间社会组织 and 研究人员报告，内容审核方面存在不一致和严重失误。

89. 报告指出，各公司很难在冲突环境中准确把握删除内容的尺度，在某些情况下过度审查，在另一些情况下则不够关注或显示出偏见。¹²⁶ 民间社会组织还抱怨说，公司容易受到国家的压力，被要求删除在冲突局势中代表不同观点的内容。¹²⁷

90. 令人关切的是，各公司没有分配足够的资源和专业知 识，用于审查所有相关语言的内容并了解冲突环境中的当地情况。¹²⁸ 此外，虽然各公司将自动化和人工流程结合起来进行内容审核，但尚不清楚这些流程如何平衡，自动化流程运行效果如何以及根据哪些参数运行，或用于人工审核的本地专业知识水平如何。

91. 除了删除或屏蔽内容之外，社交媒体公司还实行了旨在调整其用户可获得的冲突相关内容的策略，例如警告、标示、事实核对、提供其他来源的建议或减少内容的传播。

92. **货币化。**有报告记录，在社交媒体平台上，虚假信息、宣传和仇恨言论被大量用于创收，从而对持续操纵信息起到了激励作用。¹²⁹ 尽管公司政策声称会限制被认为适合投放广告的内容类型，但还是出现了这种货币化现象。¹³⁰ 由于认识到这个问题，欧盟委员会 2022 年《反虚假信息行为守则》加强版纳入了专注于去货币化举措的承诺，谷歌、Meta、微软、TikTok 和推特都签署了该守则。¹³¹

93. 一些社交媒体公司已经采取行动，限制某些与冲突有关的内容的货币化，特别是与乌克兰战争有关的内容(因实施制裁而受影响)。¹³² 然而，研究表明，

¹²⁶ 立即普及组织、“第十九条”组织和监督委员会提交的材料。

¹²⁷ 7amleh 阿拉伯社会媒体促进中心和监督委员会提交的材料；另见 Access Now, “Sheikh Jarrah: Facebook and Twitter systematically silencing protests, deleting evidence”, 7 May 2021, 可查阅 <https://www.accessnow.org/sheikh-jarrah-facebook-and-twitter-systematically-silencing-protests-deleting-evidence/>。

¹²⁸ “第十九条”组织提交的材料。

¹²⁹ 见 Karen Hao, “How Facebook and Google fund global misinformation” in *MIT Technology Review*, 20 November 2021, 可查阅 <https://www.technologyreview.com/2021/11/20/1039076/facebook-google-disinformation-clickbait/>。

¹³⁰ 见 [https://support.google.com/youtube/answer/6162278?hl=en#zippy=%2Cguide-to-self-certification;Help Center, “Content monetization standards”](https://support.google.com/youtube/answer/6162278?hl=en#zippy=%2Cguide-to-self-certification;Help%20Center,%20Content%20monetization%20standards), 可查阅 <https://help.twitter.com/en/rules-and-policies/content-monetization-standards>; Meta 商业广告政策, 可查阅 https://www.facebook.com/policies_center/ads; TikTok 广告政策, 可查阅 <https://ads.tiktok.com/help/article?aid=9552>; Snap 广告策略, 可查阅 <https://www.snap.com/en-US/ad-policies>。

¹³¹ 欧洲联盟委员会,《反虚假信息行为守则》加强版, 第二节。

¹³² 见 YouTube 频道货币化政策, 可查阅 <https://support.google.com/youtube/answer/1311392?hl=en>; Meta, “Meta’s ongoing efforts regarding Russia’s invasion of Ukraine”, 26 February 2022, 可

各公司的广告政策不够全面或执行不够充分，也没有定期更新以反映全球冲突的发展。¹³³ 事实上，这些公司对乌克兰冲突的显著关注引起了人们的疑问，即是否在积极主动地解决其他冲突环境中的货币化问题。

94. **透明度和补救措施。**各数字平台一直在努力提供有效的透明度，无论是在和平还是在冲突的环境中。2022 年，数字化权利排名评估发现，14 个使用最广泛的数字平台的透明度报告在提供基本背景和精细数据方面存在不足，也未能提供足够的资料说明人权尽责调查情况以及算法和定向广告系统的开发和部署情况，包括用于策划、排名或推荐内容的系统。¹³⁴ 有关政府提出的删除内容要求的透明度也很有限，包括要求的数量、来源、目标以及公司是否予以合作。

95. 监督委员会已经注意到，在 Meta 公司执行其内容政策(按国家和语言细分)以及关于政府提出的删除内容的要求方面，需要更大的透明度。¹³⁵ 这些与 Meta 有关的问题浮现的原因是该公司决定设立监督委员会。许多其他公司很少或根本不提供有关其运营的信息，更鲜有提供公开的上诉和审查渠道。

96. 2022 年欧盟委员会《反虚假信息行为守则》加强版对《数字服务法》进行补充与后者的监管要求保持一致，其中包括承诺提高透明度、便利研究界对平台数据的访问，以及向用户公布更多关于推荐系统设计的信息。¹³⁶ 这些措施是重要的进展，但还需要定期更新透明度报告，具体说明处于风险中或受冲突影响国家的动态。

97. 此外，安全、直接地提供上诉机制和其他通信渠道，然后及时作出回应，对于用户就冲突时期限制言论，包括在保全证据方面提出质疑或关切至关重要。《行为守则》加强版包括承诺提供及时、勤勉和客观的透明上诉机制。¹³⁷

查阅 <https://about.fb.com/news/2022/02/metaspending-efforts-regarding-russias-invasion-of-ukraine/#latest>; Sinéad McSweeney, “Our ongoing approach to the war in Ukraine”, 16 March 2022, 可查阅 https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine; 推特“国家媒体”政策，可查阅 <https://business.twitter.com/en/help/ads-policies/ads-content-policies/state-media.html>; Snap 广告策略，第 3.9 节，可查阅 <https://www.snap.com/en-US/ad-policies>; Upvoted, “Supporting Ukraine and our community”, 2 March 2022, 可查阅 <https://www.redditinc.com/blog/supporting-ukraine-and-our-community>; Rafael Frankel, “An update on the situation in Myanmar”, Meta, 7 December 2021, 可查阅 <https://about.fb.com/news/2021/02/an-update-on-myanmar/>。见 Meta 公司提交的材料。

¹³³ 全球虚假信息指数提交的材料。

¹³⁴ 见 Afef Abrougui and others, “Key findings from the 2022 RDR big tech scorecard”, Ranking Digital Rights programme at New America, 可查阅 <https://rankingdigitalrights.org/mini-report/key-findings-2022/>; Svea Windwehr and Jillian C. York, “Thank you for your transparency report, here’s everything that’s missing”, Electronic Frontier Foundation, 13 October 2020, 可查阅 <https://www.eff.org/deeplinks/2020/10/thank-you-your-transparency-report-heres-everything-thats-missing>。

¹³⁵ 监督委员会提交的材料。

¹³⁶ 欧洲联盟委员会，《反虚假信息行为守则》加强版，第六节，承诺 18-19。

¹³⁷ 同上，承诺 24。

98. **加密、匿名和账户安全。**使用加密、匿名和其他隐私协议可以防止收集用户数据、审查和未经同意以用户为目标提供定制内容，从而加强用户自主权和安全。¹³⁸与此同时，采用加密技术的平台正被用来传播仇恨言论和其他煽动性内容。例如，有报告表明，在缅甸，随着脸书上的内容审核程度提高，Telegram上的虚假信息、宣传和煽动变得更加明显，该平台上可公开访问的亲军事频道参与了公开私人信息。¹³⁹利益攸关方必须继续评估纳入加密技术的平台上的社区审核方法和报告机制，既注重用户自主权，又符合国际人权法和国际人道法。¹⁴⁰

99. **保全证据。**社交媒体平台应特别注重保全冲突期间侵权行为的证据，因为此类数据对问责和司法程序至关重要。¹⁴¹有报告称，社交媒体公司选择删除与冲突有关的内容，包括战争罪的证据，因为这些内容违反了他们关于图形或暴力图片的政策，并错误地采用了最快的删除途径，而没有努力将材料存档。¹⁴²

六. 结论和建议

100. 在数字时代，信息环境已成为一个危险的、不断扩大的冲突场所。国家和非国家行为体借助新技术和社交媒体平台，将信息武器化，以制造混乱、助长仇恨、煽动暴力、鼓动公众不信任和毒害信息环境。人类承受的痛苦和社会结构遭到的破坏远甚于战争的危急情势。

101. 在这一背景下，特别报告员得出了六个广泛的结论，为向利益攸关方提出具体建议提供指导。

102. 第一，信息权不应被视为合法的战争目标。这是一项“生存权”，在危机和冲突时期，人们的生命、健康、福祉、安全和保障有赖于这项权利。为了民众的安全、保障、尊严和自由而维护和促进这一权利是一项人权和公益。如果不能获得多种信息来源，民主社会就不可能繁荣。表达自由包括信息权，是公众信任的基础，有助于预防和解决冲突，促进和平、和解与发展。

103. 第二，打击虚假信息对于保障人权和恢复公众信任至关重要，但必须以有效而不是适得其反的方式进行。对批评声音的审查、对独立媒体的攻击和对互联网的干扰，无助于减少虚假信息，反而大大削弱意见和表达自由，使信息环境恶

¹³⁸ 见大众媒体维护中心、纪念人权维护中心、网络自由项目和 OVD-Info 联合提交的材料以及得克萨斯大学奥斯汀分校媒体参与中心提交的材料。

¹³⁹ 缅甸自由表达组织和立即普及组织提交的材料。

¹⁴⁰ 得克萨斯大学奥斯汀分校媒体参与中心提交的材料。

¹⁴¹ 见 Human Rights Center, University of California at Berkeley School of Law, and OHCHR, Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law (United Nations publication, 2022), 可查阅 https://www.ohchr.org/sites/default/files/2022-04/OHCHR_BerkeleyProtocol.pdf。

¹⁴² 见监督委员会提交的材料。

化。所有国家都必须明确承诺维护意见和表达自由权，各国为打击虚假信息而采取的任何行动都应以国际人权法为依据。

104. 第三，数字技术和社交媒体创造了一种新的模式，暴露了国际法律标准中的模糊性、不确定性和潜在漏洞，一些国家和非国家行为体正在肆无忌惮地利用这些漏洞而不受惩罚，损害人权和人道主义保护。应加强人权原则与国际人道法的适用，以便重新解释“战争诈术”的可允许限度，从而保护平民及其尊严和生存所需的信息权。

105. 禁止战争宣传应作狭义解释，以确保不侵犯抗议和批评的权利。联合国人权事务高级专员办事处应制定准则，供各国和公司使用。

106. 应重新审议人权的域外适用问题，以考虑到对言论和信息自由权的跨境数字威胁。

107. 第四，虽然社交媒体公司对乌克兰战争的应对行动在许多方面值得赞扬，但它们必须做更多的工作，以确保在世界各地连贯一致地适用政策和业务做法，并确保及时根据公司运营所在的各种冲突环境的当地具体情况，加强人权尽责调查和影响评估。

108. 第五，必须增强权利持有人和民间社会的权能，建立社会抵御信息操纵的能力。在脆弱局势中，应更多地关注媒体信息和数字素养，特别是青年人、妇女、老年人和其他边缘化群体的媒体信息和数字素养，健康的社区关系，立足于社区的事实核对，以及打击仇恨、暴力和极端主义的教育方案。

109. 最后，必须在所有利益攸关方的协作下，以综合方式应对数字生态系统的挑战。必须采取一种多层面、多利益攸关方的办法，让民间社会和传统媒体与国家、国际组织和数字公司一起充分参与。

A. 对各国的建议

110. 尊重、保护和实现意见和表达自由权的义务使各国有责任确保一个健康的信息环境。根据这一义务，各国必须避免制造、支持、鼓励或传播虚假信息，破坏信息环境。

111. 在武装冲突期间，各国不得以可能直接或间接导致对平民造成伤害，包括导致国际法界定的战争罪、危害人类罪或其他违反国际人道法行为或煽动敌意、暴力或歧视行为的方式，在境内或跨境使用、传播或鼓励第三方传播信息。

112. 各国不得干扰互联网或电信，因为这必然对获取信息造成过度限制。一般制裁应避免限制民众使用互联网或安全通信方式的后果。必要时，各国应提供豁免，使生活在制裁之下的民众能够使用互联网和自由交流信息。

113. 各国不应禁止或限制虚假信息、宣传和“错误新闻”或“假新闻”，除非它们满足《公民及政治权利国际公约》第十九条第三款所规定的合法性、必要性和合理目的的要求，或等同于第二十条所述的煽动。各国必须禁止构成煽动歧视、

敌意或暴力或其他国际罪行的鼓吹仇恨言论。除非根据《拉巴特行动计划》提供的指导，否则应避免对言论进行刑事定罪。

114. 各国必须确保所有克减措施都严格必要，并与特殊情况相称，不带歧视性，有时间限制，而且在范围上符合危机的危急情势。此外，在紧急情况下限制言论的措施应根据《公民及政治权利国际公约》的程序宣布为克减措施，以便人权事务委员会进行审查。

115. 各国应优先采取非法律措施打击虚假信息和宣传，首先是履行其主动披露官方数据的义务，鼓励进行可信的事实核对，促进获得多样、可靠的信息来源，确保媒体、数字和信息素养，并为民间社会营造一个有利和包容的环境，以采取措施打击信息操纵。

116. 各国应履行其确保信息权的义务，为此应提高自身透明度，在线上 and 线下主动披露官方数据。所有国家都必须通过和实施关于获取信息的全面法律，或使现行法律、政策和做法符合国际和区域标准。此类法律应避免以国家安全为由，就信息权规定过于宽泛的例外情况。

117. 信息权包括获取各种信息的权利，不分国界，并可借助个人选择的任何媒体。各国应尊重和保护个人接收外国新闻和宣传的权利，除非根据国际人权标准对此类信息加以限制。

118. 各国应始终尊重、保护和促进媒体的独立、自由、多元化和多样性，包括性别多样性。它们应充分遵守国际人道法规定的义务，在武装冲突中将所有外国和本国记者(根据国际人权法界定)作为平民保护。媒体的行动自由和独立报道自由应该得到严格尊重。

119. 国家在武装冲突期间强迫媒体、社交媒体平台或民间社会组织只传播当局制作或批准的信息并不合法。根据国际法，即使在紧急状态期间，也没有理由实行全面信息封锁并施以严厉的刑事处罚。

120. 各国应确保根据《关于调查潜在非法死亡的明尼苏达规程》，迅速、有效、独立和公正地调查所有袭击记者的事件。联合国应设立一个独立的任务组，负责支持国际和国家在防止、调查和起诉袭击记者行为方面的努力。

121. 国际刑事法院应审查在冲突局势中持续杀害记者的问题，以便在国家当局不愿意或无法起诉的情况下起诉战争罪。

122. 各国不应要求各平台就内容采取不符合国际人权标准的措施。国家对社交媒体的监管应鼓励各公司确保有效的透明度、人权尽责调查和用户的正当程序权利。

B. 对各公司的建议

123. 各公司应根据国际人权法，并在适用情况下根据国际人道法标准，制定在冲突环境中运作的全面政策、流程和架构，为处理信息操纵问题提供可预测、连贯一致和有效的框架，确保用户安全，并建立补救机制。各项政策应以用户使用平台时所用的语言向所有用户发布。

124. 各公司应加强人权尽责调查，并在冲突发生前、发生后或正在发生的情况下，及时启动强化的风险管理战略，配备充足的资源、语言和背景专业知识，并调动民间社会参与。尽责调查程序应包括对公司的业务、产品和服务(包括商业模式本身)对冲突动态和人权的影响进行稳健的分析。

125. 各公司应根据国际人权法和国际人道法标准调整内容审核，尽一切努力维护表达自由和获取信息的机会，同时防止传播可能煽动暴力或违反国际人权法和国际人道法其他原则的内容。

126. 各公司应确保冲突环境中的内容审核包括严格的人工审核，并融入相关语言以及当地和区域背景方面的专业知识。应与可靠的事实核对组织和民间社会建立伙伴关系，以补充内部专门知识。

127. 端到端加密必须作为享有意见和表达自由的一个重要方面受到保护。各公司还应仔细评估在冲突环境中可能影响用户的账户安全风险，并提供强化安全选项。

128. 各公司不仅应制定政策，而且应有效实施政策，限制和跟踪与武装冲突有关的有害内容的货币化情况。

129. 各公司应提高冲突局势中的透明度，包括为此定期公布专门针对具体局势的透明度报告。这类报告应包括有效评估公司政策对人权的影响所需的具体细节和背景。各公司应便利研究人员获取公司关于在冲突期间数字平台使用情况的数据。

130. 各公司应根据国际证据标准，安全地保存武装冲突期间犯下的战争罪或其他侵犯人权行为的所有潜在证据，并制定与适当的国家或国际司法机构分享证据的程序。

131. 最后，特别报告员重申，需要进行更多的研究、分析和多方利益攸关方对话，以便就处理虚假信息和其他形式信息操纵的概念、政策、战略和准则达成共识。在推出各种举措审查这些问题之际，特别报告员再次呼吁采取强有力的基于人权的办法，并欢迎会员国、国际组织、公司和民间社会提出建议，说明她的任务如何能够促进它们所作的努力。