LIBERTAD DE EXPRESIÓN, DISCURSOS DE ODIO Y LAS TIC: EL PROBLEMA DE LOS TRES CUERPOS. SOLUCIONES DESDE EL DERECHO INTERNACIONAL Y EUROPEO EN LA ERA DE LA DESINFORMACIÓN

Francisco Jiménez García*

«El régimen de la información hace que las personas sean completamente transparentes. La dominación en sí misma nunca es transparente…*La sala de máquinas de la transparencia es oscura* (y…) la verdad se desintegra en polvo informativo arrastrado por el viento digital». *Infocracia* (2023) Byung-Chul Han

«La desinformación y el odio no deben generar la máxima visibilidad y enormes beneficios» Principios globales de las Naciones Unidas para la integridad de la información (2024).

Resumen

La libertad de expresión, los discursos de odio y la desinformación en el marco de la era digital copan las agendas de los Estados y de las Organizaciones internacionales. Las tensiones inherentes entre estos principios y los medios de actuación han dado lugar a diversas respuestas. Ante la regulación normativa internacional del derecho a la libertad de expresión y sus excepciones, se ha optado por el soft-law *como medio para prevenir y erradicar los discursos de odio y sus graves consecuencias. Frente a la criminalización*

* Francisco JIMÉNEZ GARCÍA, Catedrático de Derecho Internacional Público y Relaciones Internacionales en la Universidad Rey Juan Carlos (francisco.jimenez@urjc.es). Este estudio se ha realizado en el marco del Proyecto de Investigación: Hacia un Convenio internacional integral sobre el uso delictivo de las TIC: Ciberterrorismo y discurso de odio en un marco de libertad de expresión y responsabilidad de la Agencia Estatal de Investigación, referencia PID2022-136943OB-I00, del Ministerio de Ciencia, Innovación y Universidades, y es un trabajo desarrollado dentro del Grupo de Investigación de Alto Rendimiento en Libertad, Seguridad y Ciudadanía en el Orden Internacional (INTER-CIVITAS), de la Universidad Rey Juan Carlos. Las páginas webs fueron visitadas por última vez el 7 de abril de 2025.

exclusiva de los delitos de odio que inciten a la violencia y la discriminación conforme a ciertos factores contextuales (Principios del Plan de Rabat), *en el marco europeo se ha optado por primar y ampliar el marco contextual del riesgo inminente para sancionar penalmente estos discursos. Por otra parte, la retórica divisoria y amenazante de la desinformación y el odio ha generado un reglamentario efecto Bruselas y una singular cláusula de* cibersolidaridad*, enfocados principalmente a las empresas cibertecnológicas y a las amenazas exteriores respectivamente, cuyas referencias han sido asumidas de forma asertiva por los Pactos para el Futuro y Global Digital de las Naciones Unidas.*

Palabras clave

Libertad de expresión, discursos de odio, desinformación, ciberespacio, feudalismo digital, diligencia debida, efecto Bruselas, cláusula de cibersolidaridad, «sumideros de odio».

Abstract

Freedom of expression, hate speech, and disinformation in the digital age are currently and urgently appearing on the agendas of States and international organizations. The inherent tensions between these principles and means of action have led to various responses. In contrast to the international normative regulation of the right to freedom of expression and its exceptions, soft law *has been chosen to prevent and eradicate hate speech and its serious consequences. While the exclusive criminalization of hate crimes that incite violence and discrimination according to certain contextual factors* (Rabat Plan Principles) *has been promoted, the European framework has opted to prioritize and expand the contextual framework of imminent risk to criminally sanction hate speech. On the other hand, the divisive and threatening rhetoric of disinformation and hate has generated a regulatory Brussels effect and a particular* cyber-solidarity *clause, mainly focused on cyber technology companies and external threats respectively, whose references have been assertively adopted by the United Nations' Pacts for the Future and Global Digital Compact.*

Keywords

Freedom of expression, hate speech, disinformation, cyberspace, digital feudalism, due diligence, Brussels effect, cyber-solidarity clause, «sink of hate».

I. EL DISCURSO DE ODIO Y LA DESINFORMACIÓN COMO RETÓRICAS DIVISORIAS QUE AFECTAN A LA PAZ Y SEGURIDAD MUNDIALES. EL CARÁCTER BIFRONTE DE LAS NUEVAS TECNOLOGÍAS: VENTAJAS E INCONVENIENTES

El odio es una historia fácil de contar frente a la complejidad de la facticidad que avala la veracidad de la realidad informada. Por otra parte, no todo discurso de odio es penalmente responsable pues el derecho no puede prohibir el odio, no puede castigar al ciudadano que odia; «hemos de apartarnos de la tentación de construir el juicio de tipicidad trazando una convencional y artificiosa línea entre el discurso del odio y la ética del discurso»[(1)], más si analizamos esta problemática desde una perspectiva multicultural, transfronteriza y global digital. El problema de los tres cuerpos plantea lo difícil que es predecir el movimiento de tres cuerpos que pertenecen a un mismo sistema orbital a lo largo del tiempo bajo la influencia gravitacional de cada uno de ellos. Este dilema científico constituye el título de una exitosa novela de ficción del escritor chino Liu Cixin que ha dado origen a una serie televisiva en la plataforma de *streaming* Netflix. Si la tensión orbital entre el derecho fundamental a la libertad de expresión y la previsión de sus limitaciones, en particular, las relativas a los discursos de odio, resulta una constante en el desarrollo de la convivencia de las sociedades, la irrupción de las nuevas tecnologías de la información y la comunicación (TIC), así como de la inteligencia artificial (IA) –donde el medio cobra tanta importancia como el contenido–, ha agravado el equilibrio gravitacional entre ambas consideraciones hasta el punto que se ha vuelto a insistir en la estrecha relación entre los discursos de odio y el origen o auge de los totalitarismo ahora enmarcados en las «punzantes» campañas de desinformación contra las democracias liberales (*Sharp Power*[(2)]). Como se ha destacado desde el

(1) STC (Pleno) 35/2020, de 25 de febrero de 2020, F.J.5 asumiendo el razonamiento ínsito en el F.J.2 de la Sentencia 4/2017 recurrida dictada por el Tribunal Supremo el 18 de enero. Según el TS, «No todo mensaje inaceptable o que ocasiona el normal rechazo de la inmensa mayoría de la ciudadanía ha de ser tratado como delictivo por no hallar cobertura bajo la libertad de expresión. Entre el odio que incita a la comisión de delitos, el odio que siembra la semilla del enfrentamiento y que erosiona los valores esenciales de la convivencia y el odio que se identifica con la animadversión o el resentimiento, existen matices que no pueden ser orillados por el juez penal con el argumento de que todo lo que no es acogible en la libertad de expresión resulta intolerable y, por ello, necesariamente delictivo».

(2) Daniel MARTÍNEZ CRISTÓBAL, «The current perspective on sharp power: China and Russia in the era of (dis)information», *Revista Electrónica de Estudios Internacionales*, núm. 42, 2021; Francisco JIMÉNEZ GARCÍA, *Conflictos Armados y Derecho Internacional Humanitario*, Ommpress Madrid, 2023, p.78; Christopher WALKER y Jessica LUDWIG, «The meaning of sharp power. How authoritarian states project influence», *Foreign Affairs*, 16 de noviembre de 2017 y «The long arm of

Consejo de Europa, la relación entrelazada entre el discurso de odio y la desinformación es preocupante, especialmente en tiempos de crisis. La desinformación y la información falsa se difunden deliberadamente con el objetivo de desencadenar discursos de odio, desacreditar a las autoridades y a los medios de comunicación y, en última instancia, socavar los valores democráticos. Además, el medio digital ha añadido nuevos problemas relacionados con la rápida difusión, el anonimato, la desinhibición y la sensación de impunidad agravados por la dificultad que supone cooperar con las plataformas de redes sociales, hacerlas responsables y tomar medidas proactivas(3).

No en vano, la Asamblea General de las Naciones Unidas, por un lado, ha denunciado cómo la Federación de Rusia ha intentado justificar su ilícita agresión territorial contra Ucrania sobre la base de la supuesta eliminación del neonazismo, subrayando que el uso del neonazismo como pretexto para justificar la agresión territorial menoscaba gravemente los intentos genuinos de combatir tal fenómeno; y, por otro, ha expresado su profunda preocupación por el auge, en particular en los países democráticos, de partidos extremistas de carácter racista o xenófobo. A tal efecto, ha exigido a los partidos políticos comprometidos con el respeto de las libertades, los derechos humanos, la democracia, el Estado de derecho y la buena gobernanza, que condenen todos los mensajes que difundan ideas basadas en el supremacismo, la superioridad o el odio raciales(4). El propio Consejo de Seguridad, en su Resolución 2686 (2023), ha reconocido que el discurso de odio, el racismo, la discriminación racial, la xenofobia, las formas conexas de intolerancia, la discriminación de género y los actos de extremismo pueden contribuir «a provocar el estallido, el recrudecimiento y la recurrencia de los conflictos y debilitar las iniciativas encaminadas a prevenir y resolver los conflictos y eliminar sus causas profundas, así como las actividades de reconciliación, reconstrucción y consolidación de la paz»(5).

Como ya indicara Hannah Arendt, lo que convence a las masas no son los hechos, ni siquiera los hechos inventados, sino sólo la consistencia del sistema del que son presumiblemente parte. *La propaganda totalitaria medra en esta huida de la realidad a la ficción, de la coincidencia a la consistencia*. Según la filósofa alemana, la autocoacción del pensamiento ideológico arruina todas las relaciones con la realidad y el éxito se produce cuando los hombres pierden el contacto con sus

---

the strongman. How China and Russia use sharp power to threaten democracies», *Foreign Affairs*, 12 de mayo de 2021; Joseph S. Nye JR., «How sharp power threatens soft power. The right and wrong ways to respond to authoritarian influence», *Foreign Affairs*, 24 de enero de 2018.

(3) Cfr. Consejo de Europa, *Study on Preventing and Combating Hate Speech in Times of Crisis*. Steering Committee on Anti-Discrimination, Diversity and Inclusion (CDADI), 2023, pp.32-33.

(4) Cfr. las siguientes resoluciones adoptadas por la Asamblea General el 17 de diciembre de 2024: Resolución 79/160 sobre Combatir la glorificación del nazismo, el neonazismo y otras prácticas que contribuyen a exacerbar las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia; Resolución 79/161 sobre el Llamamiento mundial para la adopción de medidas concretas para la eliminación del racismo, la discriminación racial, la xenofobia y las formas conexas de intolerancia y para la aplicación y el seguimiento generales de la Declaración y el Programa de Acción de Durban; Resolución 79/175 sobre El derecho a la privacidad en la era digital; Resolución 79/180 sobre la Lucha contra la intolerancia, los estereotipos negativos, la estigmatización, la discriminación, la incitación a la violencia y la violencia contra las personas por motivos de religión o creencia.

(5) S/RES/2686 (2023), p.3.

semejantes tanto como con la realidad que existe en torno de ellos; porque, junto con estos contactos, los hombres pierden la capacidad tanto para la experiencia como para el pensamiento. El objeto ideal de la dominación totalitaria consiste en negar la distinción entre el hecho y la ficción (es decir, la realidad de la experiencia) y la distinción entre lo verdadero y lo falso (es decir, las normas del pensamiento)(6). En la misma línea, Buyng-Chul Han, acerca de la incidencia de la digitalización en la crisis de la democracia (la «infocracia»), ha destacado que el nuevo nihilismo no supone que la mentira se haga pasar por la verdad o que la verdad sea difamada como mentira. Más bien socava la distinción entre verdad y mentira. La mentira solo es posible cuando la distinción entre la verdad y la mentira permanece intacta. Las noticias falsas no son mentiras, sino que atacan a la propia facticidad: «desfactifican la realidad». Las opiniones pueden ser muy dispares; pero son legítimas, siempre que respeten la verdad factual. La libertad de expresión, en cambio, degenera en farsa cuando pierde toda referencia a los hechos y a las verdades fácticas(7).

Desde las Naciones Unidas, y en particular por parte de su secretario general, Antonio Guterres, se ha denunciado que el discurso de odio es una señal de alarma: cuanto más fuerte suena, mayor es la amenaza de genocidio. Antecede y promueve la violencia y la intolerancia. Si bien su efecto devastador, por desgracia, no es nada nuevo, sin embargo, su escala e impacto se ven ahora aumentados por las nuevas tecnologías de la comunicación. El discurso de odio –también en Internet– se ha convertido en una de las formas más habituales de extender una retórica divisoria a escala mundial, poniendo en peligro la paz en todo el mundo(8). La propaganda del odio y los medios de comunicación han estado presentes en las principales crisis humanitarias de la historia, desde las invasiones napoleónicas y los innovadores documentales nazis de Leni Riefenstahl (*El triunfo de la voluntad* y *Olympia*), hasta el papel jugado por las denominadas «emisoras del odio», la Radio Libre de las Mil Colinas y el periódico Kangura, en el genocidio ruandés(9). En la actualidad, las redes sociales se han convertido en un instrumento estratégico de

(6) Cfr. Hannah Arendt, *Los Orígenes del Totalitarismo*, traducción de Guillermo SOLANA, Alianza Editorial, 2014, pp. 487 y 635. El énfasis es añadido.

(7) Cfr. Byung-Chul Han, *Infocracia*, traducción Joaquín Chamorro Mielke, Taurus, Madrid, 2023, pp.73-75. Entre nosotros, la democracia digital se ha presentado como: «la aparición de "esferas públicas desorganizadas" (Habermas); o, lo que es lo mismo, la pérdida de un mundo público común conocido, y su sustitución por un consumo fragmentado guiado por la lógica del enjambre, la privatización ("My daily Me") o las "cámaras de eco", el lugar virtual en el que nos encontramos con los afines en gustos, opiniones o posicionamientos políticos. Se trata, pues, de un espacio balcanizado en el que predomina la polarización, la emocionalidad y la bronca y el ruido, y donde las pasiones dominantes son el resentimiento o el odio, pero también la acrítica aceptación –incluso sumisión fanática– a determinadas posturas. Pero tienen también otros tres rasgos que consideramos dignos de mención: 1) "crean comunidad" entre los afines; es decir, la pérdida de los tradicionales vínculos comunitarios en el mundo real encuentra un equivalente funcional en la red; 2) son reactivas y expresivas, más que dialógicas o argumentativas; y 3) esta nueva realidad, por muy "virtual" que sea opera de facto como una "realidad paralela", es también objeto de observación por parte de los medios de comunicación tradicionales, que en muchos casos dan cuenta de lo que en ella ocurre como parte de su labor informativa cotidiana», Fernando Villaspín Oña, «Las principales amenazas a la democracia liberal», *Anales de la Real Academia de Ciencias Morales y Políticas (2019-2020)* p.337.

(8) El discurso de odio se extiende por todo el mundo en: https://www.un.org/es/hate-speech.

(9) Juana del Carpio Delgado, «Discurso de odio en el Derecho penal internacional: su consideración, o no, de persecución como crimen de lesa humanidad» en Eulalia W. PETIT DE GABRIEL

guerra y exterminio. Su carácter encriptado y garantista de la privacidad de los usuarios –en este contexto, la publicidad, elemento esencial para la sanción del discurso o delito de odio, opera en un terreno sinuoso y casuístico– dificulta, además, la prueba sobre la autoría de los artífices de crímenes contra la humanidad.

Así, en el caso de la persecución de los rohinyás en Myanmar, la Misión Internacional Independiente de Investigación, designada por el Consejo de Derechos Humanos de las Naciones Unidas, destacó el uso de las redes sociales, en particular *Facebook* (ahora *Meta*), para la difusión de discursos de odio por parte de los responsables de los crímenes cometidos contra los rohinyás, incluido el genocidio(10). Este asunto no solo está siendo conocido por jurisdicciones internacionales –tanto por parte de la Corte Internacional de Justicia (CIJ) que, en enero de 2020, ordenaba medidas provisionales solicitadas por Gambia para que Myanmar evitara la comisión de actos de genocidio y la destrucción de las pruebas relacionadas con el caso(11), como por parte de la Corte Penal Internacional (CPI)(12)–, sino que, en virtud de la dimensión probatoria requerida por la jurisdicción internacional, se ha proyectado al ámbito nacional mediante la solicitud de prueba presentada el 5 de junio de 2020 por Gambia contra *Facebook* ante el Tribunal Federal del Distrito de Columbia, Estados Unidos. Gambia demandó acceder a los contenidos electrónicos producidos, elaborados y publicados por los individuos y agentes estatales cuyas cuentas habían sido suspendidas. La solicitud de prueba también incluía todos los documentos relacionados con cualquier investigación interna llevada a cabo por *Facebook* sobre las violaciones de sus políticas de contenidos por parte de estos individuos y agentes estatales.

Sobre ambas solicitudes, el Tribunal de Distrito, tras constatar la amplia difusión de los contenidos y la intención manifiesta de sus autores de alcanzar al mayor número de personas –llegaron a tener una audiencia de 12 millones de seguidores–, concluyó que el material eliminado por *Facebook* no estaba amparado por las reglas de confidencialidad de la *Stored Communications Act* (Ley de Comunicaciones Almacenadas o SCA). Además, consideró que la información solicitada por Gambia era relevante para probar la intención genocida de las autoridades de Myanmar implicadas y accedió a la petición de obligar a *Facebook* a presentar toda la documentación relacionada con su investigación interna, normalmente protegida por las normas del secreto profesional(13).

(Dir.), *Valores (y Temores) del Estado de Derecho: Libertad de Expresión vs. Delitos de Opinión en Derecho Internacional***,** Aranzadi, 2023, pp.73-109.

(10) UN. Doc. A/HRC/42/50, pars.67-75.

(11) Corte Internacional de Justicia: *Aplicación de la Convención para la Prevención y la Sanción del Delito de Genocidio* (Gambia c. Myanmar: 7 Estados intervinientes), en: https://www.icj-cij.org/case/178.

(12) *Decision Pursuant to Article 15 of the Rome Statute on the Authorization of an Investigation into the Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar*, 24th November 2019. Sobre el estado de este asunto: https://www.icc-cpi.int/bangladesh-myanmar.

(13) Información sobre este caso, incluida la sentencia, puede consultarse en https://globalfreedomofexpression.columbia.edu/es/cases/gambia-v-facebook/. Asimismo, resulta de sumo interés el riguroso trabajo de Libia Arenal Lora, «Limitaciones y alcance de la responsabilidad de las empresas proveedoras de servicios en el discurso de odio *online*. El caso de Meta en la incitación al genocidio rohingya», *Cuadernos de Derecho Transnacional* 15 2023-2, pp. 141-166 (DOI: 10.20318/cdt.2023.8053). Según esta autora, la decisión del Tribunal de Distrito, «por un lado, bienvenida, puede traer un efecto contraproducente porque si las empresas se ven abordadas judicialmente con peticiones masivas de divulgación de documentos eliminados, podrían decidir limitar unilateralmente esta práctica y, de esta manera, difi-

La Asamblea General de la ONU ha articulado sus últimas resoluciones en función de estos tres elementos, subrayando la falta de uniformidad existente en las normas relativas a estas materias. Así en la Resolución 79/160 (2024) expresa su preocupación por el aumento del uso de las tecnologías digitales para promover y propagar el racismo, el odio racial, la xenofobia, la discriminación racial y las formas conexas de intolerancia. A este respecto, ha exhortado a los Estados parte en el Pacto internacional de derechos civiles y políticos a que contrarresten la difusión de esas ideas, respetando al mismo tiempo sus obligaciones en virtud de los artículos 19 y 20 del Pacto que garantizan el derecho a la libertad de expresión y en los que se enuncian los fundamentos para restringir legítimamente el ejercicio de tal derecho. Igualmente, reconoce la necesidad de promover el uso de las TIC, especialmente Internet, para contribuir a la lucha contra el racismo, la discriminación racial, la xenofobia y las formas conexas de intolerancia y asume la función positiva que los medios de comunicación pueden desempeñar en la lucha contra estas manifestaciones graves de discurso de odio y de intolerancia, promoviendo una cultura de tolerancia e inclusión representando la diversidad de la sociedad multicultural.

Por su parte, respecto a las publicaciones en Internet y la responsabilidad sobre su uso por terceros, el Tribunal Europeo de Derechos Humanos (TEDH) ha resaltado la doble dimensión de esta tecnología. Por un lado, esta nueva tecnología se ha convertido no solo en uno de los principales medios para el ejercicio del derecho a la libertad de expresión, sino también en el protector o «guardián» de la libertad de expresión, en particular, mediante el surgimiento de lo que el Tribunal denomina «periodismo ciudadano», ya que el contenido político ignorado por los medios tradicionales a menudo se difunde a través de sitios web a un gran número de usuarios, que luego pueden ver, compartir y comentar la información. Sin embargo, los beneficios de esta herramienta de información, una red electrónica que sirve a miles de millones de usuarios en todo el mundo, conllevan una serie de riesgos, diferentes a los que se derivan de los medios impresos, especialmente en lo que se refiere a la capacidad de almacenar y transmitir información. Los discursos difamatorios y otros tipos de discursos manifiestamente ilícitos, incluidos los discursos de odio y los discursos que incitan a la violencia, pueden difundirse como nunca antes, en todo el mundo, en cuestión de segundos, y a veces permanecen disponibles en línea durante largos períodos. Según el TEDH, en virtud de estas consideraciones, las restricciones a la libertad de expresión en el marco de las redes sociales y las nuevas tecnologías deben tener en cuenta los siguientes aspectos: en primer lugar, el contexto de los comentarios; en segundo lugar, las medidas aplicadas por la empresa solicitante para prevenir o eliminar los comentarios difamatorios; en tercer lugar, la responsabilidad de los propios autores de los comentarios como alternativa a la responsabilidad de la empresa solicitante; y, en cuarto lugar, las consecuencias de los procedimientos internos para la empresa demandante[(14)].

---

cultar aún más la lucha contra el discurso del odio; esto si no llega antes una reforma legislativa, por otro lado necesaria, que vaya en la línea de abordar los conflictos entre los requisitos competitivos de privacidad y la protección de datos, la moderación de contenido y la libertad expresión, la cooperación con iniciativas de responsabilidad legal y la protección de derechos humanos», par.53.

(14) Cfr. SSTEDH (Gran Sala) en los asuntos *Delfi AS v. Estonia*, demanda nº 64569/09, Sentencia16 de junio de 2015, pars.110-117; y *Sanchez c. Francia,* demanda n.º 45581/15, sentencia 15 de mayo de 2023, pars. 158-162. Acerca de la jurisprudencia nacional e internacional sobre libertad de expresión resulta de sumo interés la siguiente página web de la Universidad de Columbia: https://globalfreedomofexpression.columbia.edu/.

También se ha insistido en la necesidad de una educación de contenidos digitales fundamentada en los derechos humanos y los valores democráticos, incluyendo la adopción de sistemas educativos que elaboren los materiales necesarios para ofrecer el relato más exacto de la historia (*la alfabetización digital en valores democráticos*)[15]. Aún no se percibe socialmente el discurso de odio como un problema, ni se considera una prioridad contrarrestarlo. Como se recuerda en la Declaración y en el Programa de Acción de Durban, «la educación a todos los niveles y a todas las edades, inclusive dentro de la familia, en especial la educación en materia de derechos humanos, es la clave para modificar las actitudes y los comportamientos basados en el racismo, la discriminación racial, la xenofobia y las formas conexas de intolerancia y para promover la tolerancia y el respeto de la diversidad en las sociedades. Afirmamos además que una educación de este tipo es un factor determinante en la promoción, difusión y protección de los valores democráticos de justicia y equidad»[16].

No obstante, no existe una definición ni una regulación convencional internacional de esta realidad, ni del discurso del odio ni de la desinformación, lo que genera una doble ambigüedad. La vaguedad y la falta de consenso en torno a su significado pueden utilizarse indebidamente para permitir la infracción de una amplia gama de expresiones lícitas. En un Estado democrático deben tolerarse los riesgos inherentes a la discusión política y permitir tanto un debate amplio, robusto y desinhibido con la incorporación de todas las opiniones, so pena de socavar la propia calidad democrática de la deliberación al quedar excluidas *de iure* determinadas posiciones ideológicas, incluidas las de los grupos vulnerables. Un uso indebido de los delitos de odio puede incluso silenciar a algunas de las personas que están en mejores condiciones de contrarrestar los argumentos de odio: los defensores de los derechos humanos y los periodistas. Sin embargo, no tomarse en serio esta cuestión por parte de los gobiernos y las empresas puede suponer no hacer frente de forma efectiva a la violencia o a la discriminación contra los grupos vulnerables, incluyendo el silenciamiento de los marginados[17].

Por otra parte, existen dos modelos de afrontar la criminalización de los delitos de odio. Siguiendo la distinción efectuada por J.A. Díaz López, por un lado, se encontraría el modelo de la «discriminación selectiva» (*discriminatory selection model*), en el que las motivaciones del autor pasan a un segundo plano, y el concepto de crimen de odio –y, por lo tanto, el fundamento de la sanción penal– se vincula al hecho de que genera efectos discriminatorios en el colectivo al que pertenece la víctima. Este modelo restringe la sanción (o su agravación) a los casos en los que el autor forma parte del grupo mayoritario y la víctima pertenece a un colectivo tradicionalmente discriminado, ya que, si la víctima no forma parte de dicho colectivo, no se producirían verdaderos efectos discriminatorios. Desde esta perspectiva, un sector de la doctrina considera que

(15) Resolución 79/160 aprobada por la Asamblea General el 17 de diciembre de 2024 sobre Combatir la glorificación del nazismo, el neonazismo y otras prácticas que contribuyen a exacerbar las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia. La Unión Europea, a través del Observatorio Europeo de Medios Digitales (EDMO), ha elaborado en 2024 las Directrices para iniciativas eficaces de alfabetización mediática: https://edmo.eu/areas-of-activities/media-literacy/raising-standards/.

(16) Par. 95 de la Declaración. Se puede consultar en la siguiente dirección: https://www.un.org/es/fight-racism/background/durban-declaration-and-programme-of-action.

(17) Términos empleados en el Informe de 2019 preparado por el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión (A/74/486), pp.4-5.

estos delitos deberían denominarse de «discriminación» y no de «odio». Frente a ese modelo, se alza el de la «animosidad» (*animus model*). Aquí, se entiende que los crímenes de odio son aquellos en los que el autor actuó guiado por motivos discriminatorios, por su prejuicio hacia determinada condición personal de la víctima, hacia categorías universales, como la etnia, y no hacia grupos históricamente discriminados(18).

La Estrategia y Plan de Acción de las Naciones Unidas para la Lucha contra el Discurso del Odio (2019) proporciona una definición de trabajo de lo que podemos entender por discurso de odio: «cualquier forma de comunicación de palabra, por escrito o a través del comportamiento, que sea un ataque o utilice lenguaje peyorativo o discriminatorio en relación con una persona o un grupo sobre la base de quiénes son o, en otras palabras, en razón de su religión, origen étnico, nacionalidad, raza, color, ascendencia, género u otro factor de identidad»(19). Por su parte, el ciber-discurso de odio es aquel que se produce en línea, esto es, a través de las TIC: páginas web, apps, correos electrónicos y una amplia gama de redes sociales y videojuegos y, sin lugar a dudas, otras formas de comunicación que irán surgiendo(20). En la Recomendación CM/Rec(2022)16 del Comité de Ministros del Consejo de Europa sobre la lucha contra el discurso de odio(21), se precisa en su artículo 3 distintas manifestaciones de lo que se ha denominado la pirámide del discurso de odio que varían en su gravedad, el daño que causan y su impacto en los miembros de determinados colectivos en diferentes contextos: desde las actitudes y los actos prejuiciosos hasta el genocidio, pasando por los actos de discriminación y violencia. Tomando como referencia el Convenio Europeo de Derechos Humanos (CEDH) y la jurisprudencia pertinente del TEDH, se efectúa la

(18) Esta clasificación está extraída del *Informe de Delimitación Conceptual en Materia de Delitos de Odio* encargado por la Comisión de Seguimiento del Acuerdo para cooperar institucionalmente en la lucha contra el racismo, la xenofobia, la LGBTIfobia y otras formas de intolerancia, financiado por la Secretaría de Estado de Migraciones del Ministerio de Inclusión, Seguridad Social y Migraciones y elaborado por Juan Alberto Díaz López, p.35.

(19) Cfr. https://www.un.org/en/genocideprevention/documents/Action_plan_on_hate_speech_ES.pdf. Por su parte, la Recomendación General n.º 15 relativa a la lucha contra el discurso de odio y memorándum explicativo de la Comisión Europea contra el Racismo y la Intolerancia (ECRI), del Consejo de Europa (2015), define el discurso de odio como «el fomento, promoción o instigación, en cualquiera de sus formas, del odio, la humillación o el menosprecio de una persona o grupo de personas, así como el acoso, descrédito, difusión de estereotipos negativos, estigmatización o amenaza con respecto a dicha persona o grupo de personas y la justificación de esas manifestaciones por razones de "raza", color, ascendencia, origen nacional o étnico, edad, discapacidad, lengua, religión o creencias, sexo, género, identidad de género, orientación sexual y otras características o condición personales». Del mismo modo, se reconoce que el discurso del odio puede «adoptar la forma de negación, trivialización, justificación o condonación públicas» de los crímenes contra la humanidad. Para, finalmente, señalar que también «puede tener por objeto incitar a otras personas a cometer actos de violencia, intimidación, hostilidad o discriminación contra aquellos a quienes van dirigidas, o cabe esperar razonablemente que produzca tal efecto», CRI(2016)15. El Comité de Ministros adoptó el 20 de mayo de 2022 la CM/Rec (2022) 16 sobre la lucha contra el discurso de odio. El Comité de Expertos ADI/MSI.-DIS preparó la Recomendación, que describe un enfoque integral para abordar el discurso de odio en un marco de derechos humanos. En ambas recomendaciones se precisa que, dado que todos los seres humanos pertenecen a la misma especie, el Comité de Ministros rechaza, al igual que la Comisión Europea contra el Racismo y la Intolerancia (ECRI), las teorías basadas en la existencia de diferentes «razas». Sin embargo, en este documento, el término «raza» se utiliza para garantizar que las personas que son percibidas de forma general y errónea como «pertenecientes a otra raza» no queden excluidas de la protección prevista por la legislación y la aplicación de políticas para prevenir y combatir el discurso de odio.

(20) Iñigo Gordon Benito, «Ciberodio. Un estudio de derecho penal comparado», *Cuadernos de RES PUBLICA en Derecho y Criminología,*2024-4, pp. 14-35 (https://doi.org/10.46661/respublica.9398*).*

(21) https://rm.coe.int/cm-recommendation-on-combating-hate-speech-esp-/1680ae4164.

siguiente clasificación: i) discurso de odio prohibido por la legislación penal, esto es, crímenes o delitos de odio cuyas máximas expresiones serían el apartheid y el genocidio[22]; ii) discurso de odio que no alcanza el nivel de gravedad exigido para incurrir en responsabilidad penal, pero que, no obstante, está sujeto a la legislación civil o administrativa; iii) tipos de expresión ofensivas o perjudiciales que no son lo suficientemente graves como para ser legítimamente restringidas en virtud del CEDH, pero que, no obstante, exigen respuestas alternativas, tales como: contranarrativas y otras contramedidas; medidas que fomenten el diálogo y el entendimiento intercultural, incluso a través de los medios de comunicación y las redes sociales; y actividades educativas, de intercambio de información y de sensibilización pertinentes. Frente a estos fenómenos debemos crear efectivos «sumideros de odio» que absorban, capturen y eliminen tales manifestaciones de la atmósfera social reduciendo su presencia en el aire de sociedades sólidas en el respeto de la diversidad.

Como indicábamos al principio, el discurso de odio no solo se concibe como un ataque a individuos o colectivos vulnerables, sino que también se concibe como una herramienta para atacar a las democracias liberales bajo el concepto de *desinformación*: información falsa o inexacta, que tiene intención de engañar y que se comparte con el fin de causar un daño grave. Los procesos abiertos recientemente, como el caso de la plataforma digital *X* en Brasil[23] o la detención de Pavel Durov, cofundador y CEO de Telegram*,* en París[24], se fundamentaron, entre otras razones, en la contribución de estas plataformas digitales a la difusión de desinformación en Internet dirigida a socavar la democracia, así como en su falta de colaboración con las autoridades estatales en la persecución de importantes delitos. En la Comunicación Conjunta de la Comisión Europea al Parlamento Europeo y al Consejo, *Sin lugar para el odio: una Europa unida contra esta lacra*[25], se dedica un aparatado a examinar la protección de la democracia frente al odio. En este documento se afirma que la desinformación y, en particular, la manipulación de información e injerencia por parte de agentes extranjeros a menudo tienen por objeto azuzar la polarización y el odio. Por su parte, el secretario general de las Naciones Unidas, insistiendo en que no es un fenómeno nuevo y que no existe una definición jurídica de tal concepto[26], ha destacado que tal fenómeno plantea múltiples desafíos de diferentes maneras:

---

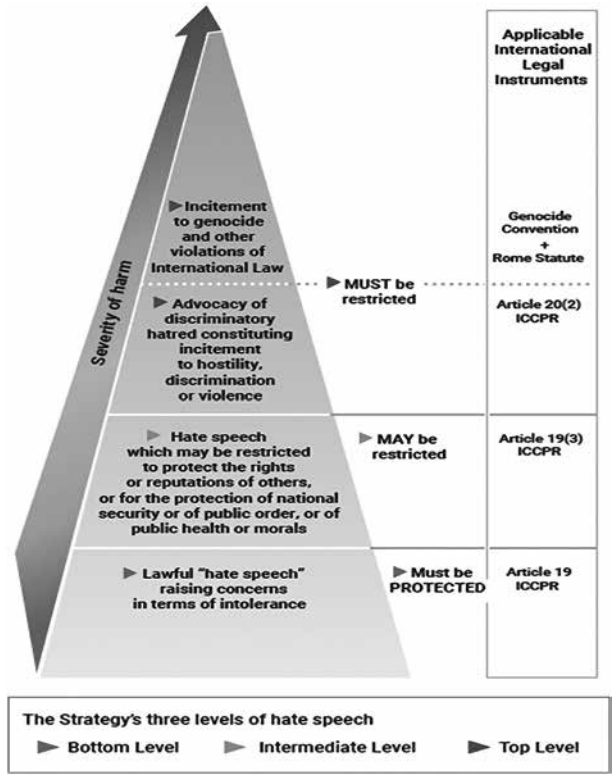
(22) En el caso del Derecho penal español, se prevén los delitos de odio en el artículo 510 del Código Penal (CP) y la agravante por motivo discriminatorio del artículo 24.4 CP. *Vid.* sobre este particular, la Circular 7/2019, de 14 de mayo, de la Fiscalía General del Estado, sobre pautas para interpretar los delitos de odio tipificados en el artículo 510 del Código Penal. *BOE núm. 124*, de 24 de mayo de 2019.

(23) Supremo Tribunal Federal de Brasil determina la suspensión de X, antiguo Twitter, en todo el territorio nacional: https://noticias.stf.jus.br/postsnoticias/stf-determina-suspensao-do-x-antigo-twitter-em-todo-o-territorio-nacional-2/

(24) https://elpais.com/internacional/2024-08-28/la-justicia-francesa-imputa-al-fundador-de-telegram-y-lo-deja-en-libertad-bajo-fianza.html

(25) JOIN/2023/51 final.

(26) En particular, señala que: «No existe una definición clara del término "desinformación", ni tampoco una postura o comprensión comunes respecto a su significado. Se ha utilizado por la Unión Internacional de Telecomunicaciones y la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) para describir contenidos falsos o engañosos que pueden provocar un daño específico, independientemente de las motivaciones, la conciencia o los comportamientos. La Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión define igualmente la desinformación como "información falsa que se difunde intencionadamente para causar un grave perjuicio social». Los estudios y aportaciones examinados para el presente informe señalan los siguientes elementos característicos de la desinformación: información inexacta, que tiene intención de engañar y que se compar-



Source: Based on ARTICLE 19, "Hate Speech" Explained: A Toolkit, p.19

«La pandemia de enfermedad por coronavirus (COVID-19) es un buen

te con el fin de causar un daño grave». Por su parte, según el Código de buenas prácticas de la Unión Europea en materia de desinformación, ésta puede definirse como «información verificablemente falsa o engañosa» que, de forma acumulativa, (a) «se crea, presenta y divulga con fines lucrativos o para engañar deliberadamente a la población», y (b) «puede causar un perjuicio público», entendido como «amenazas contra los procesos democráticos políticos y de elaboración de políticas, así como contra los bienes públicos, como la protección de la salud, el medio ambiente o la seguridad de los ciudadanos de la UE». Según el Código, el concepto de «desinformación» no incluye la publicidad engañosa, los errores de información, la sátira y la parodia ni las noticias y los comentarios claramente identificados como partidistas, y no se entenderá en perjuicio de la aplicación de obligaciones jurídicas vinculantes, códigos de autorregulación publicitaria y normas sobre publicidad engañosa. Tras evaluar su primer período de aplicación, la Comisión publicó en mayo de 2021 unas directrices detalladas sobre la manera de reforzar el Código (disponible en https://digital-strategy.ec.europa.eu/en/library/guidance-strengthening-code-practice-disinformation), en las que se reclamaba que se solventasen las deficiencias del Código de 2018 y se proponían soluciones para hacerlo más eficaz. En 2022, se adoptó el Código de buenas prácticas reforzado, disponible en https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/protecting-democracy/strengthened-eu-code-practice-disinformation_es. El 13 de febrero de 2025**,** la Comisión y el Comité Europeo de Servicios Digitales aprobaron la integración de este Código como Código de Conducta en materia de Desinformación en el marco de la Ley de Servicios Digitales (https://digital-strategy.ec.europa.eu/es/policies/code-practice-disinformation).

> ejemplo de las consecuencias potencialmente enormes de la *desinformación relacionada con la salud* para sociedades enteras, que incluyen la posible pérdida de muchas vidas. La difusión de *desinformación en contextos electorales* puede erosionar la confianza pública en la credibilidad de los procesos, lo que atenta contra el derecho a la participación política. La *desinformación puede incluir el fanatismo y el discurso de odio contra las minorías, las mujeres y los llamados "otros"*, con lo que supone de amenaza no solo para las personas directamente afectadas, sino también para la inclusión y la cohesión social. Puede amplificar las tensiones y divisiones en tiempos de emergencia, crisis, momentos políticos clave o conflictos armados. En efecto, la desinformación puede afectar a toda la gama de derechos humanos al perturbar la capacidad de las personas para tomar decisiones informadas sobre políticas relacionadas, por ejemplo, con el medio ambiente, la delincuencia, la migración y la educación, entre otras cuestiones de interés y preocupación públicos»[(27)].

II. **LA LIBERTAD DE EXPRESIÓN EN LA ERA DIGITAL**

1. LA LIBERTAD DE EXPRESIÓN COMO GARANTÍA DE UNA DEMOCRACIA PLURAL Y DEL DESARROLLO SOSTENIBLE. EL SURGIMIENTO DE UN FEUDALISMO DIGITAL Y LA NECESIDAD DE UN ECOSISTEMA INFORMATIVO QUE FOMENTE UN ENTORNO DIGITAL DIGNO DE CONFIANZA

La fuerza gravitacional de los discursos de odio aparece delimitada por el campo normativo del derecho a la libertad de expresión e información que se concibe como un pilar fundamental de la sociedad internacional. Como ha afirmado el TEDH, la libertad de expresión constituye una de las condiciones previas al funcionamiento de la democracia. No solo actúa como garantía frente a las injerencias del Estado (derecho subjetivo), sino que también representa un principio fundamental y objetivo para la vida democrática. La libertad de expresión no es un fin en sí misma, sino un medio para el desarrollo de una sociedad democrática y pluralista[(28)]. Esto significa especialmente que toda formalidad, condición, restricción o

---

(27) Informe del secretario general de 2022, *Contrarrestar la desinformación para promover y proteger los derechos humanos y las libertades fundamentales*, A/77/287, p.3. El énfasis es añadido. En junio de 2024, además, el secretario general presentaba los *Principios Globales de la ONU para la integridad de la información: claves para combatir la desinformación Naciones Unidas, «Integridad de la información. El ecosistema de la información en la era digital»,* disponible en https://www.un.org/es/information-integrity. *Vid.,* también, Naciones Unidas, «Governing AI for Humanity: Final Report»*, 2024,* https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.. Acerca del debate de la desinformación en el marco penal español véase, José León Alanpont, «La inteligencia artificial al servicio de la desinformación: un nuevo reto para el derecho penal», en Francisco Jiménez García (Dir.), Sandra López de Zubiría Díaz y Berta Alam Pérez (Coords.),, *Seguridad y Responsabilidad Penal e Internacional en el Uso de las TIC y la Inteligencia Artificial,* Iustel, Madrid, 2024, pp.95-127.

(28) Cfr. TEDH, asunto *Etxeberria Barrena Arza Nafarroako Autodeterminazio Bilgunea y Aiarako y otros c. España*, demandas nº 35579/03, 35613/03, 35626/03 y 35634/03, Sentencia 9 de junio de 2009, par. 63.

sanción impuesta en la materia debe ser proporcionada al fin legítimo que se persiga[29]. Asimismo y como ha resaltado la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, Irene Khan, la libertad de expresión es un factor esencial que propicia el impulso de los Objetivos del Desarrollo Sostenible (ODS), pues, además de empoderar a las personas, las comunidades y la sociedad civil, facilita la efectividad de otra serie de derechos que sustentan el desarrollo sostenible, como los derechos a la salud, a la educación, al agua y a un medio ambiente limpio. Permite a los Gobiernos estar mejor informados y responder mejor a las necesidades de su población y a la sociedad civil, a los medios de comunicación y a los ciudadanos exigir responsabilidades a los Gobiernos y las empresas, lo que asegura una verdadera democracia. El vínculo entre el desarrollo sostenible y la libertad de expresión no se limita a la información, sino que también reviste una importancia esencial la voz, esto es, el derecho a expresar opiniones, a debatir, discutir, criticar, cuestionar, protestar y exigir. Para lograr un verdadero desarrollo, «es preciso escuchar y prestar atención a las voces de los más desfavorecidos, y la sociedad civil y los medios de comunicación deben tener la libertad y el espacio necesarios para utilizar la información y la voz para pedir cuentas a los poderosos»[30].

Este derecho es reconocido tanto a nivel universal como regional[31], aunque su alcance presenta distintas dimensiones. Un ejemplo de ello es la marcada diferen-

---

(29) Cfr. TEDH (Pleno), asunto *Handyside c. Reino Unido*, demanda nº. 5493/72, Sentencia de 7 de diciembre de 1976, par.49.

(30) Informe de 2023 sobre el *Desarrollo sostenible y libertad de expresión: las razones de la importancia de la voz*, pp.3-5, A/HRC/53/25. En su Informe previo sobre el «Fortalecimiento de la libertad de los medios de comunicación y de la seguridad de los periodistas en la era digital», la Relatora Especial destacaba que la libertad de opinión y expresión no solo resultaba respaldada por la Agenda 2030, sino que dos sus indicadores para medir los avances en relación el objetivo 16 (Paz, Justicia e Instituciones Sólidas) insistían en la necesidad de «garantizar el acceso público a la información» y «proteger las libertades fundamentales (A/HRC/50/29, pp.4-5).

(31) Según el art.10 del CEDH, «1. Toda persona tiene derecho a la libertad de expresión. Este derecho comprende la libertad de opinión y la libertad de recibir o de comunicar informaciones o ideas sin que pueda haber injerencia de autoridades públicas y sin consideración de fronteras. El presente artículo no impide que los Estados sometan a las empresas de radiodifusión, de cinematografía o de televisión a un régimen de autorización previa. 2. El ejercicio de estas libertades, que entrañan deberes y responsabilidades, podrá ser sometido a ciertas formalidades, condiciones, restricciones o sanciones, previstas por la ley, que constituyan medidas necesarias, en una sociedad democrática, para la seguridad nacional, la integridad territorial o la seguridad pública, la defensa del orden y la prevención del delito, la protección de la salud o de la moral, la protección de la reputación o de los derechos ajenos, para impedir la divulgación de informaciones confidenciales o para garantizar la autoridad y la imparcialidad del poder judicial». Por su parte, el art.13 de la Convención Interamericana de Derechos Humanos dispone que «1. Toda persona tiene derecho a la libertad de pensamiento y de expresión. Este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección. 2. El ejercicio del derecho previsto en el inciso precedente no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar: a) El respeto a los derechos o a la reputación de los demás, o b) La protección de la seguridad nacional, el orden público o la salud o la moral públicas. 3. No se puede restringir el derecho de expresión por vías o medios indirectos, tales como el abuso de controles oficiales o particulares de papel para periódicos, de frecuencias radioeléctricas, o de enseres y aparatos usados en la difusión de información o por cualesquiera otros medios encaminados a impedir la comunicación y la circulación de ideas y opiniones. 4. Los espectáculos públicos pueden ser sometidos por la ley a censura previa con el exclusivo objeto de regular el acceso a ellos para la protec-

cia entre la amplia protección constitucional que la Primera Enmienda otorga a la libertad de expresión en el sistema de los Estados Unidos de América[32] y el *sistema de garantías de Estrasburgo* que modula este derecho en función de la protección de ciertos valores europeos. Frente a la ponderación que requieren la dignidad humana, los derechos fundamentales y los valores propios de una Europa consciente de su pasado de crímenes de odio –como el Holocausto– y de guerras, en EEUU se fomenta el individualismo frente al posible intervencionismo estatal pues, siguiendo la metáfora del «modelo de la fortaleza» planteado por Bollinger, «crear una excepción a la Primera Enmienda, por muy justificada que esté en relación con el caso concreto, conlleva el riesgo de abrir una grieta en los muros de la protección constitucional o de inclinar la pendiente de la censura»[33]. No resulta tampoco ajena a esta dialéctica, la diferente filosofía existente entre la Sección 230 de la *Communications Decency Act* norteamericana, que otorga una amplia inmunidad a las plataformas digitales sobre los contenidos emitidos salvo que se demuestre un control editorial sobre ellos[34], y el denominado *efecto Bruselas* que pivota sobre el criterio de diligencia debida y responsabilidad de los operadores digitales que posteriormente examinaremos.

Por otra parte, en un sistema de libre mercado entre operadores desiguales respecto a una materia tan esencial como la libertad de expresión, opinión e información, la posición del Estado como garante de la paz y la justicia en sociedades

---

ción moral de la infancia y la adolescencia, sin perjuicio de lo establecido en el inciso 2. 5. Estará prohibida por la ley toda propaganda en favor de la guerra y toda apología del odio nacional, racial o religioso que constituyan incitaciones a la violencia o cualquier otra acción ilegal similar contra cualquier persona o grupo de personas, por ningún motivo, inclusive los de raza, color, religión, idioma u origen nacional». La Carta Africana de Derechos Humanos y de los Pueblos se limita a declarar en su art. 9 que «Todo individuo tendrá derecho a recibir información. 2. Todo individuo tendrá derecho a expresar y difundir sus opiniones, siempre que respete la ley». Finalmente, el art.11 Carta de Derechos Fundamentales de la UE proclama que «1. Toda persona tiene derecho a la libertad de expresión. Este derecho comprende la libertad de opinión y la libertad de recibir o de comunicar informaciones o ideas sin que pueda haber injerencia de autoridades públicas y sin consideración de fronteras. 2. Se respetan la libertad de los medios de comunicación y su pluralismo».

(32) Basta recordar que EEUU formuló una reserva al Convenio Internacional sobre eliminación de todas las formas de discriminación racial conforme a la cual, «la Constitución y las leyes de los Estados Unidos se contempla una amplia protección de las libertades individuales de expresión y asociación. En consecuencia, los Estados Unidos no aceptan ninguna obligación en virtud del presente Convenio, en particular de los artículos 4 y 7, de restringir esos derechos mediante la adopción de leyes o de cualesquiera otras medidas, en la medida en que aquéllos estén protegidos por la Constitución y las leyes de los Estados Unidos» (*BOE núm. 249*, de 18 de octubre de 1995). Según Jon-Mirena Landa Gorostiza, «La Corte Suprema relega así a la zona libre de intervención penal el discurso del odio (*hate speech*) y liquida los modelos de legislación penal anti-odio "con palabras", mientras que afirma y legitima las figuras agravatorias de delitos base sin encontrar ahí tacha ninguna desde el punto de vista de la libertad de expresión» «Delitos de odio y estándares internacionales: una visión crítica a contracorriente», *Revista Electrónica de Ciencia Penal y Criminología*. 2020, núm. 22-19, p.16. *Vid*. también Héctor Gómez Peralta, «El discurso de odio y los límites de la libertad de expresión en Estados Unidos», *Revista Mexicana de Ciencias Políticas y Sociales,* vol.68, 2023, en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-19182023000300281.

(33) Cita extraída de Rafael Alcácer Guirao, «Víctimas y disidentes. El "discurso del odio" en EE. UU. y Europa». *Revista Española de Derecho Constitucional*, 2015-1, p.69.

(34) Valerie C. Brannon y Eric N. Holmes, «*S*ection 230: An Overvie», *Congressional Research Service,* 4 de enero de 2024, en: https://crsreports.congress.gov/product/pdf/R/R46751. Véase también, la entrada *Responsabilidad de las plataformas digitales* en el blog de Javier Casal Tavasci, https://protecciondata.es/responsabilidad-plataformas-digitales/.

democráticas cobra una singularidad particular que debe navegar entre la no vigilancia o intrusismo censor, directo o indirecto, y la conciliación de los derechos de la ciudadanía, en particular, la que no tienen una posición predominante. Se ha advertido por numerosos autores que el auge de las empresas tecnológicas y el control por parte de estas de la soberanía de los datos (*dataísmo*) va en detrimento de la soberanía estatal y la democracia social, favoreciendo el surgimiento de un feudalismo digital que reemplaza el concepto de ciudadano por el de siervo digital. Este feudalismo digital tiene graves consecuencias en términos de equidad y justicia social. Su poder persigue sustituir la política, moldear la opinión pública e, incluso, eliminar la competencia y reforzar su propia posición dominante en el mercado.(35). La imagen más descriptiva de este escenario feudal ha sido la toma de posesión del 47º presidente de los Estados Unidos de América rodeado por los dirigentes de la oligarquía tecnológica mundial que tienen su sede matriz de operaciones en este Estado(36).

Como ha advertido Víctor Abramovich, desde una posición «igualitaria» de libertad de expresión que no asume una desconfianza ciega en el papel de los Estados, es verdad que la intervención gubernamental «puede obturar el debate libre de ideas y opiniones y se justifica imponer límites y resguardos, por ejemplo, para que no reprima el discurso político disidente. Pero, en ocasiones, ante el papel hegemónico de algunos jugadores privados del ecosistema de la comunicación, la acción distributiva de los Estados contribuye a asegurar la discusión equilibrada y el pluralismo informativo a través de la inclusión de sectores y perspectivas sistemáticamente silenciadas. Ante estructuras desiguales de comunicación el Estado puede ser un amigo de la libertad de expresión. No sólo puede regular, sino que en ocasiones está obligado a hacerlo para revertir injusticias expresivas o bien injusticias políticas. De allí que la agenda de la intervención estatal comprende varios temas relevantes, como las regulaciones sobre la concentración de propiedad de medios, las políticas para cerrar las brechas de acceso a internet y a las tecnologías de la información, las políticas sobre medios públicos y comunitarios, entre otras cuestiones»(37).

La ONU comparte, en cierta medida, esta reflexión. En los «*Principios globales de las Naciones Unidas para la integridad de la información. Recomendaciones para la acción de diversas partes interesadas. El ecosistema de la información en la era digital*», aprobados el 24 de junio de 2024, se afirma que las grandes empresas tecnológicas, muchas de ellas con sede en lugares donde la regulación de la tecnología es limitada, ejercen un poder inmenso en la medida en que se benefician de enormes cantidades de datos recopilados sobre el comportamiento de los usuarios, lo que les permite moldear los flujos transnacionales de información y

(35) Josep Burgaya, *La Manada Digital. Feudalismo Hipertecnológico en una Democracia sin Ciudadanos*, El Viejo Topo, Barcelona, 2021; Ramón Blecua, «Feudalismo digital en un orden inestable», *Política Exterior*, 2020, vol. 34, Núm. 196, pp. 146-153; José Ramírez, «Navegando en la era del Feudalismo Digital: Blockchain como escudo para la soberanía de datos», *In Solidm*, 2023-1; Aram AHARONIAN, «El nuevo feudalismo digital», 18 de octubre de 2023 en https://esferacomunicacional.ar/el-nuevo-feudalismo-digital/.

(36) Iker Seisdedos «El día en que la oligarquía tecnológica tomó posesión con Donald Trump»., *El País*, 20 de enero de 2025.

(37) Víctor Abramovich, «Dilemas jurídicos en la restricción de los discursos de odio» *Sur - Revista Internacional de Derechos Humanos*, vol. 19, núm.32, 2022, pp.89-90.

controlar las experiencias digitales a escala mundial. Para corregir este desequilibrio de poder, se necesita un marco que dé prioridad tanto a la transparencia como a la supervisión independiente. En tal sentido se indica que:

> «Los usuarios merecen tener el dominio de sus datos y experiencias en internet, con vías claras de reclamación y rectificación. Se necesitan mecanismos de rendición de cuentas para abordar la responsabilidad de las empresas tecnológicas por las consecuencias del diseño y uso de sus productos y servicios sobre los derechos humanos y la cohesión social, incluso en situaciones de crisis y conflicto. Esto requerirá una evaluación crítica y transparente de la arquitectura de la plataforma que permita identificar las características que erosionan la integridad de la información y atentan contra los derechos humanos. Es preciso aplicar estrategias que prevengan y mitiguen dicha erosión, salvaguardando al mismo tiempo la libertad de expresión y el acceso a la información. *La desinformación y el odio no deben generar la máxima visibilidad y enormes beneficios. Nuevos modelos de negocio comercialmente viables que no dependan de la publicidad programática dirigida podrían fomentar la innovación, aumentar el empoderamiento de los usuarios y servir al interés público*. Este planteamiento polifacético puede crear un ecosistema informativo más equilibrado que respete los derechos de los usuarios y fomente un entorno en internet digno de confianza»[38].

2. LA LIBERTAD DE EXPRESIÓN Y SUS EXCEPCIONES COMO «PRUEBAS ESTRICTAS». VARIACIONES CONTEXTUALES SOBRE LA PROHIBICIÓN DE LOS DELITOS DE ODIO EN FUNCIÓN DEL RESULTADO LESIVO O DE LA INTENCIÓN DISCRIMINATORIA.

Volviendo al contenido del derecho a la libertad de expresión, los artículos 19 de la Declaración universal de los derechos humanos de 1948 y del Pacto internacional de las Naciones Unidas sobre los derechos civiles y políticos de 1966 (PIDCP), protegen el *derecho a tener opiniones sin ser molestados y garantizan el derecho a la libertad de expresión, es decir, el derecho a buscar, recibir y difundir informaciones e ideas de toda índole, sin limitación de fronteras, por cualquier medio*[39]. El Comité de Derechos Humanos, órgano de supervisión del cumplimiento del mencionado Pacto de 1966, ha subrayado que estas libertades son «condiciones indispensables para el pleno desarrollo de la persona ... [y] constituyen la piedra angular de toda sociedad libre y democrática» y forman la base para el pleno goce de una amplia gama de otros derechos humanos. Igualmente, ha subrayado que las restricciones, incluso cuando estén justificadas, «no pueden poner en peligro el derecho propiamente dicho»[40].

El carácter excepcional de las limitaciones se describe en el artículo 19.3 del Pacto cuando reconoce que los Estados pueden restringir la libertad de expresión

(38) Pp.10-11. El énfasis es añadido. Pueden consultarse en la siguiente dirección: https://www.un.org/sites/un2.un.org/files/integridad_informacion_principios_universales.pdf

(39) Para una mayor comprensión de este derecho vid. el Informe de 2019 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión (A/74/486).

(40) Comité de Derechos Humanos, Observación General núm. 34 (2011) relativa a la libertad de opinión y de expresión, CCPR/C/GC/34, pars. 2 y 21.

sólo cuando lo disponga la ley y sea necesario para proteger los derechos o la reputación de los demás, la seguridad nacional o el orden público, o la salud o la moral públicas. Se trata de excepciones que han de ser estrictamente definidas. En este sentido, el criterio de la legalidad implica que la restricción debe ser establecida por leyes que sean precisas, públicas y transparentes, debiéndose evitar proporcionar a las autoridades una discreción indeterminada o ilimitada, que se deben notificar apropiadamente a aquellos cuyo discurso está siendo regulado. Según el Comité, las medidas restrictivas deben estar sujetas a comentarios públicos y a un proceso legislativo o administrativo periódico, garantizándose en todo caso su control por parte de las autoridades judiciales independientes. Respecto a la legitimidad de la restricción, se exige que debe estar justificada para proteger uno o más de los intereses especificados en el artículo 19.3 del PIDCP y, en todo caso, debe obedecer a los criterios de necesidad y proporcionalidad. El Estado debe demostrar que la restricción es necesaria para proteger un interés legítimo y que es el medio menos restrictivo para lograr el objetivo perseguido. El Comité de Derechos Humanos se ha referido a esas restricciones como «pruebas estrictas», según las cuales las «restricciones solamente se podrán aplicar para los fines con que fueron prescritas y deberán estar relacionadas directamente con la necesidad específica de la que dependen»[41].

Respecto a este apartado, se ha insistido por parte de varios órganos de control que las leyes contra la blasfemia no cumplen con la condición de legitimidad del artículo 19.3 del Pacto por lo que los Estados debería derogar las leyes relativas a la blasfemia, debido al riesgo que representan para el debate sobre las ideas religiosas y el papel que dichas leyes juegan al permitir que los gobiernos muestren preferencia por las ideas de una religión sobre otras religiones, creencias o sistemas no basados en las creencias[42]. También se ha indicado por parte del Comité que el Pacto no autoriza las prohibiciones penales de la expresión de opiniones erróneas o interpretaciones incorrectas de acontecimientos pasados. No deben imponerse nunca restricciones al derecho a la libertad de opinión y, en cuanto a la libertad de expresión, las restricciones no deberían exceder de lo autorizado en el párrafo 3, o de lo prescrito en el artículo 20[43].

Por su parte, el artículo 20.2 del PIDCP dispone que los Estados parte están obligados a prohibir por ley «toda apología del odio nacional, racial o religioso que constituya *incitación a la discriminación, la hostilidad o la violencia»*. En la Convención internacional sobre la eliminación de todas las formas de discriminación racial, adoptada un año antes que el Pacto, su artículo 4 prescribe que los Estados parte, entre otras cosas, deben declarar como acto punible conforme a la ley toda difusión de ideas basadas en la superioridad o en el odio racial, toda incitación a la discriminación racial, así como todo acto de violencia o toda incitación a cometer tales actos contra cualquier raza o grupo de personas de otro color u origen étnico. También deben «declarar ilegales y prohibir las organizaciones, así como las actividades organizadas de propaganda y toda otra actividad de propaganda, que pro-

(41) Cfr. CCPR/C/GC/34, par. 22 e Informe de 2019 del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, cit. pp.5-6.

(42) Informe del Relator Especial sobre la libertad de religión o de creencias, A/HRC/31/18, pp.17-18

(43) Cfr. A/74/498, p.11 y CCPR/C/GC/34, par. 49.

muevan la discriminación racial e inciten a ella y reconocerán que la participación en tales organizaciones o actividades constituye un delito penado por la ley»[44]. Una diferencia fundamental entre el PIDCP y la Convención contra la discriminación racial salta a la vista: el primero *permite,* en su artículo 19.3, establecer limitaciones por ley a la libertad de expresión para los fines mencionados, mientras que el segundo *obliga,* en su artículo 4, a establecer tales limitaciones, y, además, no de cualquier forma, sino a través del establecimiento de una norma penal[45]. También, hemos de tener en cuenta que la Convención contra la discriminación racial rebajó considerablemente el estándar de la incitación sancionable al obligar a los Estados parte a declarar como «actos punibles conforme a la ley» no solo los actos de violencia o incitación a la violencia y a la discriminación raciales, sino también la «difusión de ideas basadas en la superioridad o en el odio racial», esto es, la simple difusión de ideas racistas sin exigir adicionalmente ni intencionalidad ni incitación[46].

El Comité de Derechos Humanos ha reconocido que el artículo 20 del Pacto presenta un lenguaje de emociones difícil de definir (odio, hostilidad) y una prohibición muy específica del contexto (apología de la incitación) que en todo caso están sujetas a las mismas reglas que las limitaciones a la libertad de expresión previstas en el artículo 19. En la misma dirección, el Comité para la Eliminación de la Discriminación Racial ha precisado que «la tipificación como delito de las formas de expresión racista se debe reservar para los casos más graves, que puedan probarse más allá de toda duda razonable, mientras que los casos menos graves deben tratarse por otros medios que no sean el derecho penal, teniendo en cuenta, entre otras cosas, la naturaleza y la amplitud de las repercusiones para las personas y los grupos destinatarios. La aplicación de sanciones penales debe regirse por los principios de legalidad, proporcionalidad y necesidad»[47]. Para calificar los actos de difusión e incitación como actos punibles conforme a la ley, este Comité halló que los Estados debían tener en cuenta una serie de factores a la hora de determinar si una expresión concreta entra dentro de una categoría prohibida, como el «contenido y la forma» del discurso, el «clima económico, social y político» en cuestión en el momento de la expresión, la «posición o condición del orador», el «alcance del discurso» y sus objetivos. El Comité recomendó que los Estados parte en la Convención consideraran «el riesgo o la probabilidad inminente de que el discurso en cuestión tenga por resultado la conducta deseada o pretendida por el emisor»[48].

---

(44) En su Resolución 79/160, la Asamblea General ha reiterado a los Estados que ha formulado reservas al artículo 4 de la Convención contra la discriminación racial a que, con carácter prioritario, consideren seriamente la posibilidad de retirar las reservas formuladas a esta disposición. Las reservas pueden consultarse en la siguiente dirección: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-2&chapter=4&clang=_en

(45) Beatriz Barreiro Carril, «Los derechos humanos de los migrantes ante el discurso de odio: Encaje de las normas y estándares internacionales en España», *Revista Internacional de los Estudios Vascos*, 69, 2024-1 (http://doi.org/10.61879/riev691zkia202401)

(46) Göran Rollnert Liern, «El discurso del odio: una lectura crítica de la regulación internacional». *Revista Española de Derecho Constitucional*, núm. 115, 2019, p. 100 (https://doi.org/10.18042/cepc/redc.115.03).

(47) Comité para la Eliminación de la Discriminación Racial, Recomendación General núm. 35 (2013), CERD/C/GC/35 par.12.

(48) *Ibid*., pars. 15 y 16.

Por su parte, en el *Plan de Acción de Rabat* sobre la prohibición de la apología del odio nacional, racial o religioso que constituye incitación a la discriminación, la hostilidad o la violencia de 2013[49], los términos principales se definen como sigue: El «odio» y la «hostilidad» se refieren a emociones intensas e irracionales de oprobio, enemistad y detestación hacia el grupo destinatario; el término «apología» debe entenderse como la intención de promover públicamente el odio hacia el grupo destinatario; y el término «incitación» se refiere a declaraciones sobre grupos nacionales, raciales o religiosos que crean un «*riesgo inminente de discriminación, hostilidad o violencia contra personas pertenecientes a esos grupos»*. También, en el Plan de Acción de Rabat se señalaron un total de seis factores para determinar la gravedad necesaria para tipificar como delito la incitación: a) El «contexto social y político que prevalecía en el momento en que se formuló y difundió el discurso»; b) La condición del orador, «específicamente la posición del individuo u organización en el contexto de la audiencia a la que va dirigido el discurso»; c) La intención, de manera que la «negligencia y la imprudencia no son suficientes para cometer una ofensa en virtud del artículo 20 del Pacto», que dispone que la mera distribución o circulación no equivale a apología o incitación; d) El contenido y la forma del discurso, en particular «el grado en que el discurso fue provocativo y directo, así como la forma, el estilo y la naturaleza de los argumentos utilizados»; e) La extensión o el alcance del acto del discurso, como la «magnitud y el tamaño de su audiencia», incluyendo si se trata de «un solo folleto o emisión en los principales medios de comunicación o a través de Internet, la frecuencia, la cantidad y el alcance de las comunicaciones, si la audiencia tenía medios para actuar sobre la incitación»; y, f) La probabilidad, incluida la inminencia, que significa que se «pueda identificar algún grado de riesgo de daño», en particular, mediante la determinación (por parte de los tribunales, como se sugiere en el Plan de Acción) de una probabilidad razonable de que el discurso logre incitar a la acción real contra el grupo destinatario[50].

Así pues, el discurso de odio sancionado penalmente por el Derecho internacional es aquél que contextualmente incita a una acción lesiva real contra un grupo determinado y supone un peligro inminente para la sociedad en que se produce. Según el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, si bien algunos de los conceptos anteriores pueden solaparse, «los siguientes elementos son fundamentales para determinar si una expresión constituye incitación al odio: el peligro real e inminente de violencia resultante de la expresión; la intención del autor de incitar a la discriminación, la hostilidad o la violencia; y un examen cuidadoso por parte del poder judicial del contexto en que se expresó el odio, *habida cuenta de que el derecho internacional prohíbe algunas formas de expresión por sus consecuencias, y no por su contenido*, porque lo que es sumamente ofensivo en una comunidad puede no serlo en otra»[51]. El

(49) A/HRC/22/17/Add.4

(50) Véase también, *Odio por motivos de religión o de creencias*. Informe de la Relatora Especial sobre la libertad de religión o de creencias, Nazila Ghanea A/HRC/55/47, pp.7-10

(51) A/67/357, par.46. Se añade que: «En consecuencia, cualquier examen del contexto debe incluir varios factores, como la existencia de tensiones recurrentes entre comunidades religiosas o raciales, la discriminación del grupo de que se trate, el tono y el contenido del discurso, la persona incitadora y los medios usados para difundir el discurso». El énfasis es añadido.

discurso prohibido comprendería «sólo aquellas expresiones que conllevan un *peligro claro, actual* y *particularizado*, pues están en condiciones de determinar comportamientos violentos inminentes, o un clima ostensible de hostigamiento, o de persecución en perjuicio de un determinado sector de la población por sus características»[52].

No obstante, estas condiciones pueden ser moduladas en atención al contexto del orden público y los valores en materia de derechos humanos que predominen en un determinado sistema jurídico. El TEDH, aplicando su *test de Estrasburgo* (legalidad, finalidad legítima y necesidad en una sociedad democrática), ha reiterado que, dado que la tolerancia y el respeto de la igual dignidad de todos los seres humanos constituyen los cimientos de una sociedad democrática y pluralista, puede considerarse válido penalizar o, incluso, impedir formas de expresión que propaguen, fomenten, promuevan o justifiquen el odio basado en la intolerancia, siempre que las «formalidades», «condiciones», «restricciones» o «sanciones» impuestas sean proporcionadas al objetivo legítimo perseguido[53]. Para Víctor Luis Gutiérrez Castillo, en el ámbito europeo, en particular, en el marco de la jurisprudencia del TEDH, los ataques que se cometen al injuriar, ridiculizar o difamar a grupos vulnerables o a los individuos que los conforman pueden ser suficientes para limitar la libertad de expresión. Podría afirmarse, pues, «que en Europa el umbral de la limitación de la libertad de expresión e incluso la sanción contra el discurso del odio se sitúa debajo de "la incitación indirecta" (…) El Tribunal también ha considerado justificada la sanción penal de discursos que pueden suscitar "rechazo", "hostilidad" u "odio" hacia un grupo en concreto, considerando este tipo de discurso incompatible con los valores del Convenio»[54]. El Tribunal de Estrasburgo, en atención al contexto, ha establecido que las obligaciones positivas derivadas del CEDH incluyen el *desmantelamiento de los entornos donde prospera la discriminación, incluso en ausencia de daños físicos directos e individuales*. La estereotipación negativa de un grupo en entornos de manifiesta hostilidad pública (como el caso de la comunidad LGBTI en Rusia), cuando alcanza cierto umbral, puede afectar no solo al sentido de identidad, autoestima y confianza del grupo, sino también a la vida privada de los miembros del grupo, quienes, por lo tanto, aunque no sean directamente objeto de las declaraciones controvertidas, pueden ser considerados víctimas conforme al artículo 34 del Convenio[55].

Otros autores consideran que la jurisprudencia del TEDH se ha caracterizado por su carácter fragmentario[56], pues, en ocasiones, sus decisiones han sido consideradas extralimitadas afectando gravemente la libertad de expresión, pues tienden

(52) Cfr. Víctor Abramovich, *op.cit.*, p.92.

(53) Cfr. STEDH (GS), *Sanchez c. Francia*, 15 de mayo de 2023, *demanda n.º* 45581/15, pars.148-151.

(54) Víctor Luis Gutiérrez Castillo, «El control europeo del ciberespacio ante el discurso de odio: análisis de las medidas de lucha y prevención», *Araucaria. Revista Iberoamericana de Filosofía, Política, Humanidades y Relaciones Internacionales*, 45, 2020, pp.299-300.

(55) Cfr. TEDH, asunto *Nepomnyashchiy y otros c. Rusia*, Sentencia de 30 de mayo de 2023, demandas nos. 39954/09 y 3465/17, pars.55-64.

(56) Cfr. Carmen Quessada Alcalá, «El discurso de odio *on line* y la jurisprudencia del Tribunal Europeo de Derechos Humanos en materia de libertad de expresión: ¿diferentes vías, mismos límites?», en Susana Almeida y Andrés Rousset, *Os Sistemas Europeu e Interamericano de Proteção dos Direitos Humanos: Uma leitura Comparada,* Aranzadi, 2024, p.949.

a validar injerencias en este derecho incluso cuando la conducta expresiva en cuestión no es capaz de poner en peligro los valores fundamentales del CEDH[57], mientras que en otros supuestos han sido criticadas por exigir un «daño individualizado» directo en lugar de reconocer la «degradación colectiva» que tales discursos provocan en sociedades escoradas contra los derechos de ciertos colectivos, como los determinados por la orientación sexual y la identidad de género. En estos contextos sociales, aislar la intención del efecto o riesgo pernicioso debilita el enfoque acerca de que los Estados deben combatir el «discurso prejuicioso» que fomenta la discriminación sistémica[58].

III. LOS DISCURSOS DE ODIO Y LA LIBERTAD DE EXPRESIÓN EN EL CIBERESPACIO

1. LOS PROBLEMAS ESPACIALES, PERSONALES Y JURISDICCIONALES DEL CIBERESPACIO COMO QUINTO ELEMENTO DE LA SOBERANÍA ESTATAL. LA INTENCIONALIDAD CONTEXTUALIZADA Y LA PRIVACIDAD PERCIBIDA EN EL DELITO DE ODIO *ONLINE*

El ciberespacio presenta unas características que refuerza la difusión (efecto amplificador[59], permanente y renovable en el tiempo[60]) de los discursos de odio y dificulta la efectividad de las medidas adoptadas para su disuasión. Desde la perspectiva de la soberanía estatal, la interconexión global del ciberespacio, entendido como el quinto elemento en el que esta se proyecta o busca proyectarse (juntos con los espacios terrestre, aéreo, marítimo y exterior), rompe las limitaciones físicas de las fronteras territoriales como ha señalado Sun Yirong. Esto se debe a que las TIC tienen una estructura de múltiples capas que permite separar, por un lado, la ubicación geográfica de la *infraestructura cibernética* y, por otro, las *actividades cibernéticas*. Aunque los Estados conservan la jurisdicción y el control sobre su territorio, la jurisdicción sobre una infraestructura física no conduce necesariamente a su jurisdicción sobre las actividades cibernéticas generadas. Esta separabilidad crea un complicado problema de jurisdicción. Un ejemplo de ello son los centros de datos globalizados que brindan servicios en la nube: la mayoría de las actividades cibernéticas se llevan a cabo con datos almacenados o procesados en ubicaciones extraterritoriales. Las tecnologías actuales no solo pueden disociar el almacenamiento y el acceso en dos ubicaciones geográficas conocidas, sino

(57) Cfr. Renato Lopes Militão, «O discurso de ódio no sistema da Convenção Europeia dos Direitos Humanos» en Susana Almeida y Andrés Rousset, *Os Sistemas Europeu e Interamericano de Proteção dos Direitos Humanos: Uma leitura Comparada,* Aranzadi, 2024, pp.957-981.

(58) Cfr. Sarthak Gupta, «Context, Content, and the 'Threshold of Severity': ECtHR's Jurisprudence on Satire vs Hate», *EJIL:Talk! Blog of the European Journal of International Law*, March 27, 2025 en relación con la Sentencia del TEDH en el asunto *Yevstifeyev y otros contra Rusia*, demandas nos. 226/18 y otras 3, de 3 de diciembre de 2024.

(59) STEDH, asunto *Cicad v. Suiza,* demanda, n.º 17676/09, de 7 de junio de 2016, par. 60

(60) STEDH (GS), asunto *Delfi SA c. Estonia*, demanda n.º 64569/09, par. 110.

que también permiten un almacenamiento distribuido en múltiples lugares alrededor del mundo, fuera del control de los usuarios(61).

Al problema espacial, se añade el problema personal en cuanto que al Estado se le exigiría no solo controlar sus actividades cibernéticas sino también la de los operadores privados que actúen desde su territorio o jurisdicción. Esto, en muchos casos, resulta especialmente complicado, no solo para aquellos Estados que carecen de las capacidades técnicas necesarias, sino también para los más avanzados tecnológicamente. Además, el carácter eminentemente extraterritorial del discurso de odio en línea plantea serios desafíos en la determinación de las jurisdicciones competentes para juzgar este tipo de delitos, lo que genera importantes lagunas, la existencia de «paraísos digitales» y la posibilidad de recurrir al *forum shopping* como estrategia de defensa frente a ejercicios exorbitantes de jurisdicción. Como ha puesto de relieve Libia Arenal Lora, esta situación jurisdiccional «constituye un enorme desafío para abordar los conflictos que puedan surgir relativos al foro competente y al derecho aplicable en materia de delimitación de las conductas y de la responsabilidad de los emisores de mensajes de odio y de las empresas proveedoras de servicios, más aún, si tenemos en cuenta que la regulación de estas conductas se desarrolla fundamentalmente en los derechos internos, y que pueden ser muy dispares en cuanto a su contenido y alcance»(62).

Junto a los operadores privados, Internet ha generado una nueva tecnología puntura de vigilancia diseñada para permitir el acceso integral de las autoridades a información que, de otro modo, podría permanecer encriptada. Numerosas empresas se han visto implicadas en actividades de *vigilancia digital selectiva* de periodistas a lo largo de la última década. El *Proyecto Pegasus*, por ejemplo, ha evidenciado que la vigilancia digital selectiva funciona, en última instancia, como medio de intimidación, agrava los riesgos a los que están expuestos los periodistas y sus fuentes, y daña al periodismo crítico(63). Así pues, si bien se concibe la tecnología como una herramienta neutra, las posibilidades de injerencia en los ámbitos estatales y personales son considerables. Como han destacado los Relatores Especiales sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia, existe una noción y percepción pública –tan persistente como perjudicial– de que la tecnología es intrínsecamente neutra y objetiva. Algunos han señalado que esta presunción de objetividad y neutralidad tecnológica sigue siendo una de las más arraigadas, incluso entre los propios creadores de tecnologías. Sin embargo, la tecnología nunca es verdaderamente neutra, ya que refleja los valores e intereses de quienes influyen en su diseño y uso y está profundamente condicionada por las mismas estructuras de desigualdad que operan en la sociedad(64).

El discurso de odio en línea puede producirse y compartirse con facilidad, a un bajo coste y de forma anónima. Además, a veces permanece disponible en línea durante largos períodos pudiendo ser «resucitado» en cualquier momento, incluso

(61) Sun Yirong, «The future of due diligence in cyberspace», *New York University Journal of International Law and Politics*, vol. 54, 2022-2, disponible en SSRN: https://ssrn.com/abstract=4629508.

(62) Libia Arenal Lora, *op.cit*. par.44.

(63) A/HRC/50/29, p.10.

(64) A/HRC/56/68, p.3.

con más popularidad y virulencia que en su momento. Como ha destacado nuestro Tribunal Supremo y compartido el Tribunal Constitucional, quien hoy incita a la violencia en una red social «sabe que su mensaje se incorpora a las redes telemáticas con vocación de perpetuidad. Además, carece de control sobre su zigzagueante difusión, pues desde que ese mensaje llega a manos de su destinatario éste puede multiplicar su impacto mediante sucesivos y renovados actos de transmisión»(65). También resulta problemático, aunque no sea exclusivo del medio *online*, cómo demostrar la intencionalidad en la incitación al odio, la discriminación o la violencia. Si se exige una intencionalidad clara y manifiesta o por el contrario cabe inferirse tal intencionalidad del contexto social y digital. Conforme al denominado «triángulo del odio», la intención exige la activación de una relación triangular entre el objeto, el sujeto del acto y la audiencia(66). En el referido asunto *Sanchez c. Francia*, el TEDH, remitiéndose a la jurisprudencia de la Corte Penal Internacional, dedujo la intención de los discursos racistas y xenófobos valorando el contexto inmediato, teniendo en cuenta que los comentarios se publicaron en el «muro» de *Facebook* de un político durante una campaña electoral con una amplia audiencia y habida cuenta de su contenido y tono general, así como de la virulencia y vulgaridad de algunos de los términos utilizados(67).

A tal efecto, Göran Rollnert Liern ha destacado cómo el requisito o presupuesto de la intencionalidad de la conducta como la probabilidad e inminencia de que se desencadene una acción discriminatoria, hostil o violenta previsto en la normativa internacional, ha experimentado un proceso de contextualización en el sentido de que el ambiente social y digital circundante es decisivo en la valoración tanto de la intención como del grado de riesgo existente. Para este autor, tanto la jurisprudencia del TEDH como la Recomendación 15 del CERI, al equiparar «la intención de incitar con la expectativa razonable de que el discurso tenga un efecto incitador supone reconocer "que la intención de incitar […] no es imprescindible […]", pudiendo sancionarse un discurso del odio no intencional, imprudente… Intencionalidad y riesgo inminente no son (…) requisitos que deben cumplirse acumulativamente para la sanción penal del discurso del odio sino condiciones alternativas, de manera que la presencia de un riesgo inminente permite prescindir de la intencionalidad (…) Y aquí entra en juego el contexto tanto para determinar la existencia de la intención de incitar –cuando la llamada a cometer los actos no sea inequívoca– como para valorar la existencia del riesgo»(68). Por otra parte, la fugacidad que, en ocasiones, propician las redes sociales podría dificultar la prueba del criterio de riesgo inminente. A tal efecto algunos autores, basados en precedentes judiciales, defienden la utilidad de la doctrina de las «amenazas verdaderas» que no requiere de la inminencia del riesgo, sino la intención intimidatoria del emisor, es

(65) Cfr. STC (Pleno) 35/2020, de 25 de febrero de 2020, F.J.5 en relación con el F.J.2 de la STS 4/2017 de 18 de enero.

(66) Cfr. Libia Arenal Lora, *op.cit,* par.39 con referencia a J. Temperman, «Blasphemy versus incitement: an international law perspective», en C. Beneke, C. Grenda y D. Nash (eds.). *Profane: Sacrilegious Expression in a Multicultural Age*, Oakland: University of California Press, pp. 401-425.

(67) Cfr. STEDH, *Sánchez c. Francia*, cit. par. 176.

(68) Cfr. Göran Rollnert Liern, «El discurso del odio: una lectura crítica….», *op. cit.* pp.97-98. *Vid*. también, Jon-Mirena Landa Gorostiza, *op.cit*. par. 24.

decir, la de provocar temor en el destinatario sin la necesidad de que se produzca el resultado violento[69].

Otra cuestión de especial transcendencia es el requisito de la publicidad de los discursos, pues los medios y redes digitales plantean el delicado problema de cómo trazar la línea divisoria entre lo privado y lo público. El Informe explicativo del Protocolo adicional a la Convención sobre ciberdelincuencia del Consejo de Europa, que excluye las comunicaciones o expresiones privadas del ámbito de aplicación del Protocolo, identifica como criterio de privacidad determinante «la intención del emisor de que el mensaje en cuestión será recibido solo por el destinatario predeterminado», intención subjetiva que puede establecerse a partir de «factores objetivos» como «el contenido del mensaje, la tecnología usada, las medidas de seguridad aplicadas y el contexto en el que se envía el mensaje». Según el Informe, el acceso abierto del material a cualquier persona (en un chat, en un grupo de noticias o en foros de discusión) encajará en la conducta típica de «poner a disposición del público», incluso cuando se requiera contraseña, siempre que la misma se proporcione a cualquiera o al que cumpla ciertos criterios; no obstante, la naturaleza de la relación entre los participantes en la comunicación deberá tenerse en cuenta para determinar si hubo difusión pública o puesta a disposición del público o si, por el contrario, se trató de una comunicación privada penalmente atípica[70].

A tal efecto, se ha indicado como factor decisivo de este elemento la audiencia potencial y efectiva de los discursos. Para el autor anteriormente mencionado, si el criterio fundamental para diferenciar las comunicaciones privadas no punibles de la difusión pública tipificada penalmente es, según el citado Informe Explicativo, la intención subjetiva del emisor de limitar el mensaje a un determinado destinatario, la difusión pública solo será sancionable cuando sea intencional y voluntaria. En consecuencia, si la publicidad no ha sido buscada intencionadamente por el emisor porque es el destinatario quien ha difundido un mensaje privado, será este último el responsable penalmente de la difusión. No obstante, en el asunto *Sanchez c. Francia*, la Gran Sala del TEDH hizo responsable al propietario del «muro» en *Facebook* no solo de sus propios comentarios, sino también de los añadidos por terceros, al considerar que le correspondía llevar a cabo una vigilancia o verificación mínima sobre la licitud de dichos mensajes. En concreto, señaló que, si bien el responsable del «muro» había publicado un mensaje de advertencia sobre la necesidad de respetar la legalidad de los contenidos establecida por la propia red social, no había eliminado los comentarios impugnados y, sobre todo, no se había tomado la molestia de comprobar el contenido de los comentarios entonces accesibles al público[71].

Por otra parte, la propia percepción, o falta de conciencia, sobre el entorno de la red social por los usuarios afecta a la publicidad intencional de la conducta. En este sentido, el profesor Rollnert Liern se pregunta,

> «¿Hasta qué punto los usuarios tienen conciencia de la posible repercusión pública de sus mensajes en las redes o las perciben como prolongación virtual de un espacio de comunicación informal y desenfadado con un círcu-

(69) Cfr. Libia Arenal Lona, *op.cit.* par. 41.

(70) Cfr. Explanatory Report to the Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems, p.6. Puede consultarse en https://rm.coe.int/1680989b1c

(71) STEDH, par. 194.

> lo limitado de amistades en el que son permisibles expresiones que no harían en un ámbito público? Las redes sociales convierten inmediatamente en públicos actos y comportamientos individuales antes limitados a ambientes y redes. La falta de conciencia de actuar en un espacio público de comunicación hace que expresiones exaltadas y de odio "que en principio solo se permitiría a sí mismo en un entorno reducido (…) saltan al debate público digital sin que se haya reflexionado sobre la diferencia cuantitativa y cualitativa que le da su difusión en la red"»[(72)].

2. A LA ESPERA DE UNA REGULACIÓN INTERNACIONAL, EL *SOFT-LAW* COMO ALTERNATIVA REGULADORA SOBRE COMPORTAMIENTOS RESPONSABLES Y DERECHOS DIGITALES

El ciberodio y la desinformación se desenvuelven esencialmente en el marco regulatorio antes expuesto. A la espera de un mayor desarrollo y conocimiento del futuro de las tecnologías y sus potencialidades, los sistemas jurídicos, incluyendo el internacional, han optado, como viene siendo habitual en estos procesos de transición, por el *soft-law* o derecho asertivo. No obstante, como derecho vinculante (*hard law*), por el momento y exclusivamente referido a la dimensión racista y xenófoba del fenómeno del discurso de odio, en el marco europeo –aunque abierto a todos los Estados de la comunidad internacional–, contamos con el Protocolo adicional al Convenio sobre la ciberdelincuencia relativo a la penalización de actos de índole racista y xenófoba cometidos por medio de sistemas informáticos, adoptado por el Consejo de Europa en 2003, que cuenta con 31 ratificaciones de los 46 Estados miembros, entre ellos España[(73)], y 5 de Estados no miembros del Consejo de Europa (Benín, Marruecos, Ruanda, Paraguay y Senegal[(74)]), frente a las 78 ratificaciones del Convenio matriz sobre la ciberdelincuencia[(75)]. Con un contenido casi idéntico, en el ámbito del continente africano, hallamos la Convención de la Unión Africana de 2014 sobre ciberseguridad y protección de datos personales de 2014[(76)], conocida como la Convención de Malabo, que entró en vigor desde el 8 de junio de 2023.

En el marco de las Naciones Unidas, también se ha adoptado el 24 de diciembre de 2024, la *Convención de las Naciones Unidas contra la ciberdelincuencia*[(77)] que no hace ninguna mención a los ciberdelitos de odio –se optó muy oportuna-

---

(72) Göran Rollnert Liern, «Redes sociales y discurso del odio: perspectiva internacional», *Revista de los Estudios de Derecho y Ciencia Política*, núm. 31 2020, p.7 https://roderic.uv.es/rest/api/core/bitstreams/aa61a063-beb9-460a-ac4c-325d678e211e/content

(73) *BOE núm. 26*, de 30 de enero de 2015.

(74) https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=189. Las reservas y declaraciones de los Estados parte pueden consultarse en la siguiente dirección: https://www.coe.int/en/web/conventions/full-list?module=declarations-by-treaty&numSte=189&codeNature=0

(75) https://www.coe.int/en/web/conventions/full-list?module=signatures-by-treaty&treatynum=185

(76) https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

(77) Con el largo y descriptivo subtítulo de *Fortalecimiento de la cooperación internacional para la lucha contra determinados delitos cometidos mediante sistemas de tecnología de la información y las comunicaciones y para la transmisión de pruebas en forma electrónica de delitos graves,* UN. Doc. A/RES/79/243.

mente en no tipificar los denominados delitos cibernéticos de contenido al carecer la mayoría de los propuestos de una definición y regulación específicas[78]–, aunque en su artículo 4 se dispone que al dar «efecto a otros convenios, convenciones y protocolos aplicables de las Naciones Unidas en los que sean partes, los Estados partes velarán por que los delitos tipificados con arreglo a esos convenios, convenciones y protocolos se consideren también delitos en el derecho interno cuando se cometan mediante la utilización de sistemas de tecnología de la información y las comunicaciones». Tras un tenso debate sobre la inclusión o no de una cláusula transversal de respecto a los derechos humanos[79], finalmente se adoptó el artículo 6 que establece que los Estados parte velarán por que el cumplimiento de sus obligaciones con arreglo a la presente Convención se ajuste a sus obligaciones en virtud del derecho internacional de los derechos humanos y que nada de lo dispuesto en la Convención se interpretará en el sentido de que permita la supresión de los derechos humanos o las libertades fundamentales, incluidos los derechos relacionados con las libertades de expresión, de conciencia, de opinión, de religión o creencia, de reunión pacífica y de asociación, de conformidad y en consonancia con el derecho internacional de los derechos humanos aplicable[80].

El Protocolo adicional al Convenio europeo sobre la ciberdelincuencia incluye como delitos a tipificar en los ordenamientos nacionales la difusión, la amenaza y los insultos con motivación racista y xenófoba, así como la negación, minimización burda, aprobación o justificación del genocidio o de crímenes contra la humanidad cuando se comentan intencionadamente y sin derecho. Salvo en el caso de las amenazas, en el resto de los supuestos previstos, se permite a los Estados, como así lo han hecho en su amplia mayoría (España se ha limitado a la clásica reserva

(78) Así se pretendía tipificar, entre otros, los siguientes delitos cuando fueran facilitados por las TIC: los insultos a los valores religiosos (Irán), los delitos relacionados con actividades extremistas, con actividades terroristas, incitación a cometer actos subversivos o armados, incitación o coacción al suicidio (Federación de Rusia, también en nombre de Belarús, Burundi, China, Nicaragua y Tayikistán), la propaganda de la conflictividad, la sedición, el odio o el racismo (Egipto). Cfr. Recopilación de propuestas y contribuciones presentadas por los Estados Miembros respecto de las disposiciones sobre criminalización, las disposiciones generales y las disposiciones sobre las medidas procesales y la aplicación de la ley de una convención internacional integral sobre la lucha contra la utilización de las tecnologías de la información y las comunicaciones con fines delictivos, A/AC.291/9 y A/AC.291/9/Add.1. Sobre los trabajos de este Comité véase, Eulalia W. Petit de Gabriel, «Libertad de expresión y delitos de opinión a la luz de la futura Convención Internacional Integral sobre la Lucha contra la Utilización de las Tecnologías de la Información y las Comunicaciones con Fines Delictivos» en Juana del Carpio Delgado y María Holgado González (Dirs.) y Alejandro L. de Pablo Serrano (Coord.), L*a Libertad de Expresión Asediada. Delitos de Odio, Delitos de Opinión, Censuras de Gobiernos y Empresas*, Aranzadi, 2023, pp. 405-454.

(79) Un conjunto de Estados, encabezados por Rusia e Irán, rechazaban que en esta Convención se incluyeran cláusulas relativas a los derechos humanos, frente a la propuesta de otros Estados, incluyendo la UE, defensores de la necesidad de la introducción de una cláusula de salvaguardia de derechos humanos y en contra de toda discriminación o persecución basada en características individuales. Sobre esta cuestión *vid*. el Check-in de la Convención Cibernética realizado por *The Global Initiative*: https://globalinitiative.net/announcements/cyber-convention-check-in/.

(80) Además su artículo 24.1, en el marco de *Condiciones y salvaguardias*, precisa que cada Estado parte «velará por que la instauración, ejecución y aplicación de las facultades y procedimientos previstos en el presente capítulo se sometan a las condiciones y salvaguardias previstas en su derecho interno, que deberán asegurar la protección de los derechos humanos, de conformidad con las obligaciones que haya asumido en virtud del derecho internacional de los derechos humanos y que deberán integrar el principio de la proporcionalidad».

relativa a la soberanía sobre Gibraltar), que exijan para su penalización que tales conductas inciten al odio, la discriminación o la violencia. E, incluso, cabe la posibilidad de que se reserven el derecho a no tipificar tales comportamientos aun cuando se prevea la necesidad de adoptar otros recursos eficaces.

Sin referencia expresa a los delitos/discurso de odio, en el marco del Consejo de Europa, el Convenio Marco del Consejo de Europa sobre inteligencia artificial y derechos humanos, democracia y Estado de derecho, adoptado el 5 de septiembre de 2024 en Vilna (Lituania), prevé por una parte, en su artículo 10, que cada Parte adoptará o mantendrá medidas con miras a garantizar que las actividades dentro del ciclo de vida de los sistemas de IA respeten la igualdad, incluida la igualdad de género, y la prohibición de discriminación, según lo dispuesto en la legislación internacional y nacional aplicable. Igualmente, y durante el ciclo vital de los sistemas de IA, se comprometen a adoptar o mantener medidas destinadas a superar las desigualdades para lograr resultados justos y equitativos en consonancia con sus obligaciones nacionales e internacionales aplicables en materia de derechos humanos. Por otra parte, en su artículo 16, en el marco de gestión de riesgos e impactos, se dispone que cada Parte, teniendo en cuenta los principios de dignidad y autonomía humana, transparencia y supervisión, rendición de cuentas y responsabilidad, igualdad y no discriminación, privacidad y protección de datos personales, fiabilidad e innovación segura, adoptará o mantendrá medidas para la identificación, evaluación, prevención y mitigación de los riesgos que plantean los sistemas de inteligencia artificial, teniendo en cuenta, entre otras consideraciones, la gravedad y la probabilidad de impactos reales y potenciales sobre los derechos humanos, la democracia y el Estado de Derecho[(81)].

Así pues, la regulación de los discursos de odio queda sometida fundamentalmente al *soft-law*. Además de los documentos ya citados, como el Plan de Acción de Rabat; la Recomendación General n.º 15 de la ECRI (2015); o la Recomendación CM/Rec(2022)16 del Comité de Ministros del Consejo de Europa sobre la lucha contra el discurso de odio, podríamos destacar otros instrumentos adoptados en el marco de la Unión Europea (UE). Tras la regulación normativa prevista en la Decisión marco 2008/913/JAI del Consejo de 28 de noviembre de 2008[(82)], relativa a la lucha contra

(81) https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection

(82) *DOUE L 328* de 6 de diciembre de 2008. Conforme a su artículo 1: «1) Cada Estado miembro adoptará las medidas necesarias para garantizar que se castiguen las siguientes conductas intencionadas: a) la incitación pública a la violencia o al odio dirigidos contra un grupo de personas o un miembro de tal grupo, definido en relación con la raza, el color, la religión, la ascendencia o el origen nacional o étnico; b) la comisión de uno de los actos a que se refiere la letra a) mediante la difusión o reparto de escritos, imágenes u otros materiales; c) la apología pública, la negación o la trivialización flagrante de los crímenes de genocidio, crímenes contra la humanidad y crímenes de guerra tal como se definen en los artículos 6, 7 y 8 del Estatuto de la Corte Penal Internacional, dirigida contra un grupo de personas o un miembro de tal grupo definido en relación con la raza, el color, la religión, la ascendencia o el origen nacional o étnico cuando las conductas puedan incitar a la violencia o al odio contra tal grupo o un miembro del mismo; d) la apología pública, la negación o la trivialización flagrante de los crímenes definidos en el artículo 6 del Estatuto del Tribunal Militar Internacional adjunto al Acuerdo de Londres, de 8 de agosto de 1945, dirigida contra un grupo de personas o un miembro de tal grupo definido en relación con la raza, el color, la religión la ascendencia o el origen nacional o étnico cuando las conductas puedan incitar a la violencia o al odio contra tal grupo o un miembro del mismo. 2. A los efectos de lo dispuesto en el apartado 1, los Estados miembros podrán optar por castigar únicamente las conductas que o bien se lleven a cabo de forma que puedan dar lugar a perturbaciones del orden público o que sean amenazadoras, abusivas o insultantes. 3. A los efectos de lo dispuesto en el apartado 1, la referencia a la religión

determinadas formas y manifestaciones de racismo y xenofobia mediante el Derecho penal, y la Directiva 2010/13/UE de servicios de comunicación audiovisual del Parlamento Europeo y del Consejo(83), se aprobó en el año 2014 las Directrices de la UE sobre derechos humanos relativas a la libertad de expresión en Internet y fuera de Internet y en 2016 el *Código de Conducta para la lucha contra la incitación ilegal al odio en Internet* que implicaba un compromiso por parte de las empresas digitales que lo suscribieron de combatir el discurso de odio *online* en sus respectivas plataformas ofreciendo procedimientos efectivos y equipos especializados de supervisión, revisión y denuncia de sus contenidos (los denominados informadores de confianza o comunicantes fiables, *trusted flaggers*), incluyendo su supresión en el plazo de 24 horas desde la recepción de la solicitud de retirada(84). En 2021, la Comisión presentaba una Comunicación al Parlamento Europeo y al Consejo, «Una Europa más integradora y protectora», con el objetivo de ampliar la lista de delitos de la UE a la incitación al odio y los delitos motivados por el odio(85) y el 20 de enero de 2025, el Código de conducta revisado de 2016 («Código de conducta +») se integró en el marco regulador de denominada Ley de Servicios Digitales(86). No obstante, la Comisión también ha constatado la reiterada infracción, por transposición solo parcial y de manera incorrecta, de algunos Estados miembros de la Decisión marco 2008/913/JAI lo que le ha obligado a iniciar distintos procedimientos por incumplimiento(87).

También podemos citar el *Código de buenas prácticas sobre desinformación reforzado* que fue firmado y presentado el 16 de junio de 2022 por 34 corporaciones digitales que se han sumado al proceso de revisión del Código de 2018. Asimismo, el 13 de febrero de 2025, la Comisión y el Comité Europeo de Servicios

---

tiene por objeto abarcar, al menos, las conductas que sean un pretexto para dirigir actos contra un grupo de personas o un miembro de tal grupo definido en relación con la raza, el color, la ascendencia o el origen nacional o étnico. 4. Los Estados miembros podrán hacer, en el momento de la adopción de la presente Decisión Marco o posteriormente, una declaración en virtud de la cual la negación o la trivialización flagrante de los crímenes a los que hace referencia el apartado 1, letras c) y d), sean punibles solo si los crímenes a los que hacen referencia dichas letras han sido establecidos por resolución firme de un tribunal nacional de dicho Estado miembro o un tribunal internacional, o mediante resolución firme exclusiva de un tribunal internacional».

(83) *DOUE L 95* de 15 de abril de 2010. El art.9.1 c) de la referida Directiva dispone que los Estados miembros velarán por que las comunicaciones comerciales audiovisuales realizadas por prestadores sujetos a su jurisdicción se abstengan de incluir o fomentar cualquier discriminación por razón de sexo, raza u origen étnico, nacionalidad, religión o creencia, discapacidad, edad u orientación sexual.

(84) El Código de conducta+ fue firmado y presentado para su integración bajo la DSA por Dailymotion, Facebook, Instagram, Jeuxvideo.com, LinkedIn, servicios de consumo alojados por Microsoft, Snapchat, Rakuten Viber, TikTok, Twitch, X y YouTube. https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online. Este Código, que se basa en el Código de conducta adoptado en 2016 , refuerza la forma en que las plataformas en línea abordan el contenido considerado como discurso de odio ilegal según la legislación de la UE y las leyes de los Estados miembros. Facilita el cumplimiento y la aplicación efectiva de la DSA en este ámbito específico. Tras su integración, la adhesión al Código de conducta+ puede considerarse una medida adecuada de mitigación de riesgos para los signatarios designados como Plataformas en línea muy grandes (VLOP) y Motores de búsqueda (VLOSE) según la DSA.

(85) COM/2021/777 final

(86) https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online.

(87) Entre otros, a Rumanía, Bulgaria, Estonia e Irland**a**. Véanse los últimos paquetes de procedimientos de infracción publicados por la Comisión: https://ec.europa.eu/commission/presscorner/detail/es/inf_24_4561 y https://ec.europa.eu/commission/presscorner/detail/es/inf_23_1808.

Digitales aprobaron su integración como código de conducta en materia de desinformación en el marco de la denominada Ley de Servicios Digitales[88]. Los firmantes se comprometen a tomar medidas en varios dominios, tales como: desmonetizar la difusión de desinformación; garantizar la transparencia de la publicidad política; empoderar a los usuarios; mejorar la cooperación con los verificadores de datos; y proporcionar a los investigadores un mejor acceso a los datos[89].

El *soft-law* también ha servido para la creación de un entorno programático ético digital conducente a un catálogo de derechos digitales. La UE ha proclamado en su Declaración europea sobre los derechos y principios digitales para la década digital la necesidad de crear un entorno digital protegido y seguro que le permita combatir y responsabilizar a quienes traten de socavar, en la UE, la seguridad en línea y la integridad del entorno digital o fomenten la violencia y el odio por medios digitales[90]. En España no solo nuestro Tribunal Supremo y Tribunal Constitucional han reconocido, mediante una interpretación teleológico-sistemática, el *derecho a la intimidad en el entorno digital* y el *derecho a la autodeterminación digital* en el que la intimidad pasa a ser concebida como un bien jurídico que se relaciona con la libertad de acción del sujeto, con sus facultades positivas de actuación para controlar la información relativa a su persona y su familia en el ámbito público[91], sino que también se ha aprobado, en 2021, por parte del Gobierno una Carta de Derechos Digitales que, como se indica en la misma, «no trata de crear nuevos derechos fundamentales sino de perfilar los más relevantes en el entorno y los espacios digitales o describir derechos instrumentales o auxiliares de los primeros. Se trata de un proceso naturalmente dinámico dado que el entorno digital se encuentra en constante evolución con consecuencias y límites que no es fácil predecir». Asimismo, se indica que el objetivo de la Carta es descriptivo, prospectivo y asertivo y carece de carácter normativo.

La Carta de Derechos Digitales no prevé ninguna disposición expresa sobre el ciberodio, aun cuando exhorte a que en el ámbito de la protección de las personas menores de edad en el entorno digital se introduzca estudios para prevenir la discriminación y los discursos de odio y que en materia de libertad de expresión se incide sobre la necesidad de la autorregulación y el establecimiento de códigos de conducta que prevean sistemas de comunicación, reclamación, mecanismos y herramientas de alerta y detección, así como sistemas voluntarios de mediación o

(88) https://digital-strategy.ec.europa.eu/es/policies/code-practice-disinformation

(89) Para más información *vid.*: https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation. Reconociendo la importancia de que el Código sea a prueba de futuro, los signatarios acordaron establecer un marco para una mayor colaboración a través de un grupo de trabajo permanente. El Código también viene con un marco de seguimiento reforzado basado en elementos de información cualitativa e indicadores de nivel de servicio que miden la eficacia de su aplicación. Los signatarios han creado un Centro de Transparencia, que ofrece al público una visión clara de las políticas que ponen en marcha para implementar sus compromisos y lo actualiza periódicamente.

(90) El Parlamento Europeo, el Consejo y la Comisión han proclamado solemnemente la Declaración conjunta sobre los Derechos y Principios Digitales para la Década Digital, *DOUE C 23* de 23 de enero de 2023.

(91) Fidel Ángel Cadena Serrano, «La autodeterminación informativa y el derecho penal», Diario La Ley, Nº 9754, Sección Comentarios de jurisprudencia, 15 de Diciembre de 2020 y Mariano Yzquierdo Tolsada «Comentario de la Sentencia del Tribunal Supremo de 28 de julio de 2022 (593/2022): El derecho a la propia imagen no presenta singularidad en la plataforma Youtube respecto a otras plataformas (Facebook, Twitter o Instagram)», en https://www.boe.es/biblioteca_juridica/comentarios_sentencias_unificacion_doctrina_civil_y_mercantil/abrir_pdf.php?id=COM-D-2022-8

arbitraje para resolver discrepancias. Igualmente se prevé la responsabilidad de los prestadores de servicios intermediarios en los supuestos de la comisión de tipos penales, bien por exceder del alcance típico de la prestación de su servicio, bien por no haber actuado con diligencia para bloquear o retirar contenido cuando tengan conocimiento efectivo de que es ilícito[92].

En el ámbito internacional, además de la Recomendación sobre la ética de la inteligencia artificial aprobada por la UNESCO el 23 de noviembre de 2021[93], las Naciones Unidas aprobaba el 17 de diciembre de 2024 una resolución sobre el *Derecho a la privacidad en la era digital*. Mediante esta resolución, la Asamblea General recomienda *pautas de comportamiento responsable* tanto a los Estados como a las empresas que operan en el mundo digital. En primer término, manifiesta su preocupación por la difusión de información errónea y desinformación, en particular, en las plataformas de medios sociales, que se pueden concebir e implementar de manera que induzcan a error, difundan el racismo, la xenofobia, los estereotipos negativos y la estigmatización, violen y conculquen los derechos humanos, incluido el derecho a la privacidad, frenen la libertad de expresión, incluida la libertad de buscar, recibir y difundir información, e inciten a todas las formas de violencia, odio, intolerancia, discriminación y hostilidad. Por otra parte, la Asamblea General opta, respecto a las empresas tecnológicas, por la autorregulación y medidas de cumplimiento y vigilancia que garanticen que sus modelos de negocio y sus procesos de diseño y desarrollo, sus operaciones comerciales y sus prácticas de recopilación y procesamiento de datos estén en consonancia con los *Principios Rectores sobre las Empresas y los Derechos Humanos: puesta en práctica del marco de las Naciones Unidas para «proteger, respetar y remediar»*, y pongan de relieve la importancia de llevar a cabo la diligencia debida en materia de derechos humanos con respecto a sus productos, en particular, con respecto al papel de los algoritmos y los sistemas de clasificación[94].

IV. DEL EFECTO REGLAMENTARIO DE BRUSELAS A LOS ASERTIVOS PACTOS PARA EL FUTURO Y GLOBAL DIGITAL DE LAS NACIONES UNIDAS. HACIA UNA CLÁUSULA DE CIBERSOLIDARIDAD EN LA UNIÓN EUROPEA

La debida diligencia algorítmica ha alcanzado su mayor desarrollo en el marco de la Unión Europea constituyendo uno de los fenómenos regulatorios de lo que se

(92) En su punto IV, se prevé el derecho al «pseudonimato», esto es y de conformidad con la Carta, de acuerdo con las posibilidades técnicas disponibles y la legislación vigente, se permitirá el acceso a los entornos digitales en condiciones de *pseudonimidad*, siempre y cuando no sea necesaria la identificación personal para el desarrollo de las tareas propias de dicho entorno. El diseño de la citada *pseudonimidad* asegurará la posibilidad de reidentificar a las personas previa resolución judicial en los casos y con las garantías previstas por el ordenamiento jurídico. La Carta de Derecho Digitales puede consultarse en la siguiente dirección: https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf. Normativamente, resulta de especial interés la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, *BOE núm. 294* de 6 de diciembre de 2018.

(93) https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial

(94) A/RES/79/175, pp.7 y 11.

ha denominado el *efecto Bruselas*[95]*. Se trata de una batería de normas que exigen de las empresas tecnológicas la puesta en práctica de una especie de compliance plus digital* que se extiende a todo el ciclo de vida de la actividad cibernética. Supone la adopción de medidas responsables por parte de los prestadores de plataformas en línea ante cualquier riesgo sistémico en la Unión que se derive del diseño o del funcionamiento de su servicio y los sistemas relacionados con este, incluidos los sistemas algorítmicos, o del uso que se haga de sus servicios. Y entre los riesgos sistémicos, se prevé: cualquier efecto negativo real o previsible para el ejercicio de los derechos fundamentales previstos en la Carta de derechos fundamentales de la UE, en particular, los relativos a la dignidad humana, al respeto de la vida privada y familiar, a la protección de los datos de carácter personal, a la libertad de expresión e información, incluida la libertad y el pluralismo de los medios de comunicación, a la no discriminación, a los derechos del niño y a un nivel elevado de protección de los consumidores. Asimismo, cualquier efecto negativo real o previsible sobre el discurso cívico y los procesos electorales, así como sobre la seguridad pública[96]. Respecto a la reducción de los riesgos se hace mención, en concreto, a la rápida retirada de los contenidos notificados o el bloqueo del acceso a ellos en el caso de la incitación ilegal al odio o la ciberviolencia[97].

Estas obligaciones de diligencia debida se refieren a medidas preventivas para detectar y evaluar riesgos, así como a acciones de alerta temprana para interrumpir o minimizar efectos adversos reales o potenciales, de monitoreo de la efectividad de las medidas adoptadas, de comunicación, de previsión de recursos y medidas de reparación, incluyendo sanciones. Se ha de garantizar la existencia de un procedimiento justo, a disposición del público, accesible, predecible y transparente para tramitar las reclamaciones. Este *efecto Bruselas* se iniciaría con la Recomendación (UE) 2018/334 de la Comisión sobre medidas para combatir eficazmente los contenidos ilícitos[98] que, además de identificar los contenidos ilícitos dignos de persecución, hace hincapié en la responsabilidad y diligencia debida de las empresas proveedoras de servicios en la red, con el fin de que implementasen medidas dirigidas a detectar y eliminar el contenido odioso, así como a asumir políticas más transparentes, prevenir reapariciones, o a configurar tecnología de aprendizaje automático para esta labor, entre otras medidas. Y concluye, por el momento, con la denominada Ley de Servicios Digitales (DSA, por sus siglas en inglés), esto es, el Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales[99] y, con la

(95) Anu Bradford, *The Brussels Effect: How the European Union Rules the World***,** OUP USA, 2020.

(96) Cfr. art.34 de la denominada Ley de Servicios Digitales.

(97) *Ibid.* art, 35. c).

(98) *DOUE L 63* de 6 de marzo de 2018.

(99) *DOUE L 277*, de 27 de octubre de 2022. *Vid*. también, Reglamento de Ejecución (UE) 2023/1201 de la Comisión de 21 de junio de 2023 relativo a las disposiciones detalladas para la tramitación de determinados procedimientos por parte de la Comisión con arreglo al Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo («Ley de Servicios Digitales»). *DOUE L 159*, de 22 de junio de 2023. También se ha de tener en cuenta, el Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo de 14 de septiembre de 2022 sobre mercados disputables y equitativos en el sector digital, *DOUE L 265*, de 12 de octubre de 2022; el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial, *DOUE L 1689*, de 12 de julio de 2024

Directiva (UE) 2024/1760 del Parlamento Europeo y del Consejo de 13 de junio de 2024 sobre diligencia debida de las empresas en materia de sostenibilidad[100].

Esta normativa no solo impone a las empresas[101] –en particular, y en condiciones más estrictas, a las denominadas plataformas en línea de gran tamaño (VLOP) y a los motores de búsqueda en línea de gran tamaño (VLOSE)[102]– medidas para combatir los contenidos ilícitos (noticias falsas, propaganda, incitación al odio, acoso y abuso a menores), reaccionar rápidamente ante ellos y permitir que los usuarios los denuncien, respetando la libertad de expresión, sino que también prevé procedimientos que facultan a la Comisión para acceder a sus bases de datos y algoritmos, así como para obtener explicaciones al respecto cuando sea necesario a fin de garantizar el cumplimiento efectivo de la normativa. Asimismo, prevé mecanismos de control internos y externos: el establecimiento de puntos de contacto; la creación de los denominados «alertadores fiables»[103]; la designación de coordinadores de servicios digitales nacionales que ayuden a la Comisión a supervisar y hacer cumplir las obligaciones previstas en esta normativa[104]; la constitución de una Junta Europea de Servicios Digitales (órgano consultivo independiente compuesta por los coordinadores de servicios digitales de los Estados miembros y presidida por la Comisión Europea); y la creación del Centro Europeo para la Transparencia Algorítmica[105], cuya sede se encuentra en Sevilla.

---

(100) *DOUE L 1760*, de 5 de julio de 2024. La Directiva (UE) 2024/1760 entra en vigor el 25 de julio de 2024 y se prevé un plazo de transposición por los Estados miembros de forma escalonada desde el 26 de julio de 2026 hasta el 26 de julio de 2029.

(101) Con excepción de las pequeñas empresas y microempresas que empleen a menos de 50 personas y con un volumen de negocios anual inferior a 10 millones de euros

(102) Esto es, empresas con una media de más de 1000 empleados y un volumen de negocios mundial neto superior a 450 000000 EUR. Sobre estas empresas *vid*: https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses

(103) Según la Ley de Servicios Digitales, la condición de alertador fiable debe ser otorgada por el coordinador de servicios digitales del Estado miembro donde el solicitante esté establecido y debe ser reconocida por todos los prestadores de plataformas en línea incluidos en el ámbito de aplicación del Reglamento. Esta condición de alertador fiable solo debe otorgarse a entidades, y no personas físicas, que hayan demostrado, entre otras cosas, que poseen conocimientos y competencias específicos para hacer frente a los contenidos ilícitos y que trabajan de manera diligente, precisa y objetiva. Estas entidades pueden ser de carácter público, como por ejemplo, en el caso de los contenidos terroristas, las unidades de notificación de contenidos de internet de las autoridades policiales nacionales o de la Agencia de la Unión Europea para la Cooperación Policial («Europol») o pueden ser organizaciones no gubernamentales y organismos privados o semipúblicos, como las organizaciones que forman parte de la red INHOPE de líneas directas para denunciar materiales relacionados con abusos sexuales a menores y organizaciones comprometidas con la notificación de manifestaciones racistas y xenófobas ilegales en línea. Para que el valor añadido de este mecanismo no se reduzca, debe limitarse el número total de alertadores fiables designados de conformidad con el presente Reglamento. En particular, se alienta a las asociaciones del sector que representen los intereses de sus miembros a solicitar la condición de alertadores fiables, sin perjuicio del derecho de las entidades privadas o las personas físicas a celebrar acuerdos bilaterales con los prestadores de plataformas en línea. Los alertadores fiables deben publicar informes fácilmente comprensibles y detallados sobre las notificaciones enviadas de conformidad con el Reglamento. Dichos informes deben contener información como el número de notificaciones categorizadas por el prestador de servicios de alojamiento de datos, el tipo de contenido y las medidas adoptadas por el prestador.

(104) https://digital-strategy.ec.europa.eu/en/policies/dsa-dscs. En España desempeña esta función Comisión Nacional de los Mercados y la Competencia (CNMC). El 25 de febrero de 2025, el Gobierno español aprobó el anteproyecto de ley que impulsa la gobernanza democrática en servicios digitales y crea el registro de medios de comunicación dependiente de la CNMC.

(105) https://algorithmic-transparency.ec.europa.eu/index_en

Además, los Estados miembros deben prever sanciones disuasorias, proporcionadas y eficaces en caso de incumplimiento de las medidas adoptadas, incluyendo una declaración pública en la que figuren la empresa responsable y la naturaleza de la infracción. Cuando se impongan sanciones pecuniarias, éstas se basarán en el volumen de negocios mundial neto de la empresa sin que el límite máximo de la sanción pecuniaria sea inferior al 5 % del volumen de negocios mundial neto de la empresa en el ejercicio financiero anterior a la decisión sancionatoria. Por otra parte, y con el fin de garantizar que las víctimas de efectos adversos tengan un acceso efectivo a la justicia y a una indemnización, debe exigirse a los Estados miembros que establezcan normas que regulen la responsabilidad civil de las empresas por los daños y perjuicios ocasionados, siempre que la empresa, de forma deliberada o por negligencia, no haya evitado o mitigado los efectos adversos potenciales, o eliminado los efectos reales o minimizado su alcance, y la persona física o jurídica haya sufrido daños y perjuicios como consecuencia de ello.

Es verdad que este exceso de regulación y estos procedimientos de comunicación, requerimiento, valoración y reacción, en particular cuando se producen de forma automática, así como la implantación de filtros masivos y generalizados, pueden tener un efecto contraproducente que, en última instancia, conduzcan a la censura y a una quiebra de la libertad de expresión. Se puede alcanzar, como han indicado algunos autores[(106)], un indeseado efecto *overblocking* y de vigilancia masiva que limite de forma desproporcionada el derecho a la libertad de expresión en contra de los postulados examinados del Derecho internacional. El régimen de responsabilidad de los proveedores de servicios y de otras empresas tecnológicas en función del criterio del conocimiento sobre los datos notificados o por los requerimientos efectuados puede llevar a una profunda perturbación del *ius comunicationis,* tanto desde la perspectiva activa del emisor de expresar libremente sus informaciones, ideas y opiniones como desde la perspectiva pasiva de la colectividad receptora de poder recibirlas y conocerlas. Como ha indicado Alfonso Galán Muñoz, si no se garantizan ambas facetas del acto comunicativo realizado en la red no se podrá alcanzar el pluralismo comunicativo y el contraste de ideas que permita configurar una opinión pública realmente libre[(107)].

Por otra parte, la UE ha adoptado medidas para contrarrestar las *amenazas digitales provenientes del exterior*. Bajo el marco del concepto de *ciberdiplomacia* y el controvertido régimen horizontal de sanciones cibernéticas, que han concluido en la adopción de una singular «cláusula de *cibersolidaridad*»[(108)], se pretende pro-

(106) Alfonso Galán Muñoz, «Redes sociales, discurso terrorista y derecho penal. Entre la prevención, las libertades fundamentales y ¿los negocios?» en Alfonso Galán Muñoz y Carmen Gómez Rivero, *La Represión y Persecución Penal del Discurso Terrorista*, Tirant lo Blanch, Valencia, 2022, pp.302-305.

(107) *Ibid.*, p.297.

(108) Cfr. Reglamento (UE) 2025/38 del Parlamento Europeo y del Consejo, de 19 de diciembre de 2024, por el que se establecen medidas destinadas a reforzar la solidaridad y las capacidades en la Unión a fin de detectar ciberamenazas e incidentes, prepararse y responder a ellos y por el que se modifica el Reglamento (UE) 2021/694 (*Reglamento de Cibersolidaridad*), *DOUE L 38*, de 15 de enero de 2025. Véase también, Decisión (PESC) 2019/797 del Consejo relativa a medidas restrictivas contra los ciberataques que amenacen a la Unión o a sus Estados miembros; Las nuevas directrices en ciberdiplomacia emitidas el 8 de junio de 2023 (disponible en https://data.consilium.europa.eu/doc/document/ST-10289-2023-INIT/en/pdf) que representan un paso importante hacia un dominio cibernético más seguro y estable. Estas directrices buscan abordar la creciente voluntad y capacidad de

teger las democracias de los Estados miembros frente a las amenazas y los efectos nocivos de la desinformación, la manipulación de la información y la injerencia, en particular, por parte de agentes extranjeros. Este objetivo se ha convertido en una prioridad estratégica para la UE, como queda de manifiesto en el Plan de Acción para la Democracia Europea (PADE)[109], y en la creación de la red de la Comisión contra la desinformación y del sistema de alerta rápida del Servicio Europeo de Acción Exterior, que conectan los centros de referencia en los Estados miembros y las instituciones de la UE, así como, en contextos electorales, de la Red Europea de Cooperación Electoral. El propio Reglamento de Cibersolidaridad (2025) establece que «debe crearse un Mecanismo de Emergencia en materia de Ciberseguridad para ayudar a los Estados miembros, previa solicitud de estos, a prepararse, responder a incidentes de ciberseguridad significativos y a gran escala, atenuar sus repercusiones e iniciar la recuperación de los incidentes de ciberseguridad significativos y a gran escala, y apoyar a otros usuarios en su respuesta a incidentes de ciberseguridad significativos y a incidentes de ciberseguridad equivalentes a gran escala; y debe crearse un Mecanismo Europeo de Revisión de Incidentes de Ciberseguridad para revisar y evaluar incidentes de ciberseguridad significativos o a gran escala específicos». Las acciones previstas en este Reglamento deben llevarse a cabo respetando debidamente las competencias de los Estados miembros y deben complementar, sin duplicar, las actividades que llevan a cabo la red de CSIRT[110], la Red europea de organizaciones de enlace para la gestión de cibercrisis (EU-CyCLONe, por sus siglas en inglés) o el Grupo de Cooperación establecido en virtud de la Directiva (UE) 2022/2555 (conocida como SRI-2) relativa a las medidas destinadas a garantizar un elevado nivel común de ciberseguridad en toda la Unión.

Una manifestación de estas medidas ha consistido en la adopción de medidas restrictivas del artículo 215 TFUE que afectan a determinados medios de comunicación. En concreto, mediante la Decisión 2014/512/PESC del Consejo, de 31 de julio de 2014, relativa a medidas restrictivas motivadas por acciones de Rusia que desestabilizan la situación en Ucrania[111], se decidió suspender cualquier licencia o autorización de radiodifusión, acuerdo de transmisión y distribución celebrado con las personas jurídicas, entidades u organismos enumerados en el anexo IX de la misma. Afectadas por esta Decisión, fueron la agencia de noticias estatal rusa

---

actores estatales y no estatales para perseguir sus objetivos estratégicos a través de actividades cibernéticas maliciosas. Bruselas. 6.12.2023. JOIN(2023) 51 final; y la Comunicación conjunta al Parlamento Europeo y al Consejo: Sin lugar para el odio: una Europa unida contra esta lacra (JOIN/2023/51 final).

(109) COM(2020) 790.

(110) La red de CSIRT de la Unión Europea está formada por los CSIRT designados por los Estados miembros de la UE y por el CERT-EU («miembros de la red de CSIRT»). La Comisión Europea participa en la red en calidad de observador. La red de CSIRT de la UE se creó en 2016 mediante un conjunto de normas de ciberseguridad de la UE conocidas comúnmente como Directiva NIS. En 2023, la red de CSIRT se vio reforzada aún más por la Directiva NIS2 con el fin de contribuir al desarrollo de la confianza y promover una cooperación operativa rápida y eficaz entre los Estados miembros. ENISA (Agencia de la Unión Europea para la Ciberseguridad) está impulsando la red de CSIRT proporcionando el equipo de la Secretaría, las infraestructuras y las herramientas para permitir una cooperación eficaz y un funcionamiento diario continuo y el intercambio de información. Para más información, https://csirtsnetwork.eu/.

(111) *DOUE, L 229* en su versión modificada por la Decisión (PESC) 2022/327 del Consejo, de 25 de febrero de 2022 (*DOUE, L 48*) y Decisión (PESC) 2022/351 del Consejo de 1 de marzo de 2022 (*DOUE L 65* de 2 de marzo de 2022).

Sputnik y la cadena de televisión *RT France*. Esta última presentó recurso de anulación ante el Tribunal General de la UE alegando, entre otras cuestiones, la violación del artículo 11 de la Carta de Derechos Fundamentales de la Unión sobre la libertad de expresión y de información(112).

En su sentencia de 27 de julio de 2022, la Gran Sala del Tribunal General convalidó dicha suspensión aplicando el *test de Estrasburgo* (legalidad, necesidad democrática y proporcionalidad) y fundamentando su decisión en la situación extraordinaria y de extrema urgencia, «casi-bélica» o de guerra híbrida, en la que se encuentra la UE tras la invasión el 24 de febrero de 2022 de Ucrania por parte de Rusia, en particular, por el peligro que supone el mantenimiento de esta agresión para las fronteras de la Unión(113). Dada la peligrosidad de los contenidos emitidos y la estrecha vinculación de este canal con el Gobierno ruso, se entendió que *RT France* no podía invocar la «protección reforzada» que le otorga la libertad de expresión reconocida en el artículo 11 de la Carta de derechos fundamentales de la UE. No en vano, se retiene el argumento de que la empresa de comunicación sancionada se encontraba bajo el control permanente, directo o indirecto, de los dirigentes de la Federación de Rusia. Incluso, en este contexto y en función de la campaña de propaganda a favor de la agresión rusa realizada por este canal, se ha comparado su actividad, en esencia, con la del Ministerio de Defensa ruso(114). Para reforzar tal conclusión, se invoca el artículo 20 del Pacto internacional de derechos civiles y políticos de las Naciones Unidas que dispone que «toda propaganda en favor de la guerra estará prohibida por la ley» precisando que el alcance de esta prohibición incluye no solo la incitación a una guerra futura, «sino también las declaraciones expresadas de forma continuada, reiterada y concertada en favor de una guerra en curso contraria al Derecho internacional, en particular, si emanan un medio de comunicación que se encuentra bajo el control, directo o indirecto, del Estado agresor»(115).

Además de los problemas que plantea la naturaleza política de las medidas restrictivas para imponer limitaciones a los derechos fundamentales(116) (¿la suspensión de un medio de comunicación sin orden judicial o de otra autoridad independiente y sin un procedimiento garantista?(117)), se ha discutido hasta qué punto

---

(112) STJUE de 27 de septiembre de 2022, asunto *RT Francia c. Consejo,* caso T-125/22, ECLI:EU:T:2022:483. Véase también el asunto *Kiselev c. Consejo*, Sentencia del Tribunal General (Sala Novena) de 15 de junio de 2017, ECLI:EU:T:2017:392.

(113) Cfr. par.98 de la Sentencia *RT Francia c. Consejo.*

(114) Cfr. pars. 71-71.

(115) Cfr. par. 210.

(116) Cfr. Francisco Jiménez García, «Medidas restrictivas en la Unión Europea: entre las sanciones y el unilateralismo europeo» en Carmen Martínez Capdevila y Enrique J. Mártinez Pérez (Dirs.), *Retos para la Acción Exterior de la Unión Europea*, Tirant lo Blanc, Valencia, 2017, pp.509-534.

(117) Como han indicado Lorena Tristante Pellicer y Germán M. Teruel Lozano, «debe destacarse que la medida fue adoptada sin un procedimiento mínimamente garantista, algo en lo que ha insistido el Tribunal de Estrasburgo para valorar la legitimidad de injerencias de esta severidad. En este caso, la prohibición en cuestión fue adoptada sin existir un deber normativo específico de neutralidad editorial; con una base jurídica precaria, como ya se analizó, que, como mucho, ofrecía una base competencial pero no cumplía con las exigencias mínimas del principio de seguridad jurídica en tanto que no definía ni infracciones ni consecuencias jurídicas. Para colmo, la decisión es adoptada por un órgano político y no por una autoridad independiente; y la restricción se acuerda *inaudita* parte y con una sucinta motivación, por mucho que el Tribunal General haya terminado considerando que, dadas

la medida de suspensión de toda actividad (prohibición absoluta de difundir cualquier contenido a través de cualquier medio) a esta empresa puede considerarse una medida proporcional que no atenta contra la esencia del derecho a la libertad de expresión en el orden público europeo garante del pluralismo político. Algunos autores han destacado que, en primer lugar, la UE no está en guerra con Rusia y Ucrania no es un Estado miembro de la UE y, en segundo lugar, han calificado estas medidas de «un agudo paternalismo informativo» bajo la premisa de que los ciudadanos de la UE no están en condiciones de analizar críticamente esa propaganda, teniendo acceso a una amplia gama de medios de comunicación (en línea) y a diferentes canales de información periodística. La UE también ignora que combatir la desinformación con censura es un error pues el verdadero antídoto contra la desinformación no es la prohibición de los medios de comunicación, «sino la promoción de un ecosistema mediático vibrante, pluralista, profesional, ético y viable, totalmente independiente de quienes detentan el poder»[(118)].

Incluso tal decisión podría contravenir la doctrina del TEDH que ha reiterado que:

> «No puede haber democracia sin pluralismo. La democracia se nutre de la libertad de expresión. Es esencial para la democracia permitir que se propongan y debatan programas políticos diversos, incluso aquellos que pongan en tela de juicio la organización actual de un Estado, siempre que no perjudiquen la democracia. Para garantizar un verdadero pluralismo en el sector audiovisual en una sociedad democrática, no basta con prever la existencia de varios canales o la posibilidad teórica de que los operadores potenciales accedan al mercado audiovisual. Además, es necesario permitir un acceso efectivo al mercado para garantizar la diversidad del contenido global de los programas, reflejando en la medida de lo posible la diversidad de opiniones encontradas en la sociedad a la que se dirigen los programas (….) Si bien la libertad de expresión viene acompañada de excepciones, éstas deben interpretarse restrictivamente, y la necesidad de restringirla tendrá que acreditarse de modo convincente»[(119)].

Asimismo, el TEDH ha afirmado que una suspensión temporal no vulnera el artículo 10 cuando se permiten otras alternativas para la difusión[(120)] y que es preferible una limitación, en lugar de bloquear por completo una página web, medida que convierte en inaccesible una gran cantidad de información afectando a los derechos de numerosos usuarios y provocando un efecto colateral importante[(121)].

---

las circunstancias, no se violó el derecho a la defensa del canal afectado» «Desinformación y libertad de expresión: el bloqueo europeo de canales rusos ante la invasión de Ucrania a la luz de la Sentencia del Tribunal General (Gran Sala) de 27 de julio de 2022, T-125/22, *RT France c. Consejo de la UE»* 71 *Estudios de Deusto,* 2023-2, pp. 324-325.

(118) Cfr. Dirk Voorhoof, «EU silences Russian state media: a step in the wrong direction», 8 de mayo de 2022, en *INFORRMS BLOG*, 2022, https://inforrm.org/2022/05/08/eu-silences-russian-state-media-a-step-in-the-wrong-direction-dirk-voorhoof/; Ronan Ó Fathaigh y Dirk Voorhoof, Dirk, «Freedom of expression and the EU's ban on Russia today: A dangerous Rubicon crossed», 27 *Communications Law,* 2022-4, pp. 186-193. Puede consultarse en SSRN: https://ssrn.com/abstract=4322452 o http://dx.doi.org/10.2139/ssrn.4322452

(119) Cfr. STEDH de 5 de abril de 2022, asunto *NIT SRL c. República de Moldavia*, demanda n° 28470/12 , par.185.

(120) Cfr. *ibid.* par.223.

(121) Cfr. STEDH de 18 de diciembre de 2012, asunto *Ahmet Yildirim c. Turquía*, demanda n.º 3111/10, par.66.

Según el Tribunal de Estrasburgo, el bloqueo total del acceso a un sitio web es una medida comparable a la prohibición de un periódico o una emisora de televisión, pues esta medida amplía el alcance del bloqueo no solo al contenido ilegal original, sino también a cualquier tipo de información presente en la web, constituyendo entonces una interferencia arbitraria(122).

Frente a este efecto reglamentario, de diligencia debida y sancionador de Bruselas, las Naciones Unidas han vuelto a apostar por el *soft-law*, las recomendaciones y los códigos de conducta en el marco programático de la Agenda 2030 y sus ODS. En el *Pacto para el Futuro,* aprobado el 22 de septiembre de 2024, los líderes mundiales decidían integrar la perspectiva de los derechos humanos en los procesos regulatorios y normativos de las tecnologías nuevas y emergentes y exhortar al sector privado a que respete los derechos humanos y defienda los principios éticos en el desarrollo y el uso de estas tecnologías(123). Por su parte, el Pacto Global Digital que aparece anexo al anterior y bajo la rúbrica *«Confianza y seguridad digitales»*, proclama que se deben contrarrestar y abordar urgentemente todas las formas de violencia, incluida la violencia sexual y de género, que se producen o amplifican por el uso de la tecnología, todas las formas de discurso de odio y discriminación, las informaciones erróneas y la desinformación, el ciberacoso y la explotación y los abusos sexuales de menores. Y, a tal efecto, los Estados se comprometen a establecer y mantener estrictas medidas de mitigación de riesgos y vías de recurso que también protejan la privacidad y la libertad de expresión(124). Se estima necesario fomentar un espacio digital inclusivo, abierto y seguro que respete, proteja y promueva los derechos humanos, estableciendo salvaguardias adecuadas, incluso ejerciendo la diligencia debida y creando mecanismos eficaces de supervisión y recurso(125).

Igualmente, los líderes mundiales se comprometen a implantar planes de estudio para la *alfabetización mediática e informativa* a fin de que todos los usuarios tengan las destrezas y los conocimientos necesarios para interactuar sin riesgo y con espíritu crítico con los contenidos y con los proveedores de información, y para mejorar la resiliencia frente a los efectos nocivos de las informaciones engañosas y la desinformación(126). Con relación al ámbito empresarial, de nuevo se exhorta a tomar medidas eficaces conformes a los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas y otros marcos pertinentes(127), entre ellas, la incorporación de salvaguardias en los procesos de entrenamiento de los modelos de inteligencia artificial, la identificación del material generado por la inteligencia artificial y la certificación de la autenticidad de los contenidos y su origen, así como etiquetas, marcas de agua digitales y otras técnicas(128).

---

(122) Cfr. SSTEDH, asunto *Vladimir Kharitonov c. Rusia*, demanda n.º 10795/14 de 23 de junio de 2020. pars. 37-38 y asunto *OOO Flavus y otros c. Rusia*, demanda *nos. 12468/15*, 23489/15 y 19074/16 de 23 de junio de 2020, pars. 36-43.

(123) Doc. UN A/RES/79/1, par.54

(124) Par.30 del Pacto Global Digital.

(125) *Ibid.* par. 23.

(126) *Ibid.* par. 35

(127) *Ibid.* par. 25.

(128) *Ibid.* par. 36 c).

V. **REFLEXIÓN FINAL**

Ante el dilema del problema de los tres cuerpos, como hemos visto, no existe una solución unívoca ni universal. Frente a la regulación normativa internacional de la libertad de expresión y sus estrictas excepciones –incluida la incitación a la discriminación, el hostigamiento y la violencia– como piedra angular para una sociedad democrática plural y un desarrollo sostenible de los derechos y objetivos del nuevo milenio, la regulación del discurso y de los delitos de odio se mueve particularmente en el mundo del *soft-law* y los códigos de conducta. Como ha declarado el secretario general de las Naciones Unidas, un uso indebido del concepto indefinido del discurso de odio puede no solo perjudicar a sus víctimas, sino también silenciar a algunas de las personas que están en mejores condiciones de contrarrestar los argumentos de odio: los defensores de los derechos humanos y los periodistas. Una verdadera libertad de expresión resulta necesaria para conjurar los peligros de las posiciones dominantes de los nuevos señores feudales tecnológicos y de la vigilancia digital selectiva y panóptica de los poderes estatales.

Pero también es verdad que el discurso de odio y la desinformación constituyen una señal de alarma de graves amenazas, incluidas las bélicas y genocidas, que no siempre es percibida por la sociedad ni sus gobernantes ni contrarrestada con la debida prioridad. El soporte digital y las posibilidades que ahora ofrecen las TIC y la IA, hacen que el medio sea tan peligroso como el contenido de estos odiosos comportamientos pues, como concluye Byung-Chul Han, «la verdad se desintegra en polvo informativo arrastrado por el viento digital». Odio y desinformación anteceden y promueven la violencia y la intolerancia y se han convertido en una de las formas más habituales de extender una retórica divisoria a escala global, poniendo en peligro la paz mundial. El mayor éxito de estos fenómenos sería el debilitamiento del derecho a la libertad de expresión pues significaría la quiebra de uno de los principales logros del Derecho internacional contemporáneo. La criminalización debe reservarse para los casos más graves de incitación, mientras que la fortaleza de las sociedades democráticas se ha de medir por su capacidad para adoptar medidas e instrumentos de distinta índole –incluidas acciones educativas y programas de alfabetización mediática e informativa– que permitan absorber y neutralizar otras manifestaciones de los discursos de odio («sumideros de odio»). En este proceso, resultará necesario establecer reglas para proteger, respetar y remediar los derechos fundamentales, incluida una normativa de debida diligencia y sobre responsabilidad de los operadores privados sin incurrir en un reglamentarismo excesivo y paralizante.

VI. **BIBLIOGRAFÍA**

Víctor Abramovich, «Dilemas jurídicos en la restricción de los discursos de odio» *Sur - Revista Internacional de Derechos Humanos*, vol. 19, núm. 32, 2022, pp.87-99.

Aram Aharonian, «El nuevo feudalismo digital», 18 de octubre de 2023 en https://esfera-comunicacional.ar/el-nuevo-feudalismo-digital

Rafael Alcácer Guirao, «Víctimas y disidentes. El “discurso del odio” en EE. UU. y Europa». *Revista Española de Derecho Constitucional*, 2015-1, pp.45-86

Libia Arenal Lora, «Limitaciones y alcance de la responsabilidad de las empresas proveedoras de servicios en el discurso de odio *online*. El caso de Meta en la incitación al genocidio rohingya», *Cuadernos de Derecho Transnacional* 15, 2023-2, pp. 141-166.

Hannah Arendt, *Los Orígenes del Totalitarismo*, traducción de Guillermo Solana, Alianza Editorial, 2014.

Beatriz Barreiro Carril, «Los derechos humanos de los migrantes ante el discurso de odio: Encaje de las normas y estándares internacionales en España», *Revista Internacional de los Estudios Vascos*, 69, 2024-1 (http://doi.org/10.61879/riev691zkia202401).

Ramón Blecua, «Feudalismo digital en un orden inestable», *Política Exterior*, vol. 34, Núm. 196, 2020, pp. 146-153.

Anu Bradford, *The Brussels Effect: How the European Union Rules the World***,** OUP USA, 2020

Valerie C. Brannon y Eric N. Holmes, «*S*ection 230: An Overview», *Congressional Research Service,* 4 de enero de 2024, en: https://crsreports.congress.gov/product/pdf/R/R46751

Josep Burgaya, *La Manada Digital. Feudalismo Hipertecnológico en una Democracia sin Ciudadanos*, El Viejo Topo, Barcelona, 2021.

Juana del Carpio Delgado, «Discurso de odio en el Derecho penal internacional: su consideración, o no, de persecución como crimen de lesa humanidad» en Eulalia W. Petit de Gabriel (Dir.), *Valores (y Temores) del Estado de Derecho: Libertad de Expresión vs. Delitos de Opinión en Derecho Internacional***,** Aranzadi, 2023, pp.73-109.

Juan Alberto Díaz López, *Informe de Delimitación Conceptual en Materia de Delitos de Odio* Comisión de Seguimiento del Acuerdo para cooperar institucionalmente en la lucha contra el racismo, la xenofobia, la LGBTIfobia y otras formas de intolerancia, Secretaría de Estado de Migraciones del Ministerio de Inclusión, Seguridad Social y Migraciones, Madrid, 2020.

Alfonso Galán Muñoz, «Redes sociales, discurso terrorista y derecho penal. Entre la prevención, las libertades fundamentales y ¿los negocios?» en Alfonso Galán Muñoz y Carmen Gómez Rivero, *La Represión y Persecución Penal del Discurso Terrorista*, Tirant lo Blanch, Valencia, 2022, pp. 255-309.

Héctor Gómez Peralta, «El discurso de odio y los límites de la libertad de expresión en Estados Unidos», *Revista Mexicana de Ciencias Políticas y Sociales,* vol. 68, 2023, en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-19182023000300281.

Iñigo Gordon Benito, «Ciberodio. Un estudio de derecho penal comparado», *Cuadernos de RES PUBLICA en Derecho y Criminología,* 2024-4, pp. 14-35 (https://doi.org/10.46661/respublica.9398

Sarthak Gupta, «Context, Content, and the 'Threshold of Severity': ECtHR's Jurisprudence on Satire vs Hate», *EJIL:Talk! Blog of the European Journal of International Law*, March 27, 2025.

Víctor Luis Gutiérrez Castillo, «El control europeo del ciberespacio ante el discurso de odio: análisis de las medidas de lucha y prevención», *Araucaria. Revista Iberoamericana de Filosofía, Política, Humanidades y Relaciones Internacionales*, núm. 45, 2020, pp. 291-310.

Byung-Chul Han, *Infocracia*, traducción Joaquín Chamorro Mielke, Taurus, Madrid, 2023.

Francisco Jiménez García, *Conflictos Armados y Derecho Internacional Humanitario*, Ommpress, Madrid, 2023.

Francisco Jiménez García, «Medidas restrictivas en la Unión Europea: entre las sanciones y el unilateralismo europeo», en Carmen Martínez Capdevila y Enrique J. Mártinez Pérez (Dirs.), *Retos para la Acción Exterior de la Unión Europea*, Tirant lo Blanc, Valencia, 2017, pp. 509-534

Jon-Mirena Landa Gorostiza, «Delitos de odio y estándares internacionales: una visión crítica a contracorriente», *Revista Electrónica de Ciencia Penal y Criminología*, núm. 22-19, 2020.

José León Alanpont, «La inteligencia artificial al servicio de la desinformación: un nuevo reto para el derecho penal», en Francisco Jiménez García (Dir.), Sandra López de Zubiría Díaz y Berta Alam Pérez (Coords.), *Seguridad y Responsabilidad Penal e Internacional en el Uso de las TIC y la Inteligencia Artificial,* Iustel, Madrid, 2024, pp. 95-127.

Renato Lopes Militão, «O discurso de ódio no sistema da Convenção Europeia dos Direitos Humanos» en Susana Almeida y Andrés Rousset, *Os Sistemas Europeu e Interamericano de Proteção dos Direitos Humanos: Uma leitura Comparada,* Aranzadi, 2024, pp. 957-981

Daniel Martínez Cristóbal, «The current perspective on sharp power: China and Russia in the era of (dis)information», *Revista Electrónica de Estudios Internacionales*, núm. 42, 2021.

Joseph S. Nye JR., «How sharp power threatens soft power. The right and wrong ways to respond to authoritarian influence», *Foreign Affairs*, 24 de enero de 2018.

Ronan Ó Fathaigh y Dirk Voorhoof, «Freedom of expression and the EU's ban on Russia today: A dangerous Rubicon crossed», 27 *Communications Law,* 2022-4, pp. 186-193, SSRN: https://ssrn.com/abstract=4322452 o http://dx.doi.org/10.2139/ssrn.4322452

Eulalia W. Petit de Gabriel, «Libertad de expresión y delitos de opinión a la luz de la futura Convención Internacional Integral sobre la Lucha contra la Utilización de las Tecnologías de la Información y las Comunicaciones con Fines Delictivos» en Juana del Carpio Delgado y María Holgado González (Dirs.) y Alejandro L. de Pablo Serrano (Coord.), L*a Libertad de Expresión Asediada. Delitos de Odio, Delitos de Opinión, Censuras de Gobiernos y Empresas*, Aranzadi, 2023, pp. 405-454.

Carmen Quessada Alcalá, «El discurso de odio *on line* y la jurisprudencia del Tribunal Europeo de Derechos Humanos en materia de libertad de expresión: ¿diferentes vías, mismos límites?», en Susana Almeida y Andrés Rousset, *Os Sistemas Europeu e Interamericano de Proteção dos Direitos Humanos: Uma leitura Comparada,* Aranzadi, 2024, pp. 943-956.

José Ramírez, «Navegando en la era del Feudalismo Digital: Blockchain como escudo para la soberanía de datos», *In Solidm*, 2023-1.

Göran Rollnert Liern, «El discurso del odio: una lectura crítica de la regulación internacional». *Revista Española de Derecho Constitucional*, núm. 115, 2019, Pp. 81-109 (https://doi.org/10.18042/cepc/redc.115.03).

Göran Rollnert Liern, «Redes sociales y discurso del odio: perspectiva internacional», *Revista de los Estudios de Derecho y Ciencia Política*, núm. 31, 2020, https://roderic.uv.es/rest/api/core/bitstreams/aa61a063-beb9-460a-ac4c-325d678e211e/content

Lorena Tristante Pellicer y Germán M. Teruel Lozano, «Desinformación y libertad de expresión: el bloqueo europeo de canales rusos ante la invasión de Ucrania a la luz de la Sentencia del Tribunal General (Gran Sala) de 27 de julio de 2022, T-125/22, *RT France c. Consejo de la UE»* 71, *Estudios de Deusto,* 2023-2, pp. 303-329.

Fernando Villaspín Oña, «Las principales amenazas a la democracia liberal», *Anales de la Real Academia de Ciencias Morales y Políticas (2019-2020)* pp. 327-344.

Dirk Voorhoof, «EU silences Russian state media: a step in the wrong direction», 8 de mayo de 2022, en *INFORRMS BLOG*, 2022, https://inforrm.org/2022/05/08/eu-silences-russian-state-media-a-step-in-the-wrong-direction-dirk-voorhoof/.

Christopher WALKER y Jessica LUDWIG, «The meaning of sharp power. How authoritarian states project influence», *Foreign Affairs*, 16 de noviembre de 2017.

Christopher WALKER y Jessica LUDWIG, «The long arm of the strongman. How China and Russia use sharp power to threaten democracies», *Foreign Affairs*, 12 de mayo de 2021.

Sun YIRONG, «The future of due diligence in cyberspace», *New York University Journal of International Law and Politics*, Vol. 54, 2022-2, pp.753-764. SSRN: https://ssrn.com/abstract=4629508.