



Asamblea General

Distr. general
18 de junio de 2020
Español
Original: inglés

Consejo de Derechos Humanos

44º período de sesiones

15 de junio a 3 de julio de 2020

Tema 9 de la agenda

Racismo, discriminación racial, xenofobia y formas conexas de intolerancia, seguimiento y aplicación de la Declaración y el Programa de Acción de Durban

La discriminación racial y las tecnologías digitales emergentes: un análisis de los derechos humanos

Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia*

Resumen

En el presente informe, la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia, E. Tendayi Achiume, analiza diferentes formas de discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes, incluidas las dimensiones estructurales e institucionales de esa discriminación. También perfila las obligaciones de los Estados en materia de derechos humanos y la responsabilidad de las empresas en la lucha contra esta discriminación.

* El presente informe se presenta con retraso para poder incluir en él la información más reciente.



I. Introducción

1. Las tecnologías digitales emergentes han alterado fundamentalmente la forma en que vivimos nuestras vidas y, como tales, sus repercusiones en los derechos humanos han sido objeto de importantes análisis por parte de los procedimientos especiales del Consejo de Derechos Humanos¹. En los informes realizados se aborda la forma en que esas tecnologías afectan a una amplia diversidad de derechos humanos, incluidos los derechos a la libertad de opinión y expresión, la libertad de reunión pacífica y de asociación y los derechos humanos de las personas que se encuentran en la extrema pobreza. La Alta Comisionada de las Naciones Unidas para los Derechos Humanos ha aportado su contribución al análisis de las tecnologías digitales emergentes y el derecho a la privacidad². El objetivo de la Relatora Especial con el presente informe es promover un análisis de una solidez equiparable en el ámbito en que entrecruzan las tecnologías digitales emergentes y los principios de igualdad racial y no discriminación enunciados en el derecho internacional de los derechos humanos

2. El informe abarca temas como el racismo, la intolerancia, la discriminación y otras formas de exclusión y diferenciación perjudiciales por motivos de raza, color, ascendencia u origen nacional o étnico, conforme a su definición en la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial. También se incluye la discriminación contra los pueblos indígenas. En el presente informe, la Relatora Especial insta a que, tratándose de la gobernanza de los derechos humanos de las tecnologías digitales emergentes, se adopte un enfoque que se base en el criterio de la igualdad. A tal efecto, es preciso superar las estrategias de “daltonismo racial” o “neutralidad racial”³. Un análisis basado en el daltonismo racial de las condiciones jurídicas, sociales, económicas y políticas asume una postura de imparcialidad que supone evitar un análisis racial o étnico explícito en aras de tratar a todos los individuos y grupos de la misma manera, aun cuando la situación de esos individuos y grupos no sea la misma, debido a, entre otras cosas, estructuras históricas de subordinación racial. En el contexto de las tecnologías digitales emergentes es preciso que los altos cargos públicos, las Naciones Unidas y otras organizaciones multilaterales, así como el sector privado, presten una detenida atención a sus repercusiones racializadas y étnicas. En el presente informe, la Relatora Especial pone de relieve las formas interseccionales de discriminación, incluida la de género y la discriminación por motivos religiosos, y destaca la incapacidad de los Estados y otras partes interesadas para detectar y abordar formas complejas de discriminación en los ámbitos donde se produce una intersección entre la raza, el origen étnico, el género, la condición de persona con discapacidad, la orientación sexual y otros motivos conexos.

3. La Relatora Especial solo aborda brevemente los efectos racialmente discriminatorios de las tecnologías digitales emergentes en los migrantes, los refugiados y otros no ciudadanos, porque estos grupos serán el tema sobre el que versará otro informe de la Relatora Especial que elevará a la Asamblea General.

4. Una de las principales conclusiones del informe es que las tecnologías digitales emergentes agudizan y amplifican las desigualdades existentes, muchas de las cuales obedecen a motivos de raza, etnia y origen nacional. Los ejemplos destacados en el informe suscitan preocupación por las diferentes formas de discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes. En algunos casos, esta discriminación es directa y está explícitamente motivada por la intolerancia o los prejuicios. En otros, la discriminación es el resultado de efectos dispares en determinados grupos según su raza, etnia u origen nacional, aun cuando no medie ninguna intención explícita de discriminar. También hay otros casos en que se conjugan formas directas e indirectas de discriminación y pueden tener un efecto global o sistémico tan importante como para someter a determinados grupos a estructuras de discriminación racial que condicionan el acceso y el disfrute de los derechos humanos en todos los ámbitos de sus vidas.

¹ Véase www.ohchr.org/Documents/HRBodies/SP/List_SP_Reports_NewTech.pdf.

² Véase A/HRC/39/29.

³ Véase <https://qz.com/1585645/color-blindness-is-a-bad-approach-to-solving-bias-in-algorithms>.

5. En el contexto de la pandemia de la enfermedad del coronavirus (COVID-19), por ejemplo, los primeros informes han puesto de relieve los efectos dispares de la pandemia en los grupos raciales y étnicos marginados, entre otras cosas porque esos grupos quedan excluidos de los beneficios de las tecnologías digitales emergentes, o porque estas tecnologías se implantan en formas que exponen a esos grupos a un mayor peligro de ser víctimas de violaciones de los derechos humanos. A pesar de la sensación generalizada de que las tecnologías digitales emergentes son neutrales y objetivas en su funcionamiento, la raza y el origen étnico determinan el acceso a los derechos humanos y su disfrute en todos los campos en los que estas tecnologías están actualmente generalizadas. Los Estados tienen la obligación de prevenir, combatir esta discriminación racial y corregir esta situación, y los agentes privados, como las empresas, tienen la responsabilidad conexas de hacer otro tanto.

6. Entre las tecnologías digitales emergentes, la Relatora Especial se centra en el informe sobre las tecnologías en red y predictivas, de las que muchas incorporan tecnologías de macrodatos e inteligencia artificial, en las que adquiere cierto protagonismo la adopción de decisiones algorítmica (y asistida por algoritmos). Gran parte del análisis actual de la discriminación racial y las tecnologías digitales emergentes en relación con los derechos humanos ha arrojado luz sobre un conjunto específico de cuestiones: los incidentes de odio en Internet y la utilización de plataformas digitales para coordinar, financiar y fomentar el apoyo a comunidades racistas y a sus actividades. En el informe, la Relatora Especial va un paso más allá, al hacer que los principios de igualdad racial y no discriminación se tengan en cuenta al analizar los efectos estructurales e institucionales de las tecnologías digitales emergentes, que investigadores, defensores de derechos humanos y otros han calificado de alarmantes. Entre los motivos de preocupación figura la preeminencia de las tecnologías digitales emergentes en la determinación de los resultados cotidianos en el empleo, la educación, la atención de la salud y la justicia penal, lo que acarrea el riesgo de una discriminación sistematizada a una escala sin precedentes. En un reciente informe de la Agencia de Derechos Fundamentales de la Unión Europea se destacan ejemplos de estos motivos de preocupación en la Unión Europea y se formulan valiosas recomendaciones para la adopción de la respuesta requerida⁴.

7. En cuanto “tecnologías de clasificación que diferencian, jerarquizan y categorizan”, los sistemas de inteligencia artificial son en esencia “sistemas de discriminación”⁵. Los algoritmos de aprendizaje automático reproducen los sesgos incorporados en los conjuntos de datos a gran escala capaces de imitar y reproducir los sesgos implícitos de los humanos, incluso en ausencia de reglas algorítmicas explícitas que generen estereotipos⁶. Los conjuntos de datos, como producto de diseño humano, pueden estar sesgados debido a “desviaciones, lagunas y premisas erróneas”⁷. También pueden adolecer de “problemas de señal”, no representación demográfica o infrarrepresentación debido a la desigualdad de las formas en que se crearon o reunieron los datos⁸. Además de los datos inexactos, ausentes y mal representados, los “datos sucios” incluyen datos que han sido manipulados intencionalmente o distorsionados por sesgos⁹. Esos conjuntos de datos pueden propiciar la discriminación o la exclusión de determinadas poblaciones, en particular de minorías articuladas por identidad de raza, etnia, religión y género.

8. Aun cuando no sea lo que se pretenda, la discriminación indirecta puede ser el resultado de la utilización de criterios inocuos y genuinamente pertinentes que también funcionan como indicadores indirectos de raza y origen étnico. Otros motivos de preocupación son el uso y la dependencia de modelos predictivos que incorporan datos históricos —datos que a menudo reflejan sesgos discriminatorios y perfiles inexactos—,

⁴ Véase <https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making>.

⁵ Sarah Myers West, Meredith Whittaker y Kate Crawford, “Discriminating systems: gender, race and power in AI” (Nueva York, AI Now Institute, 2019), pág. 6.

⁶ Véase https://philmachinelearning.files.wordpress.com/2018/02/gabriellejohnson_algorithmic-bias.pdf.

⁷ Véase <https://foreignpolicy.com/2013/05/10/think-again-big-data>.

⁸ *Ibid.*

⁹ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3403010.

incluso en contextos como el de las fuerzas del orden, la seguridad nacional y la inmigración. A un nivel más fundamental, el diseño de las tecnologías digitales emergentes requiere que los desarrolladores tomen decisiones sobre la mejor manera de alcanzar sus objetivos, y esas decisiones tienen siempre consecuencias distributivas diferentes¹⁰. Una de las preocupaciones fundamentales que la Relatora Especial expresa en el informe es la de las opciones que afectan de manera dispar a los derechos humanos de las personas y los grupos por motivos de raza, etnia y otros motivos conexos.

9. En lo que respecta en particular a la clase, los estudios ponen de relieve que incluso cuando los encargados de la formulación de políticas, los funcionarios públicos y los científicos han procurado automatizar la adopción de decisiones con la intención de tomar decisiones más eficientes y equitativas, se ha demostrado que los sistemas que utilizaban para lograr esos fines reforzaban la desigualdad y daban lugar a resultados punitivos para las personas que vivían en la pobreza¹¹. Dado que las comunidades racial y étnicamente marginadas suelen vivir en número desproporcionadamente mayor en condiciones de pobreza, los principios de igualdad y no discriminación deberían ser fundamentales en los análisis de derechos humanos y tecnologías digitales emergentes en sistemas de bienestar social y otros sistemas socioeconómicos. En un importante informe presentado recientemente por el Relator Especial sobre la extrema pobreza y los derechos humanos se describe el auge del estado de bienestar digital en los países en que los sistemas de protección y asistencia social están propulsados por tecnologías digitales emergentes y en formas que tienen graves consecuencias negativas para los derechos humanos en materia de bienestar social¹². Como se explica en una sección posterior, la evaluación del Relator Especial es que los estados de bienestar digitales, tal como existen hoy en día, se describen mejor como “estados de bienestar digitales discriminatorios” porque tienden a permitir que la raza y el origen étnico (entre otros motivos) configuren el acceso a los derechos humanos de forma discriminatoria. Es necesario tomar medidas urgentes para poner freno a estos patrones discriminatorios.

10. En la preparación del informe, la Relatora Especial aprovechó las valiosas aportaciones de: las reuniones de grupos de expertos organizadas por el Instituto de Estudios Mundiales de la Universidad de Ginebra, el Promise Institute for Human Rights de la Facultad de Derecho de la Universidad de California en Los Ángeles (UCLA) y el Center for Critical Internet Inquiry de esa misma universidad; las investigaciones de la Clínica de Ciberderecho de la Facultad de Derecho de Harvard del Centro Berkman Klein para Internet y Sociedad, y el Centro de la Facultad de Derecho de la Universidad de Nueva York sobre la Raza, la Desigualdad y el Derecho; una serie de entrevistas con los investigadores; y las comunicaciones recibidas de una serie de partes interesadas en respuesta a una convocatoria pública de presentación de comunicaciones. Las comunicaciones no confidenciales se podrán consultar en el sitio de la Relatoría en la Web¹³.

II. Factores que propician la discriminación y la desigualdad en las tecnologías digitales emergentes

11. Todo análisis de los derechos humanos en relación con las tecnologías digitales emergentes debe encarar en primer lugar la cuestión de las fuerzas sociales, económicas y políticas que conforman su diseño y uso, así como de los intereses y prioridades humanas individuales y colectivas en juego que contribuyen al diseño y uso racialmente discriminatorio de esas tecnologías.

12. La percepción pública suele ser que la tecnología es intrínsecamente neutra y objetiva; algunos han señalado que esta presunción de objetividad y neutralidad tecnológica es una característica que aún sigue siendo una de las más destacables incluso entre los

¹⁰ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899.

¹¹ Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (Nueva York, Picador, 2018).

¹² Véase A/74/493.

¹³ Véase www.ohchr.org/EN/Issues/Racism/SRRacism/Pages/Callinformationtechnologies.aspx.

creadores de tecnologías. Ahora bien, la tecnología nunca es neutra, pues refleja los valores e intereses de quienes influyen en su diseño y uso, y está fundamentalmente conformada por las mismas estructuras de desigualdad que operan en la sociedad¹⁴. Por ejemplo, en un examen realizado en 2019 de 189 algoritmos de reconocimiento facial realizados por 99 desarrolladores de todo el mundo se puso de manifiesto que “las probabilidades de que muchos de estos algoritmos identificasen incorrectamente una fotografía de un rostro negro o con rasgos de Asia Oriental eran entre 10 y 100 mayores que si se tratase de la fotografía de una persona blanca”. Al buscar en una base de datos para encontrar un determinado rostro, las tasas de error al seleccionar las imágenes entre las mujeres negras eran considerablemente mayores que en el caso de otros grupos demográficos¹⁵. Ya no cabe duda de que las tecnologías digitales emergentes tienen una capacidad sorprendente para reproducir, reforzar e incluso agudizar la desigualdad racial en el seno de las distintas sociedades y entre ellas. En varios estudios académicos importantes se demuestra de manera concreta que el diseño y la utilización de las tecnologías ya están teniendo este preciso efecto en diversos contextos¹⁶. Se necesitan más investigaciones y financiación para desentrañar plenamente la forma en que incluso los procesos inductivos que constituyen el núcleo de algunas técnicas de inteligencia artificial, como el aprendizaje automático, contribuyen a socavar valores como la igualdad y la no discriminación¹⁷.

13. En los campos y sectores generadores de tecnologías digitales emergentes se ha demostrado que esa fe equivocada en la neutralidad u objetividad de los números y su poder para superar el racismo ha contribuido a resultados discriminatorios¹⁸. Incluso en el campo que se ha desarrollado para promover la equidad, la rendición de cuentas y la transparencia en el diseño y la utilización de las tecnologías digitales emergentes hay que prestar una mayor atención a las estructuras societales más amplias de discriminación e injusticia¹⁹. De hecho, entre las mayores dificultades para superar el problema del uso y el diseño racialmente discriminatorio de las tecnologías digitales emergentes se sitúan los enfoques que plantean esta cuestión como un problema exclusivamente o en gran medida tecnológico al que deben encontrar solución los científicos computacionales y otros profesionales del sector creando datos y algoritmos sin sesgos. La tecnología es un producto de la sociedad, sus valores, sus prioridades e incluso sus desigualdades, incluidas las relacionadas con el racismo y la intolerancia. El determinismo tecnológico, esto es, la idea de que la tecnología influye en la sociedad pero que en sí misma es en gran medida neutral y está aislada de las fuerzas sociales, políticas y económicas, solo sirve para blindar a las fuerzas que dan forma a las tecnologías digitales emergentes y sus efectos contra toda detección y reforma. El “tecnochovinismo”, esto es, una confianza excesiva en la creencia de que la tecnología puede resolver los problemas de la sociedad²⁰, suscita un efecto similar, y puede complicar el cuestionamiento y el cambio de los valores e intereses que determinan las formas que adopta la tecnología y los productos tecnológicos.

14. Aunque sigue siendo muy necesario someter a examen la calidad de la ingeniería y exigir responsabilidades para garantizar los principios de igualdad y no discriminación, la garantía de estos y otros principios de derechos humanos debe comenzar por reconocer que el *quid* de la cuestión es político, social y económico, y no solo un problema meramente tecnológico o matemático. La desigualdad y la discriminación, incluso en las circunstancias

¹⁴ Langdon Winner, *The Whale and The Reactor: A Search for Limits in an Age of High Technology* (Chicago, University of Chicago Press, 1986), pág. 29.

¹⁵ Véase www.scientificamerican.com/article/how-nist-tested-facial-recognition-algorithms-for-racial-bias.

¹⁶ Véase, por ejemplo, Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York, Penguin, 2016); Ruha Benjamin, *Race After Technology* (Cambridge, Reino Unido, Polity Press, 2019); y Safiya Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York, New York University Press, 2018).

¹⁷ Véase, por ejemplo, Gabrielle M. Johnson, “Are algorithms value-free? Feminist theoretical virtues in machine learning”, *Journal of Moral Philosophy* (de próxima publicación).

¹⁸ Véase, por ejemplo, West, Whittaker and Crawford, “Discriminating systems”.

¹⁹ Véase, por ejemplo, www.tandfonline.com/doi/full/10.1080/1369118X.2019.1593484 y http://sorelle.friedler.net/papers/sts_fat2019.pdf.

²⁰ Véase, por ejemplo, Meredith Broussard, *Artificial Unintelligence: How Computers Misstand the World* (Cambridge, Massachusetts, Massachusetts Institute of Technology, 2018).

en que son producto del diseño y la utilización de las tecnologías digitales emergentes, no se “curarán” con una modelización más perfecta de la igualdad y la no discriminación. Concretamente, esto supone que el pensamiento y la acción que tratan de combatir la discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes, tanto en el sector privado como en el público, no deben ser el dominio exclusivo o casi exclusivo de los tecnólogos. Al contrario, ese pensamiento y esa acción y debe ser más holística, como han sostenido investigadores y otras personas expertas en tecnologías digitales emergentes²¹. Los gobiernos y el sector privado deben hacer todo lo posible por adoptar enfoques con la participación de expertos en las dimensiones políticas, económicas y sociales de la discriminación racial en todas las etapas de la investigación, el debate y la adopción de decisiones para mitigar el diseño y la utilización racialmente discriminatorios de las tecnologías digitales emergentes. Las comunidades raciales y minorías étnicas afectadas deben tener un papel protagonista en la adopción de decisiones en los procesos pertinentes.

15. Las empresas privadas ejercen una enorme influencia en el diseño y el uso de las tecnologías digitales emergentes. Entre las plataformas digitales, siete “superplataformas”, a saber, Microsoft, Apple, Amazon, Google, Facebook, Tencent y Alibaba, representan dos tercios del valor total de mercado de las 70 plataformas más grandes del mundo²². A pesar del alcance mundial de sus tecnologías digitales emergentes, las empresas que ejercen una mayor influencia sobre ellas se concentran predominantemente en Silicon Valley, en los Estados Unidos de América, mientras que la participación de Europa es del 3,6 %, la de África del 1,3 % y la de América Latina del 0,2 %²³. Así, por ejemplo, Google acapara alrededor del 90 % del mercado de búsquedas en Internet²⁴. Facebook representa dos tercios del mercado mundial de medios sociales y es la primera plataforma de medios sociales en más del 90 % de las economías del mundo. Amazon tiene una cuota de casi el 40 % del mercado mundial de comercio electrónico minorista. Consiguientemente, los valores culturales, económicos y políticos específicos de Silicon Valley determinan fundamentalmente la forma en que muchas de las tecnologías digitales emergentes funcionan a nivel mundial, incluso en ámbitos muy alejados de esta pequeña región de América del Norte.

16. Más allá del dominio del mercado, las empresas sirven de intermediarios fundamentales entre los gobiernos y sus naciones, y además tienen la capacidad para transformar significativamente la situación de los derechos humanos. La tecnología creada por poderosas empresas del Norte global se inscribe en un contexto político, económico, social y de gobernanza muy concreto. Puede tener efectos muy perjudiciales en otros contextos, como los del Sur global. Un ejemplo es el papel que Facebook tuvo en Myanmar²⁵. También son preocupantes las condiciones no reguladas, y en algunos casos expoliadoras, en las que extraen datos de personas y naciones del Sur global algunas sociedades con fines de lucro del Norte global sin que se puedan depurar responsabilidades²⁶.

17. Los sectores de las tecnologías digitales emergentes, como los de Silicon Valley, se caracterizan por una “crisis de diversidad” en lo que respecta al género y la raza²⁷, especialmente en los niveles directivos más altos. Según un importante estudio en este campo, “actualmente, los sistemas de IA a gran escala se desarrollan casi exclusivamente en un puñado de empresas tecnológicas y un pequeño conjunto de laboratorios

²¹ Véase, por ejemplo, www.odproject.org/2019/07/15/critiquing-and-rethinking-fairness-accountability-and-transparency.

²² *Informe sobre la Economía de la Información 2013: La economía de la nube y los países en desarrollo* (publicación de las Naciones Unidas, número de venta: S.13.II.D.6), pág. xvi.

²³ *Ibid.*, pág. 2.

²⁴ *Ibid.*, pág. xvii.

²⁵ Véase A/HRC/39/64.

²⁶ Véase, por ejemplo, <https://privacyinternational.org/long-read/3390/2020-crucial-year-fight-data-protection-africa>.

²⁷ Véase www.technologyreview.com/2018/02/14/145462/were-in-a-diversity-crisis-black-in-ai-founder-on-whats-poisoning-the-algorithms-in-our; Noble, *Algorithms of Oppression*; y West, Whittaker and Crawford, “Discriminating systems”.

universitarios de élite, espacios que en Occidente tienden a estar caracterizados por la fuerte presencia de varones blancos, acomodados y con un perfil técnico”. También hay espacios que arrastran un historial de problemas de discriminación, exclusión y acoso sexual²⁸. El estudio constata además que “se trata de un problema que va mucho allá y no se limita a uno o dos malos actores, sino que es indicativo, por un lado, de una relación sistemática entre patrones de exclusión en el campo de la IA y la industria que impulsa su producción y, por otro, de los sesgos que se manifiestan en la lógica y la aplicación de las tecnologías de IA²⁹. Es probable que la tecnología desarrollada en esos campos que excluyen de manera desproporcionada a las mujeres y a las minorías raciales, étnicas y de otro tipo reproduzcan esas desigualdades una vez implantada. Desarrollar tecnologías que encajen dentro de complejas realidades sociales y de los sistemas existentes exige comprender los contextos sociales, jurídicos y éticos, lo que solo puede hacerse incorporando perspectivas diversas y representativas, así como conocimientos técnicos en distintas disciplinas³⁰.

18. Las fuerzas económicas y del mercado ejercen una poderosa influencia en el diseño y la utilización de las tecnologías digitales emergentes, que, a su vez, tienen un efecto transformador en los mercados, incluso en el propio capitalismo³¹. Por un lado, ciertos actores con influencia económica buscan intencionalmente promover la discriminación y la intolerancia. Un ejemplo podría ser el caso de personajes con gran patrimonio que financian plataformas digitales que defienden la ideología supremacista³². Por otro, las fuerzas más poderosas del mercado puede que persigan primordialmente rentabilizar estas tecnologías sin una intención explícita de propagar ideas racistas o intolerantes. Ahora bien, la evidencia demuestra que los productos rentables pueden generar discriminación racial. Cuando las economías están estructuradas por la desigualdad racial y étnica, como ocurre en todo el mundo, la maximización de los beneficios suele ser coherente con la desigualdad racial y étnica y, en muchos casos, la refuerza o la agrava.

19. En gran medida, las desigualdades en el acceso y el disfrute de los beneficios de las tecnologías digitales emergentes se deben a: a) las desigualdades geopolíticas en el plano internacional; y b) los patrones de desigualdad racial, étnica y de género dentro de cada país³³.

20. A nivel internacional, los países del Sur global carecen de la infraestructura digital que existe en el Norte global: la suscripción activa a servicios de banda ancha en el Sur global es menos de la mitad que el Norte global³⁴. En África, tan solo el 22 % de la población utiliza Internet, mientras que en Europa es el 80 %³⁵. En los llamados países menos adelantados, solo una de cada cinco personas está conectada a Internet, mientras que son cuatro de cada cinco en los llamados países desarrollados³⁶. Aunque la tecnología ha sido sumamente útil en las respuestas nacionales contra la pandemia de COVID-19, sus beneficios no se han distribuido de manera uniforme. Los países menos adelantados no solo son los más vulnerables a las consecuencias humanas y económicas de la COVID-19, sino también los menos preparados digitalmente para acceder a la información sobre salud pública en línea y para hacer uso de las plataformas digitales de escolarización, trabajo y compras³⁷.

21. Los Estados Unidos y China son los países que dominan la economía digital global. Esos dos países representan el 90 % del valor de capitalización de mercado de las 70 plataformas digitales más grandes del mundo, entre ellas plataformas de medios sociales y

²⁸ West, Whittaker y Crawford, “Discriminating systems”, pág. 6 (se omite la nota a pie de página).

²⁹ *Ibid.*, pág. 7.

³⁰ *Ibid.*

³¹ Véase, por ejemplo, Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York, Public Affairs, 2019).

³² Véase, por ejemplo, <https://journals.sagepub.com/doi/full/10.1177/1536504218766547>.

³³ Véase, por ejemplo, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3403010.

³⁴ UNCTAD, 2018b, *Informe sobre el Comercio y el Desarrollo, 2018: El poder, las plataformas y la quimera del libre comercio* (publicación de las Naciones Unidas, núm. de venta: S.18.II.D.7, Nueva York y Ginebra).

³⁵ *Ibid.*

³⁶ Véase *Informe sobre la Economía Digital 2019*, pág. 13

³⁷ Véase https://unctad.org/en/PublicationsLibrary/dtinf2020d1_en.pdf.

de contenidos, plataformas de comercio electrónico, servicios de búsqueda en Internet, ecosistemas móviles y plataformas industriales basadas en la nube³⁸. Las predicciones actuales parecen indicar que las tecnologías digitales emergentes están a punto de ampliar aún más la brecha digital entre los países que disponen de ella y los que carecen de la capacidad para aprovecharlas.

22. Las brechas digitales también existen dentro de los países. Por ejemplo, a pesar del predominio de los Estados Unidos en la economía digital mundial, las minorías raciales y étnicas de ese país tienen un acceso desigual a los beneficios que ofrecen las tecnologías digitales emergentes. En muchos casos, son objeto de las más importantes vulneraciones de los derechos humanos relacionadas con las tecnologías digitales emergentes, como se verá en la parte III. Según un estudio realizado en 2019 por el Pew Research Center, los adultos negros e hispanos de los Estados Unidos son los grupos con menores probabilidades de poseer una computadora o de tener Internet de alta velocidad en su casa³⁹. Mientras que el 82 % de los blancos declaran poseer una computadora de escritorio o un portátil, solo el 58 % de los negros y el 57 % de los hispanos lo hacen⁴⁰. También existen diferencias raciales y étnicas sustanciales en la adopción de la banda ancha: los blancos tienen entre un 13 % y un 18 % más de probabilidades de declarar que tienen una conexión de banda ancha en casa que los negros o los hispanos⁴¹. Esta brecha digital articulada en torno al eje de raza y etnia es considerable. Sin embargo, como han sostenido varios investigadores, las actuaciones para promover la inclusión digital de las minorías raciales y étnicas no deben llevarse a cabo de manera que las expongan a nuevas violaciones de derechos, incluso como resultado de las preocupaciones por la privacidad y la vigilancia⁴². En el caso de China, en la parte III se reseñan una serie de ejemplos de las graves consecuencias que acarrearán para los derechos humanos el diseño y la utilización de las tecnologías digitales emergentes. Estas preocupaciones se ven amplificadas por la creciente influencia en el Sur global de las tecnologías digitales emergentes de China.

23. Los pueblos indígenas también son víctimas de una exclusión discriminatoria de los beneficios de las tecnologías digitales emergentes⁴³. Las estimaciones en el Canadá indican que aproximadamente la mitad de la población predominantemente indígena del Norte no tienen posibilidad de contratar una conexión de alta velocidad como la que disfrutaban sus homólogos del Sur⁴⁴. La inclusión digital de los indígenas también es baja en Australia, especialmente fuera de las ciudades, ya que en 2011 solo el 6 % de los residentes de algunas comunidades aborígenes de las zonas remotas disponen de una computadora⁴⁵. En 2015, los aborígenes seguían teniendo un 69 % menos probabilidades de tener algún tipo de conexión a Internet que los no aborígenes⁴⁶.

III. Ejemplos de discriminación racial en el diseño y el uso de las tecnologías digitales emergentes

A. Intolerancia explícita y conducta motivada por prejuicios

24. Los actores que tratan de difundir el discurso racista y la incitación a la discriminación y la violencia han recurrido a las tecnologías digitales emergentes, en las que las plataformas de medios sociales desempeñan un papel fundamental. La Relatora

³⁸ Véase *Informe sobre la Economía Digital 2019*, pág. xvi.

³⁹ Véase www.pewresearch.org/fact-tank/2019/08/20/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites.

⁴⁰ *Ibid.*

⁴¹ *Ibid.*

⁴² Véase <https://journals.uic.edu/ojs/index.php/fm/article/view/3821>.

⁴³ Véase <https://unesdoc.unesco.org/ark:/48223/pf0000260781>.

⁴⁴ Véase <https://rightscon2018.sched.com/event/EHqs/addressing-the-digital-divide-in-indigenous-communities-in-north-america>.

⁴⁵ Véase www.creativespirits.info/aboriginalculture/economy/internet-access-in-aboriginal-communities.

⁴⁶ *Ibid.*

Especial ha destacado estas tendencias en informes anteriores sobre los grupos neonazis y otros grupos de supremacistas blancos que se apoyan en las plataformas de medios sociales para reclutar, recaudar fondos y coordinar sus actividades⁴⁷. Otro ejemplo destacado de la utilización explícita de las tecnologías digitales emergentes para promover prejuicios es la utilización de Facebook por grupos budistas nacionalistas radicales y agentes militares de Myanmar para exacerbar la discriminación y la violencia contra los musulmanes, en particular contra la minoría étnica rohinyá⁴⁸. En 2018, el consejero delegado de Facebook, Mark Zuckerberg, en su comparecencia ante el Senado de los Estados Unidos, declaró que los sistemas de inteligencia artificial de Facebook eran incapaces de detectar los discursos de odio en esos contextos⁴⁹. Estos no son los únicos casos: en una comunicación se destacó también el uso de Facebook para amplificar los contenidos de corte discriminatorio y la intolerancia, entre ellos contenidos que incitan a la violencia contra los grupos minoritarios religiosos y lingüísticos de la India⁵⁰.

25. Los robots —cuentas automatizadas— de los medios sociales se han utilizado para modificar las coordenadas del debate político y tergiversar la opinión pública. En una muestra de 70 países, se utilizaron *bots* en 50 en campañas de manipulación en medios sociales en 2019⁵¹. En el caso de grupos que se apoyan en las tecnologías digitales emergentes para su estrategia de promoción de la discordia y la intolerancia racial, étnica y religiosa, los *bots* son fundamentales por su capacidad para difundir en Internet discursos racistas y con fines de desinformación. Los ejemplos parecen indicar que el uso coordinado de *bots* ha sido especialmente intenso antes de las elecciones. Así, por ejemplo, antes de las elecciones suecas de 2018, los investigadores comprobaron que el 6 % de las cuentas de Twitter en las que se hablaba de política nacional eran *bots* y no cuentas auténticas y donde las entradas publicadas versaban sobre temas relacionados con la inmigración y el islam⁵². Asimismo, en la campaña electoral de 2018 en los Estados Unidos, el 28 % de las cuentas de Twitter que publicaron tuits antisemitas eran *bots* y representaron el 43 % de todos los tuits antisemitas⁵³. En la Federación de Rusia se han utilizado las tecnologías digitales emergentes para promover divisiones étnicas y raciales en los medios sociales⁵⁴, a través de cientos de cuentas y perfiles personales falsos en Twitter, Facebook y otros sitios de medios sociales. Aunque algunas entradas se dirigían contra grupos étnicos minoritarios y exigían la igualdad racial, muchos denunciaron que esos grupos actuaban para atizar las tensiones raciales. Algunas personas dieron su apoyo a grupos nacionalistas blancos, provocando actos de discriminación y violencia contra minorías raciales⁵⁵.

B. Diseño o uso discriminatorio directo o indirecto de las tecnologías digitales emergentes

26. El diseño y la utilización de las tecnologías digitales emergentes puede ser causa directa e indirecta de discriminación por motivos raciales o étnicos en el disfrute de una serie de derechos humanos.

27. Con respecto al derecho al trabajo, en una comunicación se señalaba que el Paraguay había puesto en marcha un sistema de empleo digital que permitía a los empleadores cribar y filtrar a los posibles empleados por diversas categorías, algunas de las cuales actuaban como indicadores indirectos de la raza del interesado⁵⁶. Además, el sistema

⁴⁷ Véanse A/73/312 y A/HRC/41/55.

⁴⁸ A/HRC/42/50, párr. 71 a 75.

⁴⁹ Véase www.commerce.senate.gov/2018/4/facebook-social-media-privacy-and-the-use-and-abuse-of-data.

⁵⁰ Comunicación de Avaaz.

⁵¹ Véase <https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf>.

⁵² Véase www.semanticscholar.org/paper/Political-Bots-and-the-Swedish-General-Election-Fernquist-Kaati/2af3d1e16d5553dc489d8b44321ea543d571a4a9.

⁵³ Véase www.adl.org/resources/reports/computational-propaganda-jewish-americans-and-the-2018-midterms-the-amplification.

⁵⁴ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3304223.

⁵⁵ *Ibid.*, pág. 180.

⁵⁶ Comunicación de Equal Rights Trust.

solo está disponible en español, cuando menos de la mitad de los pueblos indígenas de las zonas rurales del Paraguay hablan ese idioma. Esa limitada accesibilidad lingüística restringe con criterios étnicos y de manera efectiva la disponibilidad de este sistema para los solicitantes de empleo, aunque esa no sea la intención de las autoridades.

28. Una serie de algoritmos utilizados para seleccionar a los candidatos a un puesto de trabajo en América del Norte y Europa también han sido objeto de crítica por formular recomendaciones discriminatorias. Estos sistemas están entrenados para identificar a los candidatos a partir de conjuntos de datos existentes sobre empleados “competentes”, que contienen información sobre características protegidas, como el género, el origen étnico o la religión. La consecuencia es que esos sistemas algorítmicos reproducen y refuerzan los sesgos raciales, étnicos, de género o de otro tipo existentes cuando toman decisiones que reflejan las desigualdades existentes en el empleo. Esos sistemas hacen efectivas formas directas e indirectas de discriminación racial⁵⁷. Por otra parte, cuando esos sistemas prohíben tener en cuenta todas las informaciones protegidas relativas al estado o condición jurídica, como la raza y el origen étnico, pueden socavar las medidas especiales o de acción afirmativa que los Estados puedan haber adoptado para promover la igualdad de oportunidades de empleo.

29. No obstante, en otros casos la introducción de sistemas automatizados que no dependen directamente de insumos o procesos discriminatorios puede ser causa indirecta de discriminación de grupos marginados en su acceso al trabajo al reducir o eliminar los puestos disponibles. En una comunicación se presenta el ejemplo de un nuevo proyecto basado en la inteligencia artificial para la gestión del saneamiento en la India que eliminaría la necesidad de muchos puestos de trabajo que suelen ocupar las personas de la casta más baja, esto es, los dalits⁵⁸. Los dalits, especialmente las mujeres, a menudo solo pueden encontrar empleo en el sector del saneamiento, y algunos estados de la India han dado prioridad a los dalits en la realización de este tipo de actividades. La aplicación de sistemas de saneamiento inteligentes probablemente afectaría de manera desproporcionada a los puestos de trabajo y los medios de vida de los dalits, especialmente de las mujeres de esa casta. A la vista de la extrema marginación socioeconómica y política de los dalits en la India, la automatización del sector de los servicios de saneamiento podría reducir considerablemente el acceso al trabajo de quienes dependen del empleo en el sector de los servicios de saneamiento.

30. Las tecnologías digitales emergentes también tienen un impacto discriminatorio en el derecho a la salud. Los diez principales algoritmos de atención de la salud en el mercado de los Estados Unidos utilizan el historial de gastos médicos de los pacientes para predecir los gastos futuros, que se utilizan como indicador indirecto de las necesidades de atención de la salud⁵⁹. En un estudio reciente de un algoritmo de ese tipo utilizado por una importante empresa de servicios de salud determinó que había estado discriminando involuntaria pero sistemáticamente a los pacientes negros en los Estados Unidos⁶⁰. Con la intención de ayudar a los pacientes de alto riesgo a adherirse a los planes de gestión de la atención de salud, se descubrió que el algoritmo codificaba el sesgo racial utilizando los gastos médicos de los pacientes como un indicador indirecto de sus necesidades en atención de salud para predecir su nivel de riesgo⁶¹. El algoritmo, que los desarrolladores consideraban que era “daltónico racial”, porque la raza no era una variable prevista⁶², asignaba sistemáticamente puntuaciones de menor riesgo a los pacientes negros que estaban tan enfermos como los blancos⁶³. El algoritmo no detectó que un poco menos de la mitad de los pacientes negros también corrían el riesgo de necesitar cuidados médicos complicados como los pacientes blancos. En consecuencia, la probabilidad de que los pacientes negros

⁵⁷ Véase <https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making>.

⁵⁸ Comunicación de la Asociación para el Progreso de las Comunicaciones. Véase también www.apc.org/sites/default/files/gisw2019_artificial_intelligence.pdf.

⁵⁹ Véase www.sciencenews.org/article/bias-common-health-care-algorithm-hurts-black-patients.

⁶⁰ Véase <https://science.sciencemag.org/content/366/6464/447>.

⁶¹ *Ibid.*

⁶² Véase [www.thelancet.com/journals/lancet/article/PIIS0140-6736\(16\)30466-4/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)30466-4/fulltext).

⁶³ Véase <https://science.sciencemag.org/content/366/6464/447>.

fueran derivados hacia programas de actuación para la mejora de su salud. Los hospitales, las aseguradoras y los organismos públicos utilizan este algoritmo y otros similares a gestionar instrumentos para la gestión de la atención de salud de más de 200 millones de personas anuales en el país⁶⁴.

31. En otro ejemplo de los Estados Unidos, un reciente estudio de caso examinó un modelo predictivo desarrollado por Epic Systems Corporation, el desarrollador mundial líder en sistemas de gestión de expedientes clínicos electrónicos⁶⁵. La herramienta de inteligencia artificial de Epic se encuentra integrada directamente en el sistema de expedientes clínicos electrónicos estima la probabilidad de que un paciente no se presente utilizando la información personal del paciente, incluida la etnia, la clase, la religión y el índice de masa corporal, así como la historia clínica del paciente previa a la ausencia. Al señalar el evidente potencial de discriminación de las poblaciones de pacientes vulnerables, los investigadores señalan que “la eliminación de las características personales sensibles de un modelo es un enfoque incompleto para eliminar el sesgo”⁶⁶. Las ausencias previas, por ejemplo, probablemente se correlacionan con la condición socioeconómica por medio de la incapacidad del paciente para cubrir los gastos de transporte o el cuidado de los niños, o para ausentarse del trabajo para acudir la cita. También es probable que se correlacionen con la raza y la etnia debido a las correlaciones entre la condición socioeconómica y la raza y la etnia⁶⁷. En otro estudio reciente se puso de manifiesto que la probabilidad de que los pacientes negros fueran citados en las horas de sobrecarga de pacientes y, por lo tanto, que el tiempo de espera era mayor⁶⁸.

32. En el contexto de la vivienda, estudios realizados en los Estados Unidos han demostrado la discriminación étnica en la publicidad selectiva de Facebook. Facebook solía permitir a los anunciantes “estrechar la audiencia” al excluir a los usuarios de Facebook con ciertas “afinidades étnicas” en la categoría “demografía” de su herramienta de anuncios selectivos⁶⁹. Esta publicidad selectiva podría utilizarse para impedir que las personas de raza negra vean determinados anuncios de viviendas, lo cual está prohibido por la ley de lucha contra la discriminación de los Estados Unidos. Se calcula que Facebook tiene una cuota de mercado del 22 % de los anuncios digitales en los Estados Unidos⁷⁰, y se ha demostrado que su publicidad selectiva, que es el núcleo del modelo de negocio de la empresa⁷¹, es excluyente desde el punto de vista racial⁷². La mejor manera de entender estas prácticas es como una forma de restricción discriminatoria de la oferta de servicios digitales (*digital redlining*), que se define como “el establecimiento y mantenimiento de prácticas tecnológicas que afianzan aún más las prácticas discriminatorias contra grupos ya marginados”⁷³. Facebook también utiliza la publicidad selectiva en el contexto laboral, lo que suscita preocupaciones similares.

33. En otros casos, se niega el acceso a la tecnología —y a la información disponible a través de ella— en formas que tienen repercusiones distintas, o que está dirigida a grupos raciales, étnicos o religiosos específicos, a veces de manera discriminatoria. En 2019, varios Estados, entre ellos Bangladesh, Egipto, la India, Indonesia, el Irán (República Islámica del), Myanmar, la República Democrática del Congo, el Sudán y Zimbabwe, restringieron completamente el acceso a la Internet en determinadas regiones, con el efecto de impedir casi todas las comunicaciones dentro o fuera de esas regiones⁷⁴. Los

⁶⁴ *Ibid.*

⁶⁵ Véase www.healthaffairs.org/doi/10.1377/hblog20200128.626576/full.

⁶⁶ *Ibid.*

⁶⁷ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=346747.

⁶⁸ *Ibid.*

⁶⁹ Véase www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race.

⁷⁰ Véase www.emarketer.com/content/us-digital-ad-spending-will-surpass-traditional-in-2019.

⁷¹ Véase www.motherjones.com/politics/2019/12/facebook-agreed-not-to-let-its-ads-discriminate-but-they-still-can.

⁷² Véase www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

⁷³ Véase www.congress.gov/116/meeting/house/110251/witnesses/HHRG-116-BA00-Wstate-GillardC-20191121.pdf.

⁷⁴ Véase www.hrw.org/news/2019/12/19/shutting-down-internet-shut-critics.

investigadores han comprobado que existe una conexión entre mayor número de cierres selectivos de Internet y regiones con mayor densidad de grupos minoritarios⁷⁵.

34. Con respecto al derecho a un juicio imparcial, varios órganos judiciales de América Latina han comenzado a utilizar Prometea, un programa informático que utiliza el reconocimiento de voz y la predicción basada en el aprendizaje automático, para agilizar los procedimientos judiciales. La fiscalía y los juzgados de Buenos Aires utilizan este sistema de inteligencia artificial para automatizar la toma de decisiones judiciales en casos sencillos, como los litigios sobre licencias de taxi y las demandas de los profesores por no ser compensados por el material escolar⁷⁶. En esos casos, Prometea interpreta los hechos que se le presentan y sugiere un resultado basado en la jurisprudencia anterior en casos similares. El juez debe aprobar la decisión antes de que se haga oficial, lo que ocurre el 96 % de las veces⁷⁷. Una preocupación real es que esta alta tasa de aprobación puede muy bien ser el resultado de una presunción de objetividad y neutralidad tecnológica, como ya se ha visto anteriormente. La Corte Constitucional de Colombia utiliza Prometea para filtrar las acciones de tutela, o demandas individuales por vulneración de derechos constitucionales, y decidir sobre cuáles de ellas debe pronunciarse⁷⁸. La preocupación que suscitan Prometea y muchos otros sistemas de inteligencia artificial de ese tipo es el efecto “caja negra”: la base de su toma de decisiones es opaca, y es difícil o imposible para los jueces, otros funcionarios de los tribunales y los litigantes (e incluso las autoridades públicas que encargan esos sistemas) determinar el sesgo en el diseño, la entrada o la salida. Si bien es imposible conocer el impacto que Prometea tiene o podría tener en las minorías raciales y étnicas, es alto el riesgo de que esos sistemas refuercen o agudicen las disparidades raciales y étnicas existentes en los sistemas de justicia en los que operan.

35. En el contexto de la justicia penal, los departamentos de policía de diferentes partes del mundo utilizan las tecnologías digitales emergentes para la labor de policía predictiva, en la que los sistemas de inteligencia artificial recurren a múltiples fuentes de datos, como los antecedentes penales, las estadísticas sobre delincuencia y la demografía de los barrios⁷⁹. Muchos de esos conjuntos de datos reflejan los sesgos raciales y étnicos existentes, por lo que su funcionamiento refuerza la discriminación racial a pesar de la presunta “objetividad” de esas tecnologías o incluso de su aparente potencial para mitigar los sesgos de los agentes humanos a los que complementan o sustituyen. Asimismo, los departamentos de policía tienden a emplear de manera desproporcionada las tecnologías de predicción en las comunidades empobrecidas de minorías predominantemente étnicas o raciales.

36. El Reino Unido de Gran Bretaña e Irlanda del Norte, por ejemplo, utiliza una base de datos (Gangs Violence Matrix) que ha resultado ser discriminatoria⁸⁰. Según se informa, los agentes de policía hacen suposiciones sobre los individuos basadas en su raza, género, edad y condición socioeconómica, lo que refuerza aún más esos estereotipos⁸¹. El resultado es que el 78 % de las personas de este sistema Matrix son negras, y otro 9 % pertenecen a otros grupos étnicos minoritarios, mientras que las propias cifras de la policía muestran que solo el 27 % de los responsables de actos graves de violencia juvenil son negros. La policía también comparte esta base de datos con otros organismos, como centros de empleo, organismos de vivienda social y centros educativos, lo que da lugar a que ciertas personas quedan discriminadas por su supuesta afiliación a una banda. Según la naturaleza de la forma en que se comparte esa información, ello plantea la posibilidad de que se produzcan posibles vulneraciones del derecho a la intimidad y puede afectar a los derechos a la vivienda y el empleo de forma discriminatoria. Aquellos cuyos nombres figuran en la base

⁷⁵ Véase www.accessnow.org/cms/assets/uploads/2020/02/KeepItOn-2019-report-1.pdf.

⁷⁶ Véase www.bloomberquint.com/businessweek/this-ai-startup-generates-legal-papers-without-lawyers-and-suggests-a-ruling.

⁷⁷ Véase www.giswatch.org/2019-artificial-intelligence-human-rights-social-justice-and-development.

⁷⁸ Véase www.ambitojuridico.com/noticias/informe/constitucional-y-derechos-humanos/prometea-inteligencia-artificial-para-la.

⁷⁹ Comunicación de la Asociación para el Progreso de las Comunicaciones.

⁸⁰ A/HRC/41/54/Add.2, párr. 40.

⁸¹ Véase www.amnesty.org.uk/files/reports/Trapped%20in%20the%20Matrix%20Amnesty%20report.pdf.

de datos Matrix experimentan “múltiples episodios de controles de identidad y registros que aparentemente carecen de todo fundamento jurídico”⁸². Algunos denuncian que la policía los ha controlado y registrado 200 veces, otros que hasta 1.000 veces, y otros más que son varias veces al día. Esta situación repercute en los derechos de las personas a no sufrir injerencias en su intimidad y a no ser detenidas arbitrariamente por motivos étnicos.

37. Otro ejemplo más: en una comunicación se destaca que la policía predictiva se estaba erigiendo en la metodología utilizada en las actividades de policía local como elemento de las denominadas estrategias de prevención del delito en ciudades estadounidenses como Los Ángeles⁸³. Hasta hace poco, el Departamento de Policía de Los Ángeles había estado utilizando una tecnología llamada PredPol para examinar diez años de datos sobre delitos, incluidos los tipos, las fechas, los lugares y la frecuencia de los delitos, a fin de predecir cuándo y dónde era probable que se cometiesen delitos en las siguientes 12 horas. Estos datos, reunidos y clasificados por los agentes de policía, son tanto el producto como la causa de una vigilancia reforzada en las comunidades negras y latinas. La policía predictiva repite y agrava los sesgos existentes en el sistema de policía, al tiempo que proporciona la apariencia de objetividad debido al uso en la toma de decisiones de algoritmos supuestamente neutros. Aunque el Departamento de Policía de Los Ángeles ha suspendido el uso de PredPol, no ha desautorizado el uso de otros productos de policía predictiva que probablemente suscitarán preocupaciones parecidas.

C. Estructuras racialmente discriminatorias

38. Los ejemplos en diferentes partes del mundo ponen de manifiesto que el diseño y la utilización de las diferentes tecnologías digitales emergentes pueden combinarse de manera intencional y no intencional para generar estructuras racialmente discriminatorias que socavan de manera holística o sistemática el disfrute de los derechos humanos de determinados grupos, por motivos de raza, etnia u origen nacional, solos o en combinación con otros motivos. En otras palabras, en lugar de considerar las tecnologías digitales emergentes no solo como capaces de socavar el acceso y el disfrute de derechos humanos discretos, sino también de crear y mantener la exclusión racial y étnica en términos sistémicos o estructurales. En relación con este epígrafe, el Relator Especial pasa revista a una serie de ejemplos de estructuras existentes y potencialmente discriminatorias, haciendo hincapié en el predominio de los sistemas de datos biométricos, la vigilancia racializada y el análisis predictivo racializado en el mantenimiento de estas estructuras.

39. China utiliza la identificación y la vigilancia biométrica para rastrear y restringir los movimientos y actividades del grupo étnico minoritario uigur, lo que vulnera entre otros derechos los de los miembros de este grupo a la igualdad y la no discriminación⁸⁴. Los uigures son objeto de frecuentes controles policiales sin ningún tipo de fundamento en los puestos de control desplegados por la policía, control en el que su teléfono es escaneado, lo que supone una vulneración de su derecho a la intimidad. En el caso de los uigures, es obligatoria la recopilación de datos biométricos muy completos, entre ellos, muestras de ADN y escáneres del iris. Según informaciones dignas de crédito, el Gobierno de China, “mediante el uso combinado de tecnologías de reconocimiento facial y cámaras de vigilancia en todo el país, hace un seguimiento exclusivamente de los uigures sobre la base de su aspecto y mantiene registros de sus desplazamientos para su ulterior investigación y examen”⁸⁵. También se destaca en estas informaciones que se practica esta actividad de vigilancia y reunión de datos un gran número de minorías étnicas que se encuentran retenidas en régimen de aislamiento en “campamentos de reeducación” política con el pretexto de la lucha contra el extremismo religioso, sin que los reclusos estén imputados o encausados⁸⁶. La imagen que proyectan estos hechos son los de una discriminación étnica sistémica, apoyada y, de hecho, posibilitada por una serie de tecnologías digitales

⁸² Véase www.stop-watch.org/uploads/documents/Being_Matrixed.pdf.

⁸³ Comunicación presentada por Stop LAPD Spying Coalition.

⁸⁴ Véase CERD/C/CHN/CO/14-17.

⁸⁵ Véase www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html; y A/HRC/41/35, párr. 12.

⁸⁶ Véase CERD/C/CHN/CO/14-17.

emergentes, y que además supone una violación de una gran variedad de derechos humanos en el caso de los uigures.

40. Kenya y la India han implantado la identificación biométrica para acceder a los servicios públicos, sistemas a los que se les denomina *Huduma Namba* y *Aadhaar*, respectivamente⁸⁷. Los programas incluyen la reunión de diversas formas de datos biométricos, entre ellos huellas dactilares, patrones de retina e iris, patrones de voz y otros identificadores. Al tratar de acceder a los servicios públicos a través de estos sistemas, ciertos grupos raciales y étnicos minoritarios de ambos países se encuentran excluidos de ellos, mientras que otros se enfrentan a barreras logísticas y a largos procesos de investigación de antecedentes que, de hecho, pueden dar lugar a la exclusión de hecho del acceso a los servicios públicos a los que tienen derecho. Entre esos servicios públicos figuran las pensiones y las prestaciones de desempleo en la India, y todos los servicios oficiales esenciales en Kenya, como el voto, el registro de las partidas de nacimiento y los matrimonios civiles, el pago de impuestos y la recepción de escrituras de propiedad. El Tribunal Supremo de la India ha confirmado la ley que exige el número de *Aadhaar* para poder recibir prestaciones sociales. A pesar de que la misma sentencia prohíbe a las entidades privadas utilizar el sistema *Aadhaar* para fines no oficiales, como en la banca, el empleo y las telecomunicaciones móviles, en la práctica se sigue exigiendo ese requisito. Además, las personas con discapacidades —en particular las de las minorías étnicas y raciales— se ven discriminadas porque no se les puede tomar las huellas dactilares o realizar un escáner del iris. Aunque la ley prevé mecanismos especiales para esas personas, éstas siguen enfrentándose a obstáculos logísticos porque el personal de muchos centros no está capacitado para inscribirlas sin los datos biométricos⁸⁸. Al no contar con una protección rigurosa, a las minorías raciales y étnicas, especialmente aquellas cuya condición de ciudadanía es insegura, se ven excluidas desproporcionadamente por los sistemas de identificación digital en el acceso a los servicios públicos⁸⁹.

41. Muchos Estados están experimentando con la incorporación de las tecnologías digitales emergentes en sus sistemas de bienestar social, lo que comporta un refuerzo de las estructuras de discriminación racial⁹⁰. Australia ha puesto en marcha el sistema virtual de cumplimiento normativo, conocido coloquialmente como *robo-debt* (cálculos de cantidades adeudadas mediante una aplicación robótica)⁹¹. Este sistema automatizado de adopción de decisiones utiliza algoritmos de aprendizaje automático para identificar presuntos sobrepagos en las prestaciones sociales públicas y exige documentación a los beneficiarios que han recibido más de lo que tenían derecho a recibir en ese concepto. El sistema remitió aproximadamente unas 20.000 cartas de adeudo a la semana durante un período de seis meses en 2016 y 2017. En una investigación se estimó que entre el 20 y el 40 % de las cartas de adeudo eran falsos positivos, debido a fallos en los procesos del sistema y los datos. El Estado hizo que la carga de la prueba recayera en los beneficiarios de las prestaciones sociales de forma que estos debían demostrar que no estaban en deuda con el Estado. Dado que los australianos indígenas proporcionalmente reciben más prestaciones que los blancos, son los que soportan el mayor costo de las falencias de este sistema y son los peor equipados para impugnar las decisiones que toman dadas las barreras que enfrentan⁹². Una reciente intervención relacionada con los derechos humanos en un procedimiento judicial pone de relieve la existencia de problemas parecidos en los Países Bajos, donde la utilización de las tecnologías digitales emergentes en la prestación de asistencia social ha dado lugar a vulneraciones de los derechos humanos de las personas

⁸⁷ Véase A/74/493.

⁸⁸ Véase <https://timesofindia.indiatimes.com/city/kolkata/court-relief-in-disabled-womans-aadhaar-battle/articleshow/68961357.cms>.

⁸⁹ En el documento A/HRC/38/52 figura un análisis de derechos humanos sobre la discriminación racial en el acceso a la ciudadanía, véase A/HRC/38/52.

⁹⁰ Véase A/74/493.

⁹¹ Véase www.unswlawjournal.unsw.edu.au/forum_article/new-digital-future-welfare-debts-without-proofs-authority y www.ombudsman.gov.au/__data/assets/pdf_file/0022/43528/Report-Centrelinks-automated-debt-raising-and-recovery-system-April-2017.pdf.

⁹² Véase www.aihw.gov.au/reports/australias-welfare/australias-welfare-2019-data-insights/contents/summary.

más pobres y vulnerables de ese país⁹³. También en este caso, las minorías raciales y étnicas se enfrentan a una marginación socioeconómica desproporcionada, lo que suscita la preocupación cada vez más inquietante de que la discriminación de clase es también discriminación racial.

42. Con el recurso cada vez más frecuente de los Estados a las tecnologías digitales emergentes para calcular el riesgo y clasificar las necesidades, en países como, por ejemplo, Dinamarca, los Estados Unidos, Nueva Zelanda y el Reino Unido, una de las prioridades de los Estados debe ser llevar a cabo un control más escrupuloso de su potencial para surtir efectos desiguales en las minorías raciales o étnicas⁹⁴. Dado que la digitalización de los sistemas de bienestar social tiene lugar en sociedades donde ciertos grupos quedan marginados, discriminados y excluidos por motivos raciales y étnicos, es casi seguro que estos sistemas reforzarán esas desigualdades, a menos que el Estado adopte medidas activas de prevención. Sin una intervención urgente, los estados de bienestar digitalizados corren el riesgo de afianzarse como unos regímenes de atención social discriminatorios.

43. En algunos casos, aunque las estructuras que discriminan por motivos de raza son sectoriales, por ejemplo, la justicia penal, no obstante, socavan los derechos humanos de los afectados en su conjunto y refuerzan la opresión estructural que experimentan en la sociedad. Así ocurre en los Estados Unidos, donde las tecnologías digitales emergentes sostienen y reproducen estructuras racialmente discriminatorias en la administración de la justicia penal. En ese país, esas tecnologías son habituales, no solo en las labores de policía, sino también en el sistema de justicia, ámbito en el que se les atribuyen efectos discriminatorios para las minorías raciales y étnicas. Varios Estados de la Unión utilizan instrumentos de evaluación de riesgos basados en la inteligencia artificial en todas las fases del proceso penal. La intención de los desarrolladores de estos sistemas es proporcionar que las consecuencias judiciales se basen en datos objetivos⁹⁵, pero los algoritmos a menudo se apoyan en “datos elaborados durante períodos documentados de prácticas y políticas deficientes, con sesgo racial y a veces ilícitas”⁹⁶. Dado que estos algoritmos influyen en la imposición de la pena, pueden conculcar los derechos de toda persona a la igualdad ante la ley, a un juicio imparcial y a no ser arrestado o detenido arbitrariamente. En esas evaluaciones de los riesgos suelen tener en cuenta factores como las detenciones y condenas anteriores, los antecedentes penales de los padres, el código postal y la denominada “desorganización de la comunidad”⁹⁷. Como los autores de un estudio sostienen en sus conclusiones: “Estos factores son reflejo más de los excesos en la labor policial, los comportamientos de los agentes del orden en las comunidades negras y morenas, los grandes patrones de desventaja socioeconómica resultantes del sistema de castas raciales, que aspectos concretos de las conductas de las personas a las que se investiga”. En otras palabras, los datos predicen más la desventaja racial y la presencia policial en la comunidad de la persona inculpada que el comportamiento de una persona”⁹⁸.

IV. Un enfoque estructural e interseccional del derecho de los derechos humanos para la discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes: análisis y recomendaciones

44. Las tecnologías digitales emergentes plantean un gigantesco desafío en materia de regulación y gobernanza desde el punto de vista de los derechos humanos. En muchos casos, los datos, códigos y sistemas que provocan resultados discriminatorios y efectos conexos son complejos y están blindados a cualquier tipo de examen e incluso están protegidos por el derecho contractual y el de propiedad intelectual. En algunos contextos, ni

⁹³ Véase www.ohchr.org/Documents/Issues/Housing/AmicusConstitutionalCourtHungary_1.pdf.

⁹⁴ A/74/493, párr. 27.

⁹⁵ Véase www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

⁹⁶ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423.

⁹⁷ Comunicación del Center on Race, Inequality, and the Law (Centro sobre Raza, Desigualdad y Derecho) de la Universidad de Nueva York

⁹⁸ *Ibid.*

siquiera los programadores informáticos pueden explicar por sí mismos la forma en que funcionan sus sistemas algorítmicos. Este efecto de “caja negra” supone un escollo para que los grupos afectados puedan superar el problema que supone que recaiga normalmente sobre ellos la carga de la prueba para demostrar la discriminación en un procedimiento judicial, suponiendo, para empezar, que pueda interponerse una acción judicial⁹⁹. Por otra parte, las empresas que crean y aplican las tecnologías digitales emergentes se enfrentan a pocos requisitos legales, cuando no ninguno, para demostrar que sus sistemas respetan los principios de derechos humanos y no ocasionaran efectos discriminatorios desde el punto de vista racial.

45. El derecho internacional de los derechos humanos no será en modo alguno la panacea para los problemas que se exponen en el presente informe, pero desempeñará un papel importante en la determinación y el tratamiento de los perjuicios sociales acarreados por la inteligencia artificial y en la depuración de responsabilidades por esos daños¹⁰⁰, como así ha señalado el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión¹⁰¹. Los criterios éticos por lo que deben regirse las tecnologías digitales emergentes deben aplicarse en una forma que se ajuste al derecho internacional de los derechos humanos, por lo que los Estados deben velar por que no sean un sustituto de la labor de fomento y aplicación de las obligaciones jurídicamente vinculantes actualmente vigentes. En esta sección, la Relatora Especial explica los conceptos y doctrinas de la discriminación racial directa, indirecta y estructural con arreglo al derecho de derechos humanos y esboza las obligaciones que imponen a los Estados por lo que a las tecnologías digitales emergentes se refiere. Estas obligaciones también tienen consecuencias para los agentes no estatales, como las tecnológicas, que en muchos aspectos ejercen un mayor control sobre estas tecnologías que los propios Estados. En esta sección se incluye también una lista no exhaustiva de recomendaciones para la aplicación concreta de las normas y obligaciones expuestas.

A. Alcance de las normas que prohíben la discriminación racial en el diseño y el uso de las tecnologías digitales emergentes

46. La Relatora Especial recuerda que el derecho internacional de los derechos humanos se basa en la premisa de que toda persona, en virtud de su calidad de humana, debe gozar de todos los derechos humanos sin discriminación por motivo alguno. La prohibición de la discriminación racial ha adquirido la condición jurídica de norma imperativa de derecho internacional¹⁰² y de obligación *erga omnes*¹⁰³. En virtud del derecho internacional de los derechos humanos, los Estados han desarrollado con mayor detalle las obligaciones en materia de igualdad racial y no discriminación en varios regímenes enunciados en distintos tratados; los principios de igualdad y no discriminación están codificados en todos los tratados fundamentales de derechos humanos¹⁰⁴. El artículo 26 del Pacto Internacional de Derechos Civiles y Políticos, dispone que la ley prohibirá toda discriminación y garantizará a todas las personas protección igual y efectiva contra cualquier discriminación por motivos de raza, color, sexo, idioma, religión, opiniones políticas o de cualquier índole, origen nacional o social, posición económica, nacimiento o cualquier otra condición social. El Pacto Internacional de Derechos Económicos, Sociales y Culturales también prohíbe la discriminación por estos motivos¹⁰⁵.

⁹⁹ Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Cambridge, Massachusetts, Harvard University Press, 2015).

¹⁰⁰ Véase https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf.

¹⁰¹ A/73/348, párrs. 19 a 60.

¹⁰² Comité de Derechos Humanos, observación general núm. 29 (2001), relativa a la suspensión de obligaciones durante un estado de excepción, párrs. 8 y 13 c). Véase también A/CN.4/727, párr. 59.

¹⁰³ *Barcelona Traction, Light and Power Company, Limited, fallo, I.C.J. Reports 1970*, págs. 3 y ss., en especial pág. 32, párr. 33.

¹⁰⁴ A/HRC/32/50, párrs. 10 a 14.

¹⁰⁵ Véase el art. 2, párr. 2).

47. El artículo 1, párrafo 1, de la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial define a esta como toda distinción, exclusión, restricción o preferencia basada en motivos de raza, color, linaje u origen nacional o étnico que tenga por objeto o por resultado anular o menoscabar el reconocimiento, goce o ejercicio, en condiciones de igualdad, de los derechos humanos y libertades fundamentales en las esferas política, económica, social, cultural o en cualquier otra esfera de la vida pública. El objetivo de la Convención trasciende con creces una visión de la igualdad meramente formal. La igualdad en el marco internacional de derechos humanos tiene carácter sustantivo, y obliga a los Estados a adoptar medidas para luchar contra la discriminación racial deliberada o intencional, así como contra la discriminación racial *de facto*, no intencional o indirecta¹⁰⁶.

48. **Los Estados deben conseguir que la comprensión de lo que supone la prohibición de la discriminación racial adquiera un carácter estructural y acorde con el derecho internacional de los derechos humanos en el contexto de las tecnologías digitales emergentes. Las definiciones del derecho de los derechos humanos deben emplearse en la función crucial de moldear la forma en que los Estados determinan el significado de la discriminación racial propiciada por determinados usos de las tecnologías digitales emergentes, y deben exigir que esas definiciones sirvan de fundamento de los enfoques que adopte el sector privado. Esto significa que los Estados deben abordar no solo el racismo y la intolerancia explícitos en el uso y diseño de las tecnologías digitales emergentes, sino también, y con la misma seriedad, las formas indirectas y estructurales de discriminación racial que se derivan del diseño y uso de dichas tecnologías. Las obligaciones de combatir la discriminación racial son extensivas a las estructuras de discriminación racial y otras formas de discriminación directa e indirecta descritas en la parte III *supra*. Los Estados deben rechazar un enfoque “daltónico” de la gobernanza y la regulación de las tecnologías digitales emergentes, pues se trata de un planteamiento que no toma en consideración la marginación específica de las minorías raciales y étnicas y que conceptualiza los problemas y las soluciones relacionadas con esas tecnologías sin tener en cuenta sus posibles efectos en esos grupos.**

49. Recordando los párrafos 92 a 98 del Programa de Acción de Durban, la Relatora Especial insta a los Estados a que reúnan, compilen, analicen, difundan y publiquen datos estadísticos fiables, desglosados por motivos raciales o étnicos, a fin de hacer frente a las desigualdades raciales de carácter individual y colectivo relacionadas con el diseño y la utilización de las tecnologías digitales emergentes. La Relatora Especial insta a los Estados a que adopten un enfoque basado en los derechos humanos, tratándose de la cuestión de los datos, que se fundamente en los derechos humanos y posibilite al mismo tiempo el desglose, la autoidentificación, la transparencia, la privacidad, la participación y la exigencia de responsabilidades en la recopilación y el almacenamiento de los datos¹⁰⁷. La detección y corrección de toda formas directa e indirecta de discriminación exige recopilar datos (de conformidad con las normas de derechos humanos) que puedan poner en evidencia los efectos dispares que puedan provocar ciertas tecnologías digitales emergentes. Sin embargo, muchos Estados no reúnen o no exigen que se reúnan esos datos. De hecho, algunos países de la Unión Europea prohíben la reunión de datos desglosados que permitan detectar y corregir la discriminación por motivos étnicos o raciales¹⁰⁸. Esas prohibiciones impiden que esos Estados cumplan su obligación de prevenir y combatir la discriminación racial, por lo que deberían acometer las correspondientes reformas. Al respecto, se puede señalar como ejemplo positivo un reciente acontecimiento: la Auditoría de Disparidad Racial del Reino Unido¹⁰⁹.

¹⁰⁶ Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 32 (2009), relativa al significado y el alcance de las medidas especiales en la Convención, párrs. 6 a 7.

¹⁰⁷ Véase A/HRC/42/59.

¹⁰⁸ Véase <https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making>.

¹⁰⁹ A/HRC/41/54/Add.2, párrs. 16 a 19.

50. La eliminación de la discriminación racial, según preceptúa el derecho internacional de los derechos humanos, exige un análisis interseccional. En la definición de interseccionalidad que se enuncia a continuación queda bien reflejada su importancia:

Con la noción de “interseccionalidad” se pretende captar las consecuencias tanto estructurales como dinámicas de la interacción entre dos o más formas de discriminación o sistemas de subordinación. Aborda de manera concreta la manera en que el racismo, el patriarcado, las desventajas económicas y otros sistemas discriminatorios contribuyen a crear capas de desigualdad que estructuran la posición relativa de mujeres y hombres, razas y otros grupos. Además, encara la forma en que determinados actos y políticas crean cargas que se extienden a lo largo de estos ejes que se entrecruzan y contribuyen a crear una dinámica de desempoderamiento¹¹⁰.

51. El Comité para la Eliminación de la Discriminación Racial ha dejado claro que la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial es aplicable en materia de discriminación múltiple e interseccional¹¹¹. Asimismo, la prohibición de la discriminación racial que establece la Convención debería aplicarse en conjunción con los fines que persigue la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (art. 1), la Convención sobre los Derechos de las Personas con Discapacidad (art. 2) y la Declaración de las Naciones Unidas sobre los Derechos de los Pueblos Indígenas (art. 2), que también prohíben o condenan las formas directas e indirectas de discriminación.

52. Los Estados deben tratar de combatir al mismo tiempo otras formas de discriminación que se entrecruzan con la raza y el origen étnico, por lo que cabe entender que estos están obligados a reunir y analizar datos desglosados que permitan comprender mejor la situación de los derechos humanos de los grupos y personas sujetos a estructuras de discriminación múltiple e intersectorial. En el contexto de las tecnologías digitales emergentes implica que en las intervenciones contra la discriminación racial se tenga que prestar una especial atención al género, la discapacidad y otras categorías protegidas. En un informe reciente, el Grupo de Trabajo de Expertos sobre los Afrodescendientes ejemplifica el carácter del análisis interseccional que es esencial en esta materia¹¹².

B. Obligación de prevenir y combatir la discriminación racial en el diseño y el uso de las tecnologías digitales emergentes

53. La Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial articula una serie de obligaciones generales de los Estados que deben aplicarse en el contexto específico de las tecnologías digitales emergentes. Establece el compromiso legal de que todos los Estados partes no incurran en ningún acto o práctica de discriminación racial contra personas, grupos de personas o instituciones, y velen por que todas las autoridades e instituciones públicas, nacionales y locales actúen de conformidad con esa obligación. Los Estados partes por su parte deben aplicar por todos los medios apropiados y sin demora una política de eliminación de la discriminación racial en todas sus formas¹¹³. La Convención también exige a los Estados partes tomar medidas efectivas para revisar las políticas públicas nacionales y locales, así como modificar, dejar sin efecto o anular todas las leyes y las disposiciones reglamentarias que tengan por efecto crear las condiciones para discriminación racial o perpetuarla donde persista¹¹⁴. Asimismo, los Estados partes deberán tomar, cuando las circunstancias así lo aconsejen, medidas especiales y concretas en las esferas social, económica, cultural y en otras para asegurar el adecuado desarrollo y protección de ciertos grupos raciales o de personas pertenecientes a

¹¹⁰ Véase www.un.org/womenwatch/daw/csw/genrac/report.htm.

¹¹¹ Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 32.

¹¹² Véase A/HRC/42/59.

¹¹³ Art. 2, párr. 1, a).

¹¹⁴ Art. 2, párr. 1, c).

estos grupos, con el fin de garantizar que puedan disfrutar plenamente en condiciones de igualdad de los derechos humanos y las libertades fundamentales¹¹⁵.

54. En virtud del artículo 7 de la Convención, los Estados han emprendido la adopción de medidas inmediatas y eficaces en la enseñanza, la educación, la cultura y la información con miras a combatir los prejuicios que dan lugar a la discriminación racial. En otros informes recientes, la Relatora Especial ha expuesto de manera clara las obligaciones en materia de derechos humanos que incumben a los Estados para combatir los discursos y conductas racistas y xenófobas, incluso en Internet¹¹⁶. Estas obligaciones son aplicables también a la problemática analizada en el presente informe; en el contexto de las tecnologías digitales emergentes, los Estados han de adoptar medidas eficaces para detectar y combatir el diseño y la utilización discriminatorios desde el punto de vista racial de esas tecnologías en el disfrute de los derechos civiles, políticos, económicos, sociales y culturales¹¹⁷.

55. **Las obligaciones de los Estados de prevenir y eliminar la discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes exigen que se aborde la “crisis de la diversidad” en los sectores examinados en la parte II *supra*. Los Estados deben colaborar con las empresas privadas, incluso sobre la base de marcos jurídicamente vinculantes, a fin de elaborar las medidas especiales necesarias para garantizar que las minorías raciales y étnicas estén representadas de manera significativa en todos los aspectos de la adopción de decisiones relativas al diseño y la utilización de las tecnologías digitales emergentes. Se requiere un verdadero cambio en las relaciones de poder en varios sectores de las tecnologías digitales emergentes, y no solo un gesto simbólico con las mujeres y los grupos raciales y étnicos minoritarios. Un elemento central del cambio en las relaciones de poder—incluso en el sector privado— es una mayor participación y financiamiento de la investigación y la producción de conocimientos, cuyo objetivo específico es profundizar en la comprensión de la discriminación derivada del diseño y la utilización de las tecnologías digitales emergentes desde una perspectiva interdisciplinaria.** En este sentido, cabe remitirse a los ejemplos presentados por investigadores del Center for Critical Race and Digital Studies¹¹⁸.

56. **Los Estados deben adoptar medidas rápidas y eficaces para prevenir y mitigar el riesgo del uso y el diseño racialmente discriminatorio de las tecnologías digitales emergentes, entre otras cosas, estableciendo como condición *sine qua non* la realización de una evaluación de su impacto en la igualdad racial y en el principio de no discriminación en el disfrute de los derechos humanos antes de la adopción por las autoridades públicas de sistemas basados en esas tecnologías. Estas evaluaciones de impacto deben contemplar un aspecto tan significativo como la posibilidad de compartir el diseño y la ejecución con representantes de grupos racial o étnicamente marginados. No bastará con la adopción de un enfoque puro o primordialmente voluntario para la realización de las evaluaciones de los efectos sobre la igualdad;** en este sentido, será fundamental un enfoque que contemple su obligatoriedad. Los recientes avances en esta dirección conseguidos por el Consejo de Europa¹¹⁹, por ejemplo, son dignos de elogio. No debe descuidarse la discriminación racial ni excluir a las minorías raciales y étnicas en la toma de decisiones. A veces, la prevención de resultados racialmente discriminatorios y otras vulneraciones de los derechos humanos en el diseño y la utilización de las tecnologías digitales emergentes por parte de las autoridades públicas puede exigir la prohibición total de su utilización hasta que se mitigue suficientemente el riesgo de sus efectos perjudiciales. La decisión de la ciudad de San Francisco de prohibir a sus servicios administrativos el uso de *software* de reconocimiento facial es un buen ejemplo en este sentido.

¹¹⁵ Art. 2, párr. 2.

¹¹⁶ Véanse A/73/305, A/73/312 y A/HRC/38/53.

¹¹⁷ El artículo 5 de la Convención hace explícita la aplicación de la prohibición de la discriminación racial al acceso y al disfrute de los derechos civiles, políticos, sociales, económicos y culturales.

¹¹⁸ Véase <https://criticalracedigitalstudies.com>.

¹¹⁹ Recomendación CM/Rec(2020)1 del Comité de Ministros a los Estados miembros sobre los efectos de los sistemas algorítmicos en los derechos humanos, 8 de abril de 2020.

57. **A fin de cumplir sus obligaciones en materia de igualdad y no discriminación, los Estados están obligados a garantizar la transparencia y la rendición de cuentas en lo que respecta la utilización por el sector público de las tecnologías digitales emergentes, y a posibilitar un análisis y una supervisión independientes, entre otras cosas utilizando únicamente sistemas que puedan ser auditados.** Las recientes reformas llevadas a cabo en el Canadá para aplicar la exigencia de responsabilidades al sector público por la utilización de tecnologías digitales emergentes son ejemplos de un primer e importante paso en esta dirección¹²⁰.

58. **Los Estados deben velar por que los marcos y lineamientos éticos elaborados para instaurar una regulación y una gobernanza de carácter flexible, práctico y eficaz de las tecnologías digitales emergentes se basen en principios internacionales de derechos humanos jurídicamente vinculantes, entre ellos, los que prohíben la discriminación racial. La Declaración de Toronto sobre la protección del derecho a la igualdad y la no discriminación en los sistemas de aprendizaje automático ejemplifica la relación simbiótica que debería existir entre un derecho internacional vinculante como el de derechos humanos y los lineamientos o principios éticos para la gobernanza de la inteligencia artificial¹²¹.** En la Declaración de Toronto se subraya el carácter vinculante de la igualdad y la no discriminación en el derecho internacional de derechos humanos y formula una serie de lineamientos para la aplicación práctica de esas normas jurídicas.

Empresas privadas, Naciones Unidas y otros organismos multilaterales

59. Aunque el derecho internacional de derechos humanos es solo directamente vinculante para los Estados, se exige a estos, a fin de cumplir sus obligaciones jurídicas en la materia, que contemplen vías de recursos que sean eficaces en casos de discriminación racial imputables a agentes privados, en particular empresas¹²². En virtud de la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial, los Estados deben adoptar medidas especiales para lograr y proteger la igualdad racial en todos los ámbitos de las esferas pública y privada¹²³. Entre ellas, debería figurar una estricta supervisión regulatoria de las empresas activas en el sector de las tecnologías digitales emergentes.

60. Como se especifica en los Principios Rectores de las Empresas y los Derechos Humanos, las empresas privadas tienen la responsabilidad de respetar los derechos humanos, incluso mediante la debida diligencia en materia de derechos humanos. La debida diligencia en materia de derechos humanos exige: una evaluación del impacto real y potencial de las actividades sobre los derechos humanos; la integración de las conclusiones, así como las consiguientes medidas de actuación; el seguimiento de las respuestas, y la comunicación de la forma en que se hace frente a estos impactos¹²⁴. Como se destaca en el Proyecto sobre Empresas y Derechos Humanos en la Tecnología (Proyecto B-Tech), que aplica los Principios Rectores a las tecnologías digitales, la debida diligencia debería observarse en las fases de conceptualización, diseño y ensayo de nuevos productos, así como a los conjuntos de datos y algoritmos subyacentes en los que se apoyan¹²⁵. En la Declaración de Toronto se señalan tres elementos o medidas fundamentales para la diligencia debida de las empresas en materia de derechos humanos en relación con los

¹²⁰ Véase www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592.

¹²¹ Véase www.torontodeclaration.org/declaration-text/english.

¹²² Véase, por ejemplo, la observación general núm. 31 (2004) del Comité de Derechos Humanos, relativa a la índole de la obligación jurídica general impuesta a los Estados partes en el Pacto, párr. 8.

¹²³ Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 32, párr. 23. Véase también la observación general núm. 20 (2009) del Comité de Derechos Económicos, Sociales y Culturales, relativa a la no discriminación y los derechos económicos, sociales y culturales, párr. 11; y la recomendación general núm. 29 (2002) del Comité para la Eliminación de la Discriminación Racial, relativa a la discriminación basada en la ascendencia en el contexto del artículo 1, párrafo 1, de la Convención, epígrafe. 7.

¹²⁴ A/HRC/17/31, anexo, principio 17.

¹²⁵ Véase www.ohchr.org/Documents/Issues/Business/B-Tech/B_%20Tech_Project_revised_scoping_final.pdf.

sistemas de aprendizaje automático: a) la detección de posibles resultados discriminatorios; b) la prevención y mitigación de la discriminación y el seguimiento de las respuestas; y c) la transparencia en relación con los esfuerzos para detectar, prevenir y mitigar la discriminación. Como se destaca en un informe reciente, los enfoques preventivos de la debida diligencia en materia de derechos humanos deben construirse “en equipos multidisciplinarios que puedan reconocer los puntos ciegos en la inteligencia artificial y descubrir los sesgos sistémicos en los entornos específicos de cada contexto en todas etapas de la vida útil, comenzando desde el desarrollo del producto”¹²⁶.

61. Los Estados deben velar por que los marcos éticos de derechos humanos para las empresas activas en el campo de las tecnologías digitales emergentes estén vinculados con obligaciones internacionales vinculantes en materia de derechos humanos, incluidas las relativas a la igualdad y la no discriminación, y se basen en ellas. Existe un verdadero riesgo de que las empresas hagan profusa referencia a los derechos humanos por los beneficios que supone desde el punto de vista de las relaciones públicas proyectar una imagen ética, aun cuando no hayan tomado ninguna medida relevante para poner en práctica los principios de derechos humanos. Si bien es cierto que las referencias a los derechos humanos, e incluso a la igualdad y la no discriminación, proliferan en los documentos de gobernanza empresarial¹²⁷, estas por sí solas no garantizan la rendición de cuentas. **Asimismo, la aplicación del marco de los Principios Rectores de las Empresas y los Derechos Humanos, gracias, entre otras, a iniciativas como el Proyecto B-Tech, debe incorporar obligaciones jurídicamente vinculantes para prohibir la discriminación racial —y proporcionar vías de recurso eficaces en ese tipo de casos.**

62. Un problema inherente a los enfoques de corte ético que adoptan las empresas tecnológicas es el efecto poco mensurable que surten los compromisos éticos en las prácticas de desarrollo de programas informáticos si no están directamente vinculados a estructuras de rendición de cuentas en el lugar de trabajo¹²⁸. Desde el punto de vista de los derechos humanos, es un error confiar en que las empresas se regulen a sí mismas por suponer una dejación de responsabilidad por parte del Estado. Los incentivos para que las empresas protejan de manera significativa los derechos humanos (especialmente en el caso de los grupos marginados, que no son comercialmente dominantes) pueden ser directamente contrapuestos a los fines de lucro. Cuando hay mucho en juego, las obligaciones fiduciarias para con los accionistas tienden a pesar más que las consideraciones relacionadas con la dignidad y los derechos humanos de los grupos que no tienen medios para exigir responsabilidades a esas empresas. Además, incluso las empresas con las mejores intenciones corren el riesgo de elaborar y aplicar directrices éticas utilizando un prisma predominantemente tecnológico, a diferencia de la mayor amplitud de miras del marco de derechos humanos, cuyo enfoque abarca toda la sociedad y se fundamenta en la dignidad.

63. **Los Estados deben basarse en las prohibiciones establecidas por el derecho internacional de derechos humanos en materia de discriminación racial con el fin de garantizar que las empresas apliquen la debida diligencia en esta materia.** La Comisión Europea habría dado un nuevo y prometedor paso en lo referente a la obligatoriedad de la diligencia debida obligatoria para las empresas¹²⁹, si esta obligatoriedad permitiese una aplicación y un cumplimiento de alcance sustantivo de los derechos humanos.

¹²⁶ Véase www.institut-fuer-menschenrechte.de/fileadmin/user_upload/Publikationen/ANALYSE/Analysis_Business_and_Human_Rights_in_the_Data_Economy.pdf.

¹²⁷ Véase https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518482.

¹²⁸ Véase https://ainowinstitute.org/AI_Now_2018_Report.pdf.

¹²⁹ Véase <https://responsiblebusinessconduct.eu/wp/2020/04/30/european-commission-promises-mandatory-due-diligence-legislation-in-2021>.

C. Obligación de proporcionar vías de recurso eficaces en los casos de discriminación racial en el diseño y el uso de las tecnologías digitales emergentes

64. El sistema internacional de derechos humanos opera a partir de la premisa de que las vulneraciones del derecho internacional de derechos humanos suponen que los infractores tengan la obligación de proporcionar a las víctimas de esas violaciones un recurso adecuado y eficaz¹³⁰. Las víctimas de violaciones de los derechos humanos, incluidas las de carácter discriminatorio por motivos de raza, tienen el correspondiente derecho a una resarcimiento íntegro, incluso mediante una reparación establecida por un órgano jurisdiccional u administrativo. El artículo 6 de la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial es claro al respecto: los Estados partes asegurarán a todas las personas que se hallen bajo su jurisdicción protección y vías de recurso efectivas, ante los tribunales nacionales competentes y otras instituciones del Estado, contra todo acto de discriminación racial que, contraviniendo la Convención, viole sus derechos humanos y libertades fundamentales, así como el derecho a pedir a esos tribunales satisfacción o reparación justa y adecuada por todo daño de que puedan ser víctimas como consecuencia de tal discriminación. Este requisito se establece porque, para que los derechos cobren sentido, se debe disponer de recursos efectivos para reparar toda vulneración de éstos.

65. En el contexto de los recursos eficaces contra la discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes, los Estados deben garantizar toda la gama de vías de reparación eficaces, incluido el acceso a la justicia, la protección contra posibles violaciones y las garantías de cese y no repetición de las violaciones, combatiendo al mismo tiempo la impunidad¹³¹. Véanse los Principios y Directrices Básicos sobre el Derecho de las Víctimas de Violaciones Manifiestas de las Normas Internacionales de Derechos Humanos y de Violaciones Graves del Derecho Internacional Humanitario a Interponer Recursos y Obtener Reparaciones establecen cinco elementos principales de recurso y reparación por vulneración de los derechos humanos: la restitución, la indemnización, la rehabilitación, la satisfacción y las garantías de no repetición¹³². Cada uno de estos elementos desempeña una función diferente para garantizar un recurso eficaz y holístico, que es una vía estrechamente relacionada con la idea de justicia de transición¹³³.

66. La restitución tiene por objeto restablecer la situación original de la víctima antes de que se produjeran las violaciones manifiestas de las normas internacionales de derechos humanos¹³⁴. La indemnización entraña el pago por daños ocasionados que sean económicamente evaluables, incluidos los daños físicos y mentales, la pérdida de prestaciones sociales, los daños materiales, los daños morales y los gastos incurridos¹³⁵. La rehabilitación incluye la prestación de servicios de atención médica y psicológica, así como jurídicos y sociales¹³⁶. La satisfacción es un elemento muy amplio de la reparación y las vías de recurso. Cuando proceda, la satisfacción puede abarcar medidas para poner fin a las violaciones, revelar la verdad, restablecer la dignidad, aceptar la responsabilidad, conmemorar los daños y garantizar la imposición de sanciones contra las partes responsables¹³⁷. Por último, las garantías de no repetición son medidas de reparación y recurso que contribuyen a la no repetición. Éstas están más estrechamente relacionadas con

¹³⁰ Véanse el Pacto Internacional de Derechos Civiles y Políticos, art. 2, y la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial, art. 6.

¹³¹ Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 31 (2005), relativa a la prevención de la discriminación racial en la administración y el funcionamiento de la justicia penal, y Comité de Derechos Humanos, observación general núm. 31.

¹³² Resolución 60/147 de la Asamblea General, anexo, párr. 18.

¹³³ A/69/518, párr. 20.

¹³⁴ Resolución 60/147 de la Asamblea General, anexo, párr. 19.

¹³⁵ *Ibid.*, párr. 20.

¹³⁶ *Ibid.*, párr. 21.

¹³⁷ *Ibid.*, párr. 22.

las reformas estructurales y el fortalecimiento de las instituciones del Estado, ya que aseguran una supervisión civil suficiente y el respeto cabal de los derechos humanos¹³⁸.

67. **Los Estados deben asegurar la restitución, la indemnización, la rehabilitación, la satisfacción y las garantías de no repetición a las víctimas de la discriminación racial en el diseño y la utilización de las tecnologías digitales emergentes. Los Estados también deberían remitirse a la orientación del Relator Especial sobre la promoción de la verdad, la justicia, la reparación y las garantías de no repetición, relativa a la formulación y aplicación de medidas de reparación, y a la orientación del Mecanismo de Expertos sobre los Derechos de los Pueblos Indígenas¹³⁹.**

68. **El marco de derechos humanos vigente en materia de vías de recurso y reparación es también un medio importante para las empresas privadas comprometidas en la lucha contra la discriminación racial en la utilización y el diseño de las tecnologías digitales emergentes.** En otros contextos, los agentes privados han desempeñado un papel importante en la reparación de la discriminación racial, llegando incluso a asumir la responsabilidad por su papel en dicha discriminación¹⁴⁰. **Empresas privadas como Microsoft, Apple, Amazon, Google, Facebook, Tencent y Alibaba tienen un importante papel que desempeñar en lo tocante a la restitución, la indemnización, la rehabilitación, la satisfacción y las garantías de no repetición para las víctimas de la discriminación racial relacionada con sus tecnologías y productos.**

¹³⁸ *Ibid.*, párr. 23.

¹³⁹ Véase A/HRC/ENRIC/2019/3.

¹⁴⁰ A/74/321, párr. 62.