

数字平台治理准则

保障表达自由和信息获取的多方合作方法



联合国教育，科学及文化组织，丰特努瓦广场7号，75352巴黎 07 SP, 法国
2023年出版

© UNESCO 2023

ISBN 978-92-3-500052-8



此出版物为开放获取出版物，授权条款为Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>)。此出版物内容的使用者无条件接受遵守教科文组织开放获取储存档的一切条件和规则 (<https://www.unesco.org/zh/open-access/cc-sa>)。

原版出版物名称：*Guidelines for the Governance of Digital Platforms: Safeguarding freedom of expression and access to information through a multi-stakeholder approach*
联合国教育，科学及文化组织2023年出版

本出版物所用名称及其材料的编制方式并不意味着联合国教科文组织对于任何国家、领土、城市、地区或其当局的法律地位，或对于其边界或界线的划分，表示任何意见。
本出版物表达的是作者的看法和意见，而不一定是联合国教科文组织的看法和意见，因此本组织对此不承担责任。

封面，排版和图案设计： Luiza Maximo

插图： Plastic Horse/Grand Matter

翻译： TRADUCTEO

教科文组织印刷

印刷地：巴黎

概 要

“可信任的互联网” 准则

保障表达自由和信息权的同时应对虚假信息、错误信息、仇恨言论及阴谋论，需要多方利益攸关者的合作。

为此，联合国教科文组织作为联合国促进和保护表达自由和信息自由的领军机构，推出了《数字平台治理准则》。

该《准则》旨在促进数字平台治理以维护表达自由和信息自由为核心，并厘清了政府、数字平台、政府间组织、民间社会、媒体、学术界、技术社群及其他利益攸关方所应承担的义务、责任和角色。

《准则》是通过多利益攸关方磋商进程制定的，汇集了来自134个国家的1万多条意见。这些全球范围的磋商促进了包容性参与，确保了不同的声音都能被听到，包括那些边缘化和弱势群体的声音。

培育“可信任的互联网”是所有利益攸关方的共同责任。它呼吁我们所有人致力于维护促进表达自由和信息权的环境。

来自
134个国家的
1万条
意见



“战争起源于人之思想，故务需于人之思想中筑起保卫和平之屏障”



数字平台治理准则

保障表达自由和信息获取的多方合作方法

6 序 言

8 《准则》目标

11 引 言

14 《准则》结构

16 有利环境

24 治理系统

35 原则1、平台进行人权尽职调查。

37 原则2、平台在平台设计、内容审核及内容策展等方面遵守国际人权标准。

42 原则3、平台保持透明。

46 原则4、平台为用户提供信息和工具。

48 原则5、平台对相关利益攸关方担责。

50 针对具体情况的规定

54 总 结

56 附 录

序 言

教科文组织总干事奥黛丽·阿祖莱：

维护表达自由和信息获取： 数字平台治理多方合作方法原则

2023年9月

2023年，全球60%的人口，即47.5亿人，使用社交媒体平台来表达自我、获取信息及肯定自我。

数字领域是一个自由的空间，也是一个表达和辩论的新论坛，与我们的社会关系、身份和生活互相交织。

这些平台已成为平等和自由拥护者的扩音器——为无声者发声，为各种形式的表达提供避风港。

然而，同样是这些社交网络，在原本大有可为的同时，却常常成为隔离泡泡、错误信息的保护茧，有时还会助长阴谋论和极端暴力。

作为社交互动的虚拟空间，它们受制于旨在垄断我们注意力的算法；由于点击率优先于确定性，概率优先于证据，它们无意中助长了错误信息和仇恨言论。

然而，如果我们不再能区分虚构与现实、虚假与真实，我们社会的基础就会崩溃，民主、对话和辩论——所有这些对于应对当代重大挑战都是必不可少的——将变得无法实现。

面对这些问题的全球性质，我们需要在世界各地制定一致的应对措施，以避免法规或方法各自为政，损害人权。

教科文组织必须迎接的正是这一全球性挑战，因为这是我们的核心任务。

本组织自成立以来，一直致力于增进“各国人民间之相互认识与了解”，特别是通过本组织《组织法》所强调的“运用文字与图象促进思想之自由交流”。

这一承诺促使教科文组织于2005年发布了广播监管准则。最近，我们的193个会员国于2021年通过了《人工智能伦理问题建议书》，为这一技术的发展确立了人文主义框架。

教科文组织忠实于自己的价值观和历史，努力制定本出版物中介绍的开创性准则。这些准则旨在打击错误信息和仇恨言论，同时促进平台的透明度和表达自由。

这些工作以2021年《温得和克会议三十周年纪念宣言》为指导，教科文组织所有会员国都通过了该宣言中的各项原则。该宣言确定了三大行动支柱：倡导平台的透明度、确保媒体的生存能力以及培养用户的批判性思维。

这项努力的最终成果就是制定了这些准则，这是采纳了1万多条意见后广泛磋商的结果，使其成为联合国开展的最全面的磋商之一。仅教科文组织于2023年2月举办的“可信任的互联网”会议就汇集了来自134个国家的4000多名利益攸关方。

这些准则提出了公平、明确和共同的措施：包括土著语言在内的所有语言的在线审核员；提高平台及其融资的透明度，并加强风险评估；建立独立监管机构；促进批判性思维；支持性别平等；最重要的是，保障和加强表达自由、文化多样性及其他人权。

本书不仅关注数字平台的现实情况，还探讨了未来挑战，尤其是生成式人工智能带来的挑战。

教科文组织致力于协助会员国、民间社会及主要数字参与者以本书为依据，使平台运营完全符合我们的价值观和国际人权标准。

让我们继续专注于我们的目标：打击仇恨言论和错误信息，同时维护表达自由。这并不矛盾：在加强对自由可靠信息获取的同时，我们也加强了思想和表达自由。

借用汉娜·阿伦特的话就是：“除非事实信息得到保证，并且事实本身也无争议，否则意见自由就是一场闹剧”。

数字平台治理准则

保障表达自由和信息获取的多方合作方法

2023年10月

《准则》目标

1. 教科文组织以相关原则、公约和宣言为基础，通过多方磋商和全球对话，制定了本文件：《数字平台治理准则：保障表达自由和信息获取的多方合作方法》（《准则》）。¹
2. 《准则》的目的是在数字平台治理中保障表达自由权，包括信息获取及其他人权，同时处理根据国际人权法和标准允许限制的内容。推而广之，以人权为基础的数字平台治理将进一步促进文化多样性、文化表现形式及多样性文化内容。²《准则》概述了尊重人权的治理系统，并推进了有助于降低风险的系统性内容审核与策展程序。本《准则》强调了所有影响数字平台上表达自由和信息获取的治理系统都应遵循的总体原则——这些原则与具体的监管安排和专题重点无关，只要这些安排符合本《准则》的规定即可。
3. 《准则》认识到，各种规则和条例在每个治理系统中的适用性均须遵守各项国际人权标准，包括《公民权利和政治权利国际公约》（ICCPR）³第十九条第三款，其中规

1. 本文件原始版本为英文。

2. 教科文组织2005年《保护和促进文化表现形式多样性公约》第一条和第四条。根据该《公约》，“文化内容”指源于文化特征或表现文化特征的象征意义、艺术特色和文化价值。此外，“文化表现形式”指个人、群体和社会创造的具有文化内容的表现形式。

3. 《公民权利和政治权利国际公约》（ICCPR）。1966。 <https://www.ohchr.org/zh/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>.

定，对表达自由的任何限制都必须依法规定，必须追求该条款中规定的合法目的，并且必须是必要且相称的；以及《公民权利和政治权利国际公约》第二十条和其他国际标准，特别是联合国人权事务委员会、国际和区域人权法院对这些条约条款的权威解释，以及《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》。⁴

4. 《准则》侧重于保护和促进人权标准，并允许存在各种多元化平台（包括去中心化平台）以及具有多种内容标准和审核体系的生态系统。
5. 《准则》可作为一系列利益攸关方的资源：决策者可用于确定合法目标、人权原则以及可在决策中加以考虑的包容性和参与性程序；监管和其他治理机构可用于实施和评估政策、行为准则或法规；数字平台可用于政策和实践；以及民间社会等其他利益攸关方可用于宣传和问责工作。新闻媒体在不断努力追究有权势的行为体的责任时，也可利用本《准则》。
6. 《准则》旨在为数字平台内容管理的具体治理程序提供信息，也为其他领域已在实施的治理程序提供信息，这些领域可能对行使表达自由和信息获取以及文化内容多样性产生影响，并应根据数字环境的变化加以考虑（如选举、数据保护及反垄断法规）。根据问题和管辖范围的不同，这类治理程序可以采取互补支柱——自我监管、共同监管和法定监管——相结合的形式，以符合国际人权标准的方式构建（见下文“治理系统”部分）。这类治理程序应以公开、透明、多方合作、相称及循证的方式主导进行。为此，本《准则》应成为一份动态文件，需要定期审查更新，包括考虑从实施中吸取的经验教训以及随后的技术变革和影响。

4. 《准则》应与所有核心国际人权文书一并解读。这些文书的概要见：<https://www.ohchr.org/en/core-international-human-rights-instruments-and-their-monitoring-bodies>.

7. 本《准则》旨在以务实的方式为实现以人为本的数字治理模式做出更广泛的贡献。《准则》也是实现可持续发展所需的更广泛的行动工具包的一部分。《准则》将：
- a. 鼓励并促进全球多利益攸关方网络和共同空间的发展，以讨论和分享数字平台治理方面的良好做法，汇集不同愿景和广泛视角。
 - b. 作为所有相关利益攸关方的工具，用于倡导尊重人权的监管以及向政府和数字平台问责。
 - c. 推进基于人权的循证式政策方法。
 - d. 鼓励全球各种平台治理政策尽可能趋同，以避免互联网碎片化。
8. 《准则》力图促进联合国范围内正在进行的各种程序，如实施《我们的共同议程》中的各项提案，并为这些进程提供信息。这包括制定《全球数字契约》，⁵筹备将于2024年9月举行的联合国未来峰会，以及制定促进数字平台信息完整性的《行为守则》。⁶《准则》还将为即将举行的信息社会世界峰会（WSIS）20周年审查和互联网治理论坛（IGF）的讨论提供资料。本案文还借鉴了其他国际政府组织牵头的倡议，包括区域范围的倡议，同时旨在进一步促进这些倡议。

5. 请参阅联合国秘书长发布的《我们的共同议程政策简报5》，其中交叉引用了本《准则》制定的程序：https://digitalibrary.un.org/record/4011891/files/%5EEOSG_2023_5%5E--EOSG_2023_5-ZH.pdf。

6. 请参阅联合国秘书长发布的《我们的共同议程政策简报8》，其中交叉引用了本《准则》制定的程序：<https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-zh.pdf>。



引言

9. 1945年11月，教科文组织成立，其宗旨在于“通过教育、科学及文化来促进各国间之合作，对和平与安全作出贡献，以增进对正义、法治及所确认之世界人民均享人权与基本自由之普遍尊重”。⁷
10. 教科文组织的全球任务包括促进“运用文字与图象促进思想之自由交流”，这一使命在近80年来一直指导着本组织的工作——作为思想实验室、信息交流中心、标准制定者、国际合作的催化剂和推动者以及能力建设者。这段历史也决定了本组织在联合国系统内的任务，即保护和促进表达自由、信息获取以及记者的安全，无论是线下还是线上。
11. 教科文组织正在开展的工作和承诺是，确保数字平台治理能够保护和促进所有人的表达自由、信息和多样性文化内容获取以及其他人权，包括弱势和边缘化群体。⁸

7. 《联合国教育、科学及文化组织组织法》第1条。 <https://www.unesco.org/zh/legal-affairs/constitution>。

8. “弱势和边缘化群体”系指儿童和青少年；残疾人；移民、难民和寻求庇护者；男女同性恋、双性恋、跨性别者和间性者；以及老年人。

12. 这项工作从教科文组织几十年来在广播监管领域的工作中汲取了经验教训，因为任何涉及内容问题的政府干预措施——无论内容来源如何——都必须始终将保障多样性、表达自由和信息获取作为最终目标。《准则》还有助于实施本组织《2022-2029年中期战略》（41 C/4）。⁹
13. 2015年，教科文组织大会核准了互联网普遍性ROAM原则，其中强调了“立足人权（R）、开放（O）、人人可及（A）和多利益攸关方参与（M）”在互联网发展、增长及演进中的重要性。¹⁰ 这些原则认识到，确保线上空间继续以有利于实现可持续发展目标的方式继续发展和使用，是一项根本性的需求。
14. 自信息社会世界峰会（2003年和2005年）举行以来，包括教科文组织在内的联合国系统所采取的总体战略一直以多方合作方法为基础，制定和应用影响互联网发展和使用的共同原则、规范、规则、决策程序和计划，联合国大会在2015年的十年审查程序中也重申了这一点：

此外，我们重申多方合作与参与的价值和原则，这是信息社会世界峰会程序自启动以来的特点，认识到：各国政府、私营部门、民间社会、国际组织、技术界和学术界以及其他所有相关利益攸关方在各自作用和责任范围内的有效参与、伙伴关系与合作，特别是在发展中国家代表人数均衡的情况下，对发展信息社会一直至关重要，并将继续如此。¹¹

15. 2021年5月，在“世界新闻自由日”全球庆祝活动上开启了一项多利益攸关方程序，之后在当年11月，教科文组织第四十一届大

9. 《中期战略》的战略目标3是建设包容、公正与和平的社会，包括促进言论自由。战略目标4是通过发展和传播知识和技能以及道德标准，营造一个为人类服务的技术环境。https://unesdoc.unesco.org/ark:/48223/pf0000378083_chi。

10. 联合国教科文组织。“互联网普遍性指标。”<https://www.unesco.org/zh/internet-universality-indicators>。

11. 联合国大会。2015。“关于全面审查信息社会世界峰会成果落实情况的大会高级别会议成果文件”。70/125。https://unctad.org/system/files/official-document/ares70d125_en.pdf

会核准了《温得和克会议三十周年纪念宣言》各项原则。¹²《宣言》主张信息是一种公共产品，并在目标中设定了三个步骤来保证信息成为全人类的共享资源：数字平台的透明度、通过媒体与信息素养为公民赋能以及媒体的生存能力。在促进将信息视为公共产品的愿景过程中，教科文组织认识到，这种普遍权利既是实现人类集体愿望（包括《2030年可持续发展议程》）的手段，也是最终目的。信息赋能公民行使其基本权利，支持性别平等，让人们能够参与民主治理和可持续发展并对其产生信任，做到不让任何一个人掉队。

16. 《准则》侧重于与数字环境中的表达自由以及信息和多样性文化内容获取有关的挑战，这是对本组织在教育、科学及文化领域工作的补充。这包括但不限于教科文组织的《人工智能伦理问题建议书》，¹³其中呼吁制定国际和国家政策和监管框架，以确保新兴技术造福全人类，还有2005年《保护和促进文化表现形式多样性公约》¹⁴及其《关于在数字环境中执行〈公约〉的操作指南》。该《操作指南》促进“尊重表达、信息和通信的基本自由，促进尊重隐私权和其他人权，并将其作为创作、销售和获取多样性文化表现形式的前提条件。其中包括促进艺术自由（表达自由的必要结果）、从事数字工作的作者和艺术工作者的社会和经济权利以及所有合作伙伴与其选择的伙伴建立伙伴关系的权利”。¹⁵本《准则》的重点也是对《2022年世界文化政策会议宣言》的补充，该宣言呼吁“对数字部门，特别是主要平台进行实质性监管”，以利于网络文化多样性和人人公平获取内容。¹⁶

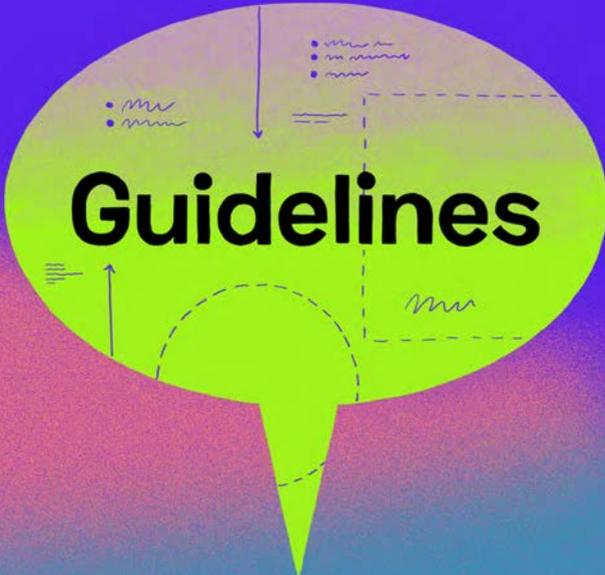
12. 教科文组织。2021。《温得和克会议三十周年纪念宣言：将信息视为公共产品》。https://unesdoc.unesco.org/ark:/48223/pf0000378158_chi。

13. 联合国教科文组织。2021。《人工智能伦理问题建议书》。https://unesdoc.unesco.org/ark:/48223/pf0000380455_chi。

14. 联合国教科文组织。2005。2005年《保护和促进文化表现形式多样性公约》。<https://en.unesco.org/creativity/convention>。

15. 联合国教科文组织。2017。《关于在数字环境中执行〈公约〉的操作指南》。https://unesdoc.unesco.org/ark:/48223/pf0000370521_chi.page=93。

16. 联合国教科文组织。2022。《教科文组织世界文化政策与可持续发展会议(MONDIACULT2022): 最终宣言》。https://www.unesco.org/sites/default/files/medias/fichiers/2022/10/6.MONDIACULT_ZH_DRAFT%20FINAL%20DECLARATION.pdf



Guidelines

《准则》结构

17. 《准则》首先描述了保障表达自由、信息获取及其他人权所需的有利环境，同时确保为数字平台用户和非用户提供一个开放、安全和可靠的环境。《准则》概述了不同利益攸关方在这方面的责任。这包括：
 - a. 国家尊重、保护和实现人权的义务。
 - b. 数字平台尊重人权的责任。
 - c. 政府间组织的作用。
 - d. 民间社会、媒体、学术界、技术界及其他利益攸关方在促进人权方面的作用。
18. 然后，《准则》采用以人权为本的多方合作方法，为数字平台治理系统制定了基本原则。这一节列出了互为补充的自我监管、共同监管和法定监管安排，以及可用于界定法定监管所涵盖公司范围的标准。

19. 然后，《准则》将媒体与信息素养以及尊重文化多样性确定为参与数字平台治理的所有利益攸关方的共同责任。
20. 最后，《准则》介绍了数字平台应在哪些领域建立各种系统和程序，以便评估风险；以便根据国际人权标准和教科文组织《2005年公约》定义的对文化多样性的尊重，对内容进行策展和审核；以便通过媒体与信息素养为用户赋能；以便通过报告机制和补救措施进行问责，以保障表达自由、信息获取及其他人权。
21. 必须强调的是，《准则》所涵盖的不同领域（如上文第17-21款所述）**必须作为一个整体加以考虑**。保障表达自由以及信息和多样性文化内容获取，需要考虑上述所有要素。



有利环境

22. 所有利益攸关方都有责任维持有利于保障表达自由、信息获取及其他人权的环境，同时确保为数字平台用户和非用户提供一个开放、安全和可靠的环境。¹⁷

23. 创造这样一个有利的环境不仅仅是一个工程问题。这也是一项需要全社会参与的工作，因此需要全社会群策群力解决。每个治理系统的所有相关利益攸关方都应采取行动，使弱势和边缘化群体、妇女和女童、土著社区以及记者、艺术家、人权维护者和环境维护者等能够行使表达自由权。社会所有成员都可以发挥作用，确保互联网安全，对暴力行为或威胁行为提出质疑，在网上交流时尊重他人的权利，尊重文化内容的多样性，并知晓社会中固有的偏见。

17. 本《准则》中的“安全的”和“安全”是指个人能够相信其人权（包括表达自由和信息获取权）受到保护的情况。

24. 儿童的地位特殊，因为他们处于独特的成长阶段，政治发言权有限或缺乏，而且童年的负面经历可能会造成终身或跨代的后果。¹⁸ 因此，虽然保护表达自由和信息获取适用于所有个人，但政府和数字平台也应认识到其在治理系统中对儿童¹⁹的具体责任。在涉及儿童参与数字环境时，每个利益攸关方都应坚持高道德标准和高专业标准，包括保护和促进儿童的表达自由和信息获取。
25. 所有参与数字平台治理的利益攸关方，都应促进并在适当情况下资助有民间社会组织、记者网络和研究人員参与的合作应对措施，以便更细致地了解根据国际人权法和标准可允许限制的内容，以及保护和支助妇女和女童、弱势和边缘化群体、记者、艺术家、人权维护者、土著社区和环境维护者的应对措施。

国家尊重、保护和实现人权的义务

26. 国家应尊重和促进人权，包括表达自由权和信息获取权。只有在《公民权利和政治权利国际公约》第十九条第三款和第二十条规定的条件下，才允许限制表达自由。国家负有保护人权免受包括数字平台在内的私人行为体无理干预的积极义务，因为国家有责任创造一个促进平台尊重人权的监管环境，并就数字平台的责任向其提供指导。
27. 此外，国家有义务对其对数字平台提出的要求做到完全透明和负责，确保法律的确信性和法律的可预测性，这是法治的基本前提。
28. 具体而言，国家应：

18. 请参阅联合国儿童权利委员会（2013年），《关于商业部门对儿童权利的影响方面国家义务的第16号一般性意见（2013年）》，第4款。另请参阅《关于与数字环境有关的儿童权利的第25号一般性意见（2021年）》。<https://www.ohchr.org/zh/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>。

19. 在大多数情况下，儿童通常被理解为18周岁以下的人。

- a. 促进互联网的普及和有意义的使用，保障网络中立性。²⁰
- b. 确保所有儿童都能以对他们有意义的方式平等有效地进入数字环境，并采取一切必要措施克服数字排斥现象。²¹
- c. 引导资源并加快努力，以弥合数字鸿沟，填补数据空白，消除弱势和边缘化群体面临的其他障碍，并实现所有妇女和女童的信息获取权。
- d. 加强公民空间，促进媒体的自由、独立和多元化，支持围绕网络言论、内容审核与策展以及平台问责制开展独立研究。
- e. 保证为记者（包括女记者）、人权维护者和举报人提供强有力的保护，并考虑支持媒体建立透明的自我监管机制，以促进和保护最高标准的专业精神。
- f. 保证为艺术家提供强有力的保护，认识到他们的作品对文化创作革新和促进文化多样性的重要性，并将其视为社会文化结构的核心。
- g. 保障数字平台用户的表达自由权、信息获取权、平等权 and 不受歧视权，保护用户的隐私权、数据保护权、结社权及公众参与权。
- h. 通过以国际人权标准为基础的法律，并确保其有效实施，以禁止、调查和起诉基于性别的网络暴力。²²
- i. 根据《公民权利和政治权利国际公约》第十九条第（三）款和第二十条的规定，确保对平台施加的任何限制始终遵循限制表达自由的高门槛，并尊重合法性、合法目的、必要性及相称性这些条件。

20. 在《表达自由与互联网联合宣言》（Joint Declaration on Freedom of Expression and the Internet）中，关于表达自由的特别国际任务指出：“落实表达自由权要求国家有义务促进互联网的普及”。2011年6月1日通过，第6(a)款。<http://www.law-democracy.org/wp-content/uploads/2010/07/11.06.Joint-Declaration.Internet.pdf>。

21. 请参阅《关于与数字环境有关的儿童权利的第25号一般性意见（2021年）》。<https://www.ohchr.org/zh/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>。

22. 请参阅《A/76/258：性别公正与意见和表达自由权 - 促进和保护意见和表达自由权特别报告员报告》。“所有限制性别仇恨言论或性别化虚假信息的法律措施都应符合合法、必要和相称这三段测试标准以及《公民权利和政治权利国际公约》第十九条第三款所列的正当目标。除非构成煽动这种最恶劣的鼓吹情况，否则应避免将其定为刑事犯罪。”<https://www.ohchr.org/zh/documents/thematic-reports/a76258-gender-justice-and-freedom-expression-report-special-rapporteur>。

- j.** 通过职业行为守则等措施，坚决阻止公职人员传播虚假信息，包括性别化虚假信息、²³错误信息以及恐吓或威胁媒体。此外，根据联合国关于仇恨言论的战略和行动计划，禁止出现相当于构成国际人权法所禁止的煽动歧视、敌意或暴力的鼓吹民族、性别、种族或宗教仇恨的言论。
- k.** 保持透明并披露所有信息，说明他们向数字平台提出的下架、删除和屏蔽内容的请求的类型、数量 and 法律依据。国家应该能够证明这如何符合《公民权利和政治权利国际公约》第十九条第三款和第二十条。
- l.** 促进媒体与信息素养，以加强与平台的积极互动，并培养网络安全技能，包括在数字空间中的安全技能，目的是增强用户的能力，特别是弱势和边缘化群体。这应包括促进对表达自由权、隐私权、平等权和诉诸司法权的了解，以及对投诉和补救手段的了解，并利用媒体与信息素养专家、图书馆、学者、民间社会组织以及信息机构获取方面的专业知识。
- m.** 确保任何涉及数字平台内容管理的监管机构，无论内容主题如何，在结构上都是独立的，不受政治和经济利益的影响，并建立了外部审查制度（请参阅本《准则》第68-73款）。这种审查制度可包括立法和司法审查，以及要求保持透明、与多利益攸关方协商、编制年度报告和定期接受外部审计的规定。这还需要就司法部门的权限和权威制定明确的规则。
- n.** 确保监管机构有足够的资源和能力，根据本《准则》目标进行评估。
- o.** 认识到一切治理系统都应借鉴人权专家、学者和民间社会组织的专门知识，以及其他治理系统中公认的良好做法。
- p.** 鼓励监管机构和司法行为体之间开展国际合作，包括三角合作和南南合作，促进良好做法和知识的交流。

23. 同上，第二十一款：“性别化的虚假信息也在增加。虽然这是性别暴力的一个子类，但自有一些明显特征，使用‘针对妇女社会性别和生理性别的虚假或误导性叙事，往往有一定程度的协同，目的是阻遏妇女参与公共领域。性别化的虚假信息结合了网络虚假信息的三个鲜明特征：虚假、恶意、协同。”

29. 国家应避免：

- a. 采取包括关闭互联网在内的措施，阻止或扰乱信息的在线和离线传播。
- b. 对数字平台强加一般监督义务或一般义务，要求其对在特定司法管辖区被视为非法的内容或根据国际人权法和标准可能允许限制的内容主动采取措施。如果数字平台本着诚信和尽职尽责行事，开展自愿调查，或采取其他措施，旨在检测、识别、删除或禁止访问《公民权利和政治权利国际公约》第二十条所禁止的内容或《公民权利和政治权利国际公约》第十九条第三款所限制的内容，则不应承担责任。
- c. 如果数字平台的工作人员在内容审核与策展工作中涉嫌或可能违反相关法规，则对其进行刑事处罚。

数字平台尊重人权的责任

30. 数字平台应遵守五项关键原则：

- a. **平台开展人权尽职调查**，评估其人权影响，包括性别和文化层面的影响，评估风险，并确定缓解措施。
- b. **平台在平台设计、内容审核及内容策展等方面遵守国际人权标准**。平台应遵循相关的国际人权标准，包括联合国《工商企业与人权指导原则》。设计应确保不歧视和平等待遇，并防止伤害；内容审核与策展政策及做法应符合人权标准，无论这些做法是通过自动化手段还是人工手段实施，都应了解当地语言和语言环境，尊重文化多样性，并为人工审核人员提供充分的保护和支持。
- c. **平台运作方式透明公开**，具有可理解和可审计的政策以及由多利益攸关方设计的绩效评估指标。这包括用于平台内容审核与策展的工具、系统及程序的透明度，包括算法决策及其产生的结果。
- d. **平台提供信息**，让用户了解所提供的不同产品、服务和工具，并就其分享和消费的内容做出明智决定。平台以用户自己的语言提供信息和支持用户的操作，同时考虑到用户的年龄和残疾情况。

- e. 在执行服务条款和内容政策时，**平台对相关利益攸关方担责**——包括用户、公众及治理系统内的行为体。针对与内容有关的决定，平台应使用户能够寻求适当及时的补救，包括内容被删除或被审核的用户以及对内容提出投诉的用户。
31. 平台应在其运营所在的每个司法管辖区运用这些原则，确保具备必要的资源和能力，以便及时有效地为用户提供服务。
 32. 为了遵循这些原则，根据国际人权标准，在某些特定领域，数字平台有责任向治理系统内的行为体报告或采取行动。本《准则》第85-129款对这些领域作了说明。

政府间组织的作用

33. 政府间组织应根据各自的任务规定，支持相关利益攸关方保证本《准则》的实施完全符合国际人权法和标准。这种支持应包括提供技术援助，监测和报告侵犯人权行为，制定相关标准，促进多利益攸关方对话，以及培育网络。
34. 政府间组织和国家监管机构可以制定参与模式，以进一步拓展和分享良好做法。这种参与可能包括分享新的见解和监管趋势，支持或建议国家监管机构完善制度标准和方法，以保障表达自由和信息获取。这类模式应努力减少互联网碎片化的风险，并提供工具，以便事先评估监管对整个互联网运作的影响。

民间社会及其他利益攸关方的作用

35. 作为用户、政策制定者、监督者或以其他任何方式参与数字平台服务的每个利益攸关方，在支持表达自由、信息获取及其他人权方面都发挥着至关重要的作用。为此，在制定、实施和评估影响数字平台内容的法规过程中，应采取多方合作方法。还应让广泛的利益攸关方参与监督，包括那些代表弱势和边缘化群体的利益攸关方，以及记者、艺术家、人权维护者和环境维护者。
36. 在了解网上有害内容和行为的性质以及抵制方面，民间社会发挥着至关重要的作用，尤其是针对所有弱势和边缘化群体、妇女和女童、记者、艺术家、人权维护者和环境维护者的有害内容和行为。在监测和报告影响人权的政府法律、政策和监管行动方面，民间社会也发挥着重要作用。他们是弥合数字治理生态系统与普通民众之间差距的关键。
37. 独立研究人员可以在确定滥用行为模式以及解决可能的根本症结方面发挥作用；研究人员还应能够对治理系统的运作方式进行独立审查。独立机构和研究人员可以为人权尽职调查提供支持，包括性别评估、审计、调查以及有关平台实践和活动的其他类型报告。研究人员应该能够收集和分析基于性别和其他相关交叉因素（如种族、民族、年龄、社会经济地位、残疾等）的分类数据。这有助于确定数字平台在不同弱势和边缘化群体方面的差异、偏见及不同影响。
38. 媒体机构、事实核查组织以及这些机构中的专业人员都是重要的利益攸关方，在履行其监督职能的同时，还在促进表达自由、信息获取及其他人权的享有方面发挥着作用。因此，有必要让媒体及其专业人员参与监管过程，承认他们在积极促进数字信息生态系统方面所发挥的积极作用。数字平台与可信新闻来源之间的建设性关系，将加强数字平台在提供符合公共利益的信息方面的作用。

39. 教育工作者和护理人员在帮助青少年和各年龄段的学习者了解更广泛的数字环境方面发挥着关键作用，包括如何寻找和理解可信信息以及如何以尊重的态度与他人进行在线交流。随着技术的日新月异，提供终身学习也同样重要。
40. 参与构建平台服务和产品的工程师、数据科学家及技术社群，在了解他们所设计和开发的产品和服务的人权、风险及伦理影响方面发挥着作用²⁴。
41. 所有这些利益攸关方都应在关于治理系统发展和运作的磋商中发挥积极作用。应促进利益攸关方之间的合作和对话。应开展建设性的讨论和审议，以交流想法、知识和观点。工作组、工作队或咨询委员会的设立，使利益攸关方有机会积极参与制定监管提案。

24. 请参阅人权高专办关于技术标准制定与人权之间关系的报告，A/HRC/53/42。



治理系统

42. 数字治理生态系统由世界各地一系列不同的利益攸关方、机构及监管安排组成。虽然现有的一些治理系统，如选举或数据保护，应根据数字时代带来的变化和挑战进行解释和考虑，但在各种背景下也正在创建新的治理系统，对数字平台进行直接监管。无论如何，这些监管机制可能会对网上表达自由以及信息和多样性文化内容获取产生深远影响。
43. 本《准则》强调了可酌情应用于涉及数字平台内容治理的各种程序的总体原则，无论这些程序的形式或领域如何。这些原则表明，一个全面的治理系统可以有效利用各种互补的监管安排，来应对数字生态系统中不同利益攸关方所面临的挑战。
44. 《准则》呼吁采用多方合作方法治理数字平台。这种方法可包括以下方面：确定所有相关利益攸关方（包括在其范围内的平台），鼓励包容性参与，确保均衡代表性，确保透明度和问责制，促进协作决策和对话，推动迭代程序，协调利益攸关方之间的实施工作，以及开展定期评估和审查等。

45. 根据具体情况，数字平台治理问责机制和合规机制可包括不同监管安排内的互补性和趋同性，如：
- a. 自我监管结构和机制，其中规则可由非国家行为体监督和执行，如全行业机构或社交媒体理事会。
 - b. 共同监管结构和机制，在某些情况下，行为守则可能被赋予法律效力而成为法规。
 - c. 法定监管框架，其中一个或多个独立监管机构就平台规则的制定做出最终决定。
46. 认识到这种环境的复杂性，本《准则》旨在适用于广泛的治理形式。必须指出的是，某些范畴可能需要法定监管框架，以应对不适合自我和共同监管机制的领域。这些框架必须始终确保官方监管机构的独立性，并根据《准则》的目标，应始终保障人权。

治理系统原则

47. 首先，**保持透明**应是一项共同的总体原则。在所有治理系统中，数字平台都应根据本《准则》以及联合国《工商企业与人权指导原则》的规定，对其平台上内容审核与策展所使用的条款、制度和程序以及任何人权尽职调查保持透明。他们应能解释其制度和程序如何履行和有效执行其服务条款，以及这些制度和程序是否符合国际人权标准。
48. 平台外部的治理系统和程序也应保持透明。任何外部监管行动的提出、公开广泛的辩论以及最终执行都应在公众监督下进行，同时公开明确地划分决策的职权范围和责任。

49. 其次，一个共同的监管原则是，不同利益之间的**相互制衡**应正式制度化。在所有形式的监管以及不同监管形式结合中，治理系统都应始终采用多方合作方法。这意味着要让所有利益攸关方广泛包容参与，因为他们最能代表不同的利益和价值观，包括不同的性别和交叉观点。多方参与在代表性以及在创建、应用、监督和审查治理程序（规则、原则及政策）方面都应富有意义。公众宣传活动、有针对性的外联活动、尊重文化多样性以及在治理程序中使用包容性语言和形式，都能促进有效参与。
50. 第三，治理程序应向所有利益攸关方**开放可及**，尤其是受拟议监管结构或监管类型影响的群体。应利用公众咨询、公开听证会和在线平台为公众提供意见和反馈的机会。在决策过程中，弱势和边缘化群体以及妇女和女童的关切应得到充分表达。
51. 治理系统应确保数字平台积极参与儿童事务，保护儿童的表达自由和其他权利，采用适当的保障措施，并在开发产品和服务时充分考虑儿童的意见。
52. 治理系统还应促进与媒体的对话，包括对独立新闻媒体的投资，并通过提供数据和支持促进媒体可持续性、多样性和多元化的行动，为媒体生态系统提供支持。
53. 第四，纳入**多样性专长**应成为所有监管安排的共同特点。治理系统要求利益攸关方通过培训和监管手段，具备理解人权框架和考虑技术发展的必要能力。他们应具备作出知情决定和应用本《准则》的能力和技术知识。应鼓励每个治理系统向公众报告和评估与新兴技术相关的风险和机遇。

54. 治理系统内的利益攸关方应在各司法管辖区之间共享监管专长和知识。国家、区域和全球治理系统应能开展合作并分享做法，以实现保障表达自由、信息获取及其他人权的目標，同时也应能处理国际人权法和标准允许限制的内容。
55. 第五，治理系统应确保数字平台在创作、制作、发行、传播、获取和享有网络文化产品与服务的过程中，参与**保护和促进文化多样性以及文化表现形式多样性**，包括确保其公平的可发现性和代表性。

问责制与合规性

56. 监管安排应确保有效和可持续，同时考虑到当地可用资源和需要关注的主要优先事项（例如，是否主要解决选举、公共卫生、广告或数据保护等方面的问题）。所有形式的监管都需要独立监督。法规制定过程应公开、透明、循证。
57. 数字平台在处理根据国际人权法和标准可以允许限制的内容时，如果被认为不符合自身政策，或未能履行保障表达自由和信息获取的职责，应根据第85-129款所述的五项原则，对治理系统内相关机构担责，并采取具有必要程序保障的适度执法措施。
58. **自我监管系统**可以与其他监管形式互补融合。这些系统应包括独立的定期强制性审计，用于评估数字平台对自我监管守则、政策或规范的遵守情况。此类审计不应直接由行业或单个数字平台提供资金，尽管对这些实体征税有助于支付此类工作的成本。审计工作也不应由可能存在或看似存在利益冲突的任何个人或实体进行。审计的条款和结果应公开征求公众意见。

59. **共同监管结构**应提供一个法律框架，为表达自由、信息获取及其他人权创造环境。在共同监管中，监管角色应由行业及其他利益攸关方与政府或官方独立监管当局或机构分担。相关公共机构的作用包括认可共同监管计划、对程序的审计以及资助该计划（可能通过对平台征税）。共同监管应允许在未实现商定目标的情况下，由国家强制实施罚款等处罚。
60. 只有当参与实施的监管机构在决策方面具有独立性时，才应考虑对数字平台进行**法定监管**，以解决可能影响表达自由的问题。这种监管应侧重于内容审核与策展的系统 and 程序，而不是确定个别内容的合法性，并且必须有法律依据（即充分定义），追求《公民权利和政治权利国际公约》第十九条第三款规定的合法目标，并且是必要且相称的。
61. 法定监管中的多方合作方法应反映在以下安排中：
- a. 包括官方独立监管机构在内的国家有关当局，通过参与性和包容性的立法程序确定监管的合法目标。
 - b. 数字平台向官方监管机构公开报告。
 - c. 民间社会组织、艺术家、独立研究人员及其他相关机构为规则制定建言献策，促进监督，并通过制度化的参与和审查实现必要的制衡。
62. 任何关于特定内容合法性的具体决定都应遵循正当程序，并接受公正独立司法机构的审查。
63. 在所有情况下，有关内容的评估都应遵循《公民权利和政治权利国际公约》第十九条第三款规定的关于表达自由合法限制的三段测试标准，²⁵以及《公民权利和政治权利国际公约》第二十条第二款的禁止构成煽动歧视、敌意或暴力的鼓吹仇恨的行为；包括

25. 联合国教科文组织。2021。《表达自由的合法限制——三段测试》。<https://www.youtube.com/watch?v=hyJaz8KZqXE>。

《拉巴特行动计划》中概述的用于界定此类内容的六项门槛测试（如适用）。²⁶

64. 每一项法定监管干预措施都应循证、相称，并包含程序性保障措施，包括确保平台能够获得做出决定所依据的所有事实和考虑因素。这一过程应涉及多个利益攸关方群体，从更广泛的角度考虑干预措施的可持续性、有效性和影响。对循证程序的呼吁不能成为拖延采取保护人权的必要监管行动的借口。
65. 包括平台在内的所有相关利益攸关方，都应有机会对违规决定提出陈述和/或申诉。应要求监管系统在指导平台实施具体措施之前，公布执法措施并就其进行协商，同时遵循适当的程序。

界定监管范围内的数字平台

66. 在界定应纳入法定监管范围的数字平台时，监管当局应确定那些在特定司法管辖区内开展相关业务并具有规模和市场份额的平台。应采用独立评估方式，评估这些平台对人权——包括对弱势和边缘化群体的人权——以及对民主机构构成的风险，来进行确定。²⁷范围的界定应保护隐私权，不应导致削弱对加密或其他隐私保护技术的保护。
67. 在确定范围内的公司时，为反映地区和司法管辖区的实际情况，可以考虑以下反映区域和司法管辖区实际情况的标准：
 - a. **规模和覆盖范围**，重点关注最有可能对大部分人口和/或弱势和边缘化群体产生影响的平台。
 - b. **市场份额**，认识到主导平台对整个信息生态系统的巨大影响力。运用《准则》时应避免惩罚初创企业和新进入者，同时确保以相称的方式

26. 联合国教科文组织。2021。《关于禁止煽动仇恨言论的拉巴特行动计划》。<https://www.unesco.org/archives/multimedia/document-5554-eng-3>。

27. 可以制定一份补充指南，用于识别平台的系统性风险，作为实施本《准则》的配套指南。

涵盖那些具有最大潜在影响力的数字平台。虽然所有平台都应遵循一般原则，但本《准则》第85-129款规定的具体报告义务可能主要适用于更有能力遵守这些义务的大型平台。

- c. 功能和特点**，认识到不同服务在内容的可见性、影响力和指向性方面的相关差异。风险可根据平台的用户群、所有权形式、商业模式、功能和特点来确定，如实时发布、病毒式传播潜力、数量、传播速度、真实性以及在多大程度上无需内容审核程序即可发布内容。²⁸



独立监管机构的特征

- 68.** 在法定监管中，官方监管机构虽然是国家行政机构的一部分，但应完全独立于政府，在履行其任务时主要对立法机构担责。²⁹ 这适用于对平台内容有合法权益的现有监管机构（如选举管理机构、广告管理机构、儿童保护实体、数据和隐私委员会、竞争监管机构等），以及任何可能新建立的专门或协调监管机构。
- 69.** 对于所有单独或联合参与平台监管的法定机构，应由直接向立法机构报告的独立机构进行定期审查³⁰。如果认为当局越权、行事不合理或以偏颇或不相称的方式行事，也应由法院对法定干预措施进行审查。
- 70.** 官方监管机构需要保持独立性，不受经济压力、政治压力或其他压力的影响。这些机构的权力和任务应由法律加以规定。它们还应遵守国际人权并促进两性平等标准。

28. “基于风险的监管是基于规则/标准制定者对与其任务相关风险的评估，以及根据风险程度所需的适当干预水平。如果行为体从事的是低风险活动，则会相应简化监管，降低合规要求。” https://www3.weforum.org/docs/WEF_Pathways_to_the_Regulation_of_Crypto_Assets_2023.pdf。

29. 世界银行表示，独立监管机构模式的关键特征是决策独立。教科文组织委托编写的广播监管指导文件（伊芙·萨洛蒙（Salomon, Eve）。《广播监管准则》（Guidelines for broadcasting regulation）。2006）还强调，“一个独立的机构（即其权力和责任由公共法律文书规定，有权管理自己的资源，其成员以独立方式任命且受法律保护免于无故解聘）更能为公共利益公正行事，避免受到政治利益或行业利益的不当影响”。有关完整的参考资料，请参阅本《准则》附录。

30. 审查应特别强调监管系统的决定如何影响人权的享有情况。

- 71.** 官方监管机构必须有足够的资金和专业知识和专业知识来有效履行职责。资金来源也必须清晰、透明，人人均能获得，且不受政府自由裁量权的约束。
- 72.** 负责平台内容问题的管理官员或官方监管机构成员应：
- a.** 通过参与式、透明、非歧视性和独立的择优程序任命。
 - b.** 对独立机构（可以是立法机关、司法机构、外部理事会或独立董事会）担责。
 - c.** 具备国际人权法和数字生态系统方面的相关专业知知识。
 - d.** 向独立机构（最好是立法机构）提交年度公开报告，并对其担责，包括告知该机构他们的合理意见。
 - e.** 公开任何可能的利益冲突，并申报任何礼物或奖励。
 - f.** 在完成任务后，在一段合理的时间内，不应受雇于曾受其监管的单位或为其提供有偿服务，以避免所谓的“旋转门”风险。
- 73.** 官方监管机构应能要求数字平台定期报告其服务条款的执行情况，并且如果认为数字平台不遵守其自身政策，或未能履行其在保障表达自由以及信息和多样性文化内容获取方面的责任，则应对其采取强制行动。他们应能建立投诉程序，提出具有约束力或不具约束力的公开建议，并有权根据国际人权标准，向促进和尊重人权的平台发出透明适当的指令。



74. 媒体与信息素养涵盖了广泛的技能，使用户能够批判性地思考他们在网上互动的信息。应通过治理系统专门应对媒体和信息素养问题，以确保包括数字平台在内的所有利益攸关方切实发挥各自作用。
75. 当治理系统内的利益攸关方拥有共同愿景，并通过共享知识和资源共同努力实现这一愿景时，媒体与信息素养将得到最有效的实现。媒体与信息素养计划应顺应现有的和新兴的媒体和信息技术的发展，以便公民能够充分利用这些媒体和技术，积极参与社会生活。
76. 媒体与信息素养计划应强调为用户赋能，确保他们具备相关技能和知识，使其能够以批判的眼光与各种形式的媒体内容以及所有信息提供者（包括中小学、大学、研究机构、图书馆、档案馆、博物馆、媒体公司、出版商、统计实体等）进行有效互动。当媒体与信息素养计划只强调保护或数字安全技能时，可能会导致对数字平台的使用施加过多限制。但是，它们应根据教科文组织和其他国际机构公布的最佳做法，优先考虑用户可以采取的具体步骤，以确定根据国际人权法和标准可以允许限制的内容。
77. 媒体与信息素养计划应促进文化多样性、社会包容和全球公民意识，旨在缩小参与创作和批判性利用媒体与信息内容的人员与未参与人员之间的“参与差距”。媒体与信息素养计划还应促进两性平等和妇女赋能，并为弱势和边缘化群体提供参与机会。

31. 请参阅教科文组织的《媒体与信息素养：政策和战略准则》（Media and information literacy: Policy and strategy guidelines）。<https://unesdoc.unesco.org/ark:/48223/pf0000225606>。

78. 各国政府应始终考虑促进用户的媒体与信息素养，包括网络安全技能，尤其是所有处于弱势和边缘化境地的群体以及妇女和女童。这使用户能够以批判的眼光看待内容和技术，驾驭以数字化转型为标志的快速发展的媒体和信息环境，促进人权，并在相关挑战面前增强应变能力。
79. 各国政府应就数字环境中的儿童权利，包括其表达自由权，传播信息并开展宣传活动，尤其要关注那些其行为对儿童产生直接或间接影响的人。它们应促进针对儿童、父母和照管者、公众以及政策制定者的教育计划，提高他们对数字产品和服务相关机遇和风险方面的儿童权利的认识。此类计划应包括有关儿童如何从数字产品和服务中获益的信息，以及有关如何培养儿童包括数字技能在内的媒体与信息素养方面的信息。
80. 平台应制定明确、公开的战略，通过媒体与信息素养，包括网络安全教育，为用户赋能，营造能够保障表达自由和信息获取的良好网络环境。数字平台内部应特别关注如何提高所有用户的数字素养，特别是弱势和边缘化群体的数字素养，并在产品开发团队中考虑到这一点。
81. 数字平台应划拨充足的资源，提高所有用户的媒体与信息素养，包括有关平台自身提供给用户的产品和服务的数字素养以及相关程序。这尤其应侧重于提高用户对特定平台呈现、策展、推荐和/或标记内容的方式的理解（也与下文“原则3”和“原则4”中概述的步骤有关），以及用户可以采取的具体步骤来自行识别根据国际人权法和标准允许限制的内容。
82. 平台应从用户赋能的角度出发，以国际标准为基础，对其产品开发团队进行媒体与信息素养（包括网络安全）方面的培训，并建立内部独立监测与评估机制。

83. 政府和数字平台都应独立于平台之外的组织和各类专家密切合作，实施媒体与信息素养计划，这些组织和专家包括但不限于：负责媒体与信息素养的公共机构、学术界、与弱势和边缘化群体合作的民间社会组织、研究人员、图书馆员、教师、专业教育工作者、记者、艺术家及文化专业人士。正如教科文组织现有诸多媒体与信息素养工具所概述的那样，应当为用户、非用户以及弱势和边缘化群体受众采取具体措施。
84. 各国政府和数字平台应开展合作，确保用户了解其线上线下各项权利，包括媒体与信息素养在享有和保护表达自由权和信息获取权方面的作用。



原则1、平台进行人权尽职调查

人权保障和风险评估

- 85.** 在任何类型的监管安排中，数字平台都应能够展示其为确保持续开展人权尽职调查而建立的系统或程序，包括人权和性别影响评估，³²以及风险缓解措施。³³应定期审查这些系统，并应公布审查情况。
- 86.** 根据联合国《工商企业与人权指导原则》等国际人权标准，平台应定期进行风险评估，以识别和应对其运营对人权的任何实际影响或潜在影响。在实施人权风险评估程序时，数字平台应考虑任何产品或服务如何影响用户行为，而不仅仅是为了获取用户或吸引用户。

32. 请参阅2021年10月18日《促进和保护意见和表达自由权特别报告员伊雷内·汗的报告》。根据联合国《工商企业与人权指导原则》的规定，“社交媒体公司应定期进行人权和性别影响评估，以识别和减轻影响妇女和性别不符者的系统性风险。这些公司应保证平台安全、具有性别包容性并符合国际人权标准，采用有效的安全政策和工具，确保有意义的透明度，包括算法的透明度，并提供适当的补救措施。” <https://www.ohchr.org/en/statements/2022/02/statement-irene-khan-special-rapporteur-promotion-and-protection-freedom-opinion>。

33. 人权影响评估应包括公司政策可能影响的所有人权。这包括公民权利和政治权利，如表达自由、信息获取和隐私，以及经济权利、社会权利和文化权利，免受暴力的权利，参与公共生活的权利等。

- 87.** 风险评估应成为数字平台内部决策的锚点，为其服务的设计和运营方法以及为应对残余风险和保障人权、非歧视和平等待遇而部署的缓解措施提供依据。此外，应明确规定风险管理责任，由最高层负责，并定期向高层决策者报告风险管理活动。
- 88.** 至少应在以下情况下进行人权和风险评估：
- a.** 在任何重大设计变更、重大政策决策（包括与广告系统相关的决策，如适用）、运营变更或推出新业务或建立新关系/合作伙伴关系之前。
 - b.** 定期保护处于弱势和边缘化境地的所有群体以及妇女和女童、记者、艺术家、人权维护者和环境维护者的权利。³⁴
 - c.** 在选举程序开始之前保护其公正性。³⁵
 - d.** 应对紧急情况、危机、冲突或运营环境的重大变化。³⁶
- 89.** 在人权尽职调查过程中，平台应确保与各种利益攸关方进行有意义的接触，以确定弱势和边缘化群体的具体风险。至关重要的是，数字平台在如何构建这些评估方面应听取专家的独立意见。
- 90.** 从一开始，平台就应创建倾听、参与和让用户参与的空间，包括那些经历过骚扰或虐待的用户及其代表，来自弱势和边缘化群体的用户，以及妇女和女童、记者、艺术家，以便为平台政策和程序提供信息。这可能包括识别和反击根据国际人权法和标准可允许限制的内容的方法，以及应对所评估风险的机会和解决方案。

34. 请参阅第126款针对具体情况的规定。

35. 请参阅第127-137款针对具体情况的规定。

36. 请参阅第138-140款针对具体情况的规定。



原则2、平台在平台设计、内容审核及内容策展等方面遵守国际人权标准。

91. 数字平台应确保将人权和正当程序考虑因素纳入设计程序的各个阶段，并纳入内容审核与策展方面的政策与做法。

设计程序

92. 数字平台新产品的设计及其内容审核与策展方面的政策，应符合联合国《工商企业与人权指导原则》以及其他既定国际人权标准中规定的企业尊重人权的责任。
93. 数字平台应确保其设计程序以及内容审核与策展方面的政策、做法和系统中实现非歧视和平等待遇。这包括解决影响妇女和女童以及弱势和边缘化群体（包括土著社区）的偏见、刻板印象、歧视性算法或歧视性内容审核做法。数字平台应确保所有用户，无论其背景或能力如何，都能充分参与并使用其服务。

内容审核与策展政策和做法

94. 内容审核与策展系统（包括自动化和非自动化部分）应可靠有效，其规模应与平台运营所在的所有辖区内被审核内容的数量相适应。这包括追求检测方法的准确性和非歧视性。内容审核与策展应符合国际人权法和标准，特别是不得侵犯表达自由和文化多样性。
95. 所有地区和语言的内容审核决策都应以透明的方式考虑到背景、影响含义的各种语言细微差别以及内容的语言特殊性和文化特殊性。
96. 在多语言环境中运营的平台，应确保（至少）以该环境中使用的所有主要语言提供人工和自动化内容审核，其规模应与内容量相适应。
97. 数字平台应确保采取迅速果断的行动，删除已知的儿童性虐待材料或恐怖行为直播，尊重所有个人的权利，包括弱势和边缘化群体的权利。尽管如此，平台仍应确保不删除此类可能对调查和起诉犯罪至关重要的内容，而是对其进行保存和安全保护，供执法机构和研究人员酌情使用。
98. 如上所述，国家有责任保障表达自由权，并确保对内容的任何限制都符合国际人权法和标准，特别是《公民权利和政治权利国际公约》第十九条第三款和第二十条。但是，数字平台应能够证明，在审核和策展内容时，所采取的任何行动都符合其服务条款和社区标准，并应向治理系统或适用的独立司法系统准确报告其责任和/或计划的执行情况。
99. 在考虑限制内容的措施时，平台应考虑到《公民权利和政治权利国际公约》第十九条第三款规定的关于表达自由合法限制的条件，以及《公民权利和政治权利国际公约》第二十条第二款的禁止构成煽动歧视、敌意或暴力的鼓吹仇恨的行为；包括《拉巴特行动计划》中概述的用于界定此类内容的六项门槛测试。

100.一旦数字平台发现根据国际人权法和标准可允许限制的内容，它们就应采取以下措施：提供替代的可靠信息；向用户说明对内容来源的关切；限制或消除对此类内容的算法放大，并适当关注反映性别偏见或性别暴力的内容；将广告收入中的内容去货币化；或删除/下架内容等。³⁷

人工内容审核

101.人工内容审核员，无论是平台直接聘用的，还是通过外包角色作为外部承包商聘用的，都应经过充分培训，精通平台上使用的语言，熟悉当地的语言和文化背景，经过评估、审查，并得到心理方面的支持。平台应进一步为内容审核员制定资金充足、人员配备齐全的支持计划，以尽量减少他们在工作中因经常接触暴力或令人不安的内容而受到的伤害。所聘用的人工审核员人数应足以应对预期处理内容的复杂性和数量。

102.平台还应明确说明是否与第三方内容审核服务提供商、外部组织或专家合作，以帮助其做出决策，尤其是在平台本身对当地情况知之甚少的国家或地区。在此过程中，平台应始终遵循尽职调查，并避免在存在安全风险的情况下透露合作伙伴信息。

使用自动化系统进行内容审核与策展

103.在适当情况下，数字平台应定期委托外部审计，并采取有约束力的后续步骤，对用于内容审核、策展及推荐机制的自动化和人工工具进行审计，以确保其精确性和准确性，并确定在不同内容类型、语言、文化和背景下可能存在的偏见或歧视；数字平台还应审查其语言能力以及在不同管辖区的一致使用情况。如第87款所述，数字平台还应定期委托有关方面对其广告系统在人权、文化

37. 所采取的措施应始终与其试图保护的合法目标相称。删除和下架内容以及暂停或封禁账户应是最后可能的手段，应在极端情况下作为极端手段使用。

多样性和性别平等方面的影响进行独立评估。这些审查的结果应予以公布。³⁸

104. 数字平台应建立相关制度和程序，以便在平台服务设计的任何方面可能导致放大国际人权法和标准所允许限制的内容时，根据本《准则》的规定进行识别并采取必要行动。
105. 平台还应确保，策展和推荐系统（包括人工和自动化工具）不会放大根据国际人权法和标准可允许限制的内容。
106. 平台还应确保，根据国际人权法和标准可允许限制的内容不会仅仅因为自动化策展机制或推荐机制的语言限制而被这些机制放大。
107. 数字平台应能向治理系统解释自动化系统的使用和影响，包括此类工具对数据收集、定向广告以及内容（包括艺术和选举相关内容）的披露、分类和/或删除的影响程度。
108. 数字平台应为用户提供调整内容策展和审核系统的选项。用户应有能力控制他们所看到的内容，并且应能轻松了解如何访问有关热门话题的各种来源和观点。还可以要求平台为用户提供选项，用于管理个人数据的收集以及内容推荐者对明确或推断的偏好做出回应的程度。
109. 数字平台不得使用直接从儿童处获得的或间接获得的个人数据、或从其他来源推断的儿童数据进行特征分析。

38. 一种选择是按照全球标准进行独立审计和评估，最好由独立机构进行核实，这样无论监管安排如何，它们都可以使用相同的报告。

通知

110. 当用户的内容被删除时，数字平台应通知用户并说明原因。这将使用户了解为什么要对其内容采取这一行动、使用的方法（通过自动化手段还是经过人工审核）以及是根据哪些平台规则采取行动。数字平台还应制定相关程序，允许用户对此类决定提出申诉（请参阅第125-128款）。这一规定可能因企业规模和用户对有关行动提出申诉的有效补救程序的程度而异。



原则3、平台保持透明。

- 111.** 数字平台应定期向公众和治理系统报告其如何遵守透明度和可解释性原则，以及它们在履行服务条款和社区标准方面的表现。这包括他们对政府要求提供信息或删除内容的回应。³⁹在实践中，这一规定的实施可能需要根据公司规模而有所不同，以限制对小型公司和初创公司造成的负担。
- 112.** 透明度应是富有意义的——提供的信息应该尽可能清晰简洁，必要时也应尽可能详细复杂。透明度不仅仅是提供法律文本或数据转储，而是为利益攸关方提供做出明智决策所需的信息。
- 113.** 本《准则》中提出的透明度标准可被视为所有公司在任何治理系统范围内都应达到的最低标准。

39. 关于数字平台透明度的指导意见请参阅教科文组织在《让阳光照耀进来：数字时代的透明度和问责制》中提出的26项高级别原则。https://unesdoc.unesco.org/ark:/48223/pf0000377231_chi。

富有意义的透明度

- 114.** 数字平台透明度机制的有效性应根据国际标准，通过定性和实证定量评估方法进行独立评估，以确定为实现富有意义的透明度而提供的信息是否达到了目的。报告应定期公布。
- 115.** 数字平台应发布信息，概述其如何确保将人权和正当程序考虑因素纳入内容审核与策展政策和做法的各个阶段。这类公开信息应包括：

与数字平台服务条款相关的透明度

- a.** 平台服务条款中规定的用于内容审核与策展的任何措施，包括被禁内容或被禁用户列表等。
- b.** 有关用于执行其服务条款和制裁用户的程序的任何信息，以及政府对删除、限制或推广内容的要求/请求。
- c.** 应在其服务条款中以易于获取的格式公布有关限制使用其服务条款的原因的信息。
- d.** 关于数字平台根据其服务条款被视为禁止的内容类型或将采取行动打击的内容类型的信息，以及所采取的措施，包括数字平台将永久关停或暂停用户账户的情况。

与内容审核与策展政策和做法的实施相关的透明度

- e.** 内容的审核与策展方式，包括自动化方式和人工审核方式，以及根据服务条款或政府要求/请求删除或封禁的内容。这应包括有关这些系统产生的实际成果、结果及影响的定量和定性信息。
- f.** 应定期以可访问的格式向用户传达内容审核与策展政策方面的任何变化。

- g.** 出于内容审核与策展目的而使用自动化手段的任何行为，包括说明自动化手段在审查过程中的作用，以及自动化手段在实现这些目的方面的益处和局限性的任何指标。
- h.** 为保护表达自由以及信息和多样性文化内容获取而实施的与内容审核与策展有关的任何保障措施（包括应政府要求），特别是涉及新闻、艺术和文化内容以及知识产权等公共利益的内容。
- i.** 关于聘用或分包的人工审核员人数的信息，关于他们在当地语言和当地环境方面的专业知识性质的信息，以及关于他们是内部员工还是承包商的信息。
- j.** 如何收集、使用、披露、存储和共享个人数据，以及如何处理用户的个人数据，包括使用哪些个人和敏感数据来做出算法决策，以便进行内容审核与策展。这还包括如何与其他实体共享个人数据，以及平台间接获取哪些个人数据，例如，通过用户特征分析或与数字生态系统其他部分的互操作性。

与用户投诉机制有关的透明度

- k.** 与有关删除、封禁或拒绝封禁内容的申诉以及用户如何访问投诉程序相关的信息。此类信息应包括关于已收到的、已处理的、已受理的以及已驳回的申诉的定量和定性信息，关于此类申诉结果的信息，以及关于从国家官员处收到的投诉和所采取行动的信息。

数字平台广告做法的透明度

- l.** 对于将广告作为其商业模式一部分的数字平台，有关政治广告和公共利益广告的信息，包括有关创作人员和广告付费者的信息，应保留在可公开访问的在线图书馆中。
- m.** 广告和数据收集的做法以及广告系统的人权和性别影响评估结果。
- n.** 允许个人了解向其展示特定广告的依据的信息。
- o.** 完全由机器生成的内容应标记为此类内容。

用于研究目的的数据访问

116. 数字平台应为经过审查的研究人员提供了解数字平台影响所必需的非个人数据和假名数据。应通过自动化手段，如应用程序接口（API）或其他允许分析上述数据的开放和可访问的技术解决方案，应要求持续提供此类数据。
117. 在涉及公共利益时，并且在确定的情况下，访问是相称且必要的，则数字平台应向记者和宣传团体提供非个人数据访问权限。需要有额外的保障措施来保护用户的隐私和个人数据，如通过共享前的去标识化和取样等不同措施来确保数据集的匿名化，以及保护企业的专有信息、商业秘密和对商业机密的遵守。
118. 平台应为数据访问建立可靠接口，并应根据性别和其他相关交叉因素（如种族、民族、年龄、社会经济地位、残疾等）提供分类数据。治理系统应确定哪些信息对研究目的有用、相称且合理。



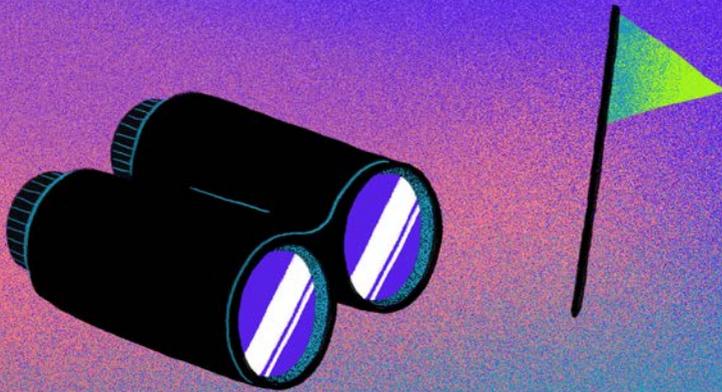
原则4、平台为用户提供信息和工具。

语言和可访问性

- 119. 平台应在其运营所在的每个国家以官方语言和主要语言提供完整的服务条款，确保能够以用户自己的语言回应用户，平等地处理用户投诉，并有能力以用户的语言对内容进行审核与策展。可以部署自动语言翻译器，以提供更大的语言可访问性，但由于其技术限制，应对其准确性进行监测。
- 120. 平台应确保报告、通知和申诉程序以用户与服务互动时使用的语言提供。

- 121.** 在儿童可能访问数字平台的情况下，这些平台应为所有儿童提供平等有效的信息获取途径，并确保保护儿童的表达自由和隐私。⁴⁰ 服务条款和社区标准应以适合儿童年龄的语言提供，并酌情从不同儿童群体的角度制定；应特别关注残疾儿童的需求，确保他们享有上一节所述的同等程度的信息获取权。
- 122.** 应始终考虑残疾人的权利，特别关注他们与平台互动的方式以及就平台提出投诉的方式。平台应实施必要调整，以提供与其服务条款、报告、通知及申诉相关的可访问信息。

40. 人权高专办。2021。《关于与数字环境有关的儿童权利的第25号一般性意见（2021年）》。
<https://www.ohchr.org/zh/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>。



原则5、平台对相关利益攸关方担责。

用户举报

- 123.** 平台应为用户和非用户或代表其利益的第三方建立举报机制，以便举报潜在的违规行为。应为弱势和边缘化群体成员建立有效便捷的投诉机制。数字平台在回应用户投诉时，还应具备了解当地背景条件的手段，确保系统设计具有文化敏感性。应为儿童建立便于快捷方便使用的特别举报机制。
- 124.** 用户举报系统应优先处理与威胁用户的内容有关的问题，确保快速响应，并在必要时提供特定的问题升级渠道或举报提交方式。在涉及侵犯人权行为，包括基于性别的暴力和骚扰时，这一点尤其重要。
- 125.** 各公司应努力防止通过虚假协同行为滥用举报系统。

用户申诉和补救

- 126.** 应建立有效的平台内和外部用户补救机制，使用户（以及受特定内容影响的非用户）能够表达他们的关切，并获得适当的补救。这应该包括一个采用当地语言的清晰、易于访问、首选、可信⁴¹且易于理解的投诉举报渠道，并将申诉结果通知用户。
- 127.** 申诉机制应遵循联合国《工商企业与人权指导原则》中概述的有效申诉机制七项原则：合法性、可及性、可预测性、公平性、透明度、权利兼容性及持续学习。
- 128.** 当用户的内容被删除、被明确标记、在评论或转发分享或广告关联方面受到限制，或在放大或推荐（有别于“有机/算法式”放大和推荐）方面受到特别限制时，数字平台应通知用户并解释申诉程序及原因。这将使用户能够了解对其内容采取这类行动的原因、使用的方法（自动化手段或人工审核）以及是根据哪些平台规则采取行动。平台还应允许用户对此类决定提出申诉并寻求适当的补救。
- 129.** 各公司应努力确保申诉和补救系统不会被虚假协同行为滥用。

41. 确保用户安全并遵守国际人权法和标准。



针对具体情况的规定

保护处于弱势和边缘化境地的所有个人、妇女和女童以及因行使表达自由权和信息获取权而可能面临风险的专业人员的权利，如记者、艺术家、人权维护者和环境维护者

130. 数字平台应为妇女和女童、来自弱势和边缘化群体的用户以及记者、艺术家、人权维护者和环境维护者提供充分的特殊保护。为了实现这一点，数字平台应：

- a.** 定期进行人权和性别影响评估，包括对其政策、审核系统和算法方法进行评估，以确定弱势和边缘化群体、妇女和女童以及记者、艺术家、人权维护者和环境维护者的系统性风险，并调整政策和做法以减轻此类风险。
- b.** 使用隐私保护技术，为外部研究人员提供平台内部数据的访问权限，以帮助识别网络性别暴力的算法放大或新兴技术引发的其他暴力趋势。
- c.** 设立专门的包容性工程团队，对他们进行专门培训，以开发审核与策展方面的算法解决方案。

- d. 开发和启动包容性结构化社区反馈机制，以应对新技术中的性别偏见、文化偏见及其他偏见。
- e. 评估其负责处理独立新闻出版商及其服务托管的新闻内容的系统和程序的人权影响。
- f. 确保独立新闻机构在数字平台上享有平等待遇。
- g. 建立相关程序，防止可能滥用举报规则和审核机制，特别是针对弱势和边缘化群体、妇女和女童以及记者、艺术家、人权维护者和环境维护者进行审查的恶意滥用。

确保选举公正的具体措施⁴²

- 131. 数字平台应认识到它们通过维护选举公正性在支持民主制度方面的作用。数字平台应建立具体的风险评估程序，保证在重大国家选举活动、重要地区选举或宪法公投（如总统制下的立法机构或国家元首选举）的筹备阶段和活动期间选举周期的完整性。
- 132. 这些评估必须保持透明，符合人权尽职调查的要求，并在所有相关选举利益攸关方的参与下进行。评估应在选举活动举行之前进行，以便采取具体措施减轻已查明的风险。鉴于针对女性选民、候选人、活动家、民选代表及选举管理官员的网络暴力有所增加，评估应包括性别方法。
- 133. 数字平台应做出合理努力，确保用户能够根据国际人权法获取各种信息和思想。尤其是，它们应确保自动化工具不会妨碍用户获取与选举相关的内容和不同观点。

42. 更多信息请参阅教科文组织《数字时代的选举：选举从业人员指南》（Elections in digital times: A guide for electoral practitioners）（2022年）[<https://unesdoc.unesco.org/ark:/48223/pf0000382102>]，以及《联合国意见和表达自由权特别报告员、欧洲安全与合作组织（欧安组织）媒体自由问题代表和美洲国家组织（美洲组织）表达自由问题特别报告员关于数字时代表达自由的联合宣言》（UN Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, and the Organization of American States (OAS) Special Rapporteur on Freedom of Expression）（2020年）https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/JointDeclarationDigitalAge_30April2020_EN.pdf。

134. 作为评估的一部分，数字平台应审查其关于政治广告的产品、政策或做法是否任意限制了候选人或政党传播信息的能力。
135. 数字平台应做出合理努力，处理选举周期内根据国际人权法和标准可允许限制的内容。应考虑促进独立事实核查、广告档案、公共警报及其他措施。根据每个司法管辖区的具体情况，可能有必要与相关的官方独立监管机构接触。
136. 数字平台应酌情对其使用的任何自动化工具的使用情况和实际影响保持透明（但不一定要公开这些工具的具体运行编码），包括这些工具影响数据采集、定向广告、内容共享、排名和/或删除的情况，尤其是与选举有关的内容。
137. 数字平台还应在选举前和选举期间与所有相关利益攸关方及其治理系统接触，以便在管理员或用户/选民提出问题时建立沟通途径。根据每个司法管辖区的具体情况，可能有必要与相关的官方监管机构接触。
138. 接受旨在影响选举周期的广告的数字平台，应将此类内容明确标识为广告。数字平台的服务条款应明确规定数字平台有责任对资金数额、提供资金的实体和广告实体保持透明，并始终如一地对此类广告适用平等的内容审核与策展规则。
139. 数字平台应跟踪政党和代表政党的个人发帖的货币化情况。
140. 平台应向公众披露有关此类广告/促销所针对的特定人群的信息。
141. 平台应将这些广告及所有相关信息保留在可公开访问和定期更新的在线图书馆中。

应对紧急情况、武装冲突及危机的具体措施

142. 作为一项人权保障措施，数字平台应开展人权尽职调查，以应对危机、武装冲突局势及其他紧急情况（包括公共卫生紧急情况）。这种尽职调查应分析公司运营、产品、服务及广告系统对危机和冲突动态的人权影响。

143. 在武装冲突和危机期间，平台应：

- a.** 确保在冲突环境中对内容的审核包括强有力的人工审查，并结合相关语言以及当地和区域背景方面的专业知识。
- b.** 促进事实核查。
- c.** 建立与相关利益攸关方（包括在受冲突影响地区和高风险地区运营的利益攸关方）进行有意义的直接接触的渠道。
- d.** 与值得信赖的合作伙伴、独立媒体组织以及其他可靠举报者发展合作。
- e.** 建立紧急情况早期预警系统和明确的升级系统，以帮助检测对个人人身安全迫在眉睫的伤害。
- f.** 实施相关政策，以便限制和跟踪与武装冲突相关的有害内容的货币化。
- g.** 保存侵犯人权或战争罪的所有潜在证据，允许适当的国家或国际问责机制查阅这些存档材料。

144. 风险评估可能要求数字平台制定相关程序，处理用户同时提出许多行动请求的情况，就像在社会动荡或大规模侵犯人权的背景下有时会发生的那样。治理系统应认可联合国机构和专家在此类情况下“加强”人权尽职调查的现有指导。



总 结

- 145.** 数字平台为个人和社会提供了大量交流、参与和学习的机会。它们为弱势和边缘化群体提供了巨大的潜力，使交流空间民主化，并提供机会，让不同的声音彼此接触、彼此听到、彼此看到。然而，由于在应对关键风险方面缺乏前瞻性，近年来这些平台的潜力逐渐被削弱。
- 146.** 《准则》的目的是在数字平台治理中保障表达自由权，包括信息获取及其他人权，同时处理根据国际人权法和标准允许限制的内容。推而广之，以人权为基础的数字平台治理将进一步促进文化多样性、文化表现形式及多样性文化内容。⁴³《准则》概述了尊重人权的治理系统，并推进了有助于降低风险的系统性内容审核与策展程序。本《准则》强调了所有影响数字平台上表达自由和信息获取的治理系统都应遵循的总体原则——这些原则与具体的监管安排和专题重点无关，只要这些安排符合本《准则》的规定即可。

43. 教科文组织2005年《保护和促进文化表现形式多样性公约》第一条和第四条。根据该《公约》，“文化内容”系指源于或表现文化特性的象征意义、艺术层面及文化价值。此外，“文化表现形式”系指由个人、群体和社会的创造力产生并具有文化内容的表现形式。

- 147.** 《准则》是通过2022年9月开始的多利益攸关方磋商进程制定的。本现行《准则》是广泛磋商的结果，教科文组织通过三次公开磋商，⁴⁴收到了1540多份意见书，其中包含1万多条意见。来自134个国家的不同利益攸关方群体参与了这一进程。这是教科文组织有史以来规模最大、最公开的磋商活动之一。
- 148.** 下一步将进一步讨论如何将《准则》付诸实施，并将其作为一种宣传工具，目标是在数字环境中保护表达自由、信息获取及其他所有人权。

44. 这三次公开磋商分别于2022年12月至2023年1月、2023年2月至3月以及2023年4月至6月举行。

附录

资源

联合国

《工商企业与人权指导原则》。(2011)

https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinessshr_en.pdf

《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》。(2012)

<https://www.ohchr.org/zh/documents/outcome-documents/rabat-plan-action>

联合国意见和表达自由权特别报告员：《基于人权的在线内容审核方法》(A human rights approach to online content moderation)。(2018)

https://www.ohchr.org/sites/default/files/Documents/Issues/Expression/Factsheet_2.pdf

联合国儿童权利委员会：《关于与数字环境有关的儿童权利的第25号一般性意见》。(2021)

<https://www.ohchr.org/zh/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

联合国秘书长报告：《打击虚假信息，促进和保护人权与基本自由》。(2022)

<https://www.ohchr.org/sites/default/files/2022-03/NV-disinformation.pdf>

联合国促进并保护言论和表达自由权问题特别报告员：《A/77/288：武装冲突期间的虚假信息与意见和表达自由》。(2022)

<https://www.ohchr.org/zh/documents/thematic-reports/a77288-disinformation-and-freedom-opinion-and-expression-during-armed>

秘书长《我们的共同议程政策简报8：数字平台上的信息完整性》

<https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-zh.pdf>

秘书长《我们的共同议程政策简报5：全球数字契约》

https://digitallibrary.un.org/record/4011891/files/%5EEOSG_2023_5%5E--EOSG_2023_5-ZH.pdf

教科文组织

A. 普德法特 (Puddephatt, A.)，2021。《让阳光照耀进来：数字时代的透明度和问责制》。巴黎：教科文组织，https://unesdoc.unesco.org/ark:/48223/pf0000377231_chi

教科文组织。2021。《表达自由的合法限制——三段测试》。[视频] <https://www.youtube.com/watch?v=hyJaz8KZqXE>

术语参考文献

煽动或描绘性别暴力的内容

- 联合国人权理事会。2018。《暴力侵害妇女及其原因和后果问题特别报告员关于从人权角度审视对妇女和女童的网络暴力的报告》。<https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/184/57/PDF/G1818457.pdf?OpenElement>
- 人权高专办。2021。《促进和保护意见和表达自由权特别报告员伊雷内·汗的报告》。<https://www.ohchr.org/en/statements/2022/02/statement-irene-khan-special-rapporteur-promotion-and-protection-freedom-opinion>

文化多样性

“文化多样性”指各群体和社会借以表现其文化的多种不同形式。这些表现形式在他们内部及其间传承。文化多样性不仅体现在人类文化遗产通过丰富多彩的文化表现形式来表达、弘扬和传承的多种方式，也体现在借助各种方式和技术进行的艺术创造、生产、传播、销售和消费的多种方式。

- 教科文组织。2005。《保护和促进文化表现形式多样性公约》。巴黎：教科文组织。https://unesdoc.unesco.org/ark:/48223/pf0000246264_chi

虚假信息 and 错误信息

- 《虚假信息与意见和表达自由》。《促进和保护意见和表达自由权特别报告员伊雷内·汗的报告》。<https://digitallibrary.un.org/record/3925306?ln=zh>

仇恨言论

- 联合国。2019。《联合国关于仇恨言论的战略和行动计划》。<https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml>

独立监管机构

独立监管机构的权力和责任由公共法律文书规定，并有权管理自己的资源；其成员以独立方式任命，并受法律保护，不得无故解雇。在这种情况下，监管机构的决定无需其他任何政府实体的事先批准，除法院或事先成立的申诉小组外，任何实体都不能推翻监管机构的决定。决策独立的制度基石是组织独立（在组织上独立于现有的部委和部门）、财务独立（专款专用、安全充足的资金来源）和管理独立（内部行政自主权和保护免受无正当理由解雇）。

- 伊芙·萨洛蒙。2016。《广播监管准则》(Guidelines for Broadcasting Regulation)。巴黎：教科文组织。
<https://unesdoc.unesco.org/ark:/48223/pf0000144292>
- A.布朗 (Brown, A.)、J.斯特恩 (J. Stern)、B.特南鲍姆 (B. Tenenbaum) 及D.根瑟 (D. Gencer) 2006。《基础设施监管体系评估手册》(Handbook for Evaluating Infrastructure Regulatory Systems)。华盛顿特区。世界银行集团。
<http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-6579-3>

监管机构

对私人行为体进行监督、监测和问责的机构。

数字平台治理准则

保障表达自由和信息获取的多方合作方法

本《准则》是通过与多方利益攸关方的广泛磋商起草的，旨在数字平台治理中保障表达自由权和信息获取权。《准则》提出了一种多方合作方法，概述了国家和数字平台在履行人权义务方面的主要责任，并界定了政府间组织、民间社会、媒体、学术界、技术界和其他利益攸关方的作用。

www.unesco.org/en/internet-trust/guidelines

internetconference@unesco.org



unesco

联合国教育、
科学及文化组织



9 789235 000528