

Directrices para la gobernanza de las plataformas digitales

Salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas



Publicado en 2023 por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 7, place de Fontenoy, 75352 París 07 SP, Francia

© UNESCO 2023

ISBN 978-92-3-300215-9



Esta publicación está disponible en acceso abierto bajo la licencia Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Al utilizar el contenido de la presente publicación, los usuarios aceptan las condiciones de utilización del Repositorio UNESCO de acceso abierto (<https://www.unesco.org/es/open-access/cc-sa>)

Título original "Guidelines for the Governance of Digital Platforms: Safeguarding freedom of expression and access to information through a multistakeholder approach"

Publicado en 2023 por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

Los términos empleados en esta publicación y la presentación de los datos que en ella aparecen no implican toma alguna de posición de parte de la UNESCO en cuanto al estatuto jurídico de los países, territorios, ciudades o regiones ni respecto de sus autoridades, fronteras o límites.

Las ideas y opiniones expresadas en esta obra son las de los autores y no reflejan necesariamente el punto de vista de la UNESCO ni comprometen a la Organización.

Diseño de cubierta y diseño gráfico: Luiza Maximo

Ilustraciones: Plastic Horse/Grand Matter

Traducción: TRADUCTEO

Imprimido por la UNESCO

Impreso en París

B R E V E R E S U M E N

Directrices para la gobernanza de las plataformas digitales

Salvaguardar la libertad de expresión y los derechos de acceso a la información mientras se aborda la desinformación, el discurso de odio y las teorías de conspiración requiere la adopción de un enfoque basado en la participación de múltiples partes interesadas.

Esta es la razón por la que la UNESCO, la principal agencia de las Naciones Unidas para la promoción y protección de la libertad de expresión y el acceso a la información, elabora Directrices para la gobernanza de las plataformas digitales.

Las Directrices describen un conjunto de deberes, responsabilidades y funciones de los Estados, las plataformas digitales, las organizaciones intergubernamentales, la sociedad civil, los medios de comunicación, el mundo académico, la comunidad técnica y otras partes interesadas para crear un entorno en el que la libertad de expresión y la información estén en el centro de las plataformas digitales.

Las Directrices se elaboraron a través de una consulta de múltiples partes interesadas que recogió más de 10.000 comentarios de 134 países. Estas consultas a escala global fomentaron la participación inclusiva, asegurando que se escuchara una diversidad de voces, incluidas las de grupos en situación de marginación y vulnerabilidad.

Cultivar una Internet de confianza es una responsabilidad compartida entre todas las partes interesadas. Nos llama a todos a mantener un entorno propicio para la libertad de expresión y el derecho a la información..



10.000
COMENTARIOS DE
134 PAÍSES



"Puesto que las guerras nacen en la mente de las mujeres y de los hombres, es en la mente de las mujeres y de los hombres donde deben erigirse los baluartes de la paz"

Directrices para la gobernanza de las plataformas digitales

Salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas

6 **Prefacio**

8 **El objetivo de las Directrices**

12 **Introducción**

16 **Estructura de las Directrices**

18 Entorno propicio

29 Principios de los sistemas de gobernanza

40 Principio 1. Aplicar el principio de debida diligencia en materia de derechos humanos

43 Principio 2. Adherir a la normativa internacional en materia de derechos humanos, incluido en el diseño de la plataforma, la moderación y la curación de contenidos

48 Principio 3. La transparencia de las plataformas

53 Principio 4. Información y herramientas a disposición de las personas usuarias

55 Principio 5. Las plataformas rinden cuentas ante las partes interesadas pertinentes

57 Disposiciones particulares

62 **Conclusión**

64 **Apéndice**

Prefacio de la Sra. Audrey Azoulay, Directora General de la UNESCO:

Salvaguardar la libertad de expresión y el acceso a la información: directrices para un enfoque de múltiples partes interesadas de la gobernanza de las plataformas digitales

Septiembre de 2023

En 2023, el 60 % de la población mundial, es decir, 4 750 millones de personas, habrá utilizado las redes sociales para expresarse, informarse y afirmarse.

El ámbito digital, espacio de libertad y nuevo foro de expresión y debate, entrelaza nuestras relaciones sociales, identidades y vidas.

Estas plataformas se han convertido en amplificadores de las personas que defienden la igualdad y la libertad, dando voz a las que no la tienen y ofreciendo un refugio a diversas formas de expresión.

Sin embargo, estas mismas redes sociales, cuyo nombre promete tanto, se convierten con demasiada frecuencia en burbujas de aislamiento, capullos de desinformación, que a veces fomentan teorías conspirativas y violencia extrema.

Como espacios virtuales de interacción social, están sometidas a algoritmos diseñados para monopolizar nuestra atención, favoreciendo inadvertidamente la desinformación y el discurso de odio al anteponer los clics a la certeza, la probabilidad a la prueba.

Sin embargo, si ya no podemos distinguir la ficción de la realidad, la falsedad de la verdad, los cimientos de nuestras sociedades se desmoronan: la democracia, el diálogo y el debate, esenciales para abordar los grandes retos contemporáneos, se vuelven imposibles.

Ante la naturaleza global de estas cuestiones, necesitamos desarrollar respuestas coherentes en todo el mundo, para evitar la fragmentación de normativas o enfoques que comprometan los derechos humanos.

Es precisamente este desafío mundial el que la UNESCO debe afrontar, ya que constituye el núcleo de nuestro mandato.

Desde su creación, nuestra Organización ha trabajado para fomentar «el conocimiento y la comprensión mutuos de las naciones», en particular mediante «la libre circulación de las ideas por medio de la palabra y de la imagen», como subraya nuestra Constitución.

Este compromiso llevó a la UNESCO a publicar en 2005 unas directrices para la regulación

de la radiodifusión. Más recientemente, nuestra Recomendación sobre la Ética de la Inteligencia Artificial, adoptada en 2021 por nuestros 193 Estados Miembros, estableció un marco humanista para la evolución de esta tecnología.

Fiel a sus valores y a su historia, la UNESCO ha trabajado para elaborar las innovadoras directrices que se presentan en esta publicación. Con ellas se pretende combatir la desinformación y el discurso de odio al tiempo que se fomenta la transparencia y la libertad de expresión en las plataformas.

Estos esfuerzos han sido dirigidos por la Declaración de Windhoek +30 de 2021, cuyos principios fueron adoptados por todos los Estados Miembros de la UNESCO. La Declaración identificó tres pilares de acción: defender la transparencia de las plataformas, garantizar la viabilidad de los medios de comunicación y fomentar el pensamiento crítico entre las personas usuarias.

Este empeño, que culmina con estas directrices, es el resultado de amplias consultas, enriquecidas con más de 10 000 comentarios, lo que lo convierte en una de las consultas más exhaustivas llevadas a cabo por las Naciones Unidas. Solo la conferencia «Por un internet confiable», organizada por la UNESCO en febrero de 2023, reunió a más de 4 000 partes interesadas de 134 países.

Estas directrices proponen medidas justas, claras y compartidas: moderadores en línea en todas las lenguas, incluidas las indígenas; una mayor transparencia de las plataformas y su financiación, con una mejor evaluación de los riesgos; el establecimiento de reguladores independientes; la promoción del pensamiento crítico; el apoyo a la igualdad de género; y, sobre todo, la salvaguardia y el fortalecimiento de la libertad de expresión, la diversidad cultural y otros derechos humanos.

Más allá de las realidades actuales de las plataformas digitales, este texto aborda también los retos del futuro, especialmente los que plantea la inteligencia artificial generativa.

La UNESCO se compromete a ayudar a los Estados Miembros, a la sociedad civil y a los principales actores digitales a adoptar este texto para que el modo de funcionamiento de las plataformas se ajuste plenamente a nuestros valores y a la normativa internacional en materia de derechos humanos.

No perdamos de vista nuestro objetivo: combatir el discurso de odio y la desinformación preservando la libertad de expresión. No se trata de una contradicción: al reforzar el acceso a una información libre y fiable, reforzamos también la libertad de pensamiento y de expresión.

En palabras de Hannah Arendt: «La libertad de opinión es una farsa a menos que se garantice la información objetiva y que los hechos mismos no sean discutibles».

Directrices para la gobernanza de las plataformas digitales

Salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas

Octubre 2023

El objetivo de las Directrices

1. Basándose en los principios, convenciones y declaraciones pertinentes, la UNESCO ha elaborado, mediante consultas a múltiples partes interesadas y un diálogo mundial, el presente documento: *Directrices para salvaguardar la libertad de expresión y el acceso a la información en la gobernanza de las plataformas digitales: un enfoque de múltiples partes interesadas* (las Directrices)¹.
2. El objetivo de las Directrices es salvaguardar el derecho a la libertad de expresión en la gobernanza de las plataformas digitales, incluidos el acceso a la información y otros derechos humanos, al tiempo que se tratan aquellos contenidos que pueden restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos. Por extensión, una gobernanza de las plataformas digitales basada en los derechos humanos promovería aún más la diversidad cultural, la expresión cultural y los contenidos culturales diversos². Las Directrices describen un sistema de gobernanza que respeta los derechos humanos y promueven procesos de mitigación de riesgos basados en sistemas para la moderación y la curación de contenidos. Estas Directrices ponen de relieve los principios rectores que deben seguirse en todos los sistemas de gobernanza que afectan a la libertad de expresión y al acceso a la información, independientemente del tipo de regulación específico y el enfoque temático, siempre que esas disposiciones estén en consonancia con las establecidas en estas Directrices.

1. La versión original de este documento está en inglés.

2. Convención de 2005 sobre la Protección y Promoción de la Diversidad de las Expresiones Culturales, artículos 1 y 4. Según la Convención, se entiende por «contenido cultural» el sentido simbólico, la dimensión artística y los valores culturales que emanan de las identidades culturales o las expresan. Además, las «expresiones culturales» son las expresiones resultantes de la creatividad de personas, grupos y sociedades, que poseen un contenido cultural.

3. Las Directrices reconocen que la aplicación de normativas y reglamentos en todo sistema de gobernanza debe adherirse a la normativa internacional en materia de derechos humanos, incluido el Artículo 19 (3) del *Pacto Internacional de Derechos Civiles y Políticos* (PIDCP)³, que establece que toda restricción a la libertad de expresión debe estar prevista por la ley, perseguir un fin legítimo y ser necesaria y proporcionada; así como el Artículo 20 del PIDCP y otras normativas internacionales, en particular las interpretaciones reconocidas de las disposiciones de estos tratados por el Comité de Derechos Humanos de la ONU, los tribunales internacionales y regionales de derechos humanos y el Plan de acción de Rabat sobre la prohibición de la apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia⁴.
4. Las Directrices se centran en proteger y promover la normativa en materia de derechos humanos y permitir la existencia de una pluralidad de plataformas, incluidas las descentralizadas, y un ecosistema que cuente con una diversidad de normas de contenido y sistemas de moderación.
5. Las Directrices pueden servir como un recurso para diversas partes interesadas: para las personas responsables de la formulación de políticas públicas en la identificación de objetivos legítimos, principios de derechos humanos y procesos inclusivos y participativos que podrían considerarse necesarios para tal fin; para los organismos reguladores y de gobernanza encargados de la implementación y evaluación de las políticas, códigos de conducta o de la regulación; para las plataformas digitales en sus políticas y prácticas; y para otras partes interesadas, como la sociedad civil, en sus esfuerzos de promoción y rendición de cuentas. Los medios de comunicación también pueden servirse de estas Directrices en sus esfuerzos por responsabilizar a los actores poderosos.
6. Las Directrices están diseñadas para fundamentar tanto los procesos de gobernanza específicos a la gestión de contenidos en plataformas digitales como los procesos de gobernanza que ya se están implementando en otras áreas que pueden tener un impacto en el ejercicio de la libertad de expresión y el acceso a la información y en la diversidad de contenidos culturales y que deben considerarse a la luz de los cambios en el entorno digital (tales como elecciones, protección de datos y reglamentos de protección de la

3. *Pacto Internacional de las Naciones Unidas sobre los Derechos Civiles y Políticos* (PIDCP). 1966. <https://www.ohchr.org/es/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

4. Las Directrices deben leerse en consonancia con todos los instrumentos internacionales fundamentales en materia de derechos humanos. Estos instrumentos pueden consultarse en: <https://www.ohchr.org/es/core-international-human-rights-instruments-and-their-monitoring-bodies>

competencia). Dependiendo de la cuestión y del contexto jurisdiccional, tales procesos de gobernanza pueden adoptar la forma de una combinación de pilares complementarios – autorregulación, corregulación y regulación estatutaria –, estructurados de conformidad con la normativa internacional en materia de derechos humanos (véase la sección «El sistema de gobernanza» más adelante). Dichos procesos de gobernanza deberán llevarse a cabo con la participación de múltiples partes interesadas, de manera abierta, transparente, proporcional y basado en evidencia. Para ello, estas Directrices deben ser un documento vivo, sujeto a revisiones y actualizaciones periódicas, que tenga en cuenta las experiencias adquiridas de su implementación, así como los cambios e impactos tecnológicos posteriores.

7. Estas Directrices están diseñadas para contribuir de forma práctica a ampliar los esfuerzos para construir un modelo de gobernanza digital centrado en el ser humano. También forman parte de un conjunto más amplio de acciones necesarias para lograr un desarrollo sostenible. Servirán para:
 - a. Fomentar y contribuir al desarrollo de redes mundiales de múltiples partes interesadas y espacios comunes para debatir y compartir buenas prácticas sobre la gobernanza de las plataformas digitales, reuniendo diferentes visiones y un amplio espectro de perspectivas.
 - b. Servir como herramienta para que todas las partes interesadas relevantes aboguen por una regulación que respete los derechos humanos y que los gobiernos y las plataformas digitales rindan cuentas.
 - c. Promover los enfoques de políticas basados en evidencia y en los derechos humanos.
 - d. Fomentar la mayor convergencia posible a nivel mundial en las políticas de gobernanza de plataformas para evitar la fragmentación de Internet.
8. Las Directrices deberán **contribuir a y estar fundamentadas en los procesos que están en curso en todas las Naciones Unidas**, como la implementación de las propuestas en «Nuestra Agenda Común», incluido el desarrollo del Pacto Digital Mundial⁵, la preparación de la Cumbre para el Futuro organizada por las Naciones Unidas en septiembre de 2024 y la creación de un Código de Conducta

5. Véase el informe de políticas 5 de Nuestra Agenda Común, publicado por el Secretario General de la ONU, con referencias cruzadas al proceso que ha conducido a estas Directrices: https://digitallibrary.un.org/record/4011891/files/%5EEOSG_2023_5%5E--EOSG_2023_5-ES.pdf

que promueve la integridad de la información en las plataformas digitales⁶. Las Directrices también contribuirán al debate en torno a la próxima revisión de 20 años de la Cumbre Mundial sobre la Sociedad de la Información (CMSI) y al Foro para la Gobernanza de Internet (FGI). Además, este texto se ha beneficiado de iniciativas lideradas por otras organizaciones gubernamentales internacionales, incluidas aquellas de alcance regional, y pretende contribuir a su desarrollo.

6. Véase el informe de políticas 8 de Nuestra Agenda Común, publicado por el Secretario General de la ONU, con referencias cruzadas al proceso que ha conducido a estas Directrices: <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-es.pdf>



Introducción

9. En noviembre de 1945, la UNESCO fue creada con la misión de «contribuir a la paz y a la seguridad estrechando, mediante la educación, la ciencia y la cultura, la colaboración entre las naciones, a fin de asegurar el respeto universal a la justicia, a la ley, a los derechos humanos y a las libertades fundamentales que sin distinción de raza, sexo, idioma o religión, la Carta de las Naciones Unidas reconoce a todos los pueblos del mundo»⁷.
10. El mandato mundial de la UNESCO, que incluye la promoción de «la libre circulación de ideas por medio de la palabra y la imagen», ha guiado la labor de la Organización durante casi 80 años, como laboratorio de ideas, centro de intercambio de información, organismo de normalización, catalizador y motor de la cooperación internacional y creador de capacidades. Esta historia también dio forma al mandato de la Organización dentro del sistema de las Naciones Unidas para proteger y promover la libertad de expresión, el acceso a la información y la seguridad de los periodistas, tanto en línea como por otras vías.
11. La labor y el compromiso constantes de la UNESCO consisten en garantizar que la gobernanza de las plataformas digitales proteja y promueva la libertad de expresión, el acceso a la información y a contenidos culturales diversos, así

7. Constitución de la Organización de las Naciones para la Educación, la Ciencia y la Cultura, Artículo 1. <https://www.unesco.org/es/legal-affairs/constitution#article-i---purposes-and-functions>

como otros derechos humanos para todos, incluidos los grupos en situación de vulnerabilidad y marginación⁸.

12. Este esfuerzo se inspira en la labor de la UNESCO en el ámbito de la regulación de la radiodifusión durante varias décadas, ya que cualquier intervención gubernamental que se ocupe de cuestiones de contenido, independientemente de la fuente del contenido, siempre debe incluir la salvaguardia de la diversidad y la libertad de expresión y el acceso a la información como objetivo final. Las Directrices también contribuyen a la implementación de la Estrategia a Plazo Medio de la Organización para 2022-2029 (41 C/4)⁹.
13. En 2015, la Conferencia General de la UNESCO respaldó los principios ROAM para la Universalidad de Internet, que destacan la importancia de los **derechos humanos**, la **apertura**, la **accesibilidad** y la participación de **múltiples partes interesadas** para el desarrollo, el crecimiento y la evolución de Internet¹⁰. Estos principios reconocen la necesidad fundamental de garantizar que los espacios en línea continúen desarrollándose y utilizándose de manera propicia para alcanzar los Objetivos de Desarrollo Sostenible.
14. El enfoque de múltiples partes interesadas para el desarrollo y la aplicación de principios, normas, reglas, procedimientos de toma de decisiones y programas compartidos que configuran la evolución y el uso de Internet ha respaldado la estrategia general sistema de las Naciones Unidas, incluida la UNESCO, desde la Cumbre Mundial sobre la Sociedad de la Información (2005), y fue reafirmado por la Asamblea General de las Naciones Unidas durante el proceso de revisión decenal en 2015:

Reafirmamos, asimismo, el valor y los principios de la cooperación y el compromiso de múltiples partes interesadas, que han caracterizado el proceso de la Cumbre Mundial sobre la Sociedad de la Información desde sus inicios, reconociendo que la participación, la asociación y la cooperación efectivas de los Gobiernos, del sector privado, de la sociedad civil, de las organizaciones internacionales, de las comunidades técnicas y académicas, y todas las

8. «Grupos en situación de vulnerabilidad y marginación» se refiere a niñas, niños y adolescentes; personas con discapacidad; personas migrantes, refugiadas y solicitantes de asilo; personas LGBTI; y personas mayores.

9. El Objetivo Estratégico 3 de la Estrategia a Plazo Medio es construir sociedades inclusivas, justas y pacíficas, también mediante la promoción de la libertad de expresión. El Objetivo Estratégico 4 es fomentar un entorno tecnológico al servicio de la humanidad, a través del desarrollo y de la difusión de conocimientos y habilidades y estándares éticos. https://unesdoc.unesco.org/ark:/48223/pf0000378083_spa

10. UNESCO. «Indicadores de Universalidad de Internet». <https://www.unesco.org/es/internet-universality-indicators>

demás partes interesadas pertinentes, dentro de sus respectivas funciones y responsabilidades (especialmente con una representación equilibrada de los países en desarrollo), han sido y siguen siendo vitales en el desarrollo de la sociedad de la información¹¹.

15. La 41.ª Conferencia General de la UNESCO respaldó los principios de la Declaración de *Windhoek+30* en noviembre de 2021, luego de un proceso de consulta a las partes interesadas que comenzó durante la celebración mundial del Día Mundial de la Libertad de Prensa, en mayo del mismo año¹². La Declaración reconoció la información como un bien público y estableció, entre otras metas, tres etapas para garantizar que la información sea considerada un recurso compartido por toda la humanidad: la transparencia de las plataformas digitales, el empoderamiento de los ciudadanos a través de la alfabetización mediática e informacional y la viabilidad de los medios. Al promover la visión de la información como un bien público, la UNESCO reconoce que este derecho universal es tanto un medio como un fin para la realización de las aspiraciones humanas colectivas, incluida la Agenda 2030 para el Desarrollo Sostenible. La información empodera a la ciudadanía para ejercer sus derechos fundamentales, apoya la igualdad de género y permite la participación y la confianza en la gobernanza democrática y el desarrollo sostenible, sin dejar a nadie atrás.
16. El énfasis de las Directrices puesto en los desafíos relacionados con la libertad de expresión y el acceso a la información y a contenidos culturales diversos en el entorno digital complementa la labor de la Organización en las áreas de educación, ciencias y cultura. Esto incluye, entre otros aspectos, la Recomendación de la UNESCO sobre Ética de la Inteligencia Artificial¹³, que aboga por políticas y marcos normativos internacionales y nacionales para garantizar que las tecnologías emergentes beneficien a la humanidad en su conjunto y la *Convención de 2005 sobre la Protección y Promoción de la Diversidad de las Expresiones Culturales de la UNESCO*¹⁴ y sus Directrices sobre la implementación de la Convención en el entorno digital. Estas Directrices promueven «el respeto de las libertades fundamentales de expresión, información y comunicación y de la privacidad y

11. Asamblea General de las Naciones Unidas. 2015. «Documento final de la reunión de alto nivel de la Asamblea General sobre el examen general de la aplicación de los resultados de la Cumbre Mundial sobre la Sociedad de la Información». 70/ 125. https://unctad.org/system/files/official-document/ares70d125_es.pdf

12. UNESCO. 2021. *Declaración de Windhoek+30: la información como bien común*. https://unesdoc.unesco.org/ark:/48223/pf0000378158_spa

13. UNESCO. 2021. «Recomendación sobre la Ética de la Inteligencia Artificial». https://unesdoc.unesco.org/ark:/48223/pf0000380455_spa

14. UNESCO. 2005. *La Convención de 2005 sobre la Protección y la Promoción de la Diversidad de las Expresiones Culturales*. <https://es.unesco.org/creativity/convention>

otros derechos humanos, como prerequisites para la creación, distribución y acceso a expresiones culturales diversas. Ello incluye promover la libertad artística, como corolario de la libertad de expresión, los derechos sociales y económicos de los autores y artistas que trabajan en el entorno digital, así como la conectividad de todas las partes implicadas con quienes ellas elijan»¹⁵. El énfasis de estas Directrices también complementa la Declaración MONDIACULT de 2022, que aboga por la «regulación sustancial del sector digital, en particular de las principales plataformas» en beneficio de la diversidad cultural en línea y del acceso justo a los contenidos para todos¹⁶.

15. UNESCO. 2017. «Orientaciones prácticas para la aplicación de la Convención en el entorno digital». https://unesdoc.unesco.org/ark:/48223/pf0000370521_spa

16. UNESCO. 2022. «Conferencia Mundial de la UNESCO sobre las Políticas Culturales y el Desarrollo Sostenible (MONDIACULT 2022) – Declaración final» https://www.unesco.org/sites/default/files/medias/fichiers/2022/10/6_MONDIACULT_ES_DRAFT%20FINAL%20DECLARATION.pdf



Guidelines

Estructura de las Directrices

- 17.** Las Directrices comienzan describiendo el entorno propicio necesario para salvaguardar la libertad de expresión, el acceso a la información y otros derechos humanos, garantizando al mismo tiempo un entorno abierto, seguro y protegido para las personas usuarias y no usuarias de las plataformas digitales. Las Directrices describen las responsabilidades de las distintas partes interesadas a este respecto. Esto incluye:
 - a.** Los deberes de los Estados de respetar, proteger y hacer cumplir los derechos humanos.
 - b.** Las responsabilidades de las plataformas digitales de respetar los derechos humanos.
 - c.** El papel de las organizaciones intergubernamentales.
 - d.** El papel de la sociedad civil, de los medios de comunicación, del mundo académico, de la comunidad técnica y de otras partes interesadas en la promoción de los derechos humanos.
- 18.** A continuación, las Directrices presentan los principios básicos del sistema de gobernanza de las plataformas digitales con un enfoque de múltiples partes interesadas y basado en los derechos humanos. Esta sección expone las formas

complementarias de autorregulación, corregulación y regulación estatutaria, así como los criterios que pueden utilizarse para definir el alcance de las compañías cubiertas por la regulación estatutaria.

19. Seguidamente, las Directrices identifican la alfabetización mediática e informacional, así como el respeto a la diversidad cultural como una responsabilidad común de todas las partes implicadas en la gobernanza de las plataformas digitales.
20. Por último, describen las áreas donde las plataformas digitales deben tener sistemas y procesos para evaluar riesgos; curar y moderar contenidos basándose en la normativa internacional en materia de derechos humanos y en el respeto de la diversidad cultural tal como la define la Convención de 2005 de la UNESCO; empoderar a las personas usuarias brindando alfabetización mediática e informacional; y rendir cuentas mediante mecanismos de presentación de informes y reparación, con el fin de salvaguardar la libertad de expresión, el acceso a la información y otros derechos humanos.
21. Es importante destacar que las diferentes áreas cubiertas por estas directrices (tal como se identifican en los párrafos 17 a 21 anteriores) **deben considerarse como un conjunto**. Para salvaguardar la libertad de expresión, el acceso a la información y la diversidad de contenidos culturales es necesario tener en cuenta todos los elementos descritos anteriormente.



Entorno propicio

22. Todas las partes interesadas comparten la responsabilidad de mantener un entorno propicio para la libertad de expresión, el acceso a la información y otros derechos humanos, garantizando al mismo tiempo que exista un entorno abierto, seguro y protegido para las personas usuarias y no usuarias de las plataformas digitales¹⁷.
23. La creación de este entorno propicio no es simplemente una cuestión de ingeniería. También es un empeño que exige el compromiso de la sociedad en su conjunto y, por lo tanto, requiere soluciones que provengan de todas las partes interesadas. Todas las partes interesadas de cada sistema de gobernanza deben tomar medidas que permitan que los grupos en situación de vulnerabilidad y marginación, las mujeres y niñas, las comunidades indígenas, así como periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente, por ejemplo, ejerzan su derecho a la libertad de expresión. Toda la sociedad tiene un papel que desempeñar para que Internet sea seguro, para cuestionar los comportamientos violentos o amenazadores, para respetar los derechos de las demás personas en los intercambios en línea, para respetar la diversidad de los contenidos culturales y para ser conscientes de los sesgos inherentes a las sociedades.

17. Las palabras «seguro» y «protegido» en estas Directrices se conciben como las condiciones en las que las personas pueden confiar en que sus derechos humanos, incluidos el derecho a la libre expresión y el acceso a la información, están protegidos.

- 24.** Todas las niñas, los niños y los adolescentes tienen un estatus especial dada su etapa única de desarrollo, voz política ausente o limitada, y el hecho de que las experiencias negativas en la infancia pueden tener consecuencias de por vida o transgeneracionales¹⁸. Así pues, mientras que la protección de la libertad de expresión y el acceso a la información se aplica a todas las personas, los Gobiernos y las plataformas digitales también deben reconocer sus responsabilidades específicas hacia todas las niñas, los niños y los adolescentes¹⁹ dentro de los sistemas de gobernanza. Todas las partes interesadas deben mantener altos estándares éticos y profesionales cuando se trata de la participación de todas las niñas y niños en el entorno digital, incluyendo la protección y promoción de su libertad de expresión y acceso a la información.
- 25.** Todas las partes implicadas en la gobernanza de las plataformas digitales deben fomentar y, cuando proceda, financiar respuestas colaborativas, en las que participen organizaciones de la sociedad civil, redes de periodistas y personas investigadoras, para obtener un conocimiento más detallado de los contenidos que podrían restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos, así como de las respuestas para proteger y apoyar a mujeres y niñas, grupos en situación de vulnerabilidad y marginación, periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente, así como a las comunidades indígenas.

Deberes de los Estados de respetar, proteger y hacer cumplir los derechos humanos

- 26.** Los Estados deben respetar y promover los derechos humanos, incluido el derecho a la libertad de expresión y el derecho de acceso a la información. Las restricciones a la libertad de expresión solo pueden permitirse en las condiciones establecidas por el Artículo 19 (3) y el Artículo 20 del PIDCP. Los Estados tienen obligaciones positivas de proteger los derechos humanos contra interferencias injustificadas por parte de actores privados, incluidas las plataformas digitales, ya que tienen la responsabilidad de crear un entorno normativo que facilite el respeto de los derechos humanos por parte de las plataformas. Asimismo,

18. Véase Comité de los Derechos del Niño de las Naciones Unidas (2013), «Observación general n.º 16 (2013) sobre las obligaciones del Estado en relación con el impacto del sector empresarial en los derechos del niño», párrafo 4. Véase también la «Observación general núm. 25 (2021) relativa a los derechos de los niños en relación con el entorno digital». <https://www.ohchr.org/es/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

19. En la mayoría de los casos, se consideran como niños y niñas a las personas menores de 18 años.

tienen el deber de orientar a las plataformas digitales en relación con sus responsabilidades.

- 27.** Además, los Estados tienen la obligación de ser totalmente transparentes y rendir cuentas sobre las condiciones que imponen a las plataformas digitales, garantizando la seguridad jurídica y la previsibilidad jurídica, que son requisitos esenciales para el estado de derecho.
- 28.** Concretamente, los Estados deben:
 - a.** Promover el acceso universal y significativo a Internet y garantizar la neutralidad²⁰.
 - b.** Garantizar que todas las niñas, los niños y adolescentes tengan un acceso equitativo y efectivo al entorno digital en formas que sean significativas para ellas y ellos, y tomar todas las medidas necesarias para acabar con la exclusión digital²¹.
 - c.** Destinar recursos y redoblar los esfuerzos para cerrar la brecha digital, colmar las lagunas de información, eliminar otras barreras a las que se enfrentan los grupos en situación de vulnerabilidad y marginación, y hacer realidad el derecho de todas las mujeres y niñas a acceder a la información.
 - d.** Fortalecer el espacio cívico y promover medios libres, independientes y plurales, y apoyar la investigación independiente sobre la expresión en línea, la moderación y curación de contenidos y la rendición de cuentas de las plataformas.
 - e.** Garantizar una sólida protección a las personas periodistas (incluidas todas las mujeres periodistas), defensoras de los derechos humanos y denunciantes de buena fe, y considerar apoyar mecanismos transparentes de autorregulación de los medios que promuevan y protejan los más altos estándares de profesionalismo.
 - f.** Garantizar una sólida protección a las personas artistas, reconociendo la importancia de su labor para la renovación de la producción cultural y la promoción de la diversidad cultural, y considerar que están en el corazón mismo del tejido cultural de la sociedad.

20. En la «Declaración Conjunta sobre la Libertad de Expresión e Internet», los mandatos internacionales especiales sobre la libertad de expresión, contenían la siguiente aclaración: «Hacer efectivo el derecho a la libertad de expresión impone a los Estados la obligación de promover el acceso universal a Internet». Adoptada el 1 de junio de 2011, párrafo 6(a). <http://www.law-democracy.org/wp-content/uploads/2010/07/11.06.Joint-Declaration.Internet.pdf>

21. Véase también la «Observación general núm. 25 (2021) relativa a los derechos de los niños en relación con el entorno digital». <https://www.ohchr.org/es/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

- g.** Garantizar los derechos de las personas usuarias de plataformas digitales a la libertad de expresión, el acceso a la información, la igualdad y la no discriminación, así como proteger los derechos de las personas usuarias a la privacidad, protección de datos, asociación y participación pública.
- h.** Adoptar leyes basadas en la normativa internacional en materia de derechos humanos y garantizar su implementación efectiva para prohibir, investigar y perseguir la violencia en línea basada en género²².
- i.** Asegurar que cualquier restricción impuesta a las plataformas respete de manera consistente el alto umbral establecido para definir las restricciones a la libertad de expresión, sobre la base de la aplicación de los Artículos 19 (3) y 20 del PIDCP, respetando las condiciones de legalidad, fin legítimo, necesidad y proporcionalidad.
- j.** Disuadir encarecidamente a las personas funcionarias públicas, incluso a través de medidas como códigos de conducta profesionales, de difundir desinformación, incluida la desinformación de género²³ y la información errónea, e intimidar o amenazar a los medios de comunicación. Además, prohibir expresiones que constituyan una apología del odio nacional, racial o religioso, que inciten a la discriminación, la hostilidad o la violencia, según lo establecido por la normativa internacional en materia de derechos humanos, en consonancia con la Estrategia y Plan de Acción de las Naciones Unidas para la Lucha contra el Discurso del Odio.
- k.** Ser transparente y divulgar toda información sobre el tipo, el número y la base legal de las solicitudes que realizan a las plataformas digitales para descargar, eliminar y bloquear contenidos. Los Estados deberían poder demostrar cómo las solicitudes se ajustan a los Artículo 19 (3) y 20 del PIDCP.
- l.** Promover la alfabetización mediática e informacional para mejorar el compromiso positivo con las plataformas y desarrollar habilidades de seguridad en línea, incluidos los espacios digitales, con el objetivo de empoderar a las personas

22. Véase «A/76/258: Justicia de género y libertad de expresión – Informe de la Relatora Especial sobre la promoción y protección de la libertad de opinión y de expresión». «Todas las medidas jurídicas orientadas a restringir el discurso de odio por razón de género o la desinformación por motivos de género deberían cumplir las tres condiciones de legalidad, necesidad y proporcionalidad, y objetivos legítimos que se establecen en el Artículo 19 (3) del Pacto. Debería evitarse la tipificación, salvo en los casos más atroces de apología que constituyan una incitación». <https://www.ohchr.org/es/documents/thematic-reports/a76258-gender-justice-and-freedom-expression-report-special-rapporteur>

23. Ídem, párrafo 21: «La desinformación de género también está aumentando. Aunque se trata de un subtipo de violencia de género, tiene algunas características particulares, ya que entraña el uso de relatos falsos o engañosos basados en el género y el sexo contra las mujeres, a menudo con cierto grado de coordinación, a fin de disuadir a las mujeres de participar en la esfera pública. La desinformación de género combina tres características definitorias de la desinformación en línea: la falsedad, la intención maliciosa y la coordinación».

usuarias, en particular los grupos en situación de vulnerabilidad y marginación. Esta iniciativa debe incluir la promoción del conocimiento sobre los derechos a la libertad de expresión, a la privacidad, a la igualdad y al acceso a la justicia, así como del conocimiento sobre los medios de reclamación y reparación. Asimismo, debe recurrir a la experiencia de personas expertas en alfabetización mediática e informacional, bibliotecas, académicos y organizaciones de la sociedad civil, y promover el acceso a instituciones informativas.

- m.** Garantizar que cualquier autoridad reguladora que se encargue de la gestión de contenidos de plataformas digitales, independientemente de la temática, tenga una estructura independiente, a resguardo de intereses políticos y económicos, y cuente con sistemas de revisión externa (véanse los párrafos 68-73 de estas Directrices). Dichos sistemas de revisión pueden incluir el escrutinio legislativo y judicial, así como requisitos de transparencia y consulta con múltiples partes interesadas, y estar sujetos a la elaboración de informes anuales y a auditorías externas periódicas. Esto también implicaría el establecimiento de normas claras sobre la competencia y la autoridad del poder judicial.
- n.** Garantizar que las autoridades reguladoras cuenten con los recursos suficientes y tengan la capacidad de realizar evaluaciones conformes a los objetivos de las Directrices.
- o.** Reconocer que los sistemas de gobernanza deben aprovechar la experiencia de las personas expertas en derechos humanos, académicos y organizaciones de la sociedad civil, así como prácticas ejemplares reconocidas de otros sistemas de gobernanza.
- p.** Fomentar la cooperación internacional, incluida la cooperación triangular y Sur-Sur, entre las autoridades reguladoras y los actores judiciales, promoviendo el intercambio de buenas prácticas y conocimientos.

29. Los Estados deben abstener de:

- a.** Imponer medidas que impidan o perturben el acceso general a la difusión de información (en línea y fuera de línea), incluidos los cortes de Internet.
- b.** Imponer una obligación general de supervisión o una obligación general para que las plataformas digitales adopten medidas proactivas vinculadas con contenidos considerados como ilegales en una jurisdicción específica, o con contenidos que puedan ser restringidos lícitamente en virtud de la normativa internacional en materia de derechos humanos. Las plataformas digitales no deben ser

consideradas responsables cuando actúen de buena fe y con la debida diligencia, cuando realicen investigaciones voluntarias o tomen otras medidas destinadas a detectar, identificar, eliminar o inhabilitar el acceso a contenidos prohibidos en virtud del Artículo 20 del PIDCP, o que hayan sido restringidos según lo estipulado por el Artículo 19(3) del PIDCP.

- c. Someter al personal de las plataformas digitales a sanciones penales por un supuesto o posible incumplimiento de la normativa en relación con su trabajo de moderación y curación de contenidos.

La responsabilidad de las plataformas digitales de respetar los derechos humanos

- 30. Las plataformas digitales deben cumplir con cinco principios fundamentales:
 - a. **Actuar con la debida diligencia en materia de derechos humanos**, evaluando el impacto en los derechos humanos, incluidas las dimensiones de género y cultural, evaluando los riesgos y definiendo las medidas de mitigación.
 - b. **Adherir a la normativa internacional en materia de derechos humanos, incluido en el diseño de la plataforma, la moderación y la curación de contenidos.** Las plataformas deben respetar la normativa internacional pertinente en materia de derechos humanos, incluidos los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos. El diseño debe garantizar la no discriminación y la igualdad de trato, así como la prevención de daños. Las políticas y prácticas de moderación y curación de contenidos deben ajustarse a la normativa en materia de derechos humanos, independientemente de que estas prácticas se apliquen a través de medios automatizados o humanos, con conocimiento de las lenguas locales y el contexto lingüístico, respetando la diversidad cultural y brindando la protección y el apoyo adecuados a los moderadores humanos.
 - c. **Ser transparentes** y abiertas sobre cómo operan. Asimismo, deben contar con políticas comprensibles y auditables, así como con sistemas de medición diseñados por múltiples partes interesadas para evaluar su desempeño. Esto incluye la transparencia sobre las herramientas, los sistemas y los procesos utilizados para moderar y curar contenidos en sus plataformas, incluso en lo que respecta a las decisiones algorítmicas y sus resultados.
 - d. **Hacer que la información sea accesible** para que las personas usuarias

comprendan los distintos productos, servicios y herramientas que ofrecen, y para que tomen decisiones informadas sobre los contenidos que comparten y consumen. Las plataformas proporcionan información, permiten que las personas usuarias interactúen en su propio idioma y tienen en cuenta su edad y discapacidades.

e. Rendir cuentas ante las partes interesadas pertinentes, incluidas las personas usuarias, el público y los actores dentro del sistema de gobernanza, en la implementación de sus condiciones de servicio y políticas de contenido. Ofrecen a las personas usuarias la posibilidad de solicitar una reparación adecuada y oportuna frente a las decisiones relacionadas con los contenidos, tanto para las personas usuarias cuyos contenidos han sido retirados o moderados como para aquellas que han presentado quejas sobre los contenidos.

31. Las plataformas deben aplicar estos principios en cada una de las jurisdicciones en las que operan, garantizando los recursos y capacidades necesarias para servir a las personas usuarias de forma eficaz y oportuna.
32. Para seguir estos principios, existen áreas específicas en las que las plataformas digitales tienen la responsabilidad de informar o actuar frente a los actores del sistema de gobernanza, de conformidad con la normativa internacional en materia de derechos humanos. Estas áreas se describen en los párrafos 85 a 129 de las presentes Directrices.

El papel de las organizaciones intergubernamentales

33. Las organizaciones intergubernamentales, en consonancia con sus respectivos mandatos, deben apoyar a las partes interesadas pertinentes para garantizar que la implementación de estas Directrices se ajuste plenamente a la normativa internacional en materia de derechos humanos. Este apoyo debe incluir una asistencia técnica, así como el seguimiento y la denuncia de las vulneraciones de los derechos humanos, el desarrollo de normativas pertinentes, la facilitación del diálogo entre las múltiples partes interesadas y el fomento de redes.
34. Las organizaciones intergubernamentales y los organismos reguladores nacionales pueden crear modalidades de compromiso para seguir desarrollando y compartiendo buenas prácticas. Este compromiso puede incluir el intercambio de ideas emergentes y tendencias reguladoras, así como el apoyo o la formulación de sugerencias a los reguladores nacionales para perfeccionar las normas institucionales y los métodos para salvaguardar la libertad de expresión y

el acceso a la información. Tales modalidades deben servir para reducir los riesgos de fragmentación de Internet, así como para proporcionar herramientas que permitan una evaluación previa de los impactos de la regulación sobre el funcionamiento de Internet en su conjunto.

El papel de la sociedad civil y de otras partes interesadas

- 35.** Toda parte interesada que esté involucrada con los servicios de una plataforma digital en calidad de persona usuaria, formuladora de políticas, organismo de control u otros, desempeña un papel importante de apoyo a la libertad de expresión, acceso a la información y otros derechos humanos. En esta perspectiva, el proceso de desarrollo, implementación y evaluación de cada regulación que afecta a los contenidos de las plataformas digitales debe basarse en un enfoque de múltiples partes interesadas. Un amplio conjunto de partes interesadas también debe participar en la supervisión, incluidos los que representan a grupos en situación de vulnerabilidad y marginación, así como periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente.
- 36.** La sociedad civil desempeña un papel fundamental a la hora de comprender la naturaleza de los contenidos y comportamientos nocivos en línea y de contrarrestarlos, especialmente los dirigidos a todos los grupos en situación de vulnerabilidad y marginación, mujeres y niñas, periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente. La sociedad civil también desempeña un papel importante a la hora de vigilar e informar sobre leyes, políticas y acciones reguladoras de los Gobiernos que afectan a los derechos humanos. Son fundamentales para tender puentes entre el ecosistema de gobernanza digital y la población en general.
- 37.** Las personas investigadoras independientes cumplen un rol importante en la identificación de patrones de comportamiento abusivo y en la manera en que se podrían solucionar las posibles causas de fondo. Las personas investigadoras también deberían poder proporcionar una supervisión independiente sobre el funcionamiento del sistema de gobernanza. Las instituciones y personas investigadoras independientes pueden apoyar la debida diligencia en materia de derechos humanos, incluidas las evaluaciones de riesgos, auditorías, investigaciones y otros tipos de informes sobre las prácticas y actividades de las plataformas. Las personas investigadoras deben ser capaces de recopilar y analizar datos desglosados en función del género y otros factores

pertinentes interrelacionados (como la raza, el origen étnico, la edad, la situación socioeconómica, la discapacidad, etc.). Esto ayuda a identificar disparidades, sesgos y impactos diferenciados de las plataformas digitales en distintos grupos en situación de vulnerabilidad y marginación.

- 38.** Los medios de comunicación, las organizaciones de verificación de hechos y los profesionales de estas instituciones son partes interesadas importantes y tienen un papel que desempeñar en la promoción del disfrute de la libertad de expresión, el acceso a la información y otros derechos humanos, al tiempo que desempeñan su función de vigilancia. Por lo tanto, es necesario involucrar a los medios de comunicación y a sus profesionales en el proceso de regulación, reconociendo su papel como participantes activos en la contribución positiva al ecosistema de la información digital. Una relación constructiva entre las plataformas digitales y las fuentes de noticias creíbles reforzará el papel de las plataformas digitales en la difusión de información de interés público.
- 39.** Las personas educadoras y cuidadoras desempeñan un papel fundamental a la hora de ayudar a la juventud y a las personas estudiantes de todas las edades a comprender el entorno digital en sentido amplio, lo que incluye cómo buscar y comprender información creíble y cómo relacionarse respetuosamente con otras personas en línea. También tienen un papel que desempeñar en el aprendizaje permanente, ya que la tecnología cambia rápidamente.
- 40.** Las personas ingenieras, expertas informáticas y la comunidad técnica involucrada en el desarrollo de servicios y productos de plataforma también desempeñan un papel en la comprensión de los derechos humanos, los riesgos y los impactos éticos de los productos y servicios que diseñan y desarrollan²⁴.
- 41.** Todas estas partes interesadas deben desempeñar un papel activo en el proceso de consulta sobre el desarrollo y funcionamiento del sistema de gobernanza. Debe fomentarse la colaboración y el diálogo entre partes interesadas. Deben celebrarse debates y deliberaciones constructivas para intercambiar ideas, conocimientos y puntos de vista. La creación de grupos de trabajo, grupos operativos o comités consultivos brinda la oportunidad de participar activamente en la elaboración de propuestas de regulación.

24. Véase el informe del ACNUDH sobre la relación entre la elaboración de normas técnicas y los derechos humanos, A/HRC/53/42.

El sistema de gobernanza

42. El ecosistema de la gobernanza digital está formado por diversas partes interesadas, organismos y formas de regulación procedentes de todo el mundo. Mientras que algunos sistemas de gobernanza existentes, como en el caso de las elecciones o la protección de datos, deben interpretarse y considerarse de acuerdo con los cambios y retos que conlleva la era digital, también se están creando nuevos sistemas de gobernanza en diversos contextos para regular directamente las plataformas digitales. En cualquier caso, estos mecanismos reguladores podrían tener profundas implicaciones para la libertad de expresión y el acceso a la información y a contenidos culturales diversos en línea.
43. Las presentes Directrices destacan principios generales que pueden aplicarse, según proceda, a diversos procesos que afectan a la gobernanza de los contenidos en las plataformas digitales, independientemente de su forma o ámbito. Indican que un sistema de gobernanza global puede aprovechar eficazmente diversas formas de regulación complementarias para abordar los retos a los que se enfrentan las distintas partes interesadas en el ecosistema digital.
44. Las Directrices abogan por un enfoque de múltiples partes interesadas para la gobernanza de las plataformas digitales. Este enfoque podría incorporar aspectos como la identificación de las principales partes interesadas (incluidas las plataformas digitales que se encuentran dentro de su ámbito de aplicación), el fomento de la participación inclusiva, la garantía de una representación equilibrada, la transparencia y la rendición de cuentas, el fomento de la toma de decisiones y el diálogo en colaboración, la facilitación de un proceso iterativo, la coordinación de los esfuerzos de aplicación entre las partes interesadas y la realización de evaluaciones y revisiones periódicas.
45. Dependiendo del contexto, los mecanismos de rendición de cuentas y cumplimiento para la gobernanza de las plataformas digitales pueden incluir la complementariedad y la convergencia dentro de diferentes formas de regulación, tales como:
 - a. Estructuras y mecanismos de autorregulación, en los que las normas pueden ser supervisadas y aplicadas por agentes no estatales, como organismos sectoriales o consejos de redes sociales.
 - b. Estructuras y mecanismos de corregulación en los que, en algunos casos, se puede

otorgar fuerza legal a los códigos de conducta para que sirvan de regulación.

c. Marcos reguladores estatutarios en los que uno o varios reguladores independientes toman las decisiones finales sobre el establecimiento de normas para la plataforma.

46. Conscientes de la complejidad de este entorno, las presentes Directrices están diseñadas para aplicarse a una amplia gama de formas de gobernanza. Es importante señalar que en algunos ámbitos pueden ser necesarios marcos reguladores estatutarios para abordar áreas inadecuadas para los mecanismos de autorregulación y corregulación. Dichos marcos deben garantizar siempre la independencia de las autoridades reguladoras oficiales y, en consonancia con el objetivo de las Directrices, deben salvaguardar siempre los derechos humanos.



Principios de los sistemas de gobernanza

47. En primer lugar, la **transparencia** debe ser un principio rector común. En todos los sistemas de gobernanza, se espera que las plataformas digitales sean transparentes en cuanto a las condiciones, sistemas y procesos que emplean para moderar y curar sus contenidos, así como en cuanto a la debida diligencia en materia de derechos humanos en consonancia con las disposiciones de estas Directrices y los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos. Deben poder explicar cómo sus sistemas y procesos cumplen con sus condiciones de servicio y su aplicación efectiva, y si estos se ajustan a la normativa internacional en materia de derechos humanos.
48. También deben ser transparentes los sistemas y procedimientos de gobernanza externos a las plataformas. Cualquier acción reguladora externa debe proponerse, debatirse abierta y ampliamente, y finalmente ejecutarse bajo supervisión pública, con una delimitación abierta y clara de las competencias y responsabilidades de las decisiones.
49. En segundo lugar, un principio regulador común es que los **pesos y contrapesos** entre los distintos intereses deben institucionalizarse formalmente. Los sistemas de gobernanza deben tener siempre un enfoque de múltiples partes interesadas en todas las formas de regulación y sus combinaciones. Esto significa prever una participación amplia e inclusiva entre todas las partes interesadas que

mejor puedan representar intereses y valores divergentes, incluidas las diversas perspectivas de género e interseccionales. La participación de las múltiples partes interesadas debe ser significativa en términos de representación y en la creación, aplicación, supervisión y revisión de los procesos de gobernanza (normas, principios y políticas). Las campañas de concienciación pública, la divulgación selectiva, el respeto por la diversidad cultural y el uso de un lenguaje y unos formatos inclusivos en los procesos de gobernanza pueden facilitar una participación efectiva.

50. En tercer lugar, los procesos de gobernanza deben ser **abiertos y accesibles** a todas las partes interesadas, en particular a los grupos afectados por una estructura o un tipo de regulación propuestos. Las consultas públicas, las audiencias abiertas y las plataformas en línea deben utilizarse para brindar al público la oportunidad de realizar aportaciones y comentarios. Las preocupaciones de los grupos en situación de vulnerabilidad y marginación, así como de las mujeres y las niñas, deben estar adecuadamente representadas en el proceso de toma de decisiones.
51. El sistema de gobernanza debe garantizar que las plataformas digitales se comprometan activamente con los niños y niñas, protejan su libertad de expresión y demás derechos, apliquen las salvaguardias adecuadas y tengan debidamente en cuenta sus opiniones a la hora de desarrollar productos y servicios.
52. Los sistemas de gobernanza también deben promover el diálogo con los medios de comunicación, incluida la inversión en medios informativos independientes, y apoyar el ecosistema mediático facilitando datos y apoyando acciones para reforzar la sostenibilidad, diversidad y pluralidad de los medios.
53. En cuarto lugar, la inclusión de **diversos conocimientos especializados** debe ser una característica común de todas las formas de regulación. El sistema de gobernanza requiere que las partes interesadas dispongan de la capacidad necesaria, a través de la formación y los instrumentos reguladores, para comprender los marcos de derechos humanos y tener en cuenta los avances tecnológicos. Deben tener las capacidades y los conocimientos técnicos necesarios para tomar decisiones informadas y aplicar las presentes Directrices. Debe alentarse a todos los sistemas de gobernanza a que informen al público y evalúen los riesgos y oportunidades asociados a las tecnologías nuevas y emergentes.

54. Las partes interesadas dentro de los sistemas de gobernanza deben compartir la experiencia y los conocimientos en materia de regulación entre jurisdicciones. Los sistemas de gobernanza nacionales, regionales y mundiales deben ser capaces de cooperar y compartir prácticas para alcanzar el objetivo de salvaguardar la libertad de expresión, el acceso a la información y otros derechos humanos, abordando al mismo tiempo los contenidos que podrían restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos.
55. En quinto lugar, el sistema de gobernanza debe garantizar que las plataformas digitales se comprometan a **proteger y promover la diversidad cultural y la diversidad de las expresiones culturales** en la creación, producción, distribución, difusión, acceso y disfrute de los bienes y servicios culturales en línea, entre otras cosas garantizando una visión y una representación imparciales.

Responsabilidad y cumplimiento

56. Las diversas formas de regulación deben ser eficaces y sostenibles, teniendo en cuenta los recursos locales disponibles y las principales prioridades que requieren atención (por ejemplo, si se deben abordar principalmente cuestiones relacionadas con las elecciones, la salud pública, la publicidad o la protección de datos, etc.). Es necesaria una supervisión independiente de todas las formas de regulación. El proceso de elaboración de la normativa debe ser abierto, transparente y basado en evidencia.
57. Las plataformas digitales que no cumplan con sus propias políticas o con su deber de salvaguardar la libertad de expresión y el acceso a la información, mientras tratan contenidos que podrían ser lícitamente restringidos en virtud de la normativa internacional en materia de derechos humanos, de conformidad con los cinco principios descritos en los párrafos 85 a 129, deben rendir cuentas ante los órganos pertinentes dentro del sistema de gobernanza y estar sujetas a medidas coercitivas proporcionales con las garantías procesales correspondientes.
58. Los **sistemas de autorregulación** pueden ser complementarios y converger con otras formas de regulación. Deben incluir auditorías periódicas independientes y obligatorias que evalúen el cumplimiento por parte de las plataformas digitales de los códigos, políticas o normas de autorregulación. Estas auditorías no deben ser financiadas directamente por la industria o las plataformas digitales individuales, si bien los gravámenes sobre estas entidades pueden ayudar a cubrir los costes de tales ejercicios. Tampoco deben ser realizadas por personas

o entidades que tengan o puedan tener un conflicto de interés. Las condiciones y los resultados de la auditoría deben estar abiertos al público para recoger posibles comentarios.

59. Las **estructuras de corregulación** deben proporcionar un marco jurídico que posibilite el entorno para la libertad de expresión, el acceso a la información y otros derechos humanos. En la corregulación, el papel regulador debe ser compartido entre la industria y demás partes interesadas y el Gobierno o las autoridades u organismos reguladores oficiales independientes. El papel de las autoridades públicas pertinentes incluye el reconocimiento del sistema de corregulación, la auditoría de los procesos y la financiación del sistema (posiblemente mediante gravámenes a las plataformas). La corregulación debe prever la posibilidad de sanciones estatales, como multas, en caso de que no se cumplan los objetivos acordados.
60. La **regulación estatutaria** de las plataformas digitales que aborden cuestiones que puedan afectar a la libertad de expresión solo debe considerarse cuando exista independencia en la toma de decisiones de las autoridades reguladoras implicadas en su implementación. Dicha regulación debe centrarse en los sistemas y procesos de moderación y curación de contenidos, en lugar de determinar la legalidad de los contenidos en particular, y debe tener una base jurídica (es decir, estar suficientemente definida), perseguir un objetivo legítimo en virtud del Artículo 19 (3) del PIDCP, y ser necesaria y proporcional.
61. El enfoque de múltiples partes interesadas de la regulación estatutaria debe reflejarse en un acuerdo por el cual:
 - a. Las autoridades estatales pertinentes, incluidas las autoridades reguladoras oficiales independientes, determinan el objetivo legítimo de la regulación mediante procesos legislativos participativos e inclusivos.
 - b. Las plataformas digitales informan públicamente a las autoridades reguladoras oficiales.
 - c. Las organizaciones de la sociedad civil, las personas artistas, las personas investigadoras independientes y otras instituciones pertinentes hacen aportaciones a la elaboración de normas, contribuyen a la supervisión y logran los controles y equilibrios necesarios mediante la participación y el escrutinio institucionalizados.

62. Cualquier decisión específica sobre la legalidad de contenidos concretos debe seguir el debido proceso y estar abierta a la revisión de un órgano judicial imparcial e independiente.
63. En todos los casos, las evaluaciones de contenido deben seguir la prueba de tres partes sobre restricciones legítimas a la libertad de expresión, tal como se establece en el Artículo 19 (3) del PIDCP²⁵, y la prohibición de la apología del odio que constituya incitación a la discriminación, la hostilidad o la violencia, tal como se establece en el Artículo 20 (2) del PIDCP; incluyendo, según corresponda, la prueba de umbral de seis puntos para definir dicho contenido, descrita en el Plan de Acción de Rabat²⁶.
64. Toda intervención regulatoria debe basarse en evidencia, ser proporcionada e incluir garantías procesales, entre otras cosas garantizando a la plataforma el acceso a todos los hechos y consideraciones sobre los que se toma la decisión. Este proceso debe implicar a múltiples grupos de partes interesadas a fin de que haya una visión más amplia de la sostenibilidad, la eficacia y el impacto de la intervención. La exigencia de un proceso basado en evidencia no puede ser una excusa para retrasar las acciones regulatorias necesarias para proteger los derechos humanos.
65. Todas las partes interesadas, incluidas las plataformas, deben tener la oportunidad de presentar alegatos y/o recurrir una decisión de incumplimiento. El sistema regulador debe estar obligado a publicar y consultar las medidas de aplicación y seguir el debido proceso antes de ordenar a una plataforma que aplique medidas específicas.

Definición de las plataformas digitales en el ámbito de aplicación de la normativa

66. A la hora de definir las plataformas digitales que deben estar en el ámbito de la regulación legal, las autoridades reguladoras deben identificar aquellas plataformas que tengan una presencia, un tamaño y una cuota de mercado relevantes en una jurisdicción específica. Estas deben determinarse mediante una evaluación independiente del riesgo que suponen para los derechos humanos, incluidos los de los grupos en situación de vulnerabilidad y marginación, así como

25. UNESCO. 2021. «Los límites legítimos de la libertad de expresión: la prueba de tres partes». <https://www.youtube.com/watch?v=4AG2y06CsC8>

26. UNESCO. 2021. «Plan de Acción de Rabat sobre la prohibición de la apología del odio». <https://www.unesco.org/archives/multimedia/document-5554-spa-2>

para las instituciones democráticas²⁷. La definición del ámbito de aplicación debe proteger el derecho a la vida privada y no debe suponer el debilitamiento de las protecciones existentes para el cifrado de información u otras tecnologías de protección de la privacidad.

67. Como reflejo de las realidades regionales y jurisdiccionales, pueden tenerse en cuenta los siguientes criterios para identificar las compañías incluidas en el ámbito de aplicación:

a. Tamaño y alcance, centrándose en las plataformas con más probabilidades de influir en una parte significativa de la población y/o en grupos en situación de vulnerabilidad y marginación.

b. Cuota de mercado, reconociendo la influencia considerable de las plataformas dominantes en todo el ecosistema de la información. La aplicación de las Directrices debe evitar penalizar a las empresas emergentes y a los nuevos operadores, garantizando al mismo tiempo que las plataformas digitales con mayor impacto potencial estén cubiertas de manera proporcional. Aunque se espera que todas las plataformas sigan los principios generales, las obligaciones específicas de información previstas en los apartados 85 a 129 de estas Directrices pueden aplicarse principalmente a las plataformas más grandes, cuya capacidad para cumplirlas es mayor.

c. Funcionalidad y características, siendo conscientes de las diferencias relevantes entre los distintos servicios en cuanto a visibilidad, influencia y orientación de los contenidos. El riesgo puede venir determinado por la base de personas usuarias de la plataforma, las formas de propiedad, el modelo de negocios, la funcionalidad y las características, tales como la publicación en tiempo real, el potencial de viralidad, el volumen, la velocidad de distribución, la verosimilitud y la medida en que los contenidos pueden publicarse sin un proceso de moderación de contenidos²⁸.

27. Se puede desarrollar una guía complementaria para identificar el riesgo sistémico de las plataformas como complemento para la puesta en práctica de estas Directrices.

28. «Las normativas basadas en el riesgo se basan en la evaluación por parte del legislador/regulador de los riesgos relevantes para su mandato y el nivel adecuado de intervención requerido de acuerdo con el nivel de riesgo. Si un actor realiza una actividad de bajo riesgo, la normativa se racionalizaría en consecuencia, previendo menores requisitos de cumplimiento». https://www3.weforum.org/docs/WEF_Pathways_to_the_Regulation_of_Crypto_Assets_2023.pdf



Características de las autoridades reguladoras independientes

- 68.** En la regulación estatutaria, las autoridades reguladoras oficiales, aunque formen parte del aparato ejecutivo del Estado, deben ser totalmente independientes del Gobierno y rendir cuentas principalmente ante el poder legislativo del cumplimiento de sus mandatos²⁹. Esto se aplica a los organismos reguladores existentes que tienen un interés legítimo en los contenidos de las plataformas (como los organismos de gestión electoral, las autoridades de publicidad, las entidades de protección de la infancia, las comisiones de datos y privacidad, los organismos de competencia, etc.), así como a cualquier nueva instancia reguladora especializada o de coordinación que pueda constituirse.
- 69.** Por lo que respecta a todos los organismos estatutarios que participan en la regulación de plataformas, ya sea de forma individual o conjunta, la revisión periódica³⁰ debe correr a cargo de un organismo independiente que informe directamente al poder legislativo. Las intervenciones estatutarias también deberían estar sujetas a revisión en los tribunales si se considera que las autoridades se han excedido en sus competencias, han actuado de forma irrazonable o han actuado de forma sesgada o desproporcionada.
- 70.** Las autoridades reguladoras oficiales deben ser independientes y estar libres de presiones económicas, políticas o de otro tipo. Su poder y mandato deben estar establecidos en la ley. También deben cumplir la normativa internacional en materia de derechos humanos y promover la igualdad de género.
- 71.** Las instituciones reguladoras oficiales deben tener los fondos y conocimientos suficientes para ejecutar sus responsabilidades de manera efectiva. Las fuentes de financiación también deben ser claras, transparentes

29. El Banco Mundial declaró que la característica clave del modelo de regulador independiente es la independencia en la toma de decisiones. Un documento de orientación sobre la regulación de la radio y teledifusión encargado por la UNESCO (Salomon, Eve. *Guidelines for broadcasting regulation*. 2006) también destacó que «una autoridad independiente (es decir, cuyas facultades y responsabilidades están establecidas en un instrumento de derecho público, que tiene facultad para administrar sus propios recursos y cuyos miembros son nombrados de manera independiente y protegidos por la ley contra el despido injustificado) está en mejores condiciones para actuar imparcialmente en el interés público y evitar la influencia indebida de los intereses políticos o de la industria». Para consultar las referencias completas, véase el apéndice de estas Directrices.

30. La revisión debe hacer especial hincapié en cómo las decisiones del sistema regulador pueden afectar al disfrute de los derechos humanos.

y accesibles para todas las personas, y no deben estar sujetas a la discreción del Gobierno.

- 72.** Las personas funcionarias o integrantes de las instituciones reguladoras oficiales que trabajen en la cuestión de los contenidos en las plataformas deberán:
- a.** Ser nombrados a través de un proceso participativo, transparente e independiente, no discriminatorio y basado en el mérito.
 - b.** Ser responsables ante un organismo independiente (que puede ser legislativo o judicial, un consejo externo o juntas directivas independientes).
 - c.** Contar con experiencia relevante en normativa internacional en materia de derechos humanos y el ecosistema digital.
 - d.** Presentar un informe público anual a un organismo independiente, idealmente legislativo, y rendir cuentas ante él, e informar incluso sobre cualquier dictamen motivado.
 - e.** Hacer público cualquier posible conflicto de intereses y declarar sobre cualquier regalo o incentivo.
 - f.** Después de completar el mandato, por un periodo razonable, no ser contratados ni prestar servicios remunerados a quienes estuvieron sujetos a su regulación, a fin de evitar el riesgo conocido como «puertas giratorias».
- 73.** Las autoridades reguladoras oficiales deben poder solicitar a las plataformas digitales informes periódicos sobre la aplicación de sus condiciones de servicio, y adoptar medidas coercitivas contra las plataformas digitales que incumplan sus propias políticas o sus responsabilidades de salvaguardar la libertad de expresión, el acceso a la información y contenidos culturales diversos. Deben poder establecer un proceso de reclamaciones y emitir recomendaciones públicas que puedan ser o no vinculantes, y estar facultadas para dar a las plataformas instrucciones transparentes y adecuadas para la promoción y el respeto de los derechos humanos, basadas en la normativa internacional en materia de derechos humanos.



- 74.** La alfabetización mediática e informacional abarca una amplia gama de competencias que permiten a las personas usuarias pensar de forma crítica sobre la información con la que interactúan en línea. La alfabetización mediática e informacional debe abordarse específicamente a través del sistema de gobernanza para garantizar que todas las partes interesadas, incluidas las plataformas, desempeñen eficazmente su papel.
- 75.** La alfabetización mediática e informacional se logrará con mayor eficacia cuando las partes interesadas del sistema de gobernanza compartan una visión común y trabajen de forma conjunta para alcanzarla mediante el intercambio de conocimientos y recursos. Los programas de alfabetización mediática e informacional deben adecuarse a la disponibilidad de los medios de comunicación y las tecnologías de la información existentes y emergentes, de modo que los ciudadanos puedan beneficiarse plenamente de su uso para participar activamente en sus sociedades.
- 76.** Los programas de alfabetización mediática e informacional deben hacer hincapié en el empoderamiento de las personas usuarias y garantizar que tengan las habilidades y los conocimientos que les permitan interactuar con el contenido de manera crítica y eficaz en todas las formas de los diversos medios de comunicación y con todos los proveedores de información, incluyendo escuelas, universidades, instituciones de investigación, bibliotecas, archivos, museos, empresas de medios de comunicación, editores, entidades estadísticas y más. Cuando los programas de alfabetización mediática e informacional solo hacen hincapié en las competencias de protección o seguridad digital, pueden dar lugar a restricciones excesivas en el uso de las plataformas digitales. Sin embargo, deberían dar prioridad a los pasos específicos que las personas usuarias pueden dar, basándose en las mejores prácticas publicadas por la UNESCO y otros organismos internacionales, para identificar el contenido que podría restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos.
- 77.** Los programas de alfabetización mediática e informacional deben promover la diversidad cultural, la inclusión social y la ciudadanía global y aspirar a reducir la «brecha de participación» entre los ciudadanos que participan en la creación

31. Véase «Media and information literacy: Policy and strategy guidelines» [Alfabetización mediática e informacional: orientaciones políticas y estratégicas]. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000225606>

y el uso crítico de los contenidos mediáticos e informativos y los que no. Los programas de alfabetización mediática e informacional también deben promover la igualdad de género y el empoderamiento de la mujer y ofrecer oportunidades de participación a los grupos en situación de vulnerabilidad y marginación.

- 78.** Los Gobiernos siempre deben promover la alfabetización mediática e informacional, incluidas las habilidades de seguridad en línea, de las personas usuarias, y en especial de todos los grupos en situación de vulnerabilidad y marginación, así como de las mujeres y las niñas. De esta manera, las personas usuarias pueden comprometerse de forma crítica con los contenidos y las tecnologías, navegar por un panorama mediático e informativo en rápida evolución marcado por la transformación digital, promover los derechos humanos y crear resiliencia frente a los desafíos relacionados.
- 79.** Los Gobiernos deben difundir información y llevar a cabo campañas de sensibilización sobre los derechos de las niñas y niños en el entorno digital, incluido su derecho a la libertad de expresión, centrándose en particular en aquellos programas cuyas acciones tengan un impacto directo o indirecto en las niñas y niños. Deben facilitar programas educativos para las niñas y niños, padres y cuidadores, el público en general y las personas responsables de la formulación de políticas a fin de mejorar su conocimiento de los derechos de las niñas y niños en relación con las oportunidades y los riesgos asociados a los productos y servicios digitales. Dichos programas deben incluir información sobre cómo las niñas y niños pueden beneficiarse de los productos y servicios digitales y desarrollar su alfabetización mediática e informacional, incluidas las competencias digitales.
- 80.** Las plataformas deben establecer una estrategia clara y pública para empoderar a las personas usuarias y promover un entorno en línea favorable que salvaguarde la libertad de expresión y el acceso a la información a través de la alfabetización mediática e informacional, incluyendo la educación en seguridad en línea. Debe existir un enfoque específico dentro de la plataforma digital sobre el modo de mejorar la alfabetización digital de las personas usuarias, en especial de los grupos en situación de vulnerabilidad y marginación, debiendo los equipos de desarrollo de productos tenerlo en cuenta.
- 81.** Las plataformas digitales deben asignar recursos adecuados para mejorar la alfabetización mediática e informacional de todas las personas usuarias, incluida la alfabetización digital sobre productos y servicios propios de la plataforma, así como los procesos relevantes para las personas usuarias. Esto debería centrarse

especialmente en mejorar la comprensión de las personas usuarias sobre las formas en que una plataforma determinada presenta, selecciona, recomienda y/o marca contenidos (también relacionado con los pasos descritos en los Principios 3 y 4 más abajo) y los pasos específicos que las personas usuarias pueden dar para identificar por sí mismas los contenidos que podrían estar restringidos lícitamente en virtud de la normativa internacional en materia de derechos humanos.

- 82.** Las plataformas deben formar a sus equipos de desarrollo de productos en alfabetización mediática e informacional, incluida la seguridad en línea, desde una perspectiva de empoderamiento de la persona usuaria y basándose en la normativa internacional, y poner en marcha mecanismos de seguimiento y evaluación tanto internos como independientes.
- 83.** Tanto los Gobiernos como las plataformas digitales deben poner en práctica programas de alfabetización mediática e informacional en estrecha colaboración con organizaciones y diversos expertos independientes de las plataformas, entre los que se incluyen las autoridades públicas responsables de la alfabetización mediática e informacional, el mundo académico, las organizaciones de la sociedad civil que trabajan con grupos en situación de vulnerabilidad y marginación, las personas investigadoras, bibliotecarias, las y los maestros, las personas educadoras especializadas, periodistas, artistas y profesionales de la cultura. Deben tomarse medidas específicas para las personas usuarias y no usuarias y escuchar a los grupos en situación de vulnerabilidad y marginación, como se indica en las numerosas herramientas de la UNESCO disponibles sobre alfabetización mediática e informacional.
- 84.** Los Gobiernos y las plataformas digitales deben colaborar y garantizar que las personas usuarias comprendan sus derechos en línea y fuera de línea, incluido el papel de la alfabetización mediática e informacional en el disfrute y la protección de los derechos a la libertad de expresión y al acceso a la información.



Principio 1. Aplicar el principio de debida diligencia en materia de derechos humanos

Salvaguardia de los derechos humanos y evaluaciones de riesgos

- 85.** En cualquier tipo de regulación, las plataformas digitales deben poder demostrar los sistemas o procesos que han implementado para garantizar que se aplica la debida diligencia en materia de derechos humanos de forma continua, incluidas las evaluaciones de impacto en los derechos humanos y de género³², así como las medidas de mitigación de riesgos³³. Estos sistemas deben revisarse periódicamente y la revisión debe hacerse pública.

32. Véase la «Declaración de Irene Khan, Relatora Especial sobre la promoción y protección de la libertad de opinión y de expresión» de 18 de octubre de 2021. En consonancia con los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos, «las empresas de redes sociales deben evaluar periódicamente el impacto en los derechos humanos y el género para identificar y mitigar los riesgos sistémicos que afectan a las mujeres y las personas no conformes con el género. Deben hacer que las plataformas sean seguras e incluyan la perspectiva de género y que, de acuerdo con la normativa internacional en materia de derechos humanos, adopten políticas y herramientas de seguridad eficaces, garanticen una transparencia significativa, incluso de los algoritmos, y proporcionen recursos adecuados». <https://www.ohchr.org/en/statements/2022/02/statement-irene-khan-special-rapporteur-promotion-and-protection-freedom-opinion>

33. Las evaluaciones de impacto en los derechos humanos deben incluir todos los derechos humanos que puedan verse afectados por las políticas de las compañías. Esto incluye derechos civiles y políticos como la libertad de expresión, el acceso a la información y la privacidad, así como los derechos económicos, sociales y culturales, el derecho a no sufrir violencia y el derecho a participar en la vida pública, entre otros.

- 86.** De acuerdo con la normativa internacional en materia de derechos humanos, incluidos los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas, las plataformas deben realizar evaluaciones de riesgo periódicas para identificar y abordar cualquier impacto real o potencial de su funcionamiento en los derechos humanos. Al implementar los procesos de evaluación de riesgos en materia de derechos humanos, las plataformas digitales deben considerar el modo en que todo producto o servicio afecta al comportamiento de la persona usuaria más allá de su objetivo de adquisición o participación.
- 87.** Las evaluaciones de riesgos deben ser un ancla para la toma de decisiones dentro de las plataformas digitales y servir de fundamento para el enfoque del diseño y funcionamiento de sus servicios, así como para las medidas que despliegan para mitigar el riesgo residual y salvaguardar los derechos humanos, la no discriminación y la igualdad de trato. Además, las responsabilidades en materia de gestión de riesgos deben estar claramente especificadas y ser asumidas por los niveles más altos, debiéndose comunicar periódicamente las actividades de gestión de riesgos a las personas responsables de la toma de decisiones.
- 88.** Como mínimo, deben realizarse evaluaciones de riesgos y en materia de derechos humanos:
- a.** Antes de realizar cualquier cambio significativo de diseño, tomar decisiones políticas importantes (incluidas las relacionadas con el sistema de publicidad, si procede), realizar cambios en el modo de funcionamiento o iniciar nuevas actividades o relaciones/asociaciones.
 - b.** Periódicamente, para proteger los derechos de todos los grupos en situación de vulnerabilidad y marginación, así como mujeres y niñas, periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente³⁴.
 - c.** Antes de cualquier proceso electoral para proteger su integridad³⁵.
 - d.** En respuesta a emergencias, crisis, conflictos o cambios significativos en el entorno operativo³⁶.
- 89.** Durante el procedimiento de debida diligencia en materia de derechos humanos, las plataformas deben garantizar un compromiso significativo con diversas

34. Véanse las disposiciones específicas del contexto, párrafo 126.

35. Véanse las disposiciones específicas del contexto, párrafos 127-37.

36. Véanse las disposiciones específicas del contexto, párrafos 138-40.

partes interesadas para identificar riesgos específicos para grupos en situación de vulnerabilidad y marginación. Es fundamental que las plataformas digitales estén abiertas a las aportaciones de personas expertas e independientes sobre cómo se estructuran estas evaluaciones.

90. Desde el principio, las plataformas deben crear espacios para escuchar, implicar y hacer partícipes a las personas usuarias, incluidas aquellas que han sufrido acoso o abuso, sus representantes y personas usuarias de grupos en situación de vulnerabilidad y marginación, así como mujeres, niñas, periodistas y artistas, para fundamentar las políticas y procesos de la plataforma. Esto podría incluir distintas formas de identificar y contrarrestar contenidos que podrían restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos, así como oportunidades y soluciones para abordar los riesgos evaluados.



Principio 2. Adherir a la normativa internacional en materia de derechos humanos, incluido en el diseño de la plataforma, la moderación y la curación de contenidos

91. Las plataformas digitales deben garantizar que las consideraciones relativas a los derechos humanos y al debido proceso se integren en cada etapa del diseño del proceso y en las políticas y prácticas de moderación y curación de contenidos.

Procesos de diseño

92. El diseño de nuevos productos y las políticas de moderación y curación de contenidos de las plataformas digitales deben ajustarse a la responsabilidad de las empresas en materia de respeto de los derechos humanos, tal como lo establecen los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos y otras normativas internacionales en materia de derechos humanos.
93. Las plataformas digitales deben garantizar la no discriminación y la igualdad de trato en sus procesos de diseño, así como en sus políticas, prácticas y sistemas

de moderación y curación de contenidos. Esto incluye abordar los prejuicios, los estereotipos y los algoritmos discriminatorios o las prácticas de moderación de contenidos que afectan a las mujeres y las niñas, así como a los grupos en situación de vulnerabilidad y marginación, incluidas las comunidades indígenas. Debe esperarse que las plataformas digitales garanticen que todas las personas usuarias, independientemente de su origen o capacidades, puedan participar plenamente e interactuar con sus servicios.

Políticas y prácticas de moderación y curación de contenidos

- 94.** Los sistemas de moderación y curación de contenidos, incluidos los componentes automatizados y no automatizados, deben ser fiables y eficaces y a una escala que se ajuste al volumen de contenidos que se moderan, en todas las jurisdicciones en las que opera la plataforma. Esto incluye buscar la precisión y la no discriminación en los métodos de detección. La moderación y la curación de contenidos deben aplicarse de conformidad con la normativa internacional en materia de derechos humanos, principalmente para no infringir la libertad de expresión y la diversidad cultural.
- 95.** Las decisiones de moderación de contenidos en todas las regiones e idiomas deben tener en cuenta, de forma transparente, el contexto, la gran variedad de matices lingüísticos que inciden en el significado y las particularidades lingüísticas y culturales de dicho contenido.
- 96.** Las plataformas que operan en entornos multilingües deben garantizar que la moderación de contenidos humana y automatizada esté disponible en las principales lenguas habladas en ese entorno (como mínimo), a una escala que se ajuste al volumen de contenidos.
- 97.** Las plataformas digitales deben garantizar una acción rápida y decisiva para eliminar todo contenido sobre abusos sexuales a menores de que se tenga conocimiento o la retransmisión en directo de actos de terror, respetando los derechos de todas las personas, incluidos los grupos en situación de vulnerabilidad y marginación. No obstante, las plataformas deben garantizar que estos contenidos, que pueden ser vitales para la instrucción y enjuiciamiento de delitos, no se eliminen, sino que se conserven y salvaguarden de forma segura para que las fuerzas del orden y las personas investigadoras puedan utilizarlos según proceda.

- 98.** Como se ha destacado anteriormente, es responsabilidad del Estado garantizar el derecho a la libertad de expresión y velar por que cualquier restricción de contenido se ajuste a la normativa internacional en materia de derechos humanos, en particular al artículo 19 (3) y 20 del PIDCP. No obstante, las plataformas digitales deben ser capaces de demostrar que cualquier acción tomada al moderar y curar contenidos se ha llevado a cabo de acuerdo con sus condiciones de servicio y las normas de la comunidad, y deben informar fielmente al sistema de gobernanza o al sistema judicial independiente, cuando proceda, sobre el desempeño en relación con sus responsabilidades y/o planes.
- 99.** Al considerar medidas que restrinjan contenidos, las plataformas deben tener en cuenta las condiciones de restricción legítima a la libertad de expresión, como se establece en el Artículo 19(3) del PIDCP, y la prohibición de la apología del odio que constituya una incitación a la discriminación, la hostilidad o la violencia, tal como se establece en el Artículo 20(2) del PIDCP, incluida la prueba de umbral que consta de seis partes para definir dicho contenido, como se destaca en el Plan de Acción de Rabat.
- 100.** Una vez que las plataformas digitales identifican contenidos que podrían restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos, deben tomar medidas como proporcionar información alternativa fiable, informar a las personas usuarias de su reserva acerca del origen de los contenidos, limitar o eliminar la amplificación algorítmica de dichos contenidos, prestando la debida atención a los contenidos que reflejen sesgos de género o violencia basada en género, desmonetizar los contenidos de los ingresos publicitarios o eliminar o retirar contenidos³⁷.

Moderadores humanos.

- 101.** Las personas moderadoras de contenido, ya sean empleadas por las plataformas directamente o contratadas como contratistas externos a través de funciones externalizadas, deben recibir una formación adecuada, dominar el idioma o idiomas utilizados en las plataformas y estar familiarizadas con los contextos lingüísticos y culturales locales, ser evaluadas, autorizadas y recibir apoyo psicológico. Además, las plataformas deben establecer programas de apoyo, bien financiados y dotados de suficiente personal, para los moderadores de contenidos con la finalidad de minimizar el daño causado por su exposición

37. Las medidas aplicadas deben ser siempre proporcionales al objetivo legítimo que pretenden proteger. La eliminación y retirada de contenidos y la suspensión o bloqueo de cuentas deben ser el último recurso posible y utilizarse en última instancia en los casos más extremos.

recurrente a contenido violento o perturbador mientras realizan su trabajo. El número de personas moderadoras empleadas debe ajustarse a la complejidad y al volumen de contenidos que deben tratar.

102. Las plataformas también deben ser explícitas sobre si colaboran con proveedores externos de servicios de moderación de contenidos, organizaciones externas o expertos que contribuyan a la toma de decisiones, especialmente en países o regiones donde la propia plataforma tiene poco conocimiento del contexto local. Al hacerlo, las plataformas siempre deben aplicar la debida diligencia y abstenerse de revelar la identidad de los socios en situaciones en las que hacerlo pueda presentar riesgos para la seguridad de dichos socios.

Uso de sistemas automatizados para la moderación y curación de contenidos

103. Cuando proceda, las plataformas digitales deben encargar auditorías externas periódicas, con medidas de seguimiento vinculantes, de las herramientas automatizadas y humanas utilizadas para la moderación de contenidos, la curación y los mecanismos de recomendación en cuanto a su precisión, exactitud y posible sesgo o discriminación en diferentes tipos de contenidos, idiomas, culturas y contextos. También deben revisar su capacidad lingüística y el uso coherente en todas las jurisdicciones. Como se indica en el párrafo 87, también deben encargar evaluaciones periódicas independientes de los impactos de sus sistemas publicitarios en los derechos humanos, la diversidad cultural y la igualdad de género. Los resultados de estas evaluaciones deben hacerse públicos³⁸.

104. Las plataformas digitales deben contar con sistemas y procesos que les permitan identificar y tomar las medidas necesarias, de acuerdo con las disposiciones de estas Directrices, cuando cualquier aspecto del diseño de los servicios de la plataforma pueda generar una amplificación de los contenidos que puedan ser restringidos lícitamente en virtud de la normativa internacional en materia de derechos humanos.

105. Las plataformas también deben garantizar que los sistemas de curación y recomendación, incluidas tanto las herramientas humanas como automatizadas, no amplifiquen contenidos que puedan ser restringidos lícitamente en virtud de la normativa internacional en materia de derechos humanos.

38. Una opción es que las auditorías y evaluaciones independientes se realicen de acuerdo con la normativa mundial y, en el mejor de los casos, sean verificadas por un organismo independiente, de modo que puedan utilizar los mismos informes independientemente del tipo de regulación.

- 106.** Las plataformas también deben garantizar que los contenidos que puedan ser restringidos lícitamente en virtud de la normativa internacional en materia de derechos humanos no sean amplificados por mecanismos automatizados de curación o recomendación, simplemente debido a las limitaciones lingüísticas de esos mecanismos.
- 107.** Las plataformas digitales deben poder explicar al sistema de gobernanza sobre el uso e impacto de los sistemas automatizados, incluida la medida en que dichas herramientas afectan a la recopilación de datos, la publicidad dirigida y la divulgación, clasificación y/o eliminación de contenido, incluido el contenido artístico y el contenido relacionado con elecciones.
- 108.** Las plataformas digitales deben ofrecer a las personas usuarias opciones para ajustar los sistemas de curación y moderación de contenidos. Las personas usuarias deben tener la posibilidad de controlar el contenido que ven y deben poder entender fácilmente cómo pueden acceder a diversas fuentes y puntos de vista en torno a los temas de tendencia. También podría exigirse a las plataformas que ofrezcan a las personas usuarias opciones para gestionar la recopilación de datos personales y el grado en que los recomendadores de contenidos responden a preferencias explícitas o inferidas.
- 109.** Las plataformas digitales no deben utilizar datos personales obtenidos directamente de niñas o niños u obtenidos indirectamente o inferidos sobre niñas o niños de otras fuentes para la elaboración de perfiles.

Notificaciones

- 110.** Las plataformas digitales deben notificar a las personas usuarias cuando se elimina su contenido e indicarles el motivo de ello. Esto permitiría a las personas usuarias comprender las razones por las que se tomó esa medida en relación con su contenido, el método utilizado (a través de medios automatizados o después de una revisión humana) y en virtud de qué normas de la plataforma se tomó dicha medida. Las plataformas digitales también deben contar con procesos que permitan a las personas usuarias apelar tales decisiones (véanse los párrafos 125-28). Esta disposición puede variar en función del tamaño de la compañía y en la medida en que existan procedimientos de reparación efectivos para que las personas usuarias puedan apelar estas decisiones.



Principio 3. La transparencia de las plataformas

- 111.** Las plataformas digitales deben informar periódicamente al público y al sistema de gobernanza sobre cómo cumplen los principios de transparencia y explicabilidad, y cómo actúan en relación con sus condiciones de servicio y las normas de la comunidad. Esto incluye sus respuestas a las solicitudes de información o de eliminación de contenidos del Gobierno³⁹. La implementación de esta disposición puede variar en la práctica en función del tamaño de la compañía, a fin de limitar la carga sobre las empresas más pequeñas y las empresas emergentes.
- 112.** La transparencia debe ser significativa: la información facilitada debe ser lo más clara y concisa posible, y tan detallada y compleja como sea necesario. La transparencia no consiste simplemente en facilitar textos jurídicos o un volcado de datos, sino en proporcionar a las partes interesadas la información que necesitan para tomar decisiones informadas.
- 113.** Las normas de transparencia presentadas en estas Directrices pueden considerarse como un mínimo que deben cumplir todas las compañías en el ámbito de cualquier sistema de gobernanza.

39. Las orientaciones sobre transparencia para plataformas digitales pueden consultarse en los 26 principios de alto nivel establecidos por la UNESCO en *Dejar entrar el sol: transparencia y responsabilidad en la era digital*. https://unesdoc.unesco.org/ark:/48223/pf0000377231_spa.

Transparencia significativa

- 114.** La efectividad de los mecanismos de transparencia de las plataformas digitales debe evaluarse de manera independiente con respecto de la normativa internacional, a través de evaluaciones cualitativas y cuantitativas empíricas, para determinar si la información proporcionada para una transparencia significativa cumplió su propósito. Los informes deben hacerse públicos periódicamente.
- 115.** Las plataformas digitales deben publicar información indicando la forma en que garantizan que las consideraciones relativas a los derechos humanos y al debido proceso están integradas en cada etapa de las políticas y prácticas de la moderación y curación de contenidos. Esta información públicamente disponible debe incluir:

Transparencia en relación con las condiciones de servicio de las plataformas digitales

- a.** Cualquier medida utilizada para moderar y seleccionar contenido establecida en las condiciones de servicio de las plataformas, incluidas, por ejemplo, listas de contenidos o usuarios prohibidos.
- b.** Toda información sobre los procesos utilizados para el cumplimiento de las condiciones de servicio y para sancionar a las personas usuarias, así como las demandas/solicitudes gubernamentales de eliminación, restricción o promoción de contenidos.
- c.** Información sobre los motivos de las restricciones impuestas en relación con el uso de sus condiciones de servicio, disponible al público en un formato fácilmente accesible en sus condiciones de servicio.
- d.** Información sobre los tipos de contenidos que se consideran prohibidos o contra los que la plataforma digital actuará en virtud de sus condiciones de servicio y las medidas adoptadas, incluidas las circunstancias en las que la plataforma digital suspenderá la cuenta de un usuario, ya sea de forma permanente o temporal.

Transparencia en relación con la implementación de las políticas y prácticas de moderación y curación de contenidos

- e. La manera en que se modera y cura el contenido, tanto por medios automatizados como por revisión humana, así como el contenido que se elimina o bloquea bajo las condiciones de servicio o de conformidad con las demandas/solicitudes del Gobierno. Esto debe incluir información cuantitativa y cualitativa sobre los efectos, resultados e impactos reales que generan estos sistemas.
- f. Todo cambio en las políticas de moderación y curación de contenidos debe comunicarse periódicamente a las personas usuarias en un formato fácilmente accesible.
- g. Todo uso de medios automatizados con el propósito de moderación y curación de contenidos, incluida una especificación sobre el papel del medio automatizado en el proceso de revisión, así como cualquier indicador de los beneficios y limitaciones del medio automatizado en el cumplimiento de tal propósito.
- h. Todo resguardo aplicado en relación con la moderación y curación de contenidos que se implemente para proteger la libertad de expresión y el acceso a la información y contenidos culturales diversos, incluso en respuesta a solicitudes del Gobierno, especialmente en relación con asuntos de interés público, incluidos los contenidos periodísticos, artísticos y culturales, y los derechos de propiedad intelectual.
- i. Información sobre el número de moderadores humanos empleados o subcontratados y la naturaleza de su experiencia y conocimiento del idioma o idiomas locales y del contexto local, así como si se trata de personal interno o contratistas.
- j. La manera en que los datos personales de las personas usuarias se recogen, utilizan, revelan, conservan y difunden y el tratamiento que se les da, incluido qué datos personales y sensibles se utilizan para tomar decisiones algorítmicas con fines de moderación y curación de contenido. Esto también incluye cómo se comparten los datos personales con otras entidades y qué datos personales obtiene la plataforma indirectamente, por ejemplo, a través de la elaboración de perfiles de usuario o la interoperabilidad con otras partes del ecosistema digital.

Transparencia en relación con los mecanismos de reclamación de las personas usuarias

- k.** Información relevante respecto a reclamaciones sobre la eliminación, el bloqueo o la negativa de bloquear el contenido y el modo en que las personas usuarias pueden acceder al proceso de reclamación. Estos datos deben incluir información cuantitativa y cualitativa de las reclamaciones recibidas, gestionadas, aceptadas y rechazadas, así como de los resultados de tales reclamaciones, e información sobre las reclamaciones recibidas por los funcionarios estatales y las medidas adoptadas.

Transparencia en relación con las prácticas publicitarias de las plataformas digitales

- l.** Para las plataformas digitales que utilizan la publicidad como parte de su modelo de negocio, la información sobre anuncios políticos, incluido el autor y quienes financian los anuncios, debe conservarse en una biblioteca de acceso público en línea.
- m.** Prácticas de publicidad y recopilación de datos y resultados de la evaluación de impacto de los sistemas publicitarios en los derechos humanos y la igualdad de género.
- n.** Información que permita a las personas comprender en qué se basa la publicidad que se les muestra.
- o.** Los contenidos generados exclusivamente por máquinas deben etiquetarse como tales.

Acceso a los datos con fines de investigación

- 116.** Las plataformas digitales deben facilitar a las personas investigadoras autorizadas el acceso a los datos no personales y a los datos seudónimos necesarios para comprender el impacto de las plataformas digitales. Estos datos deben estar disponibles, previa solicitud y de forma continua a través de medios automatizados, como interfaces de programación de aplicaciones (API) u otras soluciones técnicas abiertas y accesibles que permitan el análisis de dichos datos.

- 117.** Se espera que las plataformas digitales proporcionen acceso a datos no personales a periodistas y grupos de sensibilización cuando exista un interés público y el acceso sea proporcionado y necesario en un contexto determinado. Se deben tomar resguardos adicionales para proteger la privacidad de las personas usuarias y sus datos personales – como garantizar la anonimización de los conjuntos de datos mediante diferentes medidas, incluida la desidentificación y el muestreo antes de compartirlos –, así como la información privada de las empresas, los secretos comerciales y el respeto de la confidencialidad comercial.
- 118.** Las plataformas deben crear interfaces fiables para el acceso a los datos y proporcionar datos desglosados por género y otros factores interrelacionados pertinentes (como la raza, el origen étnico, la edad, la situación socioeconómica, la discapacidad, etc.). El sistema de gobernanza debe determinar qué es útil, proporcionado y razonable para fines de investigación.



Principio 4. Información y herramientas a disposición de las personas usuarias

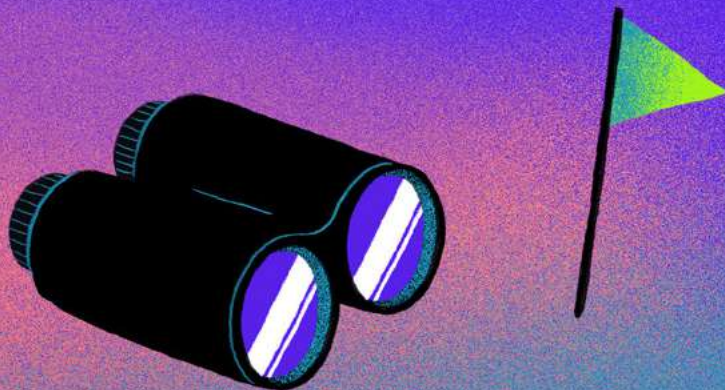
Idioma y accesibilidad

- 119.** Las plataformas deben tener sus condiciones de servicio completas disponibles en el idioma oficial y en los idiomas principales de cada país donde operan, garantizar que puedan responder a las personas usuarias en su propio idioma y procesar sus reclamaciones por igual, así como tener la capacidad de moderar y seleccionar el contenido en el idioma de la persona usuaria. Los traductores automáticos de idiomas pueden utilizarse para ofrecer una mayor accesibilidad lingüística, pero debe controlarse su fidelidad debido a sus limitaciones técnicas.
- 120.** Las plataformas deben garantizar que los informes, notificaciones y procesos de reclamación estén disponibles en la lengua en la que la persona usuaria interactúa con el servicio.
- 121.** Cuando es probable que los niños y niñas accedan a las plataformas digitales, estas deben proporcionarles un acceso equitativo y efectivo a la información,

garantizando la protección de su libertad de expresión y privacidad⁴⁰. Las condiciones de servicio y las normas comunitarias deben estar disponibles en un lenguaje apropiado para la edad de niños y niñas. Según corresponda, deben crearse teniendo en cuenta el punto de vista de un grupo diverso de niños y niñas. Se debe prestar especial atención a las necesidades de niños y niñas con discapacidad para garantizar que disfruten de los mismos niveles de acceso a la información que se establecen en la sección anterior.

- 122.** Siempre deben tenerse en cuenta los derechos de las personas con discapacidad, con especial atención a las formas en que pueden interactuar y presentar reclamaciones en relación con la plataforma. Se espera que las plataformas apliquen los ajustes necesarios para hacer accesible la información relacionada con sus condiciones de servicio, informes, notificaciones y reclamaciones.

40. ACNUDH. 2021. «Observación general núm. 25 (2021) relativa a los derechos de los niños en relación con el entorno digital» <https://www.ohchr.org/es/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>



Principio 5. Las plataformas rinden cuentas ante las partes interesadas pertinentes

Informes de parte de las personas usuarias

- 123.** Las plataformas deben establecer mecanismos de reporte para personas usuarias y no usuarias, o terceras personas que representen sus intereses, para que puedan notificar posibles incumplimientos de la política. Deben existir mecanismos de reporte eficaces y accesibles para las personas que pertenecen a grupos en situación de vulnerabilidad y marginación. Las plataformas digitales también deben contar con los medios necesarios para comprender las condiciones contextuales locales a la hora de responder a las quejas de las personas usuarias, asegurándose de que el diseño del sistema tenga en cuenta las diferencias culturales. Deben establecerse mecanismos de reporte especiales para las niñas y niños, diseñados para un uso rápido y sencillo.
- 124.** El sistema de reporte de las personas usuarias debe dar prioridad a los asuntos sobre contenidos que amenacen a las personas usuarias, garantizando una respuesta rápida y, si es necesario, proporcionar un canal de escalamiento o medios para cumplimentar el reporte. Esto es particularmente importante cuando se trata de vulneración de los derechos humanos, incluidos la violencia y el acoso por razones de género.

125. Las compañías deben esforzarse por evitar el uso indebido del sistema de reporte por comportamientos coordinados no auténticos.

Recursos de reclamación y reparación de las personas usuarias

126. Deben existir mecanismos eficaces de reparación en la plataforma y externos que permitan a las personas usuarias (y a las no usuarias, si se ven afectadas por un contenido específico) expresar sus preocupaciones y obtener la reparación adecuada. Esto debe incluir un canal de presentación de reportes claro, fácilmente accesible, privilegiado, de confianza⁴¹ y comprensible para las quejas en su idioma local, que notifique a las personas usuarias el resultado de su reclamación.

127. El mecanismo de reclamaciones debe seguir los siete principios descritos en los Principios Rectores de las Naciones Unidas sobre Empresas y Derechos Humanos para mecanismos efectivos de presentación de reclamación: legitimidad, accesibilidad, previsibilidad, equidad, transparencia, compatibilidad de derechos y aprendizaje continuo.

128. Las plataformas digitales deben informar a las personas usuarias y explicar los procesos de reclamación cuando se elimina o se etiqueta expresamente su contenido, cuando se restringe la posibilidad de realizar comentarios, de compartirlo o de compaginarlo con publicidad, cuando se limita de manera especial en términos de amplificación o recomendación (a diferencia de la amplificación y recomendación «orgánica/algorítmica»), y comunicarles los motivos para ello. Esto permitiría a las personas usuarias comprender las razones por las que se tomó esa medida en relación con su contenido, el método utilizado (medios automatizados o revisión humana) y en virtud de qué normas de la plataforma se tomó dicha medida. Las plataformas también deberían permitir a las personas usuarias recurrir estas decisiones y obtener la reparación adecuada.

129. Las compañías deben esforzarse por garantizar que los sistemas de reclamación y reparación no se utilicen de forma abusiva por comportamientos coordinados no auténticos.

41. Garantizando la seguridad de las personas usuarias y el cumplimiento de la normativa internacional en materia de derechos humanos.



Disposiciones particulares

Protección de los derechos de todas las personas en situación de vulnerabilidad y marginación, mujeres y niñas, y de aquellos profesionales que puedan estar en riesgo por el ejercicio de la libertad de expresión y acceso a la información, tales como periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente.

130. Las plataformas digitales deben implementar suficientes protecciones especiales para mujeres y niñas, personas usuarias de grupos en situación de vulnerabilidad y marginación, así como para periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente. Para lograrlo, las plataformas digitales deben:

- a.** Llevar a cabo evaluaciones periódicas de impacto en materia de derechos humanos y género en sus políticas, sistemas de moderación y enfoques algorítmicos, con vistas a identificar los riesgos sistémicos para los grupos en situación de vulnerabilidad y marginación, las mujeres y las niñas, así como los periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente, y ajustar las políticas y prácticas para mitigar dichos riesgos.
- b.** Utilizar tecnología de protección de la privacidad para proporcionar a las personas investigadoras externas acceso a los datos internos de las plataformas con el fin de ayudar a identificar la amplificación algorítmica de la violencia basada en género en línea u otras tendencias de violencia derivadas de las tecnologías emergentes.

- c.** Crear equipos de ingeniería dedicados e inclusivos que estén específicamente formados para desarrollar soluciones algorítmicas para la moderación y la curación de contenidos.
- d.** Desarrollar y poner en marcha mecanismos de retroalimentación comunitarios estructurados e inclusivos para combatir los prejuicios de género, culturales y de otro tipo en las nuevas tecnologías.
- e.** Evaluar el impacto en los derechos humanos de sus sistemas y procesos para el tratamiento de editores independientes de noticias y contenidos periodísticos alojados en sus servicios.
- f.** Garantizar la igualdad de trato de los medios de comunicación independientes en las plataformas digitales.
- g.** Establecer procedimientos de protección contra el posible uso indebido de los mecanismos de reporte y los mecanismos de moderación, especialmente el uso indebido y de mala fe destinado a censurar a grupos en situación de vulnerabilidad y marginación, mujeres y niñas, así como periodistas, artistas, personas defensoras de los derechos humanos y defensoras del medio ambiente.

Medidas específicas para la integridad electoral⁴²

- 131.** Las plataformas digitales deben reconocer su papel de apoyo a las instituciones democráticas preservando la integridad electoral. Deben implementar un proceso de evaluación de riesgos específico para la integridad del ciclo electoral en el periodo previo y durante los principales eventos electorales nacionales, elecciones regionales significativas o referéndums constitucionales (por ejemplo, elecciones legislativas o para jefes de estado en un sistema presidencial).
- 132.** Las evaluaciones deben ser transparentes, en consonancia con la debida diligencia en materia de derechos humanos, y deben realizarse con la participación de las partes interesadas electorales pertinentes. Las evaluaciones deben realizarse antes de los eventos electorales a fin de implementar medidas concretas para

42. Más información en las publicaciones de la UNESCO: «Elections in digital times: A guide for electoral practitioners» [Elecciones en la era digital: guía para profesionales electorales] (2022) [<https://unesdoc.unesco.org/ark:/48223/pf0000382102>], y «Joint Declaration on freedom of expression in the digital age of the 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» [Declaración conjunta sobre la libertad de expresión en la era digital del Relator Especial de las Naciones Unidas para la Libertad de Opinión y Expresión, el Representante de la Organización para la Seguridad y la Cooperación en Europa (OSCE) para la Libertad de los Medios de Comunicación y el Relator Especial de la Organización de los Estados Americanos (OEA) para la Libertad de Expresión.] (2020) https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/JointDeclarationDigitalAge_30April2020_EN.pdf.

mitigar los riesgos identificados. Las evaluaciones deben incluir un enfoque de género, dado el aumento de la violencia en línea contra las mujeres votantes, candidatas, activistas, representantes electas y funcionarias de gestión electoral.

- 133.** Las plataformas digitales deben hacer un esfuerzo razonable para garantizar que las personas usuarias tengan acceso a informaciones e ideas de todo tipo de acuerdo con el derecho internacional de los derechos humanos. En particular, deben garantizar que las herramientas automatizadas no obstaculicen el acceso a contenidos relacionados con las elecciones y a puntos de vista diversos.
- 134.** Como parte de la evaluación, las plataformas digitales deben revisar si sus productos, políticas o prácticas sobre publicidad política limitan arbitrariamente la capacidad de las personas candidatas o partidos para difundir sus mensajes.
- 135.** Las plataformas digitales deben hacer un esfuerzo razonable para tratar el contenido que podría ser lícitamente restringido en virtud de la normativa internacional en materia de derechos humanos durante el proceso electoral. La promoción de la comprobación independiente de hechos, los archivos de anuncios, las alertas públicas y otras medidas deben ser consideradas. El compromiso con las instituciones reguladoras oficiales independientes pertinentes puede ser necesario según las circunstancias particulares de cada jurisdicción.
- 136.** Las plataformas digitales deben, según su relevancia, ser transparentes en cuanto al uso y el impacto práctico de cualquier herramienta automatizada que utilicen, aunque no necesariamente sobre la codificación específica con la que operan estas herramientas, incluso en la medida en que estas herramientas afecten a la recopilación de datos, la publicidad dirigida y el intercambio, la clasificación y/o la eliminación de contenidos, especialmente los relacionados con las elecciones.
- 137.** Las plataformas digitales también deben interactuar con las principales partes relevantes y su sistema de gobernanza, antes y durante una elección, a fin de establecer un canal de comunicación mediante el cual el administrador o las personas usuarias/votantes puedan plantear sus inquietudes. En caso de ser requerido, se interactuará con las instituciones reguladoras oficiales relevantes de acuerdo con las circunstancias particulares de cada jurisdicción.
- 138.** Las plataformas digitales que aceptan publicidad destinada a influir en el proceso electoral deben identificar claramente esos contenidos como publicidad. Las condiciones de servicio de las plataformas digitales deben ser claras en cuanto

a la responsabilidad de la plataforma digital de ser transparente sobre el importe de la financiación, la entidad que proporciona los fondos y la entidad anunciada, y aplicar sistemáticamente las mismas normas de moderación y curación de contenidos a la publicidad.

- 139.** Las plataformas digitales deben hacer un seguimiento de la monetización de las publicaciones de los partidos políticos y de las personas que representan a los partidos.
- 140.** Las plataformas deben divulgar al público información sobre los datos demográficos específicos a los que se dirigen dichas publicidades y promociones.
- 141.** Las plataformas deben conservar estos anuncios, y toda la información relevante sobre la financiación, en una biblioteca en línea de acceso público y actualizada periódicamente.

Medidas específicas en situaciones de emergencia, conflicto armado y crisis

- 142.** Como salvaguardia de los derechos humanos, las plataformas digitales deben actuar con la debida diligencia en materia de derechos humanos para abordar las crisis, las situaciones de conflicto armado y otras emergencias, incluidas las emergencias de salud pública. Esta debida diligencia debe analizar el impacto en los derechos humanos de las operaciones, productos, servicios y sistemas publicitarios de las empresas, en contextos de crisis y conflictos.
- 143.** Durante los conflictos armados y las crisis, las plataformas deben:
 - a.** Garantizar que la moderación de contenidos en situaciones de conflicto incluya una sólida revisión humana, incorporando conocimientos especializados en las lenguas pertinentes y en los contextos locales y regionales.
 - b.** Promover la comprobación de hechos.
 - c.** Crear canales para interactuar de forma significativa y directa con las partes interesadas relevantes, incluidas las que operan en zonas afectadas por conflictos y de alto riesgo.
 - d.** Desarrollar la cooperación con socios de confianza, organizaciones de medios de comunicación independientes, entre otros abanderados de confianza.
 - e.** Establecer sistemas de alerta temprana y sistemas claros de escalada para situaciones de emergencia que ayuden a detectar daños inminentes para la seguridad física de las personas.

- f.** Implementar políticas para limitar y rastrear la monetización de contenidos nocivos relacionados con conflictos armados.
 - g.** Preservar toda la evidencia potencial de violaciones de derechos humanos o crímenes de guerra, permitiendo el acceso a este material archivado a los mecanismos de rendición de cuentas nacionales o internacionales apropiados.
- 144.** Las evaluaciones de riesgos pueden requerir que las plataformas digitales dispongan de procesos para casos en los que se produzcan muchas solicitudes simultáneas de acción por parte de las personas usuarias, como ocurre a veces en el contexto de disturbios sociales o violaciones masivas de los derechos humanos. El sistema de gobernanza debe reconocer las orientaciones existentes de los organismos y expertos de la ONU para actuar con una diligencia debida «reforzada» en materia de derechos humanos en tales escenarios.



Conclusión

145. Las plataformas digitales han empoderado a las personas y a las sociedades, otorgando enormes oportunidades para que comuniquen, participen y aprendan. Ofrecen un gran potencial para los grupos en situación de vulnerabilidad y marginación, democratizando espacios de comunicación y oportunidades para que diversas voces se relacionen entre sí, sean escuchadas y vistas. No obstante, en los últimos años el potencial de estas plataformas se ha ido deteriorando gradualmente por la falta de previsión a la hora de hacer frente a los principales riesgos.

146. El objetivo de las Directrices es salvaguardar el derecho a la libertad de expresión en la gobernanza de las plataformas digitales, incluidos el acceso a la información y otros derechos humanos, al tiempo que se tratan aquellos contenidos que pueden restringirse lícitamente en virtud de la normativa internacional en materia de derechos humanos. Por extensión, una gobernanza de las plataformas digitales basada en los derechos humanos promovería aún más la diversidad cultural, la expresión cultural y los contenidos culturales diversos⁴³. Las Directrices describen un sistema de gobernanza que respeta los derechos humanos y promueven procesos de mitigación de riesgos basados en sistemas para la moderación

43. Convención de 2005 sobre la Protección y Promoción de la Diversidad de las Expresiones Culturales, artículos 1 y 4. Según la Convención, se entiende por «contenido cultural» el sentido simbólico, la dimensión artística y los valores culturales que emanan de las identidades culturales o las expresan. Además, las «expresiones culturales» son las expresiones resultantes de la creatividad de personas, grupos y sociedades, que poseen un contenido cultural.

y la curación de contenidos. Estas Directrices ponen de relieve los principios rectores que deben seguirse en todos los sistemas de gobernanza que afectan a la libertad de expresión y al acceso a la información, independientemente del tipo de regulación específico y el enfoque temático, siempre que esas disposiciones estén en consonancia con las establecidas en estas Directrices.

- 147.** Las Directrices fueron elaboradas en base a un proceso de consulta de múltiples partes interesadas que comenzó en septiembre de 2022. Estas Directrices son el resultado de un amplio proceso de consulta en el que la UNESCO recibió, a través de tres consultas abiertas⁴⁴, más de 1 540 propuestas con más de 10 000 comentarios. Diferentes grupos de partes interesadas de 134 países participaron en este proceso. Se trata de uno de los ejercicios de consulta más amplios y abiertos jamás realizados por la UNESCO.
- 148.** Se celebrarán más adelante nuevos debates sobre cómo hacer operativas las Directrices y utilizarlas como herramienta de sensibilización, con el objetivo de proteger la libertad de expresión, el acceso a la información y todos los demás derechos humanos en el entorno digital.

44. Las tres consultas abiertas tuvieron lugar entre diciembre de 2022 y enero de 2023, entre febrero y marzo de 2023 y entre abril y junio de 2023.

Apéndice

Recursos

Naciones Unidas

Principios Rectores sobre las Empresas y los Derechos Humanos. (2011)

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

Plan de acción de Rabat sobre la prohibición de la apología del odio nacional, racial o religioso que constituya la incitación a la discriminación, la hostilidad o la violencia. (2012)

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

Relatora Especial de la ONU sobre Libertad de opinión y expresión: «A human rights approach to online content moderation» (Un enfoque de derechos humanos para la moderación del contenido en línea) (2018)

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

Convención sobre los Derechos del Niño: «Observación general núm. 25 (2021) relativa a los derechos de los niños en relación con el entorno digital» (2021)

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

Informe del Secretario General de las Naciones Unidas: Contrarrestar la desinformación para promover y proteger los derechos humanos y las libertades fundamentales (2022)

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

Relatora Especial de la ONU sobre la promoción y protección del derecho a la libertad de opinión y de expresión: «A/77/288: La desinformación y la libertad de opinión y de expresión durante los conflictos armados» (2022)

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

Informe de políticas 8 de Nuestra Agenda Común sobre la integridad de la información en las plataformas digitales, publicado por el Secretario General de la ONU. <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-es.pdf>

Informe de políticas 5 de Nuestra Agenda Común, publicado por el Secretario General de la ONU: Un Pacto Digital Global

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

UNESCO

Puddephatt, A. 2021. «Dejar entrar el sol: transparencia y responsabilidad en la era digital.» París: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000377231_spa

UNESCO. 2021. «Los límites legítimos de la libertad de expresión: la prueba de tres partes». [vídeo] <https://www.youtube.com/watch?v=4AG2y06CsC8>

Referencias sobre la terminología

Contenido que incita o retrata la violencia por razones de género

- Consejo de Derechos Humanos de las Naciones Unidas. 2018. «Informe de la Relatora Especial sobre la violencia contra la mujer, sus causas y consecuencias acerca de la violencia en línea contra las mujeres y las niñas desde la perspectiva de los derechos humanos».
<https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/184/61/PDF/G1818461.pdf?OpenElement>
- ACNUDH. 2021. «Declaración de Irene Khan, Relatora Especial sobre la promoción y protección de la libertad de opinión y de expresión».
<https://www.ohchr.org/en/statements/2022/02/statement-irene-khan-special-rapporteur-promotion-and-protection-freedom-opinion>

Diversidad cultural

Por «diversidad cultural» se entienden las múltiples formas en que se expresan las culturas de grupos y sociedades. Estas expresiones se transmiten dentro de los grupos y sociedades y entre ellos. La diversidad cultural se manifiesta no solo a través de las variadas formas en que el patrimonio cultural de la humanidad se expresa, aumenta y transmite mediante la variedad de expresiones culturales, sino también a través de diversos modos de creación artística, producción, difusión, distribución y disfrute, cualesquiera que sean los medios y tecnologías utilizados.

- UNESCO. 2005. *La Convención de 2005 sobre la Protección y la Promoción de la Diversidad de las Expresiones Culturales*. París: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000246264_spa

Desinformación e información errónea

- «La desinformación y la libertad de opinión y de expresión» Informe de la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, Irene Khan. <https://digitallibrary.un.org/record/3925306?ln=es>

Discurso de odio

- ONU. 2019. «Estrategia y Plan de Acción de las Naciones Unidas para la Lucha contra el Discurso de Odio». <https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml>

Regulador independiente

Un regulador independiente tiene sus facultades y responsabilidades establecidas en un instrumento de derecho público y cuenta con las facultades para administrar sus propios recursos. Sus miembros son nombrados de manera independiente y protegidos por la ley contra el despido injustificado. En este caso, las decisiones del regulador se toman sin la aprobación previa de ninguna otra entidad gubernamental, y ninguna entidad que no sea un tribunal o un panel de apelación preestablecido puede anular las decisiones del regulador. Los pilares institucionales para la independencia de la toma de decisiones son la independencia organizacional (separada organizativamente de los ministerios y departamentos existentes), la independencia financiera (una fuente de financiación asignada, segura y adecuada) y la independencia de la administración (autonomía sobre la administración interna y protección contra el despido sin causa justificada).

- Salomon, Eve. 2016. *Guidelines for Broadcasting Regulation (Directrices para la regulación de la radio y teledifusión)*. París: UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000144292>
- Brown, A., J. Stern, B. Tenenbaum y D. Gencer. 2006. *Handbook for Evaluating Infrastructure Regulatory Systems (Manual del Banco Mundial para evaluar sistemas reguladores de infraestructura)*. Washington, DC: Grupo Banco Mundial. <http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-6579-3>

Regulador

Un organismo que supervisa, monitorea y hace responsable a un actor privado.

Directrices para la gobernanza de las plataformas digitales

Salvaguardar la libertad de expresión y el acceso a la información con un enfoque de múltiples partes interesadas

Redactadas a través de una amplia consulta, estas Directrices tienen como objetivo salvaguardar el derecho a la libertad de expresión y el acceso a la información en el contexto de la gobernanza de las plataformas digitales. Las Directrices presentan un enfoque de múltiples partes interesadas que describe las responsabilidades clave de los Estados y las plataformas digitales para cumplir con sus obligaciones de derechos humanos, y definen roles para las organizaciones intergubernamentales, la sociedad civil, los medios de comunicación, el mundo académico, la comunidad técnica y otros actores.

<https://www.unesco.org/es/internet-trust/guidelines>
internetconference@unesco.org



unesco

Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura



9 789233 002159