

**El rol de las plataformas sociales en la incitación al odio y los riesgos que esto puede generar sobre la comisión del delito internacional de genocidio**

**Autora:**

**Elisa Delgado Forsten**

**Trabajo para optar al título de Abogada**

**Asesor:**

**Jose Alberto Toro Valencia**

**Universidad EAFIT**

**Escuela de Derecho**

**Diciembre 2021**

## **Resumen**

Los discursos de odio no son un fenómeno novedoso. Han estado presentes a lo largo de la historia del hombre. Sin embargo, al igual que sus emisores, son fenómenos que han evolucionado y se han perfeccionado, adaptándose a las nuevas formas de comunicación. Con la creación de las plataformas tecnológicas, especialmente de las plataformas sociales, los discursos y mensajes de odio han podido masificarse y globalizarse, llegando así a una gran cantidad de personas. Debido a lo anterior, está claro que los mensajes odiosos tienen una mayor capacidad de afectar a muchas más personas, ya que es un fenómeno que puede convertirse en viral. En este sentido, se han encontrado casos, como el de los Rohingya en Myanmar, en los cuales se utiliza a las redes sociales para orquestar persecuciones y hasta genocidios contra poblaciones que ostentan la categoría de especialmente protegidas.

La regulación internacional actual sobre dichos casos es aún precaria y pueden encontrarse muchas lagunas en cuanto a la protección de las víctimas. De igual manera, las mismas plataformas sociales han venido dándose sus propias regulaciones en cuanto a estos casos, sin embargo, no hay garantías suficientes que indiquen que dichas plataformas harán efectivo el cumplimiento obligatorio de sus propias normas en todos los casos por igual.

**Palabras clave:** Discurso de odio, plataformas tecnológicas, genocidio, población protegida, regulación.

## **Abstract**

The proliferation of hate speech is not a new trend, it has been present throughout history. However, it has been a phenomenon that has evolved and developed, adapting itself to the new ways of communication and interaction between people. With the emergence of the technological platforms, especially the social networks, the speeches and hateful messages are getting massive and global, being able to reach a large amount of people. That is why, it is clear that hatefull messages can affect a larger amount of people around the world and can become viral. In this sense, there are special cases, like the case of the Rohingya in Myanmar, in wich social networks like Facebook have been used to gather people who are willing to attack groups of protected people or intend to commit a genocide.

The actual international regulation about these cases is still very precarious and there are a lot of legal holes as for the victims of hate speech. In the same way, social networks are making their own regulations, nevertheless, this does not mean that the social network is going to make them mandatory for each case equally.

**Keywords:** Hate speech, technological platforms, genocide, protected groups, regulation.

## Índice

1. Introducción.....	5
2. El discurso de odio: su construcción y uso en el marco de las plataformas tecnológicas.....	9
3. Los nuevos transmisores: las plataformas sociales de base tecnológica como propagadores virales de mensajes.....	27
4. Plataformas sociales, discurso de odio y su impacto en la comisión del delito internacional de genocidio.....	39
4.1. El fenómeno del genocidio.....	40
4.2. Los antecedentes con el genocidio en Ruanda.....	43
4.3. El riesgo del discurso de odio en la práctica: El caso de los Rohingya en Myanmar.....	46
5. Retos frente a la regulación del discurso de odio difundido mediante plataformas sociales.....	59
5.1. Regulación Internacional.....	60
5.2. Autorregulación.....	66
6. Conclusión.....	72
7. Referencias.....	76

## **1. Introducción**

Las plataformas sociales digitales tales como Facebook, Instagram, Twitter, WhatsApp, entre otras, se han venido convirtiendo en medios importantes de distribución de información y, a su vez, se han instituido como grandes recolectores de datos. Lo anterior debido a que son lugares comunes donde se hace posible la masificación de mensajes que tienen un alcance global, esto significa que, mediante estas plataformas sociales, se ha creado un nuevo sistema donde los mensajes, tanto políticos como sociales, tienen la posibilidad de impactar en una mayor cantidad de la población mundial. Si bien es positivo para la globalización que mediante estas plataformas sociales se proporcionen mecanismos de interconexión que la favorezcan, se debe tener en cuenta también el contenido disruptivo que estas pueden generar.

La presente investigación se ocupa de analizar la dinámica del discurso de odio en las plataformas sociales con el fin de corroborar la manera en la que éste influye sobre la comisión del delito internacional de genocidio. Para esto, se desarrollará el concepto del discurso de odio, se proporcionará una definición de lo que puede considerarse como tal, enfocándose en el desarrollo de esta clase de discursos cuando se presentan en redes sociales y de manera masiva. Principalmente, se trabajará con la definición que clasifica el discurso de odio como los ataques directos contra alguno o varios de los grupos protegidos en razón a su raza, etnia, nacionalidad, religión, género, orientación sexual, discapacidad o enfermedad y que generalmente contienen expresiones insultantes, abusivas o con una clara intención de hacer daño a sus receptores o víctimas directas. Posteriormente, se definirá y analizará la forma de funcionamiento de las plataformas sociales. Para claridad técnica y teórica, se señala que, en este trabajo, se utilizará la palabra plataforma refiriéndose a las aplicaciones y servicios en línea que se sirven del análisis e interpretación de datos. Es decir,

se valen de la facilitación de la infraestructura necesaria para intermediar entre diferentes grupos de usuarios, desplegar tendencias monopólicas impulsadas por efectos de red y tener una arquitectura central establecida que controla las posibilidades de interacción.

Otro de los objetivos del presente trabajo es indagar sobre las posibles consecuencias de la difusión de mensajes de odio en plataformas sociales, en cuanto a la incitación a la comisión del delito internacional de genocidio, entendiéndolo cómo la intención de destruir ciertos grupos sociales o un grupo específico de personas. Esto teniendo en cuenta el proceso que se dio con el genocidio de Ruanda, donde fue utilizada toda una campaña de odio y desprestigio hacia los Tutsi a través de la radio y de medios escritos. Este caso se toma como antecedente y revela los posibles peligros de la difusión de odio y discursos violentos contra poblaciones protegidas mediante las plataformas sociales. Finalmente, se indagará sobre los posibles retos regulatorios que se presentan actualmente en relación con el control del discurso de odio en las plataformas sociales y cómo se ha tratado de solucionar el problema de la circulación del odio en línea.

Para la presente investigación, se empleará una metodología cualitativa. Esto quiere decir que se busca reconocer la diversidad de conocimientos, posturas y posiciones sobre la realidad del tema en específico mediante un proceso de indagación inductivo. Se pretende construir un concepto sobre una realidad social a partir del estudio del caso, en esta ocasión, de las implicaciones que tuvo la plataforma social Facebook sobre la comunidad Rohingya en Myanmar. Este caso presentó especial interés para la presente investigación pues se tomó como un ejemplo de la posible injerencia de las redes sociales sobre la comisión de crímenes internacionales. En esta situación en particular, se ha demostrado que la red social Facebook jugó un papel importante dentro de la difusión de mensajes de odio en contra de esta

población y la orquestación, a través de la creación de grupos, de ataques contra los Rohingya.

La plataforma social Facebook, ha sido una parte importante de la vida tanto privada como civil y pública de los habitantes de Myanmar y, además, el gobierno utiliza esta plataforma para diseminar información oficial. El problema radica en que Facebook no contó con un sistema de respuesta temprana y efectiva para detectar los mensajes de odio que estaban siendo publicados y que estaban repercutiendo sobre la matanza de los Rohingya. En este sentido, se partirá de la investigación documental, reuniendo y analizando diferente bibliografía, es decir, libros, artículos de revistas, podcasts, etc. sobre el tema de investigación y recopilando información que permita interpretar el fenómeno del discurso de odio y su conexión con el delito internacional de genocidio.

Como producto de la globalización, cada vez vivimos en un mundo más interconectado, donde se facilita la comunicación entre personas en diferentes lugares del mundo en tiempo presente. Pero no solo eso, también se da un fenómeno muy particular y es el de la capacidad que tienen los mensajes de masificarse y alcanzar niveles globales en solo cuestión de minutos. Teniendo en cuenta lo anterior, se plantea la discusión sobre si en realidad somos conscientes de lo que pueden causar nuestros mensajes, ya que estos podrían crear un efecto mariposa, es decir, que nuestras acciones en línea pueden repercutir gravemente en otros sitios.

Es precisamente ese efecto el que tienen muchos mensajes de odio que se publican en redes sociales. En la actualidad, hemos presenciado como, por ejemplo, los mensajes publicados por el expresidente Trump en Estados Unidos, desataron actos por parte de civiles seguidores de sus políticas hasta el punto de tomarse el Capitolio de dicho Estado en el año 2021.

Es así como, a medida que el uso de redes sociales va aumentando y se propicia una interacción libre entre las personas, se presenta también una proliferación de mensajes de odio, racistas o que incitan a la discriminación hacia grupos particulares e internacionalmente protegidos. Muchas veces, estos mensajes tienen un impacto profundo en los receptores, lo cual puede desencadenar en la comisión de delitos internacionales.

En este sentido, se pretende determinar si la difusión de estos mensajes de odio podría recaer, y en efecto muchas veces lo hace, sobre grupos de comunidades específicamente determinados y protegidos internacionalmente y si esto puede llegar a generar tal magnitud de odio que se cree un objetivo colectivo de eliminar al grupo específico en razón a los mensajes difundidos por las redes sociales. Es decir, si estos mensajes de odio pueden tener la potencial repercusión de la incitación al genocidio.

## **2. El discurso de odio: su construcción y uso en el marco de las plataformas tecnológicas**

Según La Real Academia de la Lengua Española, la palabra odio se define de la siguiente manera: “Antipatía y aversión hacia algo o hacia alguien cuyo mal se desea.” A pesar de que dicha definición considera el odio como un simple deseo de hacer el mal, la historia y el propio comportamiento humano nos han demostrado que el odio, muchas veces, puede conllevar a una acción que puede causar gran daño sobre la persona o cosa sobre la que recae. Tal como lo expone Bonnet (2019), el odio es un sentimiento que puede definirse como una “pasión sorda” la cual puede durar años o toda una vida ya que no se clasifica como un sentimiento pasajero, sino algo que se cultiva con el tiempo. (p. 178)

Siguiendo la línea argumentativa de Bonnet (2019), “el odio, como todas las emociones y sentimientos, parte de un juicio, entraña una manera de ver, se asienta en creencias.” (p.179). Esto quiere decir que es un sentimiento que, a diferencia de la ira, no tiende a aparecer sorpresivamente, sino que proviene de construcciones tanto personales como colectivas que son difíciles de deconstruir. Normalmente, se odia lo que se percibe como una amenaza, lo que se teme o lo que simplemente se disgusta, pero no se puede negar que, igualmente, puede existir una clase de odio irracional que no obedece a ninguna razón material.

En este sentido, el odio se presenta como una especie de sentimiento fabricado, no es algo que exista sin más o que sea inherente al ser humano, se clasifica, más bien, como una construcción humana y colectiva. Según Emcke (2017), existen ciertos aspectos característicos del odio. En primer lugar, el odio requiere de una certeza absoluta, es decir

que, para sentir odio, no es posible dudar de lo que se piensa, debe haber una seguridad intrínseca. Asimismo, el odio siempre se presenta de una manera difusa, no es preciso ni delimitado. Esto significa que la idea del individuo se ve desdibujada generalmente y no se reconoce a cada ser humano como un ser sintiente, sino que se crean “colectivos desdibujados como receptores del odio”. (Emcke, 2017, p. 14).

De la misma manera, Emcke (2017) sostiene que “el odio se fabrica su propio objeto. Y lo hace a medida.” (p.14). Por lo anterior, se afirma que el odio es un sentimiento que, como antes se había explicado, no aparece sin más, no es un sentimiento repentino ni difuso que se presente por descuido, sino que se presenta como algo preconcebido y cultivado. Emcke (2017) señala, igualmente, que un aspecto importante sobre el concepto del odio es que se genera, con dicho sentimiento, cierta idea de otredad, el odio tiene siempre una perspectiva vertical en la que se mueve siempre teniendo en cuenta la idea de lo otro como una amenaza o algo inferior. (p. 14).

El odio generalmente se presenta desde una colectividad y casi siempre, representa una ideología en particular, lleva un trasfondo histórico, cultural y social mucho más profundo. De esta manera, se puede afirmar que no solo la persona en particular está sujeta a sentir odio, sino también los grupos o colectividades y de igual manera, el odio puede recaer sobre un sujeto o sobre un grupo de personas, que históricamente, han sido colectividades marginadas por su raza, etnia, religión, nacionalidad, género, etc. Generalmente, los síntomas del odio se presentan como una clase de fanatismo, el cual supone la creencia ciega de ciertos dogmas que generalmente son la base del odio y la violencia y que, más allá de eso, justifican conductas intolerantes y violentas contra el objetivo del odio, es por esto por lo que, muchas

veces, la personas que perpetran estos actos, no sienten ninguna clase de remordimiento o arrepentimiento al respecto.

El odio muchas veces se puede camuflar sobre ciertos sentimientos diferentes como la preocupación o el miedo. Esto sucede así porque las personas intentan buscar palabras que se vean como menos contundentes y chocantes para expresar sus sentimientos, pero en la realidad, cuando se les invita a argumentar sobre la causa de sus preocupaciones, muchas veces salen a relucir sus verdaderas intenciones. El odio no obedece a ninguna clase de argumentación ni razón, generalmente se justifica con argumentos tales como la lucha de la persona o el colectivo del cual proviene el odio o discriminación contra el “otro” que generalmente es visto, no como un individuo con su propia vida, sentimientos y complejidades, sino simplemente como parte de un colectivo, el cual es percibido de una manera difusa y, al mismo tiempo, las personas que integran dicho colectivo son privadas de su propia condición de humanidad y por lo tanto, sus derechos inherentes como seres humanos son removidos y desconocidos.

Urge preguntarse si este odio envuelto en “preocupación” puede estar funcionando como sustitutivo o “válvula de escape” para canalizar experiencias colectivas de privación de derechos, marginación y falta de representación política. En este sentido, también es necesario analizar objetivamente las causas que generan esa energía que hoy se descarga en forma de odio y violencia en tantos lugares. (Emcke, 2017, p. 46).

A su vez, el odio no necesita una razón, no necesita cimentarse sobre una base real ya que, generalmente, los grupos o colectivos que dispersan el odio y violencia tienden a reducir o desconocer, conscientemente, la realidad. Pero, generalmente, el sentimiento de odio se puede explicar por el contexto en el que surge. Tal y como lo expone Emcke (2017)

generalmente el odio lleva atado consigo el desconocimiento del otro (p.62-63) y muchas veces, produce que los individuos sujetos a ese odio se invisibilicen de la sociedad, se anulen artificialmente y pierdan su condición de seres humanos sujetos de derechos, dándose así, junto con el concepto de “otredad” es decir, de la creación de otro hostil y peligroso para la sociedad, la exclusión de ese o esos otros.

Sin embargo, existen también actores diferentes a las personas que perpetran actos de odio abierta y violentamente. Estos se presentan como simples espectadores, personas que toleran las manifestaciones de odio porque, de manera menos abierta, están de acuerdo con dichas expresiones, solo que no de un modo tan extremista y más discreto. Estas personas, que no intervienen y se muestran como simples espectadores, propician la expansión de las expresiones de odio, porque precisamente, esto es lo que quieren los individuos que llevan a cabo estos actos de intolerancia, que les presten atención, sobre todo atención mediática, la cual facilita la expansión de su mensaje violento.

Igualmente, hay personas que actúan, según Emcke (2017), como *proveedores* del odio. (p.74). Son quienes suministran la carga ideológica necesaria para alimentar el odio, la cual luego se expande mediante las redes sociales o los diferentes medios de comunicación. Estas personas generalmente toman una distancia pública y actuada del odio, es decir que generalmente actúan de una manera en la que hacen creer a las otras personas que en realidad no están tratando de defender un discurso violento, pero utilizan los distintos medios para proporcionar facilidades y argumentos retóricos para defender dicho odio y violencia. Un claro ejemplo de esto son los políticos que utilizan Twitter y otras redes sociales para difundir su ideología racista, homófoba, xenófoba, etc, cómo lo hizo el ex presidente de los Estados

Unidos, Donald Trump. Esto se da ya que, como lo expresa Emcke (2017), el odio y el miedo también se ven alimentados por quienes esperan sacar beneficio. (p. 75).

Por otro lado, según Nussbaum (2014), los humanos muchas veces presentamos un sentimiento de “*asco proyectivo*” hacia algunos miembros de la sociedad, lo cual describe como un mal endémico en la condición humana.

“El proyectivo es el asco que se siente por un grupo de otros seres humanos separados conceptualmente del grupo dominante y clasificados como inferiores por su (presunta) animalidad más acusada. Se atribuye a los miembros de este grupo las propiedades de los objetos primarios de la repugnancia: sucios, grasientos, que huelen mal. Se los asocia con los fluidos sexuales, los excrementos y la descomposición. Son representados con rasgos cuasi animales, como si ocuparan una zona fronteriza entre lo verdaderamente humano (relacionado con aquello que consigue trascender el cuerpo y sus sustancias) y lo que definitivamente no lo es.” (Nussbaum, 2014, p. 223).

Es por esto por lo que algunas personas suelen negarse a compadecer a otras que perciben como diferentes y, por lo tanto, las consideran como completamente merecedoras de su odio y rechazo hasta llegar, algunas veces, a considerar que deben ser exterminadas.

De esta manera, según Nussbaum (2014), “el estigma y la exclusión son unos de los problemas fundamentales que nuestras sociedades deben combatir.” (p.200). Esto se da debido al deseo que se presenta en algunas personas de humillar a ciertos grupos o comunidades llevando a cabo practicas segregacionistas. Según lo anterior, un típico comportamiento humano muestra la tendencia que tenemos a ceder ante los que son considerados como nuestros iguales o los miembros de algún grupo o comunidad al que se

pertenezca. Esto es así ya que dichos grupos generan tal presión social en las personas, que muchos terminan cediendo ante sus demandas o ideologías, aún a costa de la verdad o de los ideales morales que pueda tener cada individuo como particular.

Probablemente una de las raíces del odio hacia otros es el propio narcisismo que caracteriza a la mayoría de la sociedad, además de una clara y contundente exclusión de lo distinto, de lo que se desvía de la norma social, de lo que es dictado como lo “normal” o bien visto para ciertos grupos que, en la mayoría de los casos, ostentan en poder en determinadas sociedades. De esta misma manera, las personas y las sociedades como tal se valen de la creación de jerarquías y tal cómo lo explica Nussbaum (2014),

“uno de los mecanismos clave de la subordinación es el asco: las personas que detentan el poder imputan a otros grupos de personas propiedades animales que inspiran normalmente repugnancia (viscosidad, pegajosidad, mal olor, relación con la descomposición o con los fluidos corporales y los excrementos) y usan luego esa presunta asquerosidad como razón para negar el contacto con tales colectivos, entre los que históricamente han sido encuadrados los afroamericanos, las mujeres, las castas inferiores, los judíos o los hombres homosexuales, entre otros.” (Nussbaum, 2014, p.221).

Es interesante como Nussbaum (2014) propone una definición de lo que es, según ella, “el mal real”:

“aquel comportamiento deliberadamente cruel y desagradable con otros individuos que no se debe simplemente a un problema de inadvertencia o descuido, o ni siquiera a una suspicacia teñida de temor, sino que entraña un deseo activo de denigrar o

humillar. Esta tendencia desempeña una función central en el odio y la discriminación grupales.” (Nussbaum, 2014, p. 201).

Partiendo del fenómeno del odio anteriormente puntualizado, surgen personas que aprovechan dicho sentir y crean discursos específicos para fomentar y esparcir el odio y violencia contra grupos particulares. Dichos discursos, generalmente, no tienen la necesidad de contar con un gran peso argumentativo o alguna razón práctica. Por lo general el objetivo de las personas que generan este tipo de discursos es viralizarse utilizando una serie de apelaciones emocionales, lo cual facilita el reclutamiento o la atracción de otras personas que se identifican con los sentimientos y emociones contenidas en dicho discurso.

Es importante puntualizar la diferencia entre un discurso de odio y otras clases diferentes de discursos que buscan informar sobre puntos de vista sin la intención de hacer alguna clase de daño a otros. La diferencia recae en que, precisamente, el discurso de odio se vale de ideologías y emociones particulares para atacar o discriminar a grupos que generalmente han sido históricamente maltratados. Aunque no existe una definición universalmente aceptada sobre este tema, se pueden distinguir algunos contenidos comunes a casi todas las definiciones, según Mathew, Dutt, Goyal & Mukherjee (2019), el discurso de odio es un ataque directo y serio contra cualquier grupo protegido según su raza, etnia, nacionalidad, religión, género, orientación sexual, discapacidad o enfermedad. (p. 175). Igualmente, como Nemes (2002) lo sostiene, el discurso de odio se refiere a expresiones que son abusivas, insultantes, intimidantes y que pretenden hostigar o incitar a la violencia, el odio o la discriminación. (p. 196). Los discursos de odio pueden generarse en diversos lugares, pero para el particular análisis, se tendrán en cuenta su construcción y uso por medio de las

plataformas tecnológicas, más particularmente, mediante las redes sociales tales como Twitter, Facebook, Instagram y similares.

Como lo señalan Patni & Kaumudí (2009), el discurso de odio se puede presentar de distintas maneras: puede ser directo o indirecto, puede ser sistemático o darse de manera aislada, respaldado por alguna clase de autoridad o amenaza o no y puede ser expresado de manera oral o escrita. (p. 750). Para efectos de esta investigación, se centrará en el concepto de discurso de odio expresado de manera escrita a través de plataformas o redes sociales. Dicho discurso de odio puede causar daño a grupos específicos y sociedades y sus repercusiones pueden ir desde causar dolor, estrés, humillación, tristeza o miedo a la víctima hasta la pérdida de la dignidad, el daño al derecho a la igualdad y a la libre expresión, ataques físicos y, muchas veces, hasta la muerte.

Según lo exponen Reichelmann, Hawdon, Costello, Ryan, Blaya, Llorent, Oksanen, Räsänen & Zych (2020), a diferencia del ciberbullying o el acoso cibernético, los cuales se centran en atacar a individuos particulares, el discurso de odio que se da en las diferentes plataformas tecnológicas se enfoca en colectivos o grupos históricamente marginados. (p. 2). Igualmente, los mensajes de odio pueden ser compartidos en la red mediante diferentes métodos comunicativos como lo son los textos, las imágenes, videos y links en general.

Las plataformas proporcionan un campo amplio para la intercomunicación, lo cual es positivo para el fenómeno de la globalización, ya que permiten una conexión internacional de tiempo presente. Sin embargo, el término plataforma puede presentar diferentes acepciones, puede considerarse, según Gillespie (2010), como una infraestructura que sirve para construir diferentes aplicaciones. (p.352). No obstante, lo anterior, en términos de Gillespie (2010), una noción más amplia del término plataforma, que en varios campos de conocimiento, es

vista, según la terminología computacional, como una construcción que lleva a la innovación y desde el campo político, como un lugar desde el cual es posible dar a conocer opiniones y puntos de vista para ser escuchado. (p. 352).

El filósofo Sur Coreano Byun Chul Han ha caracterizado a la comunicación digital como una herramienta “pobre en mirada”, en este sentido, expone que:

El smartphone es un aparato digital que trabaja con un input-output pobre en complejidad. Borra toda forma de negatividad. Con ello se olvida de pensar de una manera compleja y deja atrofiar formas de conducta que exigen una amplitud temporal o una amplitud de mirada. Fomenta la visión a corto plazo. Fomenta el corto plazo y la mirada de corto alcance y ofusca la de larga duración y lo lento. (Han, 2014, p. 42- 43).

De esta manera, los mensajes tienen una facilidad especial de expansión y recepción nunca antes vista, convirtiéndose así en el transmisor perfecto para los mensajes de odio sin reservas ni filtros. A pesar de lo anterior, hay que puntualizar que, tal y como Bonnet (2019) expone, no son las redes las culpables en sí mismas de la difusión de este tipo de mensajes y las consecuencias que estos puedan conllevar, el problema, más bien, recae en el uso que las personas les dan a éstas. (p. 182). El problema recae igualmente en que, siguiendo la línea argumentativa de Han (2014), las redes tienden a normalizar y valerse de un sistema simplista de la realidad y las interacciones sociales. (p. 42-43).

Según lo anterior, al utilizar un sistema binario como lo es el de “Me gusta” y de esta manera dejando implícito el “No me gusta”, la comunicación se convierte en algo simplificado que no da cabida alguna a los matices ni a las diferencias y propiciando así su exclusión.

A pesar de que las plataformas, en especial las redes sociales, son presentadas como una innovación importante para la interconexión entre las personas y un gran avance para la globalización, es evidente que, con la llegada de dichas, debimos sufrir muchos cambios como sociedades, tanto conductuales como intelectuales y relacionales. Hemos acogido comportamientos que alguna vez fueron considerados como inaceptables pero que ahora han sido normalizados dentro del uso de las redes. En palabras del filósofo Byun Chul Han:

“La comunicación digital deshace, en general, las distancias. La destrucción de las distancias espaciales va de la mano con la erosión de las distancias mentales. La medialidad de lo digital es perjudicial para el respeto. Es precisamente la técnica del aislamiento y de la separación, como el Ádyton, la que genera veneración y admiración.” (Han, 2014, p.14).

Según la tesis anterior, el mayor problema de los medios digitales, y en particular de las redes sociales, es que las personas sienten que ya no necesitan salir de sus propios mundos digitales, se convierten en verdaderos “Hikikomoris”, es decir, en seres aislados del mundo exterior. Es por esto por lo que la soledad se convierte en el determinante del día a día de muchos. En palabras de Han (2014) los medios digitales realmente aíslan a las personas. A su vez, esto ayuda a que los individuos cada vez se centren más en el “yo” y dejen atrás el “nosotros”, con lo cual se crea una creciente tendencia hacia el egoísmo y el narcisismo. (p.28). Según Han (2014) expone,

“La comunicación digital hace que se erosione fuertemente la comunidad, el nosotros. Destruye el espacio público y agudiza el aislamiento del hombre. Lo que domina la comunicación digital no es el “amor al prójimo” sino el narcisismo. La técnica digital

no es una “técnica del amor al prójimo”. Se muestra como una maquina narcisista del ego.” (Han, 2014, p. 75).

Es, tal vez, por este narcisismo y falta de empatía que las personas comienzan a reproducir comportamientos destructivos y es ese mismo egoísmo desmedido el que les impide situarse en el lugar de los demás y comprender sus luchas, sentimientos y diferencias. Mediante las redes sociales, se abre paso a que estas personas puedan tener un espacio para expresar sus pensamientos sin tener que contar con ningún tipo de reflexión. Es así como se crean los discursos que llevan implícitos un gran contenido de odio, los cuales se esparcen por toda la red para que otras personas los apoyen y los reproduzcan.

Esto es importante porque, en palabras de Han (2014) “el medio digital, que separa el mensaje del mensajero, la noticia del emisor destruye el nombre.” (p. 15). Es decir que, al fin y al cabo, la manera en cómo funciona el sistema permite que, a pesar de difundir mensajes de odio abierta y descaradamente, casi nunca se hace el énfasis en la persona que produce o reproduce el mensaje sino en el mensaje en sí. A pesar de que las personas creen mensajes de odio proporcionando sus propios nombres, esto se hace irrelevante ya que dicho mensaje va a ser reproducido tan incontables veces al final, se pierde la identidad de la fuente de la que provino. De esta manera, al desligarse el mensaje de odio de un individuo en especial, se pierde también la noción del respeto y la empatía y es así como también desaparece la solidaridad con las personas o grupos contra quienes se dirigen los ataques y discursos de odio discriminatorios.

De la misma manera, la comunicación intermediada por los aparatos tecnológicos como la computadora o el smartphone produce cierto cambio en nuestra manera de sentir al otro. El hecho de no tener presente a la persona con la que se está interactuando, puede causar

diferencias en nuestra percepción y en nuestros sentimientos. Es por esto por lo que, muchas veces, al no tener presente al otro, perdemos la capacidad de desarrollar sentimientos como la empatía, perdemos interés sobre las emociones del otro y no nos preocupamos por comprenderlas. Lo anterior puede verse muy claramente cuando, en un simple chat con otra persona, se presentan mal entendidos porque alguna persona interpretó una palabra o una frase de una manera diferente a la intención de lo que el emisor quería expresar.

Mediante la comunicación en línea, se corre el riesgo de mal interpretar ciertos mensajes ya que lo único que se tiene disponible es la comunicación verbal, pero se pierde el contexto, la entonación, la gesticulación y hasta la proxémica. Por esta misma razón, los mensajes se tienden a despersonalizar y de esta manera, se pierde la idea o simplemente no interesa la posibilidad de que un mensaje determinado pueda hacer daño a una persona o un grupo.

Por otro lado, hay que tener en cuenta que estas redes funcionan con unos algoritmos específicos que clasifican los gustos de cada usuario para así seguir mostrándoles publicaciones que sean de su interés. Esto es problemático también porque, tal y como lo sostiene Cortés (2018), “las plataformas en línea incentivan el consumo de informaciones y opiniones que reafirman nuestros puntos de vista –las conocidas burbujas de información.”

En este sentido, la información que consumimos se nos ofrece según nuestros intereses, es decir que, si alguien se identifica con expresiones extremistas, como por ejemplo la supremacía blanca, las redes sociales le seguirán recomendando grupos, personas e información que se identifique con esas expresiones, alentando así su odio y reafirmandolo.

En este mismo sentido, no podemos dejar de lado la influencia que ejercen los discursos de las figuras importantes o que ostentan cierta clase de poder sobre un grupo considerable de personas como lo pueden ser ciertos políticos, académicos o figuras públicas en general. Si

bien es cierto que se debe respetar el derecho a la libre expresión dentro de las redes sociales, muchas veces éstas figuras tergiversan el sentido y la valía de aquel derecho y lo convierten en algo perjudicial y muchas veces hasta peligroso para ciertos grupos o colectivos de personas. Debido a la autoridad que poseen estas personas, los mensajes que emiten generalmente cuentan con una presunción de validez y verdad entre los usuarios de ciertas plataformas sociales como por ejemplo Twitter o Facebook.

Es aquí donde también se debe prestar especial atención a las llamadas *Fakenews*, es decir, noticias que provienen de emisores poco confiables y que se propagan de una manera descontrolada, valiéndose de la ignorancia o las creencias populares, pero que en la realidad no tienen ninguna clase de sustento material. Muchas veces, esta clase de noticias son fuente importante para la creación y reproducción de mensajes de odio. En este sentido, Mathew, Dutt, Goyal & Mukherjee (2019) exponen que, como es sabido, las *Fakenews* tienden a reproducirse y propagarse de una manera más veloz que otras noticias, por lo que los autores sostienen que este tipo de noticias van muy de la mano con el discurso y la propagación del odio. (p. 174).

Junto con las llamadas *Fakenews*, también se ha visto como el fenómeno de las cuentas falsas ha venido cobrando relevancia dentro del panorama del uso de los discursos de odio. Dicho fenómeno brinda a los usuarios un cierto nivel de anonimato ya que éstos no revelan su verdadero nombre ni sus datos personales, lo cual hace casi imposible que otros usuarios puedan identificar a las personas que se encuentran detrás de dichas cuentas. Según el estudio de Mondal, Araújo & Benevenuto (2017), los usuarios de redes sociales tienden a desinhibirse y expresar sus sentimientos o ideas de una manera más abierta cuando publican

en redes sociales desde una cuenta anónima. De esta forma, las personas tienden a ser más agresivas cuando sienten que cuentan con un nivel alto de anonimato. (p. 86).

Siguiendo esta línea argumentativa, Mondal, Araújo & Benevenuto (2017) sostienen que, en lo concerniente a los discursos de odio basados en la raza o la orientación sexual, los usuarios anónimos tienden a ser más desinhibidos y agresivos. (p. 90). Lo anterior sugiere que los mecanismos de identidad débiles, cómo lo es la posibilidad de usar el anonimato, generan más discursos de odio en las plataformas sociales y estos tienden a ser, igualmente, más violentos y discriminatorios.

En este sentido y según el estudio de Reichelmann, Hawdon, Costello, Ryan, Blaya, Llorent, Oksanen, Räsänen & Zych (2020), se pudo identificar que Facebook es lugar más común para encontrar expresiones y discursos de odio. (p. 6). Esto tiene sentido ya que Facebook es la plataforma social más popular en internet. Sin embargo, se pueden encontrar, igualmente, importantes expresiones de odio en plataformas como Youtube, Instagram y Twitter. Lo anterior haciendo la salvedad de que el uso de plataformas para la diseminación de los discursos de odio varía según cada país. Al puntualizar lo anterior, salta a la vista que dichas plataformas permiten, de una manera u otra, que sus usuarios puedan esconder su identidad mediante la creación de cuentas falsas para viralizar sus discursos de odio sin sanción ni reproche alguno.

Lo anterior no significa, de ninguna manera, que los usuarios de redes sociales completamente identificables no generen mensajes de odio y discriminación contra grupos o colectivos. Muchas veces estas personas experimentan sentimientos de orgullo y se encuentran tan satisfechos con sus actos, hasta tal punto que quieren que los otros usuarios sepan sobre sus opiniones e ideologías. Muchas de estas personas pueden clasificarse como

extremistas quienes, de manera común, usan las redes sociales para reclutar personas que quieran unirse a sus causas violentas. Un buen ejemplo de lo anterior es el Estado Islámico, quienes muchas veces llegan a las personas a través de las plataformas sociales. Es así como también se puntualiza que las redes sociales son utilizadas por individuos particulares, pero también por grupos organizados y semi organizados para promover esta retorica del odio

La reiterada interacción con estos mensajes y opiniones discriminatorias y violentas en las redes sociales frecuentemente causa que los usuarios se desensibilicen y resten importancia al impacto que esto puede causar en otras personas. Al normalizar los discursos de odio, se corre el peligro de deshumanizar a las comunidades las cuales se discrimina y con esto, restarle importancia al cumplimiento de sus derechos inherentes como seres humanos que merecen ser tratados con dignidad y ostentar una igualdad material frente a otros individuos.

Según la investigación de Mathew, Dutt, Goyal & Mukherjee (2019), las publicaciones realizadas por usuarios que exponen discursos de odio y contenido discriminatorio tienden a propagarse y masificarse de una manera más rápida, más extensa y con un alcance más global que otros mensajes publicados por usuarios que no incitan al odio. (p. 174). Es decir que esta clase de publicaciones violentas recibe una atención más importante y es por esto por lo que es posible que dichos mensajes se extiendan y lleguen a una cantidad considerable de personas que, de alguna manera, coincidan en opiniones e ideologías destructivas.

La investigación de Mathew, Dutt, Goyal & Mukherjee (2019) revela, también, que las personas que se configuran como propagadoras de odio en las redes generalmente forman grupos cohesivos. (p.180). Lo anterior significa que la propagación de estos diferentes contenidos con mensajes y discursos de odio proviene de un esfuerzo colectivo organizado que ayuda a que este tipo de contenidos se disperse de manera viral, oponiéndose a los

simples esfuerzos individuales, los cuales tienen menos posibilidad de viralizarse de esta manera.

Lo anterior quiere decir que, si bien los mensajes de odio pueden provenir de individuos particulares, los discursos de odio generalmente se construyen de manera colectiva, mediante grupos que se apoyan y se legitiman entre sí. Ahora bien, el caso de los individuos que ostentan cierto poder, como lo son los políticos o figuras públicas, y que se dedican a difundir el odio y el resentimiento contra ciertos colectivos o grupos, representa, también, la idea de que los mensajes de odio se esparcen con mayor facilidad de una manera colectiva. Lo anterior porque, si bien son individuos particulares, generalmente representan a una cierta comunidad o ideología, lo cual hace que ellos igualmente, pertenezcan a un grupo o colectivo específico que se encarga de legitimar dicha clase de retórica.

Asimismo, en muchos de los casos en los que se exponen discursos de odio en las redes, especialmente en las plataformas sociales, se presenta un fenómeno interesante, muchos de los usuarios que encuentran estos mensajes de odio, optan por replicar a éstos empleando, igualmente, mensajes llenos de odio y, muchas veces, llegan a ser igual o más discriminatorios que el mensaje original. Según el estudio de Mondal, Araújo & Benevenuto (2017), se encontró que también existen los discursos de odio como respuesta al mismo odio. (p.89).

Es posible que lo anterior se deba a que muchos usuarios responden de una manera emocional a los ataques de odio, y aunque tal vez no se sientan directamente aludidos, como anteriormente se explicó, los ataques de odio que pueden darse en redes sociales pueden desencadenar diferentes situaciones y sentimientos en las víctimas que pueden ir desde el

miedo, la tristeza hasta la rabia y la ira, por esto, diferentes individuos pueden reaccionar de diferentes formas y una de esas formas de reaccionar puede ser con odio.

En síntesis, puede afirmarse que el odio es una emoción que difiere de la rabia o la ira entre otras más ya que no es un sentimiento explosivo, el odio no aparece sin más, sino que se cultiva y va creciendo con el tiempo. A su vez, la aparición de las nuevas tecnologías y de las plataformas tales como las redes sociales, han cambiado drásticamente nuestra manera de percibir el mundo y de relacionarnos con otros individuos como sociedad. La forma de comunicarnos ha pasado a ser predominantemente escrita y simbólica y esto ha causado la aparición de ciertos fenómenos que pueden llegar a ser problemáticos para la convivencia. Entre los anteriores, se encuentra el surgimiento de los discursos de odio dirigidos contra grupos que tienen una especial protección ya que han sido históricamente discriminados por su raza, etnia, identidad sexual, género, entre otras, y que se propagan a través de las redes sociales.

Estos discursos de odio difieren de otros mensajes precisamente porque tienen la intención de hacer daño a sus receptores y legitimar ciertas ideologías que generalmente son apoyadas por grupos específicos los cuales muchas veces pueden llegar a ser grupos extremistas.

Pese a que se ha denotado lo complejo que es definir y encuadrar completamente el significado concreto del término mensaje o discurso de odio, sí puede hablarse sobre unas características particulares que se encuentran presentes en toda definición de dicho concepto. Entre éstas se encuentra que los discursos de odio generalmente son ataques directos contra alguno o varios de los grupos protegidos en razón a su raza, etnia, nacionalidad, religión, género, orientación sexual, discapacidad o enfermedad y en su generalidad, contienen

expresiones insultantes, abusivas o con una clara intención de hacer daño a sus receptores o víctimas directas.

A pesar de todo lo anterior, se concluye que las plataformas sociales no son buenas o malas per se, no puede afirmarse que inciten por si solas a la dispersión de los discursos de odio. En donde hay que hacer especial énfasis es en la manera en la que los usuarios utilizan dichas redes sociales, convirtiéndolas en focos de noticias falsas y mensajes discriminatorios y violentos contra ciertas comunidades o colectivos. En este sentido, se hace énfasis en cómo las plataformas representan una especie de escenario dónde los mensajes ideológicos y discursos de odio de ciertas personas o grupos organizados pueden darse a conocer y viralizarse a través de todo el mundo. Lo anterior tiene la capacidad de causar una cadena de reacciones diferentes que van desde la aprobación de otras personas o grupos de dichos mensajes de odio, la generación de otros contenidos o mensajes de odio a partir de dicho mensaje primigenio hasta la generación de respuestas cuyo objetivo sea combatir dicho odio, pero a su vez, que terminen generando más y más odio. Por otro lado, también se presentan las reacciones de las víctimas, las cuales pueden ir desde el miedo, la tristeza, la preocupación hasta la depresión y la sensación de una real amenaza y de que su vida corre peligro.

Siguiendo esta línea argumentativa, y para lograr llegar a una visión más profunda de lo que significa la propagación de los mensajes de odio por medio de las plataformas sociales, se ocupará específicamente, en el siguiente capítulo, de brindar una definición sobre qué son las plataformas, cuáles son sus distintas características, cómo es su funcionamiento y de qué manera han logrado posicionarse como los gigantes tecnológicos que hoy representan en nuestro mundo.

### **3. Los nuevos transmisores: las plataformas sociales de base tecnológica como propagadores virales de mensajes**

Las plataformas digitales son un fenómeno que ha venido evolucionando de una manera vertiginosa en nuestra sociedad. Hoy en día, muchas de las actividades cotidianas de cada persona dependen del uso de éstas, como, por ejemplo, el manejo del dinero o algo tan simple como la comunicación con otros. Debido a lo anterior, es evidente que, con la llegada de dichas plataformas, el mundo debió sufrir un cambio de paradigma, ya que ahora, la tecnología forma parte plena e importante de la vida de cada individuo y la vida sin ella es casi inconcebible.

De esta misma manera, las plataformas digitales, especialmente las redes sociales, han representado una gran revolución en el campo de la comunicación en el sentido en el que permiten la existencia de una interconectividad a nivel mundial. Las distancias ahora pueden verse mucho más cortas y la globalización se encuentra presente y a la mano de cada individuo. Nuestra sociedad ha tenido que adaptarse a diferentes cambios a nivel comunicativo e informativo que han producido respuestas diferentes a la manera en como vemos el mundo actualmente. Con la llegada de las plataformas tecnológicas, aumentó también el grado de información al que estamos expuestos día a día y puede sostenerse que, actualmente, permanecemos conectados, la mayor parte de nuestras vidas, a las plataformas mediante nuestros Smartphones, tabletas, computadores, etc.

A lo largo de la historia del surgimiento de dichas plataformas, se ha presentado una evolución acelerada, llegando a tal punto en el que, hoy en día, redes sociales como Facebook, Instagram, Twitter, Youtube, entre otras, se han convertido en medios importantes de distribución de información y, a su vez, se han instituido como grandes recolectores de

datos. Lo anterior debido a que son lugares donde se hace posible la masificación de mensajes que tienen un alcance global, esto significa que, mediante estas plataformas sociales, se ha creado un nuevo sistema donde los mensajes, tanto políticos como sociales, tienen la posibilidad de impactar en una mayor cantidad de la población mundial.

Actualmente, según la agencia global We Are Social (2020), existen mundialmente más de 4.500 millones de usuarios de internet y a su vez, se evidencia que, hoy en día, 3.800 millones de personas hacen uso de las plataformas sociales. Lo anterior, demuestra que, tanto la internet como las redes sociales se han convertido en una parte indispensable de la vida de muchas personas alrededor del mundo.

Según un estudio realizado por Datareportal (2020), globalmente, más de 5.190 millones de personas actualmente utilizan sus dispositivos móviles para conectarse a internet y redes sociales. En promedio, las personas pasan aproximadamente 3.7 horas conectados a través de sus dispositivos móviles y se reporta que aproximadamente la mitad de ese tiempo es utilizado por los usuarios para conectarse a redes sociales y aplicaciones de comunicaciones.

Las compañías que operan las plataformas más grandes del planeta, como lo son Facebook o Google, se han convertido en potencias globales, en multinacionales que, actualmente, son actores sumamente importantes para el proceso de la globalización. Estas compañías cuentan con una gran influencia dentro de la vida pública y política de todos sus usuarios y ostentan un gran poder económico. Estas compañías ostentan tal poder ya que, como afirma Gorwa (2019), las plataformas que han creado son actores políticos fundamentales que toman decisiones políticas importantes mientras construyen lo que se ha convertido en la infraestructura global para la libertad de expresión. (p. 857) Sin embargo, no debe

confundirse a las compañías o corporaciones que crean y operan plataformas tecnológicas con las plataformas tecnológicas en sí.

Gorwa (2019) expone que es importante examinar la manera en cómo las compañías de plataformas tienen la capacidad para navegar el mundo de los complejos y contrastantes intereses de las diferentes partes interesadas en estos procesos ya que el término plataforma, se ha convertido en sí mismo en un imaginario político discursivo. Se resalta la manera en la que las plataformas intervienen en la vida cotidiana de sus usuarios, moldeando su experiencia en línea y determinando, algorítmicamente, qué información se visibiliza o, por el contrario, se invisibiliza. (p.858)

Las compañías que crean y operan las plataformas tecnológicas pueden clasificarse como los arquitectos o creadores de los espacios en línea. A diferencia de las plataformas tecnológicas, las cuales, según Helmod (2015), han tenido una creciente importancia en la sociedad y se han situado como la infraestructura dominante y el modelo económico predominante en el mundo de las redes sociales. (p.1). A pesar de lo anterior, el término plataforma puede suscitar algunas confusiones ya que su definición depende de la mirada desde la que se abarque porque es una palabra que posee diferentes connotaciones. Según esto, Gillespie (2010) define el término desde un sentido computacional como una infraestructura que sirve para crear aplicaciones. (p. 352). Sin embargo, con el creciente desarrollo tecnológico y económico que se ha venido dando en este campo, se han creado otras nociones más amplias del término plataforma. El término plataforma es entendido de diferentes maneras tanto por los científicos computacionales como por los economistas o los abogados.

En un primer momento, las plataformas no eran concebidas como tales, sino que los investigadores y desarrolladores las clasificaban simplemente como redes sociales, pero

luego, ocurrió una transición importante. De acuerdo con Gorwa (2019), la historia de las plataformas tecnológicas empieza en California en la década de los 90, cuando los desarrolladores de software comenzaron a conceptualizar su trabajo como algo más que la creación de simples programas, entendiéndolo como la creación de plataformas flexibles que permitían desarrollar códigos y desplegarlos. Luego este término fue desarrollado estratégicamente por ciertas compañías para denominarse a sí mismas como plataformas que facilitan el acceso a contenido generado por usuarios y que, por lo tanto, no es creado por ellas mismas y en ese sentido, no deben ser responsables por el mismo. (p. 856).

Sin embargo, tal como Bonina, Koskinen, Eaton & Gawer (2019) sostienen, existe una falta de claridad en cuanto al entendimiento sobre qué exactamente son las plataformas digitales, cuáles son sus principales características y cómo generan valor. En este sentido, afirman que las plataformas digitales comparten tres características básicas las cuales son: que están tecnológicamente mediadas, facilitan la interacción entre grupos de usuarios y permiten que esos grupos de usuarios realicen ciertas actividades definidas. Es por esto por lo que entienden a las plataformas tecnológicas como un conjunto de recursos digitales, ya sea que proporcionen servicios o contenidos, que facilitan las interacciones entre sus usuarios. Sin embargo, sostienen que la naturaleza especial de cada plataforma depende de la actividad específica que los usuarios lleven a cabo. (p.3).

De esta manera, Bonina, Koskinen, Eaton & Gawer (2019) afirman que las definiciones del término plataforma digital, dependen de la mirada que se le dé desde el campo de estudio específico. Así, en los estudios que se concentran en los componentes tecnológicos, se hace énfasis en sus características tecnológicas y digitales como su arquitectura basada en capas y su modularidad. En sistemas de la información, la atención se centra en las dimensiones

sociotécnicas de las plataformas digitales, por ejemplo, en su impacto en estructuras organizacionales o estándares internacionales. Por otro lado, en el campo de la economía, la discusión ha girado alrededor de las funciones de demanda y suministro de estas plataformas y como éstas son diferentes de otros tipos de mercados. Desde el punto de vista de las industrias, las plataformas digitales son presentadas en términos de características como la capitalización del mercado, propiedad, sector o industria en la que están clasificadas, modelos de gobernanza, país de origen, alcance geográfico y propósito subyacente. (p.3).

Dentro de su investigación, Bonina, Koskinen, Eaton & Gawer (2019) sugieren que las plataformas digitales son una clase de artefacto de tecnología de la información con propiedades distintivas, las cuales se prestan para alcanzar logros particulares en el campo del desarrollo. Igualmente, las plataformas digitales son un fenómeno sociotécnico que requiere una consideración especial en cuanto a cómo se da su funcionamiento en un contexto social. (p.3).

En este mismo sentido, Bonina, Koskinen, Eaton & Gawer (2019) identifican dos categorías en las que puede clasificarse el término plataforma digital: plataformas transaccionales y plataformas innovativas. El principal propósito de las plataformas transaccionales es el de facilitar, como su nombre lo dice, transacciones entre diferentes organizaciones, entidades e individuos como también el de conectar vendedores con compradores. Esta clase de plataformas puede ser subdividida teniendo en cuenta su función principal. En este sentido, se pueden encontrar las plataformas sociales como lo es Facebook, plataformas de e-commerce como Mercado Libre, portales online, app stores, etc. (p.2-5).

Por otro lado, según lo propuesto por Bonina, Koskinen, Eaton & Gawer (2019), las plataformas innovativas actúan como piezas fundamentales a partir de las cuales, otras firmas

pueden construir productos complementarios, servicios o tecnologías. De este modo, la arquitectura técnica de una plataforma innovativa está conformada por módulos o bloques que pueden ser adquiridos y combinados por desarrolladores de apps para construir apps y servicios. Los ejemplos de este tipo de plataformas van desde los sistemas operativos como Android e iOS hasta los servicios de almacenamiento en la nube como los que ofrecen Amazon, Google y Microsoft. (p.7-8).

Según lo afirmado por Andersson (2017), una plataforma es, básicamente, una infraestructura digital (basada en software o algunas veces en hardware) destinado a que los usuarios le apliquen códigos computacionales (por ejemplo, para crear aplicaciones) o para ser aplicado a una serie de usos humanos que serán delimitados y adaptados según la plataforma en cuestión. Las plataformas digitales son superficies que se prestan para la innovación técnica y a través de estas, nuevos actores pueden desarrollar servicios o productos adicionales, es decir que dichas plataformas generan nuevas oportunidades de negocio. Es por esto por lo que en el campo económico son llamadas mercados multilaterales, ya que permiten la realización de transacciones entre actores que, de otra manera, probablemente no se habrían encontrado o, al menos, habría sido mucho más difícil que lo hicieran. Esto quiere decir que, diferentes usuarios se reúnen en este espacio para reducir los costos de transacción. Las plataformas digitales automatizan los intercambios mercantiles y funcionan como mediadores en las acciones sociales. (p.376-377).

Más allá de lo anterior, una idea importante de Andersson (2017) sustenta que las compañías que ostentan la propiedad de las plataformas sociales digitales, las cuales están pensadas para permitir la creación de contenido por los mismos usuarios y, a la vez permiten que este contenido sea compartido a otros usuarios, como por ejemplo lo son Youtube, Facebook,

Twitter, Instagram, etc. También tienen una naturaleza dual. Es decir que permiten e invitan a sus propios usuarios a desplegar su creatividad dentro de la plataforma y a expresarse, pero, en últimas, su propósito los lleva a considerar que cierta participación de algunos usuarios sea delimitada de ciertas maneras. Esto quiere decir que las plataformas permiten y, simultáneamente, constriñen ciertas expresiones de sus usuarios. (p.377). Esto, sin embargo, es polémico ya que dichas plataformas, generalmente, no tienen la suficiente capacidad para controlar todo lo que sus usuarios comparten en sus redes, por esto, solo pueden observar por una pequeña parte de este mundo.

Si bien podría ser reprochable que las plataformas eliminaran unilateralmente el contenido de sus usuarios ya que cabría la posibilidad de que se estuviera violando el derecho a la libre expresión de los individuos, en algunos casos, este control se presenta como necesario y hasta deseable. Es el caso de los discursos de odio, los cuales violan los reglamentos y los términos y condiciones de la mayoría de las plataformas sociales. Es evidente que dichas plataformas no tienen la capacidad de impedir que estos mensajes sean creados y publicados, pero si deben velar por que sean eliminados lo más pronto posible y, además, deben socializar sus reglamentos y procurar que sus términos y condiciones sean conocidos y entendidos por los usuarios y así, crear una especie de gobernanza interna donde sean las mismas compañías de las plataformas las que velen por el cumplimiento de las normativas y el respeto dentro de dichos espacios virtuales.

Andersson (2017) establece, igualmente, que los individuos no son vistos solamente como usuarios, sino también como “*inputs*” o datos que sirven a los sistemas ya que su participación en las plataformas crea valor para otros usuarios. (p. 377). Desde luego, las plataformas tecnológicas también juegan un importante rol en el campo de las decisiones

automatizadas. Es decir que, en nuestra vida actual, hay decisiones que simplemente no somos conscientes de tomar ya que los algoritmos de las plataformas automáticamente deciden por nosotros mismos. Esto sucede, por ejemplo, como lo expone Zuboff (2020),

los contratos en línea, cómo los acuerdos de términos de servicio o términos de uso, también se conocen como *clickwraps* porque, como muchas investigaciones han demostrado, envuelven en esos opresivos términos contractuales a la mayoría de las personas desde el momento en que hacen clic en la casilla de “sí, acepto”, sin haber llegado nunca a leerse el acuerdo en cuestión. En muchos casos, el simple hecho de navegar por un sitio web ya nos obliga a su acuerdo de términos de servicio, aunque desconozcamos cuál es. (Zuboff, 2020, p.73)

De la misma manera, Andersson (2017) establece que, debido a la generación automática de datos importantes que poseen dichas plataformas y la inteligencia especial que puede ser extraída de dichos datos, las plataformas producen nuevas formas para converger con aquellos que figuran como los dueños y controlan dicha información y, así, poder utilizarla en su propio beneficio. Una misma compañía puede ser capaz de dominar varios mercados que, en una situación cotidiana, deberían estar nominalmente separados, pero que materialmente no lo están ya que la información y los datos han sido fusionados debido a la *plataformización*. Un ejemplo de lo anterior se puede ver con la compañía Microsoft, la cual adquirió otra plataforma llamada LinkedIn, lo cual les permite utilizar los datos y la información que contiene LinkedIn como un recurso competitivo, a pesar de lo reprochable que esto pueda ser. (p.383-384.)

Otra característica común de las plataformas, especialmente las que representan una gran dominancia global, es que, según Andersson (2017), dichas plataformas no tienen que

diversificarse adquiriendo nuevas subsidiarias en otros países, sino que lo que optan por hacer es ofrecer a otras plataformas competidoras considerablemente más pequeñas y débiles que se valgan de la propia infraestructura de las plataformas dominantes. Esto claramente permite que las grandes plataformas recolecten también la información que se produce a partir de dichas colaboraciones. Esta es una práctica muy común de Facebook, plataformas más pequeñas normalmente utilizan el login de Facebook para identificar a sus usuarios. Esto no solo conviene a dichas plataformas pequeñas, ya que pueden ganar datos e información contenida en Facebook para integrar a sus propios servicios, sino que también, Facebook puede ganar nuevos datos de los usuarios de cada plataforma, lo que hace que su propia base de datos se enriquezca. (p.384)

No debe ignorarse la gran importancia que resulta de la minería de datos que llevan a cabo la gran mayoría de las plataformas sociales para poder tener una rentabilidad de su propia actividad. En este sentido, se ha encontrado, según Zuboff (2020) que

En esta nueva lógica, la experiencia humana queda subyugada por los mecanismos de mercado del capitalismo de la vigilancia y renace en forma de *conducta*. Esas conductas son rendidas-convertidas en forma de datos preparados para ocupar su lugar en una innumerable cola de ellos que alimenta las máquinas de fabricación de predicciones, unas predicciones que son los productos finales que se intercambian en los nuevos mercados de futuros conductuales. (Zuboff, 2020, p. 141.)

Según la historia del surgimiento de dichas plataformas sociales, y refiriéndose a la minería de datos, “al principio, esas materias primas se encontraban sin más, pues eran un subproducto de las acciones de búsqueda de los usuarios. Más adelante, esos activos fueron

ya objeto de una agresiva caza y obtenidos básicamente mediante la vigilancia.” (Zuboff, 2020, p. 133.)

Debido a esta recolección de información sobre los hábitos, gustos, necesidades, etc. de cada usuario, estas plataformas cuentan con la ventaja de sugerir información y páginas que, desde el principio, están seguros de que el usuario específico abrirá y les prestará especial atención. Es así como, si un usuario muestra una predilección específica por la lectura, el algoritmo le seguirá mostrando información que haga referencia a ese campo en específico. Es aquí donde se pregunta, ¿qué pasa si los gustos o predilecciones de ciertos usuarios tienen que ver con incitaciones al odio hacia ciertos grupos? Esta pregunta es pertinente porque, según el algoritmo, a este usuario se le sugerirán ciertas personas que concuerden con su ideología y se les ofrecerá información acorde a sus intereses, lo cual podría causar que dicha persona refuerce sus opiniones potencialmente destructivas y, es más, encuentre otros grupos de personas con las cuales asociarse y reafirmar su odio.

Por otro lado, estas plataformas sociales son espacios de creación de contenido por parte de sus mismos usuarios, pero también, dicho contenido, el cual es compartido públicamente, tiene una gran capacidad de propagarse y reproducirse a una gran velocidad, lo cual supone que los mensajes de los usuarios se conviertan en una fuente de información y opiniones masiva. En este sentido las plataformas sociales juegan un papel importante en la difusión de mensajes y opiniones que pueden ser, por un lado, una simple manifestación de la libre expresión de las personas, pero por otro, pueden ir más allá y convertirse en mensajes que instigan al odio y a la discriminación hacia ciertos grupos protegidos por su especial vulnerabilidad, lo cual puede desencadenar grandes consecuencias debido a la falta de

acuerdos en el sentido de los significados de odio y cómo deberían ser tratados en dichas plataformas sociales, además de la falta de regulación legal en cuanto a lo anterior.

En este sentido, aunque estas plataformas sociales se han convertido en gigantes transnacionales, el asunto de su regulación ha suscitado grandes problemas tanto para académicos como para los gobiernos en sí. Precisamente por su característica global, se derivan los problemas de regulación a nivel nacional dentro de la jurisdicción de cada Estado. Tal cómo lo aclara Nemes (2002), el internet es una estructura global y descentralizada de computadores interconectados y vinculados. (p.198). Y, de esta manera, Nemes (2002) expone que es claro que las regulaciones nacionales específicas son sumamente difíciles de aplicar cuando el material publicado en dichas plataformas o redes sociales se origina en una jurisdicción específica, sus efectos son sentidos en otra jurisdicción diferente y el creador de dicho contenido podría residir en una tercera jurisdicción separada de las anteriores. (p.199).

En suma, puede afirmarse que las plataformas digitales son, básicamente, espacios o infraestructuras digitales creadas para que, mediante la aplicación de códigos computacionales especiales, sea posible la aparición de aplicaciones que permitan a las personas o usuarios una interacción entre sí. De esta manera, los espacios creados mediante las plataformas digitales permiten que la vida y las interacciones humanas se desarrollen de una manera un tanto diferente e innovadora ya que, por ejemplo, pueden llegar a desarrollarse ciertos mercados multilaterales donde las interacciones se simplifican debido a la facilidad de encuentro y comunicación.

De la misma manera, las plataformas digitales, especialmente las plataformas sociales, se han configurado como entes sumamente importantes en el desarrollo de nuestras vidas. De esta forma, ha sido comprobado que las plataformas sociales necesitan de la actividad y creación

de información constante por parte de sus usuarios para no desaparecer y tener cierta rentabilidad. Debido a que los usuarios de dichas plataformas ostentan un rol importante en su funcionamiento, los contenidos que crean, como por ejemplo la redacción de comentarios, opiniones o reacciones a otros contenidos, se almacenan en las bases de dichas plataformas para que luego éstas puedan usar dichos datos como productos de predicción de conducta para terceros interesados, es decir, para realizar una minería de datos.

Igualmente, los contenidos creados por los usuarios, como los mensajes que se exponen públicamente, tienen la capacidad de llegar a un nivel global ya que las plataformas sociales como Facebook o Instagram se han convertido en gigantes potencias mundiales. Lo anterior abre paso a preguntarnos si las plataformas deberían permitir la publicación de cualquier mensaje por parte de sus usuarios, como un ejercicio de la libertad de expresión, o si es deseable que dichas plataformas cuenten con ciertas políticas que restrinjan algunos mensajes que podrían ser perjudiciales y hasta peligrosos para ciertas personas o comunidades.

Pero, luego de haber definido con precisión el concepto del odio y cómo éste puede transformarse en un discurso peligroso para ciertas poblaciones amenazadas, al igual que se expuso la realidad sobre las plataformas sociales y cómo éstas son transmisoras masivas de mensajes, se debe llevar la investigación al caso de estudio particular. Para este trabajo se decidió tomar el acto de genocidio cometido en Ruanda como antecedente del uso de mensajes de odio y luego exponer el caso de la persecución de la minoría étnica de los Rohingya en Myanmar como punto de partida de lo que puede suceder cuando no se lleva control alguno sobre los mensajes, comentarios y opiniones que se esparcen en dichas redes sociales.

#### **4. Plataformas sociales, discurso de odio y su impacto en la comisión del delito internacional de genocidio**

Desde hace poco tiempo, organizaciones como la ONU, han venido notando y alarmando sobre el creciente aumento de actitudes racistas, xenofóbicas y de intolerancia en general que se presentan a nivel global. En este sentido, la ONU presentó su estrategia y plan de acción para la lucha contra el discurso de odio (2019), en la cual se consigna que

En todo el mundo, estamos presenciando una inquietante oleada de xenofobia, racismo e intolerancia, con un aumento del antisemitismo, el odio contra los musulmanes y la persecución de los cristianos. Se están explotando los medios sociales y otras formas de comunicación como plataformas para promover la intolerancia. Los movimientos neonazis y a favor de la supremacía blanca están avanzando, y el discurso público se está convirtiendo en un arma para cosechar ganancias políticas con una retórica incendiaria que estigmatiza y deshumaniza a las minorías, los migrantes, los refugiados, las mujeres y todos aquellos etiquetados como “los otros”. (Estrategia y Plan de Acción de las Naciones Unidas para la Lucha contra el Discurso de Odio, 2019, p. 1)

Sin duda, todas estas actitudes potencialmente destructivas han visto en las redes sociales una herramienta útil para difundir su odio y encontrar nuevos adeptos que estén de acuerdo con sus discursos. Es por esto por lo que organizaciones internacionales como la ONU se han preocupado por dicha situación, ya que los diferentes discursos de odio pueden presentarse como amenaza a la democracia y a la paz en los estados y posiblemente, pueden desencadenar consecuencias sumamente graves para las víctimas de dichos discursos, como lo son las persecuciones sistemáticas o hasta el genocidio.

#### **4.1. El fenómeno del genocidio**

Pero para poder entender las implicaciones que tiene el discurso de odio dentro del delito internacional de genocidio, se debe primero definir y reflexionar sobre el propio concepto de genocidio y su utilización tanto en el ámbito nacional como en el internacional. Según lo expone Shaw (2007), en sus orígenes, la idea del genocidio estaba completamente ligada a contextos de guerra. Según lo anterior, se expone que la idea principal de Lemkin, jurista al que se le atribuye la acuñación del término genocidio, era establecer el genocidio como un crimen general, subsumiendo en él muchos otros actos que constituían crímenes particulares bajo el derecho internacional humanitario y otros códigos. En este sentido, su concepto era muy diferente del de otras autoridades que hoy en día clasifican al genocidio como un crimen particular junto a los crímenes de guerra y los crímenes de lesa humanidad. Sin embargo, el mismo Lemkin decidió luego extender el concepto de genocidio más allá de los contextos de guerra argumentando que el genocidio no es solo un problema de la guerra sino también de contextos de paz. (p. 462).

Sin embargo, tal como lo expresa Shaw (2007), ya desde el proyecto de lo que se convertiría en la Convención de Genocidio de las Naciones Unidas, el genocidio comenzó a ser definido fuera de los contextos de la guerra. La resolución 96(i) de la Asamblea General de 1946 definió el genocidio como la negación del derecho de existencia de grupos humanos, tal como el homicidio es la negación del derecho a la vida de un individuo. De la misma manera, el artículo I de la Convención para la Prevención y Sanción del Delito de Genocidio de 1948 designó explícitamente al genocidio como un crimen, ya sea cometido en contextos de paz tanto como en contextos de guerra. (p.463, 464).

Según la definición de Shaw (2007), el genocidio ha sido generalmente establecido de acuerdo con una posición subjetiva en cuanto a sus perpetradores, entendida cómo la intención de destruir ciertos grupos sociales. De esta manera, muchos académicos tienden a enmarcar el significado de destrucción dentro de la destrucción física o el asesinato de los miembros de un grupo particular, sin embargo, es claro que un grupo de personas no puede clasificarse como un organismo de la misma manera en que lo puede ser un ser humano. Es por esto por lo que una colectividad es definida según las relaciones y vínculos sociales de los integrantes que conforman dicho grupo. Por lo tanto, la destrucción de una colectividad o grupo no puede ser efectuada simplemente por la destrucción corporal de sus miembros. Esa destrucción implica más la destrucción de las relaciones sociales, instituciones y formas de vida. En este sentido, Lemkin fue muy claro al expresar la diferencia entre la destrucción social, que define el genocidio, y la destrucción física de los miembros de un grupo, la cual es solo una de las formas en las que se puede dar esa destrucción que define al genocidio. De la misma manera, la idea de destrucción social debe involucrar violencia de por medio, la cual se puede dar de muchas formas, no solo con los asesinatos, o la amenaza del uso de la violencia. (p. 464).

Los detonantes o causas que llevan a un genocidio pueden ser diferentes dependiendo de cada contexto y lugar. Dichas causas pueden ser económicas, culturales, sociales, políticas y hasta militares. A pesar de lo anterior, Shaw (2007) opina que es, precisamente, cuando los conflictos políticos se tornan violentos o tienen injerencia militar, que es más probable que se dé un genocidio. Según lo anterior, en principio, cualquier acción militar que sea dirigida contra un grupo social de civiles con la intención de destruirlos, puede ser considerada cómo genocidio. En este sentido, el terrorismo también puede ser considerado como genocida. En

primer lugar, porque generalmente estos grupos terroristas utilizan una retórica genocida. Si bien dicha retórica no es suficiente para establecer una intención genocida concreta, la evidencia material, como el genocidio de Ruanda, deben servir para prestar especial atención a los desarrollos de situaciones o conflictos donde el pensamiento o intención genocida puede llevar a la acción genocida. En segundo lugar, muchas de las acciones de dichos grupos terroristas han ido más allá de la simple propaganda armada impregnada por la retórica genocida a cometer masacres genocidas reales como, por ejemplo, la campaña de violencia contra los musulmanes Chiitas en Iraq liderada por al-Qaeda. (p. 466, 470, 471).

De la misma manera, debe acotarse que el genocidio generalmente no es un fenómeno que se presente sin previo aviso o sin señales que puedan advertir lo que va a suceder. Según Faustin (2016), éste es un crimen que involucra planeación, el cual necesita de participantes los cuales se encuentran condicionados a jugar un rol específico. Las personas interesadas en planear un genocidio generalmente organizan campañas y definen el grupo específico el cuál será su blanco para, de esta manera, poder diseminar un discurso de odio específico contra ellos. Por ejemplo, en Ruanda, el genocidio contra los Tutsi fue precedido y preparado mediante la diseminación de propaganda de odio e incitación a la violencia. De la misma manera, se ha revelado que los discursos de odio jugaron un importante papel en el genocidio perpetrado contra los Tutsi. (p.117).

En este sentido, debe prestarse especial atención a los discursos de odio que se generan contra poblaciones con especial protección, ya que dichos discursos, que incitan a la violencia, pueden configurarse como señales, muchas veces muy claras, de las intenciones de los perpetradores. Por esto, debe prestarse atención a los mensajes que se masifican por medio de las redes sociales, ya que es allí donde pueden conseguir un alcance potencialmente mortal

y perfectamente se pueden dar casos donde se llame a un genocidio mediante las redes sociales.

#### **4.2. Los antecedentes con el genocidio en Ruanda**

Y precisamente, no es descabellado pensar que una de las consecuencias del creciente fenómeno de los discursos de odio en las redes sociales pueda conllevar a consecuencias más graves como la comisión de un genocidio. Esto ya que, anteriormente, se ha visto cómo se ha hecho uso de diferentes medios de comunicación y difusión para esparcir mensajes de odio contra poblaciones determinadas. Un ejemplo de lo anterior puede ser el genocidio en Ruanda cometido por incitación de parte del gobierno hacia todos los miembros de la comunidad Hutu para que exterminaran a los Tutsi. En dicha ocasión, el gobierno de Ruanda se valió del uso de medios de comunicación como el periódico y la radio para difundir mensajes de odio e intolerancia, los cuales incrementaron la tensión ya existente en el país y propiciaron la comisión de un genocidio.

Según lo explican Allen & Norris (2011), el 6 de abril de 1994, el avión en el que viajaba Juvinal Habyarimana, el presidente de Ruanda, fue derribado, por lo que él y otras figuras importantes de la época murieron. Dicho incidente disparó una ola de violencia la cual resultó en la muerte de más de 800.000 personas en aproximadamente tres meses y la mayoría de las víctimas fueron personas pertenecientes a la minoría étnica Tutsi. (p.148) Pero a lo largo de la historia se ha podido documentar cómo, mediante medios de comunicación, agentes oficiales y miembros del gobierno, esparcieron odio e incitaron a la violencia contra la población Tutsi. De acuerdo con lo anterior y refiriéndose a la revista Kangura, la cual fue utilizada para diseminar el odio contra los Tutsi en Ruanda, Faustin (2016), expone que el Tribunal Penal Internacional para Ruanda, o ICTR por sus siglas en inglés, expuso que gran

parte del contenido de dicha revista combinaba el odio en sentido étnico y el uso del miedo con un llamado a la violencia contra la población Tutsi. Por lo cual, la revista Kangura fue uno de los medios que abrieron el camino para el genocidio en Ruanda. (p. 121).

De la misma manera, se presenta la estación radial Radio Television Libre des Milles Collines (RTLM), la cual, según Faustin (2016), tuvo un rol preponderante en el genocidio de Ruanda, el cual fue señalado por la ICTR debido a que dicha estación de radio se encargó de difundir estereotipos étnicos de una manera en la que promovía el desprecio y el odio hacia la población Tutsi. Hizo, igualmente, un llamado a sus escuchas a que tomaran las armas en contra de sus enemigos, los cuales eran identificados como cucarachas. El 6 de abril de 1994 los llamados incitando a la violencia incrementaron, durante las transmisiones se viralizó un llamado explícito para que se procediera con la exterminación del grupo étnico de los Tutsi. (p. 122).

Junto con estos dos medios de comunicación anteriormente mencionados, se encontraba también la estación de radio llamada Radio Rwanda, la cual diseminaba, igualmente, una propaganda de odio para incitar a la violencia. Tal como lo explica Faustin (2016), Radio Rwanda fue usada para promover los asesinatos de la población Tutsi en la región de Bugesera. Antes de que sucedieran las masacres, dicha estación de radio transmitió anuncios advirtiendo que la población Hutu sería atacada por los Tutsi en Bugesera. En este sentido, los oficiales locales inventaron dicho anuncio transmitido mediante la radio para convencer a los Hutu de que necesitaban protegerse atacando antes que los Tutsi. Liderados por soldados de una base militar cercana, civiles pertenecientes a la población Hutu atacaron y asesinaron a miles de personas pertenecientes a la comunidad Tutsi. (p. 122).

Todos estos medios de comunicación decidieron difundir mensajes incendiarios incitando a la violencia resguardándose en el nombre del derecho a la libre expresión. De esta manera, se valieron de varios métodos que implicaban la utilización del discurso de odio para incitar a la población Hutu a cometer un genocidio. Un ejemplo de esto es la utilización, por parte de todos los medios, de lenguaje simbólico, el cual se utilizó para humillar y deshumanizar a las personas pertenecientes a la comunidad Tutsi, ya que se usaban palabras tales como “cucarachas”, “serpientes”, “subversivos” y “el enemigo”, para referirse a las personas de dicha comunidad. (Faustin, 2016, p. 123). Esta es una estrategia común que se implementa dentro de los discursos de odio, porque al comparar a la población víctima de dichos discursos con características animales o con un valor de inferioridad, se logra deshumanizar a las personas. De esta manera, cuando se presenta el llamado al genocidio, los perpetradores sienten que están cumpliendo su deber y pierden toda clase de empatía o sentimientos de arrepentimiento para con las víctimas ya que se encuentran convencidos de que están cumpliendo con su deber.

Según Faustin (2016), el discurso de odio se ha presentado como una parte integral de cualquier persecución, especialmente las organizadas por estados, y sirve principalmente para preparar a la población de un estado para realizar o tolerar ciertos crímenes. En Ruanda, antes de la comisión del genocidio, el discurso de odio y la propaganda de odio se presentaron como instrumentos políticos para convencer a la población Hutu de cometer asesinatos masivos. En este sentido, la población Hutu fue movilizada mediante el desarrollo de la idea de que los problemas por los que estaban pasando y por los que pasaba el país no se resolverían a menos que la población minoritaria Tutsi fuera exterminada. (p.123).

Finalmente, los reportes del Reporte especial de la ONU para Ruanda y la Comisión de Expertos establecida por el Consejo de Seguridad de la ONU concluyeron que en Ruanda se cometió el delito internacional de genocidio. Según Allen & Norris (2011), en el caso de Jean-Paul Akayesu, el Tribunal Penal Internacional para Ruanda fue el primero en interpretar la definición de genocidio y se dio la primera condena desde la adopción de la Convención para la Prevención y la Sanción del Delito de Genocidio de 1948. Igualmente, el primer ministro Jean Kambanda se convirtió también en el primer líder estatal en ser condenado por el delito de genocidio. De la misma manera, la ICTR encontró que las transmisiones realizadas por la estación radial RTLW después del 6 de abril de 1994 contribuyeron sustancialmente a la comisión de los actos de genocidio que se dieron en Ruanda. Por otro lado, la ICTR también condenó a Hassan Ngeze, fundador del periódico Kangura, por el delito de incitación directa y pública a cometer genocidio basándose en lo que contenían algunos artículos de dicho periódico, varios escritos y firmados por él. (pp.151-152).

#### **4.3. El riesgo del discurso de odio en la práctica: El caso de los Rohingya en Myanmar**

Uno de los casos que han dado la vuelta al mundo debido al mal manejo por parte de Facebook ante situaciones potencialmente peligrosas es el caso de la persecución del grupo minoritario musulmán de los Rohingya en Myanmar. En este caso particular, más de un millón de personas pertenecientes al grupo de los Rohingya tuvieron que escapar, desde el 2017, de Myanmar hacia Bangladesh debido a la persecución que se estaba llevando a cabo el gobierno y la población birmana contra dicho grupo.

Dicho grupo es un grupo apátrida asentado en Myanmar. Debido a la persecución que se ha dado por parte del gobierno, militares y habitantes de Myanmar, el pasado 2017 los Rohingya

tuvieron que emprender un exhaustivo éxodo hacia Bangladesh. Pero, para llevar a cabo un correcto análisis de dicha coyuntura, es preciso anotar algunos datos importantes sobre Myanmar y su conflicto étnico-religioso interno. Cómo bien lo expone Calles (2018),

La República de la Unión de Myanmar es un Estado situado en el extremo noroeste de la subregión del Sudeste Asiático, fronterizo con Tailandia, Laos, China, India y Bangladés, con salida en el sur al Golfo de Bengala y al mar de Andamán. Su capital actual es Nay Pyi Taw. (pp. 7).

En cuanto a la historia de Myanmar, de acuerdo con Calles (2018), se tiene que fue colonizada por el Estado británico entre 1824 y 1948. Durante este período se iniciaron diferentes luchas entre etnias que estaban asentadas en este país a causa de la política implementada por los británicos de introducir población extranjera cómo mano de obra, la cual, en muchas ocasiones, ocupaba los territorios de etnias autóctonas, irrespetando así las costumbres y tradiciones birmanas. (pp. 8). Posteriormente, tal como lo documenta Torres (2016), “en 1942, como consecuencia de la II Guerra Mundial, Japón invadió Birmania y expulsó a los británicos de Arakan, lo que obligó a muchos Rohingya a desplazarse al este bengalí.” (pp. 2) No está muy claro cuál es el verdadero origen del odio de los budistas contra los Rohingya, pero si se sabe que el gobierno, los militares y los ciudadanos birmanos consideran a dicho grupo minoritario cómo inmigrantes ilegales. Es por esto por lo que el gobierno se niega a concederles la nacionalidad y no poseen derechos fundamentales como el derecho al voto. En este sentido, se ha encontrado que:

La clasificación como “inmigrantes ilegales bengalíes” se debe a que durante la ocupación colonial de Gran Bretaña (1824-1948), la cual también poseía la India, así como lo que sería posteriormente Pakistán, y Bangladés, fomentó una política de

movimiento poblacional de Bangladés y la India a Birmania y de esta forma obtener más mano de obra en el país. Muchos de estos inmigrantes se asentaron en el Estado de Rakhine, pero, tal y como expone la organización Human Rights Watch, estas migraciones no eran ilegales ya que se encontraban bajo el mismo poder y orden colonial. (Calles, 2018, pp. 19-20)

Por otro lado, Calles (2018) propone que dicho conflicto estalló cuando los militares llegaron al poder ya que tomaron el budismo como punto de unión e identidad birmana. (pp. 19) Esto llevó a que se segregaran otras comunidades, especialmente las musulmanas, pero dentro de este grupo, especialmente la minoría de los Rohingya ya que previamente habían sido también considerados como inmigrantes ilegales de Bangladesh y desde el año 1982, han sido privados de tener una nacionalidad, violando así los derechos humanos.

A pesar de todo lo anterior, los Rohingya se consideran ciudadanos de Myanmar y niegan tener procedencia bengalí o ser inmigrantes ilegales. Al respecto,

afirman que habitaban en el territorio del actual Estado de Rakhine, conocida anteriormente como el Estado de Arakan, antes de la colonización británica. De acuerdo a los mismos, son descendientes de comerciantes árabes que se asentaron en la zona manteniendo su propia lengua y cultura. (Calles, 2018, pp. 20)

Grupalmente, como sociedades, tenemos sentidos de los lugares que compartimos, es decir, constantemente construimos sentidos de lugar colectivos que implican una gran variedad de significados asociados como a memorias, hábitos o emociones. Al mismo tiempo, pueden existir diferentes concepciones de lugar sobre un mismo espacio y dichas concepciones pueden habitar en armonía o, por el contrario, pueden entrar en pugna. En este sentido, se

encuentra que los Rohingya han acogido un espacio, específicamente el Estado de Rakhine, y lo han dotado de un sentido de lugar muy arraigado a sus costumbres y su manera de ver el mundo, y por este hecho, se sienten parte de dicho territorio. Pero, por el contrario, el Estado de Myanmar, los militares y los birmanos en general tienen una idea muy diferente de dicho territorio, han creado una idea de lugar donde es inaceptable que los Rohingya se asienten, ya que, según ellos, son inmigrantes ilegales y, por lo tanto, no pueden habitar dicho territorio. A pesar de que los Rohingya tienen la intención de que Rakhine sea considerado su lugar en el mundo, el Estado birmano ya tiene creado su propio sentido de lugar sobre este territorio, y es por esto por lo que afirman que tienen el derecho a excluir a dicha población, precisamente porque son el Estado y ostentan el poder.

Lo anterior denota también la existencia de un propósito territorializador por parte del Estado de Myanmar. Para lograr lo anterior, se valen del derecho, expidiendo normas como la ley de Ciudadanía de 1982, la cual privó de la nacionalidad birmana a los Rohingya. Esto se lleva a cabo ya que el gobierno, claramente, tiene el propósito de afectar y controlar a dicho grupo minoritario, delimitando y generando control sobre un área específica. Este es el resultado de la territorialización, lo cual denota que existe en Myanmar una competencia específica por las relaciones en el espacio, el cual en este caso se presenta como el Estado de Rakhine, donde el Estado no quiere ver, o simplemente es indiferente a lo que está sucediendo dentro de dicho territorio. Pero no solo es una competencia por las relaciones, es también una competencia por las relaciones de poder, en este caso, se está territorializando una problemática, y específicamente está utilizando la estrategia del vaciamiento, donde le resta toda importancia a las relaciones de los humanos que se encuentran en dicho lugar y, en este sentido, se ocupa de vaciar conceptualmente ese territorio que para los Rohingya tiene un

espacial significado cultural, por eso los desampara y hasta los priva de sus derechos fundamentales y humanos.

En este sentido, puede verse la territorialización también como una

“tecnología que permite” la división territorial del continuo social espacial en dicotomías espaciales como adentro/ afuera, facilita la polarización de una gama continua de colores en solo blancos y negros y al hacerlo facilita que se congelen o establezcan las identidades. (Castro, 2020, pp. 60)

El gobierno de Myanmar ha creado una división muy marcada de lo que es ser birmano, y ha dejado muy en claro que los Rohingya no caben dentro de esta definición. Pero esto no solo lo ha hecho refiriéndose a identidades particulares, sino también creando territorios y sometiendo a estos grupos marginados a que se asienten en éstos. El gobierno ha creado campos de concentración donde deben permanecer los Rohingya, es así como los marginan. Esto hace que toda la población cree un imaginario colectivo de siente que, si una persona se encuentra en esos espacios, es porque pertenece a la etnia de los Rohingya, y así, mediante el territorio, se crea toda una idea de identidad, pero una identidad nociva y discriminatoria debido a que, en todo el país, son vistos como inmigrantes ilegales.

Según lo expuesto en las conclusiones de la Misión Investigadora nombrada por el Consejo de Derechos Humanos de las Naciones Unidas, divulgadas en septiembre del 2019, se concluyó que hay una fuerte inferencia de intentos de genocidio continuados por parte del Estado de Myanmar en relación con los Rohingya. Igualmente, hay un gran riesgo de que se cometa un genocidio en contra de dicha población y, por lo tanto, Myanmar no está cumpliendo con sus obligaciones contraídas bajo la Convención para la Prevención y la

Sanción del Delito de Genocidio, de investigar el genocidio y de promulgar legislación efectiva que criminalice y castigue el genocidio. (Misión Investigadora nombrada por el Consejo de Derechos Humanos de las Naciones Unidas, ONU, 2019.)

Tal como lo exponen Byman & Joshi (2020), hoy en día, las plataformas sociales juegan el rol que anteriormente poseían los medios tradicionales de la radio y los periódicos, pero dichas plataformas tienen algunas características especiales que representan nuevas oportunidades para sus usuarios. Por ejemplo, es más probable que ciertas plataformas sociales posean monopolios naturales. Parte de las razones para estar en Facebook o Instagram es que muchas otras personas también son usuarios de dichas plataformas. En efecto, en muchos países como Myanmar, plataformas sociales como Facebook son consideradas como el internet, esto significa que, para los ciudadanos de dichos países, estar en línea significa simplemente usar Facebook o WhatsApp. Debido a este monopolio del mercado que poseen ciertas redes sociales en algunos países y la capacidad de penetración pública que tienen, la información que se difunde por ejemplo en Facebook en Myanmar tiene un gran impacto en dicha sociedad. Debido a lo anterior, es mucho más fácil que un número pequeño de individuos esparza rápidamente ideas y rumores y hasta discursos de odio a otros usuarios mediante las redes sociales y así, dicho mensaje se reproduce y llega a otros usuarios que lo seguirán pasando indefinidamente. Dado que la exposición repetitiva a los rumores refuerza su credibilidad y aumenta la posibilidad de que sea reproducido y utilizado para informar, la capacidad de realizar un gran número de publicaciones en las redes sociales desde múltiples cuentas diferentes representa una manera útil y económica de reforzar dichos rumores. Es de esta manera como diferentes regímenes como Azerbaiyán o Guatemala han suministrado deliberadamente noticias falsas e información mal intencionada

contra sus oponentes políticos para restarles credibilidad y legitimidad entre su población. (p. 129, 130). Esta fue, efectivamente, la misma técnica que utilizó el gobierno de Myanmar para esparcir noticias falsas sobre la minoría religiosa de los Rohingya. Noticias falsas como por ejemplo que insurgentes Rohingya asesinaron a 12 oficiales de seguridad. Dichas noticias fueron difundidas por militares y miembros del gobierno por medio de Facebook para acrecentar las tensiones ya existentes y hacer un llamado a la población para que reaccionaran con violencia contra dicho grupo minoritario.

Tal como sucedió con los momentos anteriores al genocidio en Ruanda, las autoridades y el gobierno de Myanmar se han ocupado de difundir discursos de odio utilizando los medios de comunicación. La única diferencia que se puede encontrar en este caso es que, ahora, estas personas tienen a la mano el uso de las redes sociales, a parte de los otros medios de comunicación tradicionales. De esta manera, se crea un peligro de intención de genocidio más inminente, ya que la capacidad de difusión de estos mensajes es más directa y puede llegar a los usuarios más masivamente. Aparte de esto, el hecho de que los mensajes de odio provengan de perfiles oficiales o de militares hace que el llamado cobre cierta legitimidad ya que proviene de personas que ostentan el poder en el país. Al mismo tiempo, al crear pánico por medio de estos mensajes, como por ejemplo afirmando que los Rohingya suelen atacar y asesinar a ciudadanos budistas, crean cierto caos y un sentimiento de zozobra entre la comunidad, haciéndolos pensar que deben actuar primero, antes de que la comunidad a la que el gobierno pretende victimizar, lo haga. Pero no solo eso, el hecho de que el mismo gobierno de Myanmar sea reticente a entablar una comunicación internacional con diferentes organismos que puedan ayudar en esta peligrosa situación habla mucho sobre sus intenciones. Tal como lo expresan las conclusiones de la Misión Investigadora nombrada por el Consejo

de Derechos Humanos de las Naciones Unidas, la catástrofe contra los derechos humanos continúa. El gobierno de Myanmar parece despreocupado frente a lo anterior. Esta misión cree firmemente que el monitoreo internacional y el reporte público son esenciales para garantizar una defensa de facto. Myanmar continúa en un estado de negación, de hecho, ha rechazado cualquier posibilidad de dialogo con nosotros y esto es profundamente lamentable. Esto demuestra una falta total de buena fe hacia los mecanismos de protección de los derechos humanos, incluyendo esta misión. (Misión Investigadora nombrada por el Consejo de Derechos Humanos de las Naciones Unidas, ONU, 2019.)

Pero, cómo antes ya se mencionó, el gobierno birmano y sus militares no son los únicos actores que han tenido parte en estas violaciones de derechos humanos, tal como lo afirman Zamora & Murillo (2020),

En relación con los actores que están teniendo un rol destacado en las graves violaciones contra la etnia Rohingya, cabe mencionar a otros intervinientes que de una u otra forma han propiciado el conflicto o han proporcionado las herramientas para difundir la campaña de odio contra esta etnia. Nos referimos, por ejemplo, al papel desempeñado por la red social Facebook. Esta plataforma ha servido como herramienta del Gobierno, del Ejército birmano y de los radicales budistas, para la difusión de mensajes de odio y de discriminación. Por estos hechos, está siendo ahora investigada por no haber tomado las medidas necesarias para prevenir la difusión de los mensajes de odio y no haber actuado de forma contundente contra la propaganda de los militares. (pp. 469)

En este sentido, puede afirmarse que Facebook, cómo plataforma social, es un actor que tiene plena responsabilidad debido a su mal manejo de sus propias políticas contra el discurso de

odio. Esto porque permitió que se difundieran mensajes que atentaban contra el bienestar de los Rohingya, lo cual, a su vez, ayudó a acrecentar el sentimiento anti Rohingya ya existente en Myanmar. Como lo exponen Zamora & Murillo (2020), se crearon páginas y perfiles en Facebook por parte de nacionalistas budistas, como Ashin Wirathu, que contenían mensajes violentos contra dicho grupo minoritario. Dichos mensajes tuvieron un impacto muy importante sobre la comunidad birmana y sobre los Rohingya, (pp. 470) ya que incluso las autoridades religiosas del país actuaron como incitadores del odio y la respuesta de Facebook fue completamente tardía, cuando todo el daño ya estaba hecho y los mensajes ya habían sido difundidos entre toda la comunidad.

Y todo lo que se ha visto al respecto de la violencia contra los Rohingya, indica que la intención de los perpetradores no era simplemente la de expulsarlos del país, en las palabras de la Asesora Especial de las Naciones Unidas para la Prevención del Genocidio, seamos claros: crímenes internacionales fueron cometidos en Myanmar. Los musulmanes Rohingya han sido asesinados, torturados, violados, quemados vivos y humillados por el simple hecho de ser quienes son. Toda la información recibida indica el intento de los perpetradores de realizar una limpieza étnica del norte del estado de Rakhine o, posiblemente, de destruir a los Rohingya, lo cual, si es probado, sería constituyente del crimen de genocidio. Sin embargo, ya sea que se pruebe que hubo un genocidio o no, esto no debe retrasar la urgencia que existe de actuar inmediatamente. Lo anterior es una deuda con la población Rohingya. (Dieng, 2018.)

Igualmente, la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018), expone que se ha encontrado que las redes sociales, en particular Facebook, jugó un rol determinante en cuanto

a la difusión de discursos de odio y desinformación, lo cual llevó a la incitación a la violencia contra la minoría étnica de los Rohingya. El 10 de abril del 2018, durante un testimonio en el congreso, el senador Patrick Leahy cuestionó a Mark Zuckerberg y lo cuestionó sobre por qué las publicaciones no pudieron ser removidas dentro de las 24 horas siguientes a su publicación. Igualmente, citó las repetidas solicitudes que realizaron diferentes organizaciones civiles de Myanmar al grupo de Facebook para que detuvieran la diseminación de dichos mensajes. Zuckerberg respondió prometiendo que se haría mucho más para combatir los discursos peligrosos, incluyendo la contratación de docenas más de personas que pudieran hablar birmano para revisar el contenido que se sube a la red social. Igualmente, hizo un énfasis continuo en el uso de Facebook de herramientas de inteligencia artificial para combatir este tipo de contenidos. (pp.1)

Asimismo, la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018) expone que, este caso resalta la relación entre el sector privado y los mecanismos de las Naciones Unidas y el grado en el que las compañías están dispuestas a comprometerse con los principios rectores sobre las empresas y Derechos Humanos de la ONU. Este caso ilustra cómo las compañías pueden necesitar ser más proactivas y consientes al momento de evaluar y remediar los potenciales impactos que sus productos puedan tener sobre los derechos humanos en general. (pp.2). Esto ya que, tal y cómo lo exponen dichos principios rectores, las empresas, domiciliadas en cualquier lugar del mundo, no están exentas de respetar y proteger los derechos humanos de terceros, con independencia de la voluntad de los Estados. Y, a su vez, deben hacerse responsables a las consecuencias o infracciones a derechos humanos que puedan surgir en razón a su actividad. De la misma manera, la carta de los Principios Rectores sobre las

Empresas y Derechos Humanos de la ONU del 2011 se declara, en el principio número 12, que:

De acuerdo con las circunstancias, es posible que las empresas deban tener en cuenta otras normas. Por ejemplo, las empresas deben respetar los derechos humanos de las personas pertenecientes a grupos o poblaciones específicos y deberán prestarles una atención especial cuando vulneren los derechos humanos de esas personas. Los instrumentos de las Naciones Unidas han detallado a tal efecto los derechos de los pueblos indígenas, las mujeres, las minorías nacionales, étnicas, religiosas y lingüísticas, los niños, las personas con discapacidad y los trabajadores migrantes y sus familias. (pp 16-17)

De esta manera, se evidencia cómo grupos minoritarios o étnicos como los Rohingya, tienen especial protección por su condición especial de vulnerabilidad y que esto no aplica solo para los Estados, sino que las empresas también deben velar por proteger su bienestar y respetar sus derechos. En este caso, se tiene que Facebook no cumplió con esta obligación especial al no revisar cabalmente todas las publicaciones que se crearon en birmano dónde se abogaba por el discurso de odio y la violencia contra los Rohingya. Y al no revisar dichas publicaciones, tampoco se dieron al trabajo de eliminar, en las 24 horas siguientes a su publicación, los mensajes potencialmente peligrosos, los cuales luego sirvieron como gasolina para el fuego que se estaba creando en cuanto al ataque contra dicha población. Uno de los dilemas a los que frecuentemente se enfrentan redes sociales como ésta es el impacto que las censuras de mensajes y publicaciones pueden traer al derecho a la libertad de expresión. El problema radica en saber cuándo se está frente a un discurso de odio que debe ser controlado o cuándo se está frente al ejercicio del derecho legítimo a la libertad de

expresión. Según la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018), mientras que el derecho a la libre expresión implica el derecho a estudiar, debatir y criticar ideas, opiniones y creencias, por más irrazonables que sean, el mismo no defiende el odio que incite a la hostilidad, discriminación o violencia contra un individuo o un grupo de individuos. De hecho, existen estándares internacionales que prohíben ciertos discursos que infringen en la igualdad y los derechos de los otros, incluyendo el derecho a vivir libre de discriminación y miedo, la libertad de culto y el derecho a la autodeterminación. (pp.2).

En conclusión, si se hace la pregunta sobre si Facebook tiene algún tipo de responsabilidad en la violencia y la posible incitación al genocidio del grupo musulmán de los Rohingya, se debe responder que sí. Los Estándares Comunitarios de Facebook tienen una definición muy clara de lo que es el discurso de odio:

Definimos el lenguaje que incita al odio como un ataque directo a las personas, y no a los conceptos ni a las instituciones, en función de lo que denominamos "características protegidas": raza, etnia, nacionalidad, discapacidad, religión, casta, orientación sexual, sexo, identidad de género y enfermedad grave. Definimos un ataque como lenguaje violento o deshumanizante, estereotipos dañinos, declaraciones de inferioridad, expresiones de desprecio, repulsión o rechazo, insultos, o incitaciones de exclusión o segregación. También prohibimos el uso de estereotipos dañinos, que definimos como comparaciones deshumanizantes que se usaron históricamente para atacar, intimidar o excluir a grupos específicos, y que suelen vincularse con la violencia en la vida real. (2021)

De la misma manera, expresan que no permiten “el lenguaje que incita al odio en Facebook, ya que crea un entorno intimidatorio y excluyente que, en algunos casos, puede incitar a la violencia en la vida real.” (Estándares Comunitarios de Facebook, 2021, parte 3) Todo lo anterior concuerda con la situación que tuvo lugar en Myanmar. Es decir, se dieron ataques directos contra la comunidad de los Rohingya en razón a su etnia, nacionalidad y religión mediante lenguajes de odio y que incitaban a cometer actos violentos contra dicha comunidad, y, sin embargo, Facebook no tomó las acciones pertinentes que en el momento se necesitaban para que dichos mensajes de odio no se materializaran, como efectivamente sucedió.

Este caso concreto es útil para ilustrar los peligros que se esconden detrás de los discursos de odio que se difunden en las redes sociales sin control alguno. Igualmente, muestra cómo actualmente siguen existiendo muchas dudas y lagunas en cuanto a cómo se debería manejar este problema particular, porque a partir de lo que sucedió con los Rohingya, se evidencia que efectivamente, los mensajes de odio difundidos por las redes sociales pueden ser un detonante para la comisión de un genocidio, y esto deben tenerlos en cuenta tanto los Estados como los organismos internacionales y las empresas de tecnología. Urge seguir preguntándose sobre la regulación existente, pero en clave de su real utilidad. En realidad ¿podemos depender 100% en los mecanismos de Inteligencia Artificial? O deberíamos explorar otros métodos que puedan complementar dichos mecanismos pero que también tengan en cuenta los contextos de cada territorio y sociedad a la hora de juzgar que discursos pueden ser potencialmente peligrosos y ameriten restringir el derecho a la libre expresión debido a las posibles repercusiones que puedan tener sobre poblaciones o grupos específicos.

En el siguiente capítulo se esbozan los mecanismos regulatorios vigentes y su utilidad en cuanto a prevención y sanción del genocidio. Igualmente, se analiza el discurso y uso de la Inteligencia Artificial y hasta qué punto es seguro depender de ésta para identificar discursos que puedan incitar o poner en acción la comisión del delito de genocidio. O hasta que punto es deseable repensar una solución a los problemas de los discursos de odio que pueda contar con características tan importantes como la lectura del entorno, la consideración de los hechos culturales y la ética.

## **5. Retos frente a la regulación del discurso de odio difundido mediante plataformas sociales**

Cómo ya se ha mencionado anteriormente, el tema de la regulación de los discursos de odio, y en general, de los mensajes que se publican en las redes sociales, tanto como la regulación para dichas empresas en sí es un tema que ha suscitado discusiones difíciles y aún queda bastante trecho que recorrer en este sentido. Debido a que las compañías que controlan redes sociales como Facebook se han convertido en poderosas multinacionales, es difícil hablar de una regulación que pueda aplicarles completamente. Sin embargo, la comunidad internacional se ha dado cuenta que, cada vez más, éste es un tema que se necesita abordar desde diferentes perspectivas, entre éstas, la perspectiva internacional, la nacional pero también se deben explorar las alternativas de la auto regulación que, de hecho, se está presentando actualmente.

Para este capítulo, se recopilarán las diferentes regulaciones internacionales que hasta ahora existen en el tema. También se expondrán los diferentes esfuerzos que las compañías de plataformas sociales han llevado a cabo por proveerse sus propias regulaciones en cuanto al

tema y, hasta qué punto, dichas autorregulaciones han sido satisfactorias para resolver los problemas de la publicación de discursos de odio y la incitación a la violencia.

### **5.1. Regulación Internacional:**

En primer lugar, se debe mencionar la Declaración Universal de Derechos Humanos (1948) dónde se consigna, en su artículo 1, que “todos los seres humanos nacen libres e iguales en dignidad y derechos y, dotados como están de razón y conciencia, deben comportarse fraternalmente los unos con los otros.” Éste y otros artículos de la Declaración Universal de Derecho Humanos consignan un precedente muy importante para el caso que se está analizando, pues proclaman la igualdad que tienen todos los seres humanos a disfrutar de diferentes derechos, entre éstos, el derecho a la no discriminación.

A pesar de esto, aún no existía una regulación que se encargara específicamente de hablar sobre el discurso de odio y la protección que se debía dar a los grupos protegidos en contra de éste. Fue solo hasta 1965, con la Convención Internacional sobre la Eliminación de todas las Formas de Discriminación Racial (CERD por sus siglas en inglés), que se comenzó a hablar sobre este tema y a otorgarle la importancia debida. Según Mendel (2010), es útil distinguir los cuatro diferentes aspectos de las obligaciones contraídas por el CERD en cuanto a desarrollar el significado del discurso de odio. Dichos aspectos se encuentran en su artículo 4 (a) de la siguiente manera: 1. La diseminación de ideas basadas en la superioridad racial. 2. La diseminación de ideas basadas en el odio racial. 3. La incitación a la discriminación racial y 4. La incitación a cometer actos de violencia racial. Igualmente, el artículo 4 (c) habla sobre las prohibiciones que tienen las autoridades públicas o instituciones de promover o incitar a la discriminación racial. (pp.2) Aunque el CERD, según lo expuesto por Alkiviadou (2019), no se ocupa específicamente del problema de los discursos de odio en línea, si prohíbe

ciertos tipos de expresiones. Particularmente, sostiene que los Estados deben declarar cómo ofensas criminalizables por sus propias leyes, todas las diseminaciones de ideas basadas en la superioridad racial o el odio, la instigación a la discriminación racial, así como todos los actos de violencia o incitación a dichos actos contra cualquier grupo de personas con especial protección. (pp.26-27)

De igual manera, se presenta el Pacto Internacional de Derechos Civiles y Políticos, el cual fue adoptado por la Asamblea General de las Naciones Unidas en 1966. En su artículo 20 (2) se consigna que “toda apología del odio nacional, racial o religioso que constituya incitación a la discriminación, la hostilidad o la violencia estará prohibida por la ley.” (Asamblea General de las Naciones Unidas, 1966). En distinto sentido, todos los tratados regionales de derechos humanos, la Convención Europea de Derechos Humanos, la Convención Americana sobre Derechos Humanos y la Carta Africana de Derechos Humanos y de los Pueblos consignan y protegen el derecho a la libertad de expresión, pero, sorpresivamente, solo la Convención Americana sobre Derechos Humanos hace énfasis en la prohibición del discurso de odio en su artículo 13(5) de la siguiente manera:

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.

Todos los documentos internacionales anteriormente mencionados fueron diseñados y suscritos en una era anterior a la aparición y éxito del internet, es por esto por lo que, en ninguno de los anteriores, se menciona el problema del discurso de odio en línea, sino que simplemente se centran en atacar los discursos racistas y la discriminación religiosa.

(Alkiviadou, 2019, pp. 27) Sin embargo, tal como lo menciona Alkiviadou (2019), aunque los discursos de odio existen hace mucho tiempo, el surgimiento del internet y la creación de las redes sociales han añadido nuevas dimensiones al tema, ya complejo, del discurso de odio. (pp.19)

En este sentido, se puede expresar que no existe mucha regulación internacional que se ocupe específicamente del tema de los discursos de odio en línea o que se den en redes sociales. Sin embargo, en el ámbito europeo se han presentado algunos instrumentos internacionales interesantes los cuales abren paso para el desarrollo en este campo. Según lo expresa Clough (2012), mientras que puede ser extremadamente positiva, también está dentro de la naturaleza humana el encontrar maneras de explotar la tecnología para propósitos oscuros. El uso de la tecnología en la comisión o facilitación del crimen ha existido desde la aparición misma de los avances tecnológicos. Sin embargo, fue la proliferación de las computadoras personales en los años 1980s y la disponibilidad del internet en los 1990s los que crearon el ambiente del crimen cibernético que conocemos actualmente. La Naturaleza de interconexión de la tecnología, significa que éste es un problema global. Y dada la globalidad de dicho problema, es generalmente aceptado que un cierto nivel de armonización entre países es vital si se quiere alcanzar una regulación efectiva. Fue un convenio regional, El Convenio sobre la Ciberdelincuencia, el cual emergió como el primero, y hasta ahora el único, instrumento multilateral que regula el crimen cibernético. (pp. 365-366). De acuerdo con Nemes (2002), el Consejo Europeo tomó el problema del discurso de odio muy seriamente al redactar el Convenio anteriormente mencionado. Dicho instrumento estuvo disponible para su firma en noviembre del 2001 y entró en vigor el 1 de julio de 2004. Como lo afirma Clough (2012), dicho convención adopta un modelo tradicional de cumplimiento de la ley, predicado en

Estados que procesen daños ocurridos en sus propias jurisdicciones y busca adaptar ese modelo a un ambiente digital. Provee una respuesta comprensiva al desafío de los crímenes cibernéticos, direccionándolos a asuntos de ofensas sustantivas, leyes procedimentales y la cooperación internacional. (pp. 369).

Por otro lado, según lo afirmado por Nemes (2002), el 8 de noviembre de 2001, el Consejo Europeo decidió unánimemente respaldar el Convenio sobre la ciberdelincuencia con un protocolo adicional que define y lucha contra los discursos de odio en línea específicamente. Las naciones que adopten dicho protocolo deberán criminalizar la comisión de actos racistas o xenofóbicos que se lleven a cabo mediante sistemas computacionales. El protocolo hace obligatorio para las naciones que lo ratifiquen el promulgar legislación doméstica que convierta los actos proscritos en delitos, así como promulgar provisiones que refuercen lo anterior. (pp.198) Según Rollnert (2020),

Entre estas conductas se encuentran las amenazas de comisión de delitos graves, los insultos, la difusión o puesta a disposición del público de material negacionista o justificador de genocidios o crímenes contra la humanidad y, especialmente, la difusión o puesta a disposición del público de «material racista y xenófobo. (pp.3)

Tal cómo lo afirma Rollnert (2020) dicho protocolo es el único instrumento internacional que contempla específicamente las expresiones o discursos de odio en internet. El objetivo de este protocolo es armonizar la legislación penal existente referente a la lucha contra la propaganda que pueda ser racista o xenofóbica. (pp.3) El protocolo también proporciona ciertas luces a lo que debe entenderse cómo material racista o xenófobo. Según Rollnert (2020), el protocolo dicta que:

«todo material escrito, toda imagen o cualquier otra representación de ideas o teorías, que propugne, promueva o incite al odio, la discriminación o la violencia» (artículo 2.1). Afirma el Informe que se penaliza la difusión de «ideas y teorías» en cualquier formato (escrito, imágenes o cualquier otra representación de «ideas o teorías» almacenable, procesable y transmisible por medios informáticos), no tanto porque sea «expresión de sentimientos/creencias/aversión» sino porque puede llevar a «cierta conducta» (pág. 3) en la medida que «propugne, promueva o incite al odio, la discriminación o la violencia». Dicho de otra forma, la relevancia penal del material radica, más que en lo que expresa, en las acciones que podría provocar en terceras personas, en su efecto perlocutivo. (pp. 3-4)

Igualmente, Rollnert (2020) sostiene que:

Con carácter general, el Protocolo anuda la responsabilidad penal a que las conductas sean cometidas «intencionalmente», si bien, como señala el Informe, en ciertos casos se exige una intención específica adicional; así, en la negación o justificación del genocidio o de crímenes contra la humanidad, cabe condicionar la penalización a la presencia de la «intención de incitar» al odio, discriminación o violencia; y en la cooperación y la complicidad (artículo 7), el cooperante o cómplice debe tener también la intención de que el delito sea cometido. Y respecto a la publicidad de la conducta, exigida por la propia naturaleza de los delitos de «difundir o poner a disposición del público» material racista y xenófobo o material negacionista o justificador de genocidios o crímenes contra la humanidad, el Informe define la difusión como la «divulgación (dissemination) activa» y la puesta a disposición como la acción de subir material a internet para uso de terceros, incluyendo la creación o

compilación de hiperenlaces que faciliten acceso al mismo. La exigencia de publicidad se incorpora también al delito de insultos racistas o xenófobos que serían penalmente atípicos en comunicaciones privadas. Por contra, en las amenazas racistas o xenófobas, el tipo delictivo se extiende también a las realizadas en comunicaciones privadas. (pp.4)

A pesar de todo lo anterior, se encuentra que, aunque el protocolo adicional innova en el sentido de regular los discursos de odio racistas y xenófobos en línea, deja por fuera otras categorías importantes como lo son los discursos de odio con connotaciones religiosas, por discriminación de identidad sexual, entre otros.

## **5.2. Autorregulación:**

Cómo se vio anteriormente, el avance de la regulación internacional en cuanto al tema de los discursos de odio en redes sociales y sus posibles impactos, ha sido lento y se ha concentrado en ámbitos regionales más que globales. Es por esto, por lo que se ha considerado la posibilidad de que las propias plataformas se den, ellas mismas, sus propias normas de funcionamiento y estándares para regular el comportamiento de sus usuarios. De esta manera, desde las propias redes sociales se crean definiciones de conductas y prohibiciones a conductas específicas, las cuales cuentan con sus respectivas sanciones, ya sea que se deba borrar la publicación hecha por el usuario o que se bloquee la cuenta por un tiempo determinado o indefinidamente. Las propias plataformas determinarán la gravedad de la conducta, y con esto, el alcance de la respectiva sanción.

Tal cómo lo expone Nemes (2002), el internet no tiene, por sí solo, ninguna forma especial de gobernanza. (pp. 201) Esto significa que, si se quiere resolver el conflicto de lo que se

publica en las redes sociales, es beneficioso pensar de una manera más abierta y alejarse un poco de los métodos tradicionales cómo lo son la regulación internacional o nacional. En este caso, puede existir un beneficio en la autorregulación ya que, de alguna manera, puede ser más expedita ya que, al ser el internet un lugar en constante evolución se puede beneficiar del tipo de regulaciones más flexibles, que acepten cambios y que puedan adaptarse a la velocidad en la que corren las mismas redes sociales y sus usuarios. Debe considerarse que cada red social tiene su propia singularidad y cumple un propósito específico en el mundo de sus usuarios. Por ejemplo, mientras Whatsapp, Facebook o Twitter pueden ser utilizadas en una mayor medida para la comunicación y publicación de opiniones, Instagram o YouTube pueden ser redes sociales que se ocupen más de la interacción audiovisual de sus usuarios. En este sentido, puede considerarse beneficioso que se dé una regulación que considere cada aspecto distintivo de cada red social.

Por ejemplo, si se mira el documento de los Estándares o Normas Comunitarias de Facebook, se encuentra que, desde la misma plataforma social, se provee una definición clara de lo que se entiende cómo discurso de odio de la siguiente manera:

Consideramos lenguaje que incita al odio el que ataca directamente a las personas, por encima de conceptos e instituciones, en relación con las denominadas “características protegidas”: raza, etnia, nacionalidad, discapacidad, religión, clase, orientación sexual, sexo, identidad de género y enfermedad grave. Consideramos ataques las expresiones violentas o deshumanizantes, los estereotipos dañinos, las declaraciones de inferioridad, las expresiones de desprecio, repulsión o rechazo, los insultos, o las peticiones de exclusión o segregación. También prohibimos el uso de estereotipos dañinos, que definimos como comparaciones deshumanizantes con las

que, históricamente, se ha atacado, intimidado o excluido a grupos específicos, y que suelen vincularse con la violencia fuera de internet. Consideramos que la edad es una característica protegida si se menciona con otras características de este tipo. También protegemos a los refugiados, migrantes, inmigrantes y solicitantes de asilo de ataques graves, aunque sí permitimos los comentarios y las críticas relacionadas con las políticas de inmigración. De manera similar, también ofrecemos ciertas protecciones para características, como la profesión, cuando se mencionan junto con una característica protegida. A veces, basándonos en matices dependientes de la zona geográfica, consideramos que determinadas palabras o expresiones constituyen palabras clave que se refieren a grupos con características protegidas.

De esta manera, puede afirmarse que, de hecho, la definición de discurso de odio que disponen las Normas Comunitarias de Facebook abarca mucho más que algunas convenciones internacionales como, por ejemplo, el Protocolo adicional al Convenio sobre la Ciberdelincuencia, el cual solo hace énfasis en los actos racistas y xenófobos, pero deja otras categorías importantes fuera de su alcance, como, por ejemplo, la violencia por conflictos religiosos.

Por otro lado, al observar las normas de convivencia de Twitter, encontramos que también cuentan con una definición de discurso de odio clara, la cual se muestra de la siguiente manera:

No se permite fomentar la violencia contra otras personas ni atacarlas o amenazarlas directamente por motivo de su raza, origen étnico, nacionalidad, pertenencia a una casta, orientación sexual, género, identidad de género, afiliación religiosa, edad, discapacidad o enfermedad grave. Tampoco permitimos la existencia de cuentas cuyo

objetivo principal sea incitar la violencia contra otras personas en función de las categorías antes mencionadas.

A pesar de todo lo anterior, no puede negarse que estos sistemas de autorregulación distan de ser completamente satisfactorios. Tal cómo lo expone Alkiviadou (2019), uno de los obstáculos centrales es que estas redes sociales se apoyan predominantemente en los propios usuarios para reportar los discursos de odio que encuentren en estas plataformas. (pp. 33) En este sentido, la plataforma, por ejemplo, Facebook, confía en que los usuarios reportarán los discursos de odio que, según su opinión, encuentren mientras navegan en la red social. Así, los usuarios reportarán el mensaje odioso a Facebook, el cual luego lo revisará apoyándose en las personas que la misma plataforma contrata para controlar dichas situaciones o apoyándose en su tecnología de Inteligencia Artificial. Según si alguno de estos mecanismos internos considera que el mensaje publicado coincide o no con la definición de discurso de odio, dicho mensaje será removido de la red social y sus usuarios no podrán verlo más. Pero ¿qué sucede entonces si los usuarios no reportan los discursos odiosos que puedan circular por la red?

Igualmente, siguiendo la línea argumentativa de Alkiviadou (2019), el procedimiento de remoción de contenido no permitido no demuestra, por sí solo, una eficiencia y suficiencia. Éste también necesita ser rápido, en el sentido de poder limitar el daño hecho a la o las víctimas contra las que va dirigido el mensaje. (pp. 30) De esta misma manera, se puede cuestionar el hecho de depender en una gran medida de mecanismos de inteligencia artificial para gestionar y eliminar este tipo de comentarios. Recientemente, Mark Zuckerberg se ha comprometido a mejorar la calidad de las revisiones de contenido de lo posteo por los usuarios de Facebook apoyándose, en una gran medida, de la Inteligencia Artificial. Según

la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018), el apoyo exagerado en las herramientas de inteligencia artificial como mecanismo principal para detectar y responder al problema del discurso de odio no es una respuesta adecuada, en el sentido en el que la inteligencia artificial no puede entender el contexto en el que se dan los diferentes discursos y, además, estos mecanismos pueden prescindir, por falta en la definición, de discursos que puedan ser considerados odiosos o peligrosos, es por esto por lo que también se encuentra necesario que sean los humanos los que determinen los contextos. (pp. 1).

Siguiendo la línea de la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018), los desafíos para identificar y remover la propaganda anti-Rohingya demuestran que poseer una confianza absoluta en las herramientas de inteligencia artificial no es, por si sola, una solución al problema. Mientras que las herramientas de inteligencia artificial pueden ayudar en el señalamiento de contenido cuestionable que se presente en redes sociales, no puede confiarse en dichas herramientas para que dicho contenido sea removido efectivamente. Los matices que pueden presentarse en los discursos de odio y las diferencias y variaciones lingüísticas pueden estar más allá de las capacidades actuales de la inteligencia artificial. Incluso si la misma inteligencia artificial se convierte en una herramienta más confiable en detectar y remover discursos peligrosos en el futuro, los algoritmos, modelos y fuentes de datos por si mismos nunca estarán a salvo de ser parciales. Este caso es importante porque demuestra lo que está en juego para otras comunidades si no se desarrollan soluciones sostenibles y respetuosas de los derechos humanos. (pp. 2)

Para concluir, puede afirmarse que aún quedan muchos retos en cuanto al problema de los discursos de odio en las redes, entre estos, el que más salta a la vista es la falta de acuerdo sobre una definición concreta e incluyente de lo que significa discurso de odio y cómo debe manejarse dentro de las redes sociales para que no influya sobre el derecho legítimo a la libre expresión. Existe una falta de acuerdo, internacionalmente, sobre que significa el discurso de odio y esto se puede deber a que muchas veces, dichos discursos están plasmados mediante contextos específicos de cada cultura, lo cual hace difícil que se cree una entidad internacional que se ocupe de calificarlos. Esta falta de definición ha hecho que las posibles soluciones al problema sean más complejas de imaginar y llevar a cabo y, por esto es por lo que se hace complejo buscar una solución al problema de los discursos de odio que se dan en las redes sociales, porque aunque se tiene conciencia de que pueden ser causantes de consecuencias mucho más graves como la posible incitación a un genocidio, las propias plataformas sociales deben considerar sus propias particularidades y llegar a una solución, la cual muchas veces puede ser falible como se vio en el caso de Facebook y los Rohingya.

En este sentido, en línea con la investigación del Centro Data & Society sobre Redes Sociales, Inteligencia Artificial y Discursos de Odio en Myanmar de Melanie Penagos (2018), parece prudente concluir con algunas preguntas que aún quedan en el aire. Al pensar en la ética por diseño, ¿se debería ofrecer algún tipo de entrenamiento ético a los ingenieros y desarrolladores de las herramientas de inteligencia artificial? ¿Cómo debería llevarse a cabo lo anterior? (pp. 3) ¿debería considerarse un estudio de derecho comparado para evaluar cómo cada Estado ha llevado el problema del discurso de odio en línea y que regulaciones han desarrollado? ¿la estrategia de criminalizar el discurso de odio en plataformas sociales es realmente útil para combatir dicho fenómeno?

## 6. Conclusión

Debe entenderse que, si se quiere comenzar a desarrollar mucho más a fondo el tema de los discursos de odio en línea, primero tiene que hacerse un énfasis importante en qué es lo que puede considerarse exactamente cómo tal. Para lo anterior, es útil comenzar diferenciando el sentimiento de odio con otros sentimientos que, aunque pueden parecerse, no son equivalentes, cómo por ejemplo la ira. El odio no es un sentimiento que surja sin más, no es un sentimiento pasajero, sino que es algo mucho más profundo. Con el odio, pueden surgir también otros problemas más graves, porque de dicho sentimiento se desprende el querer hacer daño a una persona o personas determinadas. Parece que, aunque no hay una definición concreta de discurso de odio a nivel internacional, sí puede hablarse de ciertas características comunes a las definiciones que pueden encontrarse en diferentes instrumentos internacionales y también en diferentes regulaciones que llevan a cabo las propias plataformas sociales.

Es así como es posible afirmar que los discursos de odio generalmente son ataques directos contra alguno o varios de los grupos protegidos en razón a su raza, etnia, nacionalidad, religión, género, orientación sexual, discapacidad o enfermedad y en su generalidad, contienen expresiones insultantes, abusivas o con una clara intención de hacer daño a sus receptores o víctimas directas. Cómo se vio en la presente investigación, dentro de la regulación tanto internacional cómo dentro del fenómeno de autorregulación de diferentes plataformas, existen definiciones de lo que significa el discurso de odio que pueden ser más o menos garantistas, esto porque, algunas veces, ciertas definiciones dejan por fuera aspectos importantes sobre la población a la que dichos discursos pueden ir dirigidos.

Debido a los cambios que han surgido con la manera en la que interactuamos y nos comunicamos, los fenómenos de los discursos de odio también han cambiado y se han adaptado. Con el surgimiento de las plataformas sociales, se han venido dando diferentes formas de discursos de odio que, predominantemente, se expresan de manera escrita o por medio de imágenes o videos. En cierta manera, con la aparición de las plataformas sociales, el fenómeno de los discursos de odio se ha agravado ya que se presenta una mayor posibilidad de que dichos comentarios se conviertan en información viral, que tenga la posibilidad de ser globalmente conocida y, por ende, afectar a más personas.

De igual forma, es importante analizar el funcionamiento concreto de las plataformas sociales. Esto debido a que, en algunos casos, tienen formas de funcionamiento que pueden acrecentar o incentivar ciertas opiniones que pueden ser destructivas u odiosas. Los usuarios de las plataformas sociales tienen la capacidad de interconectarse y crear comunidades donde pueden encontrar personas que validen sus opiniones odiosas y violentas. Los contenidos que dichos usuarios crean, como por ejemplo la redacción de comentarios, opiniones o reacciones a otros contenidos, se almacenan en los servidores de dichas plataformas para que luego éstas puedan usar dichos datos como productos de predicción de conducta. De esta manera, la misma plataforma le sugerirá al usuario información con la cual éste se identifica, creando así el riesgo de que la misma plataforma conecte al usuario con grupos o comunidades que se ocupen de difundir los mismos mensajes de odio con los cuales puede concordar su opinión.

De igual manera, aunque las propias plataformas tengan la capacidad y la posibilidad de darse su propia regulación en cuanto a lo que está permitido y lo que no dentro de cada red social, también debe analizarse en qué medida hacen que dichas regulaciones sean de obligatorio

cumplimiento. Esto ya que, por ejemplo, se tiene el caso de Facebook en Myanmar, quienes no hicieron cumplir sus propias regulaciones al no eliminar comentarios y publicaciones que se constituían, claramente, en discursos de odio contra la población minoritaria de los Rohingya. En este orden de ideas, no sirve de mucho que las plataformas se den sus propias normas si, al fin y al cabo, no serán de obligatorio cumplimiento para todos los casos por igual. Es por esto por lo que surge la pregunta: ¿contar con la autorregulación de las plataformas sociales es suficiente para solucionar el problema de los discursos de odio y la incitación a posibles genocidios? La respuesta es, evidentemente, no.

Se hace evidente que, aunque las plataformas tengan sus propios reglamentos, esto no significa que dichas normas se cumplan en todos los casos o que estos sistemas no sean falibles. Se necesita también del apoyo del sistema internacional para que regule las situaciones de los discursos de odio en línea. Pero lo anterior debe ser incluyente, debe existir un mecanismo internacional que vele por los derechos de todos los posibles grupos protegidos que puedan ser objeto de incitación al genocidio por medio de los discursos de odio. Si bien esto es deseable, es cierto que las regulaciones internacionales requieren de más convenios y pactos y obedecen, muchas veces, a intereses de algunos estados. Esto es diferente con el aspecto de la autorregulación de las plataformas sociales, las cuales no necesitan de ningún acuerdo ni negociación, sino que, simplemente, redactan sus propias normas y las hacen obligatorias para las personas que quieran ser usuarios de la red social en particular. Sería interesante pensar en una colaboración entre las plataformas sociales y las entidades internacionales donde se incluya a todos los posibles actores y las posibles víctimas de este tipo de discursos.

Igualmente, debe aclararse que el tema tratado reviste una gran importancia ya que se debe considerar la cuestión de la relación entre las empresas y el cumplimiento de los Derechos Humanos. En este caso, se presenta una dimensión particular ya que la empresa actúa como un mediador entre los violadores de derechos humanos y las víctimas. Es la empresa misma la que decide si permite o no que se den posibles violaciones a derechos humanos, pero este es un tema álgido ya que las empresas deben proteger ciertos derechos humanos, pero sin violar otros. Este es el caso del derecho a la libre expresión. En este sentido, puede que, si se centra en soluciones únicas como la Inteligencia Artificial, se terminen eliminando mensajes que no tengan nada que ver con discursos de odio o que estén denunciando el odio y que, con esto, se vea afectado, de una manera inaceptable, el derecho a la libre expresión. Es por esto por lo que debe plantearse una solución global que tenga en cuenta ambos derechos humanos.

## 7. REFERENCIAS:

Alkiviadou, N. (2019). Hate speech on social media networks: towards a regulatory framework? *Information & Communications Technology Law*, 28(1), 19-35.

<https://doi.org/10.1080/13600834.2018.1494417>

Allen, J.M., & Norris, G.H. (2011). Is genocide different dealing with hate speech in post-genocide society. *Journal of International Law and International Relations*, 7(1), 146-174.

Bonina, C., Koskinen, K., Eaton, B., & Gawer, A. (2021). Digital platforms for development: Foundations and research agenda. *Inf Syst J*, 1-34. <https://doi.org/10.1111/isj.1232634>

Bonnet, P. (2019). Apuntes sobre el discurso del odio en la sociedad contemporánea. *Desde el Jardín de Freud*, 19, 177-186. doi: 10.15446/djf.n19.76716

Byman, D., & Joshi, A. (2020). Preventing the Next Social-media Genocide. *Survival*, 62(6), 125-152. <https://doi.org/10.1080/00396338.2020.1851097>

Byung, C.H. (2014). *En el enjambre*. Herder.

Calles Gómez, S. (2018). Los Rohingya: La crisis humanitaria de una etnia olvidada. [Tesis para optar al título de Abogada, Universidad de Salamanca]. Repositorio institucional

Universidad de Salamanca.

[https://gedos.usal.es/bitstream/handle/10366/139415/TG\\_CallesGomez\\_%20Rohingyas.pdf?sequence=1&isAllowed=y](https://gedos.usal.es/bitstream/handle/10366/139415/TG_CallesGomez_%20Rohingyas.pdf?sequence=1&isAllowed=y)

Castro, M.V. (2020). Estudio preliminar derecho espacio y poder: Trayectorias de la geografía legal y su utilidad para el análisis legal. Siglo del Hombre Editores.

<http://dx.doi.org/10.15425/2017.273>

Clough, J. (2014). A world of difference: The Budapest convention on Cybercrime and the challenges of Harmonisation. *Monash University Law Review*, 40(3), 698-736.

Cortés, C. (2018). La desconfiguración de la empatía. *Revista Semana*.  
<https://www.semana.com/periodismo-cultural---revista-arcadia/articulo/elecciones-presidenciales-en-redes-sociales-colombia-empatia/69300/>

Data & Society. (2018). Social Media, Artificial Intelligence, and Hate Speech in Myanmar Case Study. [https://datasociety.net/wp-content/uploads/2018/09/Social-Media-Artificial-Intelligence-and-Hate-Speech-in-Myanmar\\_Case-Study\\_Final.pdf](https://datasociety.net/wp-content/uploads/2018/09/Social-Media-Artificial-Intelligence-and-Hate-Speech-in-Myanmar_Case-Study_Final.pdf)

Dieng, A. (2018). *Note to Correspondents: Statement by Adama Dieng, United Nations Special Adviser on the Prevention of Genocide, on his visit to Bangladesh to assess the situation of Rohingya refugees from Myanmar*. <https://www.un.org/sg/en/content/sg/note-correspondents/2018-03-12/note-correspondents-statement-adama-dieng-united-nations>

Emcke, C. (2017). *Contra el odio: Un alegato en defensa de la pluralidad de pensamiento, la tolerancia y la libertad*. Taurus.

Faustin, M. (2016). Preventing Genocide by Fighting Against Hate Speech. *International Journal of Advanced Research*, 4(3), 117-132.

Facebook. (2021). Estándares Comunitarios. <https://about.fb.com/ltam/news/tag/estandares-comunitarios/>

Gillespie, T. (2010). The politics of “platforms.” *New Media & Society*, 12(3), 347-364.  
<https://doi.org/10.1177/1461444809342738>

Gorwa, R. (2019). What is platform governance?, *Information, Communication & Society*, 22(6), 854-871. <https://doi.org/10.1080/1369118X.2019.1573914>

Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Media + Society*, 1(2), 1-11. <https://doi.org/10.1177/2056305115603080>

Mathew, B., Dutt, R., Goyal, P., & Mukherjee, A. (2019). Spread of Hate Speech in Online Social Media. 11th ACM Conference on Web Science (WebSci '19), June 30-July 3. 173-182. <https://doi.org/10.1145/3292522.3326034>

Mendel, T. (2010). Hate Speech Rules Under International Law. *Centre For Law and Democracy*. <http://www.law-democracy.org/wp-content/uploads/2010/07/10.02.hate-speech.Macedonia-book.pdf>

Misión Investigadora nombrada por el Consejo de Derechos Humanos de las Naciones Unidas, (2019). *Conclusiones*. <https://www.ohchr.org/EN/HRBodies/HRC/Pages/NewsDetail.aspx?NewsID=25002&LangID=E>

Mondal, M., Araújo, L., & Benevenuto, F. (2017). A Measurement Study of Hate Speech in Social Media. *Proceedings of HT'17, Prague, Czech Republic, July 04-07*, 85-94. <https://doi.org/10.1145/3078714.3078723>

Nemes, I. (2002). Regulating Hate Speech in Cyberspace: Issues of Desirability and Efficacy. *Information & Communications Technology Law*, 11(3), 193-220. <https://doi.org/10.1080/1360083022000031902>

Nooren, P., van Gorp, N., van Eijk, N., & Ó Fathaigh, R. (2018). Should We Regulate Digital Platforms? A New Framework for Evaluating Policy Options. *Policy and Internet*, 10(3), 1-38. <https://doi.org/10.1002/poi3.177>

Nussbaum, M.C. (2014). *Las emociones políticas: ¿Por qué el amor es importante para la justicia?* Paidós.

Organización de las Naciones Unidas. (2019). La Estrategia y Plan de Acción de las Naciones Unidas para la Lucha Contra el Discurso de Odio. [https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action\\_plan\\_on\\_hate\\_speech\\_ES.pdf](https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_ES.pdf)

Patni, R., & Kaumudi, K. (2009). Regulation of hate speech. *NUJS Law Review*, 2(4), 749-778.

Penagos, M. (2018). *Social Media, Artificial Intelligence, and Hate Speech in Myanmar Case Study*. [https://datasociety.net/wp-content/uploads/2018/09/Social-Media-Artificial-Intelligence-and-Hate-Speech-in-Myanmar\\_Case-Study\\_Final.pdf](https://datasociety.net/wp-content/uploads/2018/09/Social-Media-Artificial-Intelligence-and-Hate-Speech-in-Myanmar_Case-Study_Final.pdf)

Reichelmann, A., Hawdon, J., Costello, M., Ryan, J., Blaya, C., Llorent, V., Oksanen, A., Räsänen, P., & Zych, I. (2020). Hate Knows No Boundaries: Online Hate in Six Nations. *Deviant Behavior*, 1-12. <https://doi.org/10.1080/01639625.2020.1722337>

Rollnert, G. (2020). Redes sociales y discurso del odio: perspectiva internacional. *Revista de Internet, Derecho y Política*.

Schwarz, J.A. (2017). Platform Logic: An Interdisciplinary Approach to the Platform-Based Economy. *Policy and Internet*, 9(4), 374-394. <https://doi.org/10.1002/poi3.159>

Shaw, M. (2007). The general hybridity of war and genocide. *Journal of Genocide Research*, 9(3), 461-473. <https://doi.org/10.1080/14623520701584281>

Torres Guzmán, K.T. (2016). Crisis de los Rohingya (1991-2015): ¿Una amenaza a la seguridad nacional de Bangladesh? [Tesis para optar al título de Intercionalista, Universidad Colegio Mayor de Nuestra Señora del Rosario]. Repositorio institucional Universidad del Rosario. <https://repository.urosario.edu.co/bitstream/handle/10336/13806/TorresGuzman-KarolTatiana-2016.pdf?sequence=1&isAllowed=y>

Twitter. (2021). Las reglas de Twitter. <https://help.twitter.com/es/rules-and-policies/twitter-rules>

Zamora, F.J., & Murillo, M.C. (2020). El conflicto Rohingya y sus proyecciones jurídicas: Aspectos destacados. *Ordine internazionale e diritti umani*, 461-484.

Zuboff, S. (2020). *La Era del Capitalismo de la Vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder*. Editorial Planeta.

### **Instrumentos internacionales:**

Organización de los Estados Americanos. (1969). *Convención Americana sobre Derechos Humanos*.

[https://www.hchr.org.co/documentoseinformes/documentos/html/pactos/conv\\_americana\\_derechos\\_humanos.html](https://www.hchr.org.co/documentoseinformes/documentos/html/pactos/conv_americana_derechos_humanos.html)

Organización de las Naciones Unidas. (1966). *Pacto Internacional de Derechos Civiles y Políticos*. <https://www.ohchr.org/sp/professionalinterest/pages/ccpr.aspx>

Organización de las Naciones Unidas. (1948). *Declaración Universal de Derechos Humanos*. <https://www.un.org/es/about-us/universal-declaration-of-human-rights>