

CONTROLAT A LES XARXES!

#BCN vs ODI

El discurs de l'odi
a les xarxes socials:
un estat de la qüestió



ajuntament.barcelona.cat/bcnvsodi

Ajuntament de
Barcelona



Àrea de Drets de Ciutadania, Cultura, Participació i Transparència

Direcció de Serveis de Drets de Ciutadania i Diversitat

Barcelona, desembre de 2017

Autoria de l'informe: Alex Cabo Isasi i Ana García Juanatey

TAULA DE CONTINGUTS

INTRODUCCIÓ	4
1.UN DIAGNÒSTIC DEL PROBLEMA	5
1.1.Què s'entén per discurs d'odi? Un concepte objecte de debat	6
1.2.Les particularitats del discurs de l'odi online	8
1.3.Les conseqüències del discurs de l'odi	8
1.4.Qui són els autors del discurs de l'odi a Internet?	9
2.LA RESTRICCIÓ DEL DISCURS DE L'ODI	11
2.1.El dilema entre la llibertat d'expressió i la repressió del discurs de l'odi	11
2.2.Les particularitats jurídiques de l'odi a Internet i les xarxes socials	14
2.3.Legislació espanyola	14
3.EL ROL DE LES XARXES SOCIALS COM A INTERMEDIÀRIES	19
3.1.Atribució de responsabilitat	19
3.2.Autoregulació i implementació de filtres informàtics	20
3.3.El Codi de Conducta amb la Unió Europea	22
4.LES ESTRATÈGIES NO JURÍDIQUES PER COMBATRE EL DISCURS DE L'ODI A LES XARXES SOCIALS	24
4.1.Les estratègies de monitoreig i recerca	25
4.2.Les estratègies de pressió a les xarxes socials	26
4.3.Les estratègies orientades a canviar percepcions i actituds	26
4.3.1. Les campanyes contra el discurs de l'odi a Internet	26
4.3.2. Les estratègies contra el discurs de l'odi a Internet basades en l'educació i la formació	30
CONCLUSIÓ	33
BIBLIOGRAFIA	35

INTRODUCCIÓ

S'ha teoritzat àmpliament sobre el potencial deliberatiu i democratitzador d'Internet, i més en concret, de les xarxes socials. El seu potencial participatiu és innegable, i ha generat moltes esperances en les possibilitats de transformació política i social que ofereixen. Les xarxes socials proporcionen un ampli accés a fonts d'informació alienes al control dels governs i les grans corporacions, faciliten la generació i coordinació de xarxes d'activistes, i constitueixen un espai immillorable per a la deliberació i l'intercanvi d'idees. Les xarxes socials, són, en paraules de l'expert en tecnologies de la informació i la comunicació, Manuel Castells, una plataforma de "auto-comunicació de masses", un espai d'autonomia comunicativa idoni perquè els actors socials propiciïn la transició de la indignació a l'esperança. En aquest sentit, es considera l'aleshores denominada Primavera Àrab, el 15-M, o Occupy Wall Street com a exemples de com les xarxes socials poden contribuir a transformar la indignació en accions i moviments horitzontals, transformadors i emancipadors (Castells 2012).

Malgrat això, no convé pecar d'un excés d'optimisme, perquè la veritat és que Internet i les xarxes socials també poden ser utilitzades amb finalitats radicalment diferents a les exposades. Tal com afirma el ciberescèptic Evgeny Morozov, Internet no "conduïx necessàriament al respecte universal dels drets humans". Internet no és alliberador ni democratitzador en si, sinó que pot produir "diferents resultats polítics en diferents entorns", i per tant convé no deixar-se portar per cert "solucionisme tecnològic" o ciberutopisme (Morozov 2012).

De la mateixa manera que alguns moviments emancipadors han fet un esperançador ús de les xarxes socials, forces de signe contrari han demostrat la seva capacitat d'aprofitar aquest potencial de les xarxes per a finalitats oposades. Gràcies a Edward Snowden, es va comprovar com Estats Units utilitza les noves tecnologies per a la vigilància i l'espionatge indiscriminat. Un altre exemple del costat fosc de la Xarxa és el denominat *50 Cent Army*, el "exèrcit" de comentaristes contractat per les autoritats xineses per circular missatges propagandístics i favorables al govern, als debats a les xarxes.

Però potser l'exemple més tràgic, de com poden ser utilitzats Internet i les xarxes, és el de la crisi que es va produir a Kenya després de les eleccions de desembre de 2007. La propagació de missatges incitadors a la violència, a través de varis fòrums a Internet, està documentada com "un canal determinant a través del qual la violència va ser alimentada, emportant-se les vides de més de mil persones, i el desplaçament de més de 600.000" (Gagliardone et al. 2014: 5).

UN DIAGNÒSTIC DEL PROBLEMA

Efectivament, aquest tipus de missatges incitadors a la violència ha trobat a Internet i les xarxes socials, el canal adient per a la seva propagació. Tot i que l'abast real del problema no ha estat determinat quantitativament, més que d'una manera fragmentada i parcial, existeix una sensació compartida per periodistes, juristes, ONGs, investigadors, i usuaris de les xarxes, en general, que el problema del discurs extrem a les xarxes socials és cada dia més preocupant.

A tota Europa s'està produint una gran proliferació de missatges extremistes, associada al context de la crisi de refugiats, amb preocupants "puntes" d'odi islamòfob detectats a les xarxes arran dels atemptats de París, Brussel·les, o Niça. Després dels atemptats de París, *#matadatodoslosmusulmanes* es va convertir durant algunes hores en el tercer *hashtag* més utilitzat a Espanya (Jubany i Roiha 2016).

Segons estudis del *think tank* Demos, a Twitter es produeixen aproximadament unes 10.000 piulades per dia amb insults racistes en llengua anglesa, la qual cosa suposa una de cada 15.000 piulades (Gagliardone et al., 2014). En un altre estudi realitzat per Demos, sobre misogínia, que és un altre dels motius recurrents d'intolerància, durant tres setmanes al mes d'abril de 2016, entre usuaris britànics de Twitter, es van trobar més de 200.000 piulades amb les paraules "meuca" o "puta"; constatant-se que cada deu segons algú insulta amb aquestes paraules a una dona a Twitter.

A l'estat espanyol el problema no és menor. A continuació, alguns exemples de 2016. Moha Gerehou, president de SOS Racismo Madrid, d'origen gambià, ha estat objecte de múltiples atacs. Al juliol de 2016, se'l "subhastava" a Twitter com un esclau o una peça de caça, com a resposta a la campanya *#EstadoEspañolNoTanBlanco*. Al setembre del mateix any, va ser amenaçat de mort després de penjar a Twitter una fotografia d'una concentració en contra de la retirada de la llista de carrers de Madrid del nom de Millán Astray, fundador de la Legió¹.

A l'agost de 2016, Jordi Ballart, alcalde de Terrassa, va rebre amenaces homòfobes a través de Twitter i Facebook, arran de la seva decisió de retirar el nom d'un carrer dedicat a un membre de la Divisió Blava².

Al setembre de 2016, la transsexual Carla Antonelli, diputada de l'Assemblea de Madrid, va rebre amenaces de mort a través de Twitter, per la seva defensa dels drets del col·lectiu LGTB³. El mateix mes de setembre, l'àrbitre de futbol infantil, Jesús Tomillero, president de l'associació contra la LGTBfòbia en l'esport, va rebre insults i amenaces de mort a Twitter, per ser el primer col·legiat auto-declarat gai a Espanya⁴.

Aquests són tan sols alguns casos recents d'assetjament individual per motius d'odi i intolerància, que han arribat als mitjans de comunicació, i que constitueixen la punta de l'iceberg d'una situació que s'ha convertit en habitual a les xarxes socials. A nivell col·lectiu, l'antigitanisme

1. http://www.eldiario.es/sociedad/Nuevas-presidente-SOS-Racismo-Madrid_0_563094347.html

2. <http://www.elperiodico.com/es/noticias/sociedad/jordi-ballart-alcalde-terrassa-denuncia-amenazas-homofobas-redes-sociales-5473496>

3. http://ccaa.elpais.com/ccaa/2016/09/04/madrid/1472990074_481479.html

4. <http://www.elperiodico.com/es/noticias/deportes/arbitro-gay-jesus-tomillero-denuncia-amenazas-muerte-5376384>

prolifera amb impunitat a les xarxes, la islamofòbia es converteix en tendència de manera recurrent, a través de *hashtags* com #stopIslam, #terroristaswelcome o #musulmanerroristas; l'antisemitisme, l'homofòbia, la misogínia... Totes les ideologies intolerants troben a les xarxes socials un espai d'expressió privilegiat, que ha generat una espècie de cultura de l'odi⁵, que contamina i intoxica les xarxes amb llenguatge abusiu, denigrant o agressiu, per motius, en gran mesura, d'intolerància contra població immigrada, refugiats, musulmans, homosexuals, i altres minories.

Les institucions internacionals i europees no són alienes a aquest problema. Prova d'aquesta preocupació és la intensa activitat que ha desenvolupat la Unió Europea durant l'últim any, en la cerca de solucions per a aquest problema; malgrat que al mateix temps certes formes de racisme han anat guanyant cada vegada més terreny en el discurs polític europeu. Com a fites destacades d'aquesta dedicació destaquen la signatura del Codi de Conducta amb les empreses tecnològiques, i l'engegada d'un Grup d'Alt Nivell, en el qual ha involucrat a representants dels estats membres de la Unió, del Parlament Europeu, del Consell d'Europa, de l'Alt Comissionat de l'ONU per als Refugiats (ACNUR), de l'Organització per a la Seguretat i la Cooperació a Europa (OSCE), de l'Agència Europea de Drets Fonamentals (FRA), i d'organitzacions de la societat civil com Amnistia Internacional, la Xarxa Europea contra el Racisme (ENAR), la Plataforma Europea d'ONGs socials, i l'Associació Internacional de Lesbianes, Gais, Bisexuals i Intersex (ILGA), amb la finalitat de prevenir i combatre els crims d'odi i el discurs de l'odi, amb una especial dedicació a la contra-argumentació del discurs de l'odi *online*.

També la UNESCO o la Comissió Europea contra el Racisme i la Intolerància (ECRI) han destacat, en informes recents, el creixement a Internet i xarxes socials del discurs intolerant contra diferents minories, i la necessitat de trobar estratègies eficaces per combatre'l.

En definitiva, el problema del discurs de l'odi a Internet, i a les xarxes socials, en particular, és un assumpte que ha adquirit tal dimensió que actualment es troba en l'agenda d'un bon nombre d'organismes europeus i internacionals. Les xarxes socials s'han convertit en un espai d'impunitat per a la ira i l'odi. El masclisme, l'homofòbia, la xenofòbia, la islamofòbia, l'antigitanisme, l'antisemitisme i altres ideologies intolerants aprofiten Internet i les xarxes socials per a l'insult, la humiliació, el linxament social, l'assetjament, o l'amenaça.

1.1. Què s'entén per discurs d'odi? Un concepte objecte de debat

No hi ha una definició universalment acceptada de discurs de l'odi. En primer lloc, perquè el terreny de les definicions amb implicacions ètiques i jurídiques sempre és controvertit. I en segon lloc, perquè el propi terme "odi" ho converteix en un concepte emocional i obert a la subjectivitat. És un concepte que genera confusió i que per la seva subjectivitat, és relativament manipulable.

El mínim comú denominador de qualsevol definició de discurs de l'odi seria qualsevol expressió d'opinió o idees basada en el menyspreu i l'animadversió cap a persones o col·lectius als quals es desitja el mal. Malgrat això, aquesta definició simple englobaria un rang massa ampli d'expressions perquè el concepte tingués alguna utilitat per a l'anàlisi o la intervenció social o jurídica (Article 19 2015).

5. <http://time.com/4457110/internet-trolls/>

Es parla de discurs de l'odi per referir-se a les expressions que inciten directament a la comissió d'actes de discriminació o violència per motius d'odi racial, xenòfob, orientació sexual, o altres formes d'intolerància. Però per a alguns, el concepte s'estén també a aquelles expressions que fomenten els prejudicis o la intolerància, considerant que aquest tipus d'expressions contribueixen indirectament al fet que es generi un clima d'hostilitat que pugui propiciar, eventualment, actes discriminatoris o atacs violents (Gagliardone et al. 2015).

En el llenguatge quotidià, i especialment en els mitjans de comunicació, s'ha generalitzat l'ús del concepte de discurs de l'odi per referir-se a un conglomerat heterogeni de manifestacions que engloba des d'amenaques a individus o col·lectius fins a casos en els que algunes persones simplement expressen la seva ira contra les autoritats, de manera més o menys ofensiva (Gagliardone et al. 2015). A més s'ha estès en els mitjans de comunicació, i fins i tot en l'àmbit jurídic, l'ús del terme per referir-se a conceptes com la incitació o l'apologia del terrorisme, que encara que tenen certa relació amb l'objecte d'aquest informe, requereixen un enfocament singularitzat.

En definitiva, el terme de discurs de l'odi és un concepte discutit, i és objecte de debat polític, jurídic i acadèmic a nivell internacional. És un concepte complex, en la mesura en què posa en joc i en col·lisió diferents valors i principis cabdals dels sistemes democràtics: la igualtat, la dignitat humana, la llibertat d'expressió, etc., la concepció dels quals no és idèntica en tots els contextos soci-polítics. Per tant, el concepte de discurs de l'odi pot ser implementat políticament amb objectius ben diferents, que poden ser més o menys legítims.

Les definicions que tenen major acceptació a nivell internacional es poden agrupar en dues tendències principals: 1) aquelles que defineixen el discurs de l'odi en un sentit més ampli i expansiu, i inclouen qualsevol expressió que promogui o justifiqui l'odi per motius racistes, xenòfobs, religiosos, de gènere, orientació sexual o discapacitat⁶; i 2) aquelles que el defineixen d'una manera més restrictiva i precisa, incloent només aquelles formes d'expressió que, en contextos determinats d'instabilitat, puguin contribuir a desencadenar episodis violents contra un grup de persones per raó de la seva pertinença a un dels grups assenyalats⁷ (Gagliardone et al., 2014).

La diferència fonamental entre ambdues es podria sintetitzar en la dialèctica contingut-context, i en el major o menor risc que els missatges desencadenin accions violentes. La primera de les tendències conceptuals defineix el discurs de l'odi atenent exclusivament al seu contingut, mentre que la segona considera que el que converteix una determinada expressió en discurs de l'odi és el risc manifest que, donat el context històric-social en el qual s'emet, provoqui episodis violents. El context, efectivament, determina en gran mesura el dany que determinades expressions i manifestacions poden arribar a ocasionar, que pot oscil·lar des del dany "emocional" fins a la provocació d'escalades de violència massiva, com en els genocidis de Iugoslàvia o Rwanda en els anys 90.

La importància d'aquestes diferents maneres de definir el discurs de l'odi radica en què, encara que el concepte de discurs de l'odi no és una categoria estrictament jurídica, aquestes definicions informen les diferents legislacions i jurisprudències nacionals, i les declaracions, directives, i convenis internacionals.

6. La Recomanació No. R(97)20 del Consell d'Europa, de 30/10/1997, el va definir com "tota forma d'expressió que difongui, inciti, promogui o justifiqui l'odi racial, xenofòbia, antisemitisme o altres formes d'odi basades en la intolerància (...)"

7. Susan Benesch (2013) proposa el terme "discurs perillós" ("dangerous speech"), per a aquell discurs que té una probabilitat raonable de desencadenar episodis de violència genocida.

En qualsevol cas, a l'efecte d'aquest informe, prendrem com a referència les definicions del primer grup que conceptualitzen el discurs de l'odi en un sentit relativament ampli, perquè recullen el tipus d'expressions que, donada la seva proliferació en el context europeu, mereixen un rang més variat d'intervencions tant de tipus jurídic com de tipus social.

1.2. Les particularitats del discurs de l'odi *online*

El discurs d'odi *online*, pel qual s'ha encunyat l'expressió ciberodi, afegeix una sèrie de particularitats que el converteixen en un fenomen descontrolat amb un potencial de dany encara major. En primer lloc, la sobreabundància comunicativa. Per mitjà d'Internet i les xarxes socials, la comunicació de missatges que abans estaven limitats a l'àmbit privat, s'ha convertit en pública de manera absolutament massiva. En segon lloc, la descentralització de la comunicació, que “democratitza” les comunicacions, en el sentit de que qualsevol pot emetre un missatge amb un enorme potencial d'audiència. L'efecte multiplicador de les xarxes socials permet convertir un determinat missatge en un fenomen de transmissió exponencial, donant lloc al que es denomina metafòricament com “viralitzacions”.

La permanència dels continguts, la itinerància entre diferents plataformes, l'ús de pseudònims, l'anonimat, i la transnacionalitat són altres particularitats de l'espai *online* que afegeixen reptes en relació amb el discurs de l'odi (Gagliardone et al. 2015). La permanència i la itinerància són particularitats que poden augmentar el dany potencial del discurs de l'odi. L'anonimat, l'ús de pseudònims i la transnacionalitat són característiques que dificulten la seva persecució, i que per la sensació d'impunitat que ofereixen, encoratgen l'expressió del discurs de l'odi.

En ocasions, en el llenguatge col·loquial ens referim a Internet com un espai “virtual”, adjectiu equívoc, en la mesura en què es refereix a allò “que té existència aparent, però no real”⁸. Aquesta sensació de “virtualitat”, com si el que succeeix a Internet no fos “real” i no tingués repercussions fora de la Xarxa, també actua com a efecte desinhibidor per a l'expressió dels discursos extrems.

En aquest sentit, Jubany i Roiha (2016) adverteixen de la fal·làcia del dualisme digital, terme concebut per Nathan Jurgenson per referir-se a la falsa separació entre el món *online* i el món *offline*. La qüestió del discurs de l'odi a les xarxes socials, encara que tingui les especificitats pròpies d'aquests espais de comunicació, no és més que una exteriorització pública d'un problema subjacent en la societat en general, independentment de l'escenari en el qual es manifesti.

1.3. Les conseqüències del discurs de l'odi

Les repercussions que pot tenir el discurs de l'odi, sigui o no *online*, són variades. D'una banda, cal considerar el dany directe emocional o psicològic que poden produir les amenaces, l'assetjament, i altres atacs dirigits a individus concrets, per motius d'odi i intolerància. D'altra banda, el discurs de l'odi també produeix un dany indirecte, erosionant la dignitat de les persones i danyant la seva reputació.

El discurs de l'odi juga un paper fonamental en la perpetuació d'estereotips discriminatoris, l'estigmatització de col·lectius, la contribució a la seva marginació, etc. Constitueix un mecanisme

8. Definició de la R.A.E.

lingüístic d'importància fonamental en la deshumanització de certs col·lectius, als quals es tracta de negar la seva pertinença a la ciutadania en condicions d'igualtat. En aquest sentit, el discurs de l'odi envia un missatge que divideix i segrega la societat. Però al mateix temps, també juga un paper cohesionador per als emissors del discurs de l'odi, reforçant el seu sentiment de pertinença a un grup (Gagliardone et al. 2015).

Mitjançant la proliferació i acceptació del discurs de l'odi es produeix una reducció d'empatia cap als col·lectius deshumanitzats, i com a conseqüència, es pot generar el caldo de cultiu adequat per justificar actes discriminatoris, abusos i actes violents de diversa naturalesa. Encara que no es pugui establir, amb caràcter general, una connexió directa entre la proliferació de discurs de l'odi i els crims violents d'odi, cada vegada hi ha una consciència més clara del vincle indirecte entre tots dos fenòmens.

Si bé no tot discurs de l'odi desencadena episodis de violència, el que sí sembla evident és que rarament es produeixen episodis de crims d'odi sense una prèvia estigmatització i deshumanització de les víctimes. Per tant, no està de més recordar aquí els incidents de delictes d'odi registrats el 2015 pel Ministeri d'Interior. Durant el 2015, es van denunciar i enregistrar com a delictes d'odi 1.328 casos, que inclouen lesions, abusos sexuals, danys, tractes degradants, etc. Movimiento contra la Intolerancia maneja estimacions que situen la xifra real entorn dels 4.000 casos. Tenint en compte que el propi Ministeri d'Interior reconeix que una gran part dels delictes d'odi no es denuncien, aquestes xifres són un indicatiu de fins a quin punt no és convenient prendre's a la lleugera el discurs de l'odi, com a substrat justificador o incitador de comportaments violents (Ministeri d'Interior 2016).

1.4. Qui són els autors del discurs de l'odi a Internet?

No hi ha dubte que una gran part del discurs de l'odi a les xarxes socials procedeix d'individus directa o indirectament vinculats a col·lectius feixistes com ara Hogar Social Madrid, que compta amb més de 15.000 seguidors a Twitter, grups ultrarassistes d'equips de futbol, o partits polítics d'ultra-dreta més o menys marginals com ara Plataforma per Catalunya, Falange Española de las JONS, España 2000, Alianza Nacional, etc.

Aquests grups han trobat a les xarxes una útil eina de difusió dels seus missatges i la seva simbologia feixista, que els serveix tant per atacar les seves víctimes com per captar adeptes. Els individus associats a aquests grups solen actuar a les xarxes de manera col·lectiva, cosa que provoca una major sensació d'assetjament a les víctimes dels seus atacs. A més, per abocar les seves incitacions a la violència a les xarxes, desenvolupen diferents codis lingüístics i paraules clau, que els serveixen per evitar tant els filtres informàtics de les empreses tecnològiques, com la persecució judicial i policial de les seves conductes⁹.

No obstant, la veritat és que el fenomen del discurs de l'odi no s'esgota en aquests grups amb un marcat perfil ideològic. Hi ha altres perfils d'usuaris de les xarxes socials, que en l'argot d'Internet són coneguts com *haters* o trols, que constitueixen una part important del problema.

A Internet, es denomina *haters* a individus que es dediquen obsessivament a atacar i agredir verbalment a individus concrets (en gran mesura, persones famoses), o a col·lectius als quals

9. <http://www.genbeta.com/redes-sociales-y-comunidades/asi-es-como-los-extremistas-camufan-su-discurso-de-odio-en-las-redes-sociales>

menyspreen pel seu origen ètnic, la seva religió, etc. Les xarxes socials estan plenes d'usuaris amb una dedicació obsessiva al menyspreu i l'agressió verbal contra musulmans, gitanos, homosexuals, dones, etc. Un altre tipus de perfil assimilable seria el d'aquells usuaris amb una tendència a exterioritzar les seves reivindicacions i preocupacions, sobre qualsevol assumpte, de manera extrema i agressiva, com una manera d'atreure l'atenció i fer més visibles els seus missatges.

El fenomen del "troleig" també està àmpliament estès a les xarxes, però és molt arriscat tractar de caracteritzar-lo com un fenomen homogeni. Alguns actuen a títol individual per buscar atenció o descarregar la seva frustració. Uns altres actuen de manera col·lectiva, i s'organitzen per atacar determinats objectius. En qualsevol cas, no hi ha dubte que bona part del fenomen narcisista i sàdic del "troleig" es produeix amb el simple ànim de divertir-se, provocant a altres usuaris de les xarxes socials. De manera irreflexiva, i moltes vegades aliens al dany que poden arribar a causar, molts usuaris de les xarxes es dediquen a escampar la seva hostilitat i agredir verbalment a altres individus o a col·lectius minoritaris per divertir-se amb les reaccions que produeixen en els altres, amb l'excusa d'un pretès sentit de l'humor.

En tot cas, destriar els que estan intentant provocar per divertir-se dels que estan expressant les seves veritables opinions no sempre és fàcil. Per això, per abordar el fenomen del discurs de l'odi a les xarxes socials, cal tenir en compte la diversitat i l'heterogeneïtat de perfils dels seus autors, de cara, no només a dimensionar correctament el problema, sinó també per traçar estratègies adequades als diferents orígens del discurs extrem. Amb l'ús d'aquests termes, propis de l'argot d'Internet, es corre el risc de caricaturitzar excessivament el problema, i es produeix la tendència ja esmentada a considerar-ho com una cosa aliena a la realitat *offline*, que Nathan Jurgenson denomina fal·làcia del dualisme digital. Les obsessions hostils dels haters o el sadisme narcisista dels trols tenen el seu espai d'expressió a les xarxes, però les seves motivacions i els seus prejudicis hostils provenen de l'espai *offline*.

2. LA RESTRICCIÓ DEL DISCURS DE L'ODI

2.1. El dilema entre la llibertat d'expressió i la repressió del discurs de l'odi

Davant la creixent alarma generada, a nivell internacional, per la proliferació de discurs de l'odi, i les devastadores conseqüències que va tenir en alguns genocidis com el de Rwanda o Iugoslàvia, s'han produït importants moviments i pressions per restringir-lo. Com a resultat, els Estats han optat per diferents nivells de restricció, cadascun en funció del seu context històric i socio-polític.

La diversitat d'enfocaments jurídics d'aquesta problemàtica, i el consegüent nivell de repressió del discurs de l'odi, té el seu origen en el debat que aquesta qüestió genera sobre els límits de la llibertat d'expressió. La llibertat d'expressió, encara que és un pilar fonamental del sistema democràtic, no és un dret absolut¹⁰. Tots els ordenaments jurídics contenen tipus penals que la restringeixen. Per exemple, no existeix un dret a l'insult, la calúmnia o l'amenaça. En el cas del discurs de l'odi, la tensió es produeix entre el dret a la llibertat d'expressió i el respecte a la dignitat humana i la igualtat. Mitjançant el discurs de l'odi, es busca la deshumanització i l'exclusió social de minories vulnerables, privant-los així de dignitat i del dret a la igualtat.

En l'àmbit internacional, les diferents tendències de restricció del discurs de l'odi se solen sintetitzar, de manera una mica reduccionista, en la divisió entre la perspectiva nord-americana, més liberal i "tolerant"; i la perspectiva europea, més militant i "intransigent" amb el discurs de l'odi.

- Estats Units s'ha caracteritzat tradicionalment per una defensa a ultrança de la Primera Esmena¹¹, i pel "*compromís de sostenir la llibertat, fins i tot en circumstàncies excepcionals*" (Revenga-Sánchez 2015). S'ha articulat la defensa de la llibertat d'expressió en una concepció del debat públic com a lliure mercat d'idees, en el qual la veritat sempre acaba obtenint acceptació.
- Europa, i també altres països com Canadà, Austràlia, o Sud-àfrica, han seguit en general una línia més intransigent amb el discurs de l'odi, i s'han recolzat en la figura de l'abús de dret, per negar la possibilitat d'invocar la llibertat d'expressió per emparar la negació de l'holocaust, l'apologia del terrorisme o missatges obertament xenòfobs o racistes.

No obstant, aquesta distinció entre models d'entendre els límits de la llibertat d'expressió no és tan simple. En primer lloc, la diferència s'ha vist atenuada en els últims anys, entre altres motius, perquè el tradicional respecte nord-americà per la llibertat d'expressió ha perdut una mica de terreny com a conseqüència del gir securitari produït en el marc de la denominada "guerra contra el terror". I en segon lloc, perquè tant a Europa com als Estats Units hi ha experts i acadèmics

10. El Pacte Internacional de Drets Civils i Polítics de 1966 va establir límits a la llibertat d'expressió. L'article 19 contempla la possibilitat de restriccions per "assegurar el respecte als drets o a la reputació dels altres" i per a "la protecció de la seguretat nacional, l'ordre públic o la salut o la moral públiques"; i l'article 20 prohibeix "tota apologia de l'odi nacional, racial o religiós que constitueixi incitació a la discriminació, l'hostilitat o la violència".

11. "El Congrés no podrà fer cap llei pel que fa a l'establiment de la religió, (...) ni limitant la llibertat d'expressió, ni de premsa..."

que simpatitzen i defensen la perspectiva i els arguments contraris als que els correspondrien per la seva tradició jurídica (Phillipson 2015).

D'altra banda, al món musulmà, alguns països han pressionat perquè s'adoptin normes internacionals que prohibeixin la blasfèmia, com una forma d'odi religiós, en resposta a casos com el de les caricatures de Mahoma en el diari danès Jyllands-Posten (Phillipson 2015).

En el terreny acadèmic, el debat sobre les prohibicions del discurs de l'odi i la llibertat d'expressió és un debat que transcendeix l'àmbit jurídic i té implicacions d'ordre polític, ètic i filosòfic. S'ha enfocat el debat des de múltiples angles, i es plantegen qüestions molt diverses: Són les prohibicions del discurs de l'odi una amenaça per a la llibertat d'expressió? Es poden considerar un mal necessari? Prohibir el discurs de l'odi és un símbol necessari del compromís democràtic amb el respecte de la dignitat humana i la igualtat? O és la normalització de les prohibicions de discurs de l'odi símptoma d'una democràcia feble? Són les prohibicions del discurs de l'odi eficaces?

En el següent quadre es presenten de manera simplificada alguns dels principals arguments que se solen utilitzar a favor i en contra de les prohibicions del discurs de l'odi.

Taula 1¹²

A favor	En contra
La llibertat d'expressió no és un dret absolut. Ha de ser limitat per evitar que s'atempti contra la dignitat humana, la igualtat, la pau i la convivència social, el dret a viure sense assetjament i intimidació, etc. El Pacte Internacional de Drets Civils i Polítics així ho reconeix en els seus arts. 19.3 i 20.	La llibertat d'expressió i el debat públic són essencials per desenvolupar una democràcia plena. No han d'establir restriccions basant-se únicament en el indesitjables o ofensives que puguin ser determinades expressions. Es corre el risc d'abús per part de governs que les utilitzin com un "dret penal de l'enemic", per castigar a la dissidència i oponents polítics.
La proliferació de comentaris i expressions ofensives i extremes poden generar un clima social que acabi desembocant en episodis de violència i, en contextos conflictius o societats dividides, en atacs genocides.	En democràcies consolidades i pròsperes, no s'ha demostrat que existeixi una correlació clara entre proliferació de discurs de l'odi i un augment de crims d'odi. L'estat ja pot castigar la motivació d'"odi" dels actes criminals com a circumstància agreujant.
El discurs de l'odi causa danys psicològics directes (sentiment de ser amenaçat, humiliat, etc.) i danys indirectes, en el sentit que contribueix a la perpetuació de situacions de discriminació.	Existeixen mitjans legals per prohibir l'assetjament, les amenaces i els atacs individuals a la dignitat; i instruments més legítims, proporcionats i eficaços per combatre la discriminació: lleis de no discriminació laboral, educació en la diversitat i el pluralisme, campanyes de conscienciació, etc.

12. Quadre basat en Nineteen arguments for hate speech bans- and against them d'Eric Heinze.

A favor

Les declaracions més extremes d'odi són incitadores d'actes violents i de discriminació, i per això han de ser restringides i castigades.

En contra (continuació)

L'estat ja pot castigar actes de provocació o conspiració, en els quals es pot establir una clara connexió material amb la comissió d'un delictes; però el concepte d'incitació li atorga un mecanisme per castigar sense haver de demostrar la possibilitat que es realitzi el dany com a conseqüència de l'expressió pública d'idees.

L'estat no pot mantenir-se neutral, i les prohibicions del discurs de l'odi són un símbol i una declaració de principis del compromís amb els valors democràtics, la igualtat, i la dignitat humana.

L'estat té vies més legítimes i eficaces per posicionar-se a nivell simbòlic. Les prohibicions són poc eficaces, i fins i tot contraproductives, perquè es corre el risc de deslegitimar la promoció real de la igualtat i perquè poden convertir certs individus en "màrtirs" de la llibertat d'expressió.

En relació amb l'eficàcia o al caràcter contraproductiu de les prohibicions, és molt interessant l'estudi realitzat a Holanda per Van Spanje, J. i de Vreese, C. (2015), sobre l'impacte que la judicialització del discurs de l'odi de determinats polítics d'extrema dreta té sobre el suport als seus partits. Aquests investigadors van trobar evidències empíriques que l'inici d'una causa judicial contra el discurs anti-immigració de Geert Wilders va reforçar l'atractiu del seu partit per als votants, i suggereixen que la decisió de judicialitzar les seves declaracions islamòfobes va contribuir a la seva arrencada electoral (Van Spanje i de Vreese 2015).

En relació amb la possibilitat que es produeixin abusos polítics del concepte de discurs de l'odi, un cas paradigmàtic és de la Sud-Àfrica de l'Apartheid, on s'utilitzaven les lleis contra aquest tipus de discurs per criminalitzar les crítiques contra la dominació blanca (Gagliardone et al 2015). També a l'estat espanyol hem presenciat recentment diversos exemples d'ús del concepte de discurs de l'odi contra la dissidència i els enemics polítics, per exemple, el cas de César Strawberry, cantant de Def-Con-Dos¹³, a qui el fiscal de l'Audiència Nacional acusava d'incitar a l'odi; el de Guillermo Zapata, regidor de l'Ajuntament de Madrid, a qui es va encausar per uns acudits publicats a Twitter quan debatia sobre els límits de l'humor negre¹⁴; o el cas de la jove valenciana Maria Lluch, "Madame Guillotine" a Twitter, que va ser jutjada per fer acudits sobre víctimes d'ETA, i condemnada a 1 any de presó per considerar el Tribunal Suprem les seves piulades com a discurs de l'odi no protegit per la llibertat d'expressió¹⁵, malgrat rebaijar la condemna inicial de 2 anys per desproporcionalitat. Aquests casos, independentment del judici ètic que a cadascun li puguin merèixer, il·lustren la problemàtica de com el concepte de discurs de l'odi pot utilitzar-se per justificar el càstig penal de certes manifestacions proferides per oponents polítics.

13. http://www.eldiario.es/politica/Absuelto-Cesar-Strawberry-enaltecido-Twitter_0_538946313.html

14. <http://www.lavanguardia.com/politica/20160307/40269028704/archivada-causa-zapata-tercera-vez-tuits.html>

15. <http://www.poderjudicial.es/cgpj/es/Poder-Judicial/Noticias-Judiciales/El-Tribunal-Supremo-condena-a-un-ano-prision-a-una-joven-por-humillar-a-traves-de-twitter-a-Irene-Villa-y-a-Miguel-Angel-Blanco>

No es pot perdre de vista, per tant, la idea que qualsevol legislació, inclosa la relativa a la restricció del discurs de l'odi, és l'expressió del grup dominant que controla el contingut de la llei (Gagliardone et al. 2015).

2.2. Les particularitats jurídiques de l'odi a Internet i les xarxes socials

Per si aquest debat polític, ètic i jurídic no tingués suficients arestes, Internet i les xarxes socials afegeixen una major complexitat al problema del discurs de l'odi. En primer lloc, perquè Internet s'ha convertit en una espècie de fetitxe d'alliberament, considerat utòpicament per molts com “una força revolucionària que no hauria d'estar sotmesa a cap regulació”, i qualsevol intent de regular el que succeeix a la Xarxa, és automàticament acusat de reaccionari, il·legítim i antidemocràtic (Morozov, I., 2012).

Internet i les xarxes socials plantegen un segon problema, consubstancial al caràcter “sense fronteres” de la Xarxa: el problema de la delimitació de la qüestió jurisdiccional. Un missatge emès fent ús d'un servidor d'Estats Units pot tenir conseqüències lesives a Espanya. Quin normativa és aplicable? La d'Estats Units, més tolerant, o l'espanyola?

Aquesta extraterritorialitat suposa també un repte en matèria de cooperació judicial, en la mesura en què les diferències legislatives també afecten a matèries d'importància cabdal en l'àmbit de la cibercriminalitat, com la protecció de dades o el secret de les comunicacions.

Es plantegen, a més, problemes derivats de la pròpia configuració tècnica i l'operativitat d'Internet (localització de servidors, autenticació d'IPs, comptes robot, diferents procediments d'encryptació per ocultar la identitat dels atacs, etc.) que plantegen una altra sèrie d'incerteses i dificultats en qüestió d'obtenció de proves o de determinació de responsabilitat (Moretón Toquero 2012).

Sense entrar aquí en un terreny de gran complexitat jurídica, la transformació del model comunicatiu que suposa Internet situa el debat relatiu al discurs de l'odi davant dues alternatives. La primera, la cerca d'un diàleg internacional per aconseguir un “acord global sobre els criteris de restricció o permissivitat del discurs de l'odi a Internet”, que és poc realista. I la segona, la “imposició a Internet de les polítiques constitucionals de cada estat” (Rodríguez Izquierdo, 2015), que presenta els límits i incerteses jurídiques que s'han exposat aquí superficialment.

2.3. Legislació espanyola

La legislació espanyola relativa al discurs de l'odi a les xarxes socials està condicionada per la necessitat d'ajustar-se a les normes europees, que segueixen la tendència, cada vegada més generalitzada a nivell global, d'introduir restriccions penals al discurs de l'odi¹⁶.

16. En concret, la Decisió Marc 2008/913/JAI, de 28 de novembre, relativa a la lluita contra determinades formes i manifestacions de racisme i xenofòbia mitjançant el Dret penal, de la Unió Europea, estableix unes directrius bàsiques que han de recollir-se en tots els estats membres de la Unió. També és destacable en aquest sentit, el Protocol Addicional al Conveni sobre la cibercriminalitat, relatiu a la penalització d'actes d'indole racista i xenòfoba comesos per mitjà de sistemes informàtics de 2003, del Consell d'Europa.

L'última reforma del Codi Penal¹⁷ respon a l'exigència d'ajustar-se a aquesta regulació europea, i a la sentència del TC 235/2007, de 7 de novembre, que va imposar una interpretació restrictiva del delictes de negació del genocidi, limitant la seva aplicació als casos en què aquesta conducta impliqui una incitació a l'odi o hostilitat contra minories.

La nova regulació ha unificat els diferents delictes de discurs de l'odi a l'article 510 del Codi Penal, i els castiga amb les següents penes:

1. 1 a 4 anys de presó i multa de 6 a 12 mesos
 - a. Incitació directa o indirecta **a l'odi, a l'hostilitat, a la discriminació i a la violència**, contra grups o individus per motius racistes, antisemites o altres relatius a la seva ideologia, religió, origen, sexe, orientació sexual, malaltia o discapacitat
 - b. Producció i difusió de **materials**, amb la mateixa motivació que a l'apartat a).
 - c. **Negació o enaltiment** dels delictes de **genocidi**, lesa humanitat o contra les persones o béns protegits en cas de conflicte armat que haguessin estat comesos contra els grups assenyalats, quan això promogui o afavoreixi un **clima de violència, hostilitat o odi** contra els mateixos
2. penes de presó de 6 mesos a 2 anys i multa de 6 a 12 mesos (i presó d'1 a 4 anys quan es promogui un clima de violència, odi o discriminació)
 - a. Realització d'actes, o producció o difusió de materials, que suposin **humiliació o menyspreu** contra algun dels grups assenyalats o dels seus membres per motius discriminatoris;
 - b. **l'enaltiment o justificació** pública dels **delictes** comesos contra els grups assenyalats.

La regulació de l'article 510 estableix a més una sèrie d'agreujants i conseqüències accessòries. A l'efecte d'aquest informe, és destacable l'agreujant per haver-se dut a terme qualsevol de les conductes descrites fent ús de les "noves tecnologies"¹⁸. En aquest cas, atenent al potencial major abast de les conductes punibles, les penes s'imposaran en la seva meitat superior, i es preveu la retirada dels continguts, i el bloqueig de l'accés o la interrupció de la prestació del servei¹⁹.

Per a part de la doctrina, la reforma corregeix algunes de les deficiències de la regulació anterior. La principal aportació és que elimina la paraula "provocació" que havia generat dificultats d'aplicació de l'article, disparitat de criteris i sentències contradictòries, en interpretar-se en el sentit de l'article 18 del Codi Penal, de "incitació directa". En eliminar-se aquest terme de "provocació", es facilita que es castigui també la "incitació indirecta", que està a més recollida explícitament en el text (Gascón Conca 2015).

També considera, aquesta part de la doctrina, que la inclusió, a l'apartat 2 de l'article, de dos delictes atenuats per a aquelles conductes "que no reuneixen totes les característiques com per incitar a l'hostilitat, a la discriminació o a la violència, però que sí són ideals per lesionar la digni-

17. L.O. 1/2015, de 30 de març

18. Art. 510.3 CP

19. Art. 510.6 CP

tat dels individus”, amb penes inferiors (6 mesos a 2 anys de presó), adequa millor la regulació al principi de proporcionalitat (Gascón Conca 2015).

No obstant, per a altres autors, aquesta reforma és tècnicament pobre i reaccionària, i segueix plantejant dubtes de la seva constitucionalitat, per manca d'adequació a la justificació de la intervenció del dret penal, i per seguir patint de manca de proporcionalitat. Alguns penalistes consideren fins i tot que s'ha aprofundit en la majoria dels defectes de la regulació anterior, deixant massa a criteri dels jutges el sancionar o no conductes que aparentment encaixen en els tipus regulats per l'article 510 (Teruel Lozano 2015).

Per als crítics amb la nova regulació, existeixen respostes millors i més proporcionades que la intervenció penal. L'ordenament espanyol compta amb *“instruments suficients per castigar”*, per exemple les humiliacions, les amenaces o la provocació d'un perill cert i imminent d'actes il·lícits, *“sense que sigui necessari plantejar una resposta excepcional a través de tipus penals ad hoc anti-discriminatoris”* (Teruel Lozano 2015: 39).

Segons Rei Martínez (2015: 52), els problemes parteixen de que s'ha realitzat un enfocament de la qüestió massa inclinat cap al dret penal. Per a ell, *“una sanció administrativa i, fins i tot, l'exigència de responsabilitat civil, podrien ser molt més respectuoses”* amb la Constitució i a més, serien possiblement més eficaces que una norma penal, que per la seva rigorositat i desproporcionalitat, pot acabar resultant inaplicable per als jutges.

Efectivament, fins avui, la jurisprudència ha fet una interpretació restrictiva dels delictes de discurs d'odi. Això ha tingut com a resultat un panorama jurídic de penes molt dures i nul·la sanció del discurs de l'odi. Per a Rei Martínez (2015: 77), és necessària *“més sanció dels discursos de l'odi, però amb una duresa penal proporcionada”*. La nova regulació, malgrat la introducció dels tipus atenuats de l'apartat 2 de l'article 510, ha millorat poc en aquest aspecte, en establir penes de fins a 4 anys de presó, apartant-se fins i tot de la Decisió Marc 2008/913/JAI de la UE, que estipula penes màximes de 3 anys (Gascón Conca, 2015).

Veurem quina és la interpretació que els jutges i magistrats fan de la nova regulació. En qualsevol cas, seria desitjable que no incorreguessin en les contradiccions que es van produir en aplicar la legislació anterior, i que alguns experts, com Miguel Ángel Aguilar, fiscal coordinador de Delictes d'Odi de la Fiscalia de Barcelona, van criticar durament²⁰, quan el Tribunal Suprem va dictar dues sentències contradictòries en el termini d'una setmana.

En qualsevol cas, no s'ha d'oblidar que la repressió penal del que genèricament s'entén per discurs de l'odi no s'esgota en l'article 510 del Codi Penal. El Codi Penal contempla una circumstància agreujant genèrica per motius discriminatoris, que agreuja la responsabilitat penal de qualsevol delictes, i que, per exemple, pot aplicar-se a conductes típiques a les xarxes socials²¹, com les amenaces o l'assetjament. La consideració de que concorre l'agreujant de discriminació suposa que la pena corresponent a cada delictes s'imposarà en la seva meitat superior. A més existeix un delictes d'amenaces a col·lectius²², que castiga les amenaces, amb entitat suficient com per aconseguir atemorir, dirigides a grups ètnics, culturals, etc.

20. <http://www.elmundo.es/elmundo/2011/06/13/barcelona/1307982347.html>

21. Art. 22.4 CP

22. Art. 170.1 CP

Principals problemes per a la persecució dels delictes de discurs de l'odi

Alguns dels principals problemes, d'ordre pràctic, que dificulten una efectiva persecució dels delictes d'odi en general, i dels delictes de discurs de l'odi a les xarxes socials en particular són, segons el parer dels experts (Aguilar et al. 2015), els següents:

- Desconeixement dels fets per l'escàs nombre de denúncies.
- Minimització dels fets en alguns casos per part dels responsables de les investigacions.
- Insuficients investigacions policials i judicials.
- Dificultats per acreditar la motivació “odiosa” o discriminatòria del delicte.
- Poca formació en igualtat i no discriminació de les institucions implicades: cos judicial, funcionaris judicials, forenses, policia, seguretat privada, etc.
- Dificultats inherents a l'escenari “xarxes socials”: problemes de jurisdicció, territorialitat, identificació i obtenció de dades dels agressors, etc.

Normes administratives

Des del punt de vista del dret administratiu, existeixen algunes normatives que contemplen la possibilitat de sancionar conductes que poden ser enteses com a discurs de l'odi. En aquest sentit, la normativa més específica és la Llei 19/2007, contra la violència, el racisme, la xenofòbia i la intolerància en l'esport, que és un àmbit d'especial incidència de conductes intolerants, i en el qual els grups ultres d'extrema dreta van trobar durant anys un espai propici per deixar anar les seves actituds obertament racistes, homòfobes, i discriminatòries, amb relativa impunitat.

Pel que fa a l'objecte específic d'aquest informe, aquesta llei considera la difusió, per mitjans informàtics o tecnològics, vinculats a informació esportiva, “*de continguts que promoguin o donin suport a la violència, o que incitin, fomentin o ajudin als comportaments violents o terroristes, racistes, xenòfobs o intolerants per raons de religió, ideologia, orientació sexual (...)*” com a infracció molt greu, i la castiga amb importants sancions econòmiques.²³

Institucions especialitzades i protocols policials contra l'odi i la discriminació

Al març de 2013, el Ministeri Fiscal va posar en marxa una xarxa de **fiscals especialitzats contra l'odi**, amb la designació d'un fiscal a cada província, encarregat de coordinar l'actuació en matèria d'odi i discriminació de la Fiscalia de la província, i de coordinar aquesta actuació amb la resta de la xarxa de fiscals provincials especialitzats en la matèria.

Aquesta figura es va inspirar en l'exemple del Servei de Delictes d'Odi i Discriminació de la Fiscalia de Barcelona, constituït el 2009, i que va ser pioner en aquest àmbit, i ha estat reconegut per l'Agència de Drets Fonamentals de la Unió Europea com un exemple de bona pràctica, que hauria de ser reproduït en altres estats de la UE.

23. Art. 23.1.c) i art. 24.1.c) de la Llei 19/2007. Sancions des de 60.000,01€ a 650.000€

Entre els reptes que es planteja aquest servei, es troben els següents (Aguilar et al. 2015):

- Obtenir millor informació sobre fets que poguessin constituir conductes punibles en aquesta matèria;
- promoure la comunicació amb la Fiscalia dels col·lectius i ciutadans afectats, tant per denunciar com per compartir iniciatives i suggeriments;
- contribuir a unificar els criteris d'aplicació i interpretació del Codi Penal;
- afrontar la complexitat de les investigacions dels delictes d'incitació a l'odi, la violència i la discriminació, quan es produeixen a Internet i xarxes socials;
- recopilar recomanacions i jurisprudència emesa per organismes i tribunals internacionals en matèria de no discriminació;
- combatre els grups violents organitzats que difonen el discurs de l'odi;
- afrontar la tendència a minimitzar la importància dels delictes d'odi per part d'alguns policies, jutges i fiscals;
- millorar la coordinació amb les Forces i Cossos de Seguretat de l'Estat.

En aquest sentit, també és destacable el desenvolupament de **protocols d'intervenció policial**. Els Mossos d'Esquadra van ser els pioners en aquesta matèria en dotar-se d'un "Procediment de fets delictius motivats per l'odi o la discriminació" el 2010. El Ministeri de l'Interior va publicar el 2014 el "Protocol d'actuació per a les Forces i Cossos de Seguretat per als delictes d'odi i conductes que vulnereu les normes legals sobre discriminació", dirigit a la Policia Nacional i la Guàrdia Civil, basant-se en materials editats per l'Observatori Espanyol del Racisme i la Xenofòbia, del Ministeri d'Ocupació i Seguretat Social (Aguilar García 2015).

Aquests protocols responen a la necessitat de formar als cossos policials i dotar-los d'unes pautes en la recollida de dades estadístiques, l'elaboració dels atestats, la investigació dels fets per acreditar l'existència d'una motivació racista o intolerant, l'atenció i la protecció de les víctimes, o la interlocució amb la comunitat i les ONGs que representen a les víctimes.

El Protocol del Ministeri d'Interior contempla recomanacions específiques per als delictes d'odi comesos a través d'Internet i les xarxes socials. En concret, i donades les dificultats que l'ús de les tecnologies de la informació suposa per a la persecució i identificació dels responsables de les infraccions, recomana que les investigacions siguin realitzades per Unitats especialitzades de la Policia Judicial (Ministeri d'Interior 2015).

A més, tant les institucions públiques com algunes organitzacions no governamentals (SOS Racisme, Movimiento contra la Intolerancia, Observatori contra l'Homofòbia, Fundació Rais, etc.) ofereixen una sèrie de serveis orientats a assessorar i assistir jurídicament a les víctimes dels delictes d'odi, en general, i del discurs de l'odi, en particular. A Catalunya, els Ajuntaments de Sabadell, Terrassa, i Barcelona (Oficina per la No Discriminació, OND) compten amb diferents serveis, que ofereixen assessorament jurídic en aquest camp.

3. EL ROL DE LES XARXES SOCIALS COM A INTERMEDIÀRIES

Indiscutiblement, les pròpies xarxes socials, com a canal de comunicació dels missatges “odiosos”, juguen un paper clau en la lluita contra el discurs de l’odi. Cal recordar que la xarxa social per excel·lència, Facebook, compta amb més de 1.700 milions d’usuaris actius²⁴, que seria aproximadament l’equivalent a la suma de la població dels continents europeu i americà. Aquestes xifres massives converteixen a les xarxes socials en un actor, amb un rol d’intermediari, però amb un poder innegable.

El seu paper d’intermediàries en la comunicació digital, les converteix en el primer àrbitre en la determinació de què es pot dir i què no. En aquest sentit, convé recordar que les grans xarxes socials més usades a Europa tenen la seva seu a Estats Units, i per tant estan imbuïdes de la filosofia nord-americana, més defensora d’una llibertat d’expressió més absoluta i més tolerant amb cert discurs de l’odi. Aquesta circumstància té, com hem assenyalat en la secció anterior, repercussions legals, per exemple, en termes de col·laboració amb la justícia pel què fa a requeriments de dades. Però també influeix en les seves polítiques d’autoregulació i en la forma de portar-les a la pràctica.

En relació amb les empreses intermediàries, es plantegen diferents alternatives: l’atribució de responsabilitat a les empreses, la promoció de l’autoregulació i la implementació de filtres informàtics (Rodríguez Izquierdo 2015). Qualsevol d’aquestes vies planteja certs dilemes.

3.1. Atribució de responsabilitat

Mentre que als Estats Units els proveïdors de serveis estan pràcticament exempts de responsabilitat pels continguts publicats, l’atribució de responsabilitat als prestadors de serveis intermediaris a la Unió Europea està basada principalment en el coneixement efectiu del caràcter il·lícit dels continguts publicats en el seu servei (Rodríguez Izquierdo 2015). En principi, només seran responsables quan no retirin o impedeixin l’accés a un contingut il·lícit amb promptitud, havent tingut coneixement del seu caràcter il·lícit per mitjà d’una notificació del titular dels drets o per mitjà d’una ordre d’un òrgan competent.

La sentència Delfi del Tribunal Europeu de Drets Humans va establir la responsabilitat del portal de notícies estonià, Delfi, en considerar-lo com a proveïdor de continguts, pels comentaris publicats pels seus lectors, fins i tot sense haver rebut una sol·licitud de retirada del contingut il·lícit. No obstant, a la sentència, el Tribunal va determinar que aquesta interpretació no és aplicable a les xarxes socials, establint una diferenciació entre proveïdors de serveis tècnics, amb un rol passiu, com les xarxes socials, i els proveïdors de serveis de continguts, com una web de notícies o un mitjà de comunicació *online*²⁵.

24. <http://newsroom.fb.com/company-info/>

25. <http://blog.garrigues.com/se-confirma-la-responsabilidad-de-los-portales-de-noticias-por-los-comentarios-difamatorios-de-los-lectores/>

3.2. Autoregulació i implementació de filtres informàtics

La persecució i restricció dels delictes de discurs de l'odi per part dels estats s'enfronta a les dificultats ja assenyalades per actuar fora de les seves fronteres. Malgrat això, i aquí resideix una dels principals avantatges de les estratègies d'autoregulació, les companyies, com a operadors privats que són, tenen el dret d'establir codis ètics i condicions d'ús dels seus serveis, aplicables globalment.

Totes les grans xarxes socials, Twitter, Facebook, Instagram, etc., i fins i tot els mitjans de comunicació, els fòrums de comentaris dels quals també són un canal habitual del discurs de l'odi, compten amb polítiques o condicions d'ús que, en principi, prohibeixen el discurs de l'odi en les seves plataformes de comunicació.

Twitter, per exemple, prohibeix expressament²⁶ les amenaces, la incitació a la violència, l'assetjament, així com el comportament que incita a l'odi, per motius d'origen ètnic, nacionalitat, orientació sexual, gènere, religió, etc.

Facebook inclou en les seves condicions de servei clàusules relatives a la no publicació de contingut intimidatori, discriminatori, incitador de la violència, però reconeix expressament la incapacitat de garantir que Facebook sigui un lloc segur, i enuncien les seves normes com compromisos de l'usuari. Des de la perspectiva que Facebook constitueix una comunitat global, sotmesa a diferents legislacions nacionals, estableixen unes normes comunitàries per tal d'autoregular-se²⁷. Facebook exposa, mitjançant aquestes normes, quin contingut es pot denunciar i eliminar, destacant, a l'efecte d'aquest informe, l'eliminació del llenguatge que incita a l'odi per motius de raça, ètnia, nacionalitat, religió, orientació sexual, sexe, gènere, identitat sexual, o discapacitats o malalties greus.

La majoria de les xarxes socials, incloses les dues grans xarxes socials, Facebook i Twitter, compten amb sistemes de denúncia que permeten als usuaris posar en coneixement de les empreses la presència de contingut abusiu que vulneri les normes d'ús de les plataformes, perquè les companyies valorin la seva eliminació. Facebook adverteix expressament en les seves normes comunitàries, que no tot contingut ofensiu infringeix les seves normes, i per això, ofereix a més la possibilitat de bloquejos i filtres personalitzats per ocultar aquell contingut que l'usuari no desitja veure.

No obstant, malgrat aquesta actitud teòricament intolerant amb el discurs de l'odi per part de les empreses, la veritat és que els mecanismes de denúncia i eliminació de comentaris es porten a la pràctica de manera una mica disfuncional. Així ho testifica una recerca del projecte PRISM (Jubany i Roiha 2016), en la qual d'un total de 100 denúncies realitzades a Facebook, només 9 van obtenir com a resultat l'eliminació dels comentaris²⁸. L'erràtica política d'eliminació de comentaris per part de l'empresa, ja sigui pel mal funcionament dels algorismes que empren o perquè els equips encarregats de gestionar les denúncies estiguin superats per l'ingent volum de comentaris denunciats, va tenir com a resultat l'eliminació de missatges com "Els gossos són més discrets que els musulmans", mentre que un missatge com "Matar-los a tots sense pietat" va ser considerat per Facebook com un missatge que no vulnerava les normes de comunitat.

26. <https://support.twitter.com/articles/72688>

27. <https://www.facebook.com/communitystandards#>

28. <http://www.elperiodico.com/es/noticias/sociedad/estudio-proyecto-prism-denuncia-banalizacion-discurso-odio-internet-5284591>

Una de les principals crítiques que s'ha fet sobre els mecanismes de denúncia de les xarxes socials és la seva falta de transparència. No se sap amb certesa si el procés el realitzen mitjançant algorismes, mitjançant equips humans o amb una barreja de tots dos. D'altra banda, tampoc fan públiques dades sobre el nombre de denúncies, els motius de les denúncies, o el percentatge de denúncies que finalitzen amb comentaris o comptes eliminats (Jubany i Roiha 2016).

Facebook ha desenvolupat també un potent sistema d'intel·ligència artificial, denominat Deep-Text, suposadament capaç d'analitzar milers de posts per segon en més de 20 llengües. Malgrat això, a dia d'avui no sembla que el seu funcionament sigui un èxit. Per molt sofisticats que siguin els algorismes en els que es basen aquest i altres filtres informàtics desenvolupats per altres xarxes socials, tendeixen a cometre errors importants i no solen ser capaços, per exemple, de destriar quan s'està fent un ús crític o de denúncia d'una determinada expressió insultant. El cas de Shaun King, periodista de New York Daily News i activista de Black Lives Matter, va ser especialment simptomàtic. Malgrat que havia estat convidat mesos abans per Facebook com a ponent per parlar d'aquest moviment polític contra la violència cap a les persones negres, el seu compte de Facebook va ser suspès per publicar un correu electrònic en el qual li dirigien insults racistes. Gràcies als seus contactes a Facebook, l'activista va aconseguir que se li restablís el servei, però ell mateix es qüestionava com pot gestionar un cas similar una persona que no tingui l'accés privilegiat a determinades instàncies de la companyia²⁹.

Les pròpies xarxes socials són conscients del problema que tenen amb la gestió de l'assetjament *online*, les amenaces i el discurs de l'odi. El 2015, es va filtrar un comunicat intern de l'aleshores CEO de Twitter, Dick Costolo, en el qual reconeixia que tenien un greu problema en la gestió d'aquest assumpte, que els estava fent perdre usuaris cada dia i limitant en els seus objectius de creixement³⁰. Per aquesta raó, Twitter ha començat a considerar prioritari aquest assumpte, i l'empresa està immersa en un procés de canvis que permetrà als usuaris un control més proactiu del seu compte de Twitter. L'empresa ha buscat l'assessorament d'una comissió externa d'organitzacions de la societat civil involucrades en la prevenció d'abusos *online*³¹, i estan treballant, entre altres eines, en la implementació de filtres de qualitat per bloquejar determinades paraules, *hashtags*, insults racistes, etc. perquè no apareguin en el seu "*timeline*". El seu objectiu és trobar millors mecanismes per bloquejar a altres usuaris, i agilitzar i optimitzar els sistemes de denúncia a la companyia de comportaments abusius.

Un cas paradigmàtic, en aquesta línia de menor tolerància de Twitter amb l'odi *online*, es va produir al juliol de 2016, amb la suspensió temporal de Milo Yiannopoulos, un icona del moviment *alt-right* a Estats Units, per liderar suposadament l'assetjament misogin i racista a l'actriu afroamericana Leslie Jones. El caràcter mediàtic de Milo, i el fet que comptés amb 300.000 seguidors a Twitter, van convertir el cas en un debat sobre la llibertat d'expressió, convertint-se en tendència el *hashtag* #FreeMilo, que el convertia en una mena de màrtir de la llibertat d'expressió³².

No hi ha dubte que el fet que la persona atacada fos una celebritat de Hollywood va facilitar la visibilitat i la resolució del cas, en el qual va intervenir públicament l'actual CEO de la companyia, Jack Dorsey. El problema, més enllà de la importància de casos simbòlics com aquest, el tenen les persones "anònimes" que pateixen humiliacions i vexacions a les xarxes socials, però que no tenen la capacitat de que els seus casos arribin a tan altes instàncies. De moment, caldrà espe-

29.http://www.eldiario.es/theguardian/Facebook-temporalmente-Black-Lives-Matter_0_558544311.html

30.<http://www.lavanguardia.com/tecnologia/20160901/4122751789/twitter-acoso-trolls-redes-sociales.html>

31.<https://about.twitter.com/es/safety/council>

32.<http://www.bbc.com/mundo/noticias-36847433>

rar per comprovar si els canvis tecnològics en els que treballa la companyia suposen un canvi real en la gestió del problema de l'abús i l'odi *online*.

En qualsevol cas, per equilibrar i matisar les crítiques que se solen fer a aquestes companyies, és just recordar aquí l'enorme dificultat de la tasca a la qual s'enfronten. Twitter, per exemple, maneja uns 300 milions de piulades al dia, i han arribat fins a 600 milions³³. Fins i tot amb l'ajuda de filtres i algorismes informàtics, la tasca de moderació i autoregulació és de proporcions gegantines, i al mateix temps, molt delicada, en la mesura en què entren en joc conceptes tan sensibles per a la seva imatge com la llibertat d'expressió o la censura.

3.3. El Codi de Conducta amb la Unió Europea

Com a conseqüència de la crisi de refugiats de 2015, i davant l'evident creixement de les manifestacions xenòfobes i racistes a les xarxes socials, la Unió Europea va pressionar a les empreses tecnològiques perquè assumissin un rol més actiu en la lluita contra el discurs de l'odi.

Com a resultat de la pressió d'alguns estats i de la pròpia Unió Europea, al maig de 2016, Facebook, Twitter, Youtube i Microsoft van signar un Codi de Conducta en matèria d'incitació il·legal a l'odi a Internet³⁴. En aquest acord, les empreses tecnològiques van adquirir una sèrie de compromisos, entre els quals destaca el compromís de revisar les sol·licituds de retirada de continguts il·legals d'incitació a l'odi en el termini de 24 hores. A més, es van comprometre a establir procediments clars d'examen dels continguts denunciats, i a valorar les denúncies que rebien en els seus serveis, conformement a les seves normes d'autoregulació, però tenint en compte també, quan sigui necessari, les legislacions nacionals de transposició de la Decisió Marc europea relativa a la lluita contra determinades formes i manifestacions de racisme i xenofòbia mitjançant el Dret penal, que l'acord considera la seva base jurídica.

L'acord busca la cooperació entre les empreses signatàries i l'intercanvi de bones pràctiques amb altres serveis de comunicació social d'Internet, i es promourà l'adhesió d'aquestes a l'acord.

Les empreses signatàries també es comprometen a promoure iniciatives que desenvolupin un "contradiscurs" alternatiu, a recolzar programes educatius que fomentin el pensament crític, i a col·laborar amb les organitzacions de la societat civil en activitats formatives. Aquest compromís signat en el Codi de Conducta formalitza i compromet una línia d'actuació que ja havien anant desenvolupant les empreses tecnològiques en aquest àmbit per millorar la manera de gestionar l'assetjament i l'odi a les xarxes.

El compliment del codi de conducta i els seus resultats seran sotmesos a avaluació pel recentment creat **Grup d'Alt Nivell sobre la Lluita contra el Racisme, la Xenofòbia, i altres formes d'Intolerància**, que està format pels estats membres de la UE, representants del Parlament Europeu, del Consell d'Europa, de l'Alt Comissionat de l'ONU per als Refugiats, de l'Organització per a la Seguretat i la Cooperació a Europa, de l'Agència Europea de Drets Fonamentals, i d'organitzacions de la societat civil.

Encara que la Comissària europea de la Direcció General de Justícia, Consumidors i Igualtat de Gènere, Vera Jourová, va presentar aquest acord com un punt d'inflexió, a partir del qual

33.<http://uk.businessinsider.com/tweets-on-twitter-is-in-serious-decline-2016-2>

34.http://europa.eu/rapid/press-release_IP-16-1937_es.htm

canviarien les regles del joc³⁵, la veritat és que la signatura d'aquest codi de conducta no ha estat exempta de crítiques. En primer lloc, s'ha lamentat que el procés de negociació no hagi estat transparent i que no s'hagin tingut en compte les aportacions de les organitzacions de la societat civil. A més es lamenta que, mitjançant aquest codi de conducta, les companyies prendran el lideratge en un assumpte, el del control del discurs de l'odi *online*, que hauria de ser liderat per les autoritats públiques, sotmeses a un marc de responsabilitat democràtica. Deixar el lideratge d'aquesta tasca a companyies privades suposa un clar risc per a la llibertat d'expressió, en la mesura en què els mecanismes de revisió de contingut de les companyies no estan sotmesos a un marc de responsabilitat clar i transparent³⁶. En el fons, manifesten els crítics, es tracta de confiar en la responsabilitat de les empreses o, més aviat, en que l'interès dels seus accionistes coincideixi, d'alguna manera, amb l'interès social, el respecte a la llibertat d'expressió, i els principis d'igualtat, dignitat, etc. propis dels sistemes democràtics.

Una altra qüestió que es planteja com a problemàtica és que, en la mesura en què el presumpte discurs de l'odi sigui eliminat per les companyies, i no s'estableixi un procediment de col·laboració clar per a l'enviament d'aquest contingut a les autoritats públiques, els autors de les suposades infraccions legals no seran sotmesos a judici ni càstig penal de cap tipus³⁷.

En qualsevol cas, no hi ha dubte que les institucions europees estan seriosament preocupades amb el problema del discurs de l'odi *online*, i que estan tractant de pressionar a les companyies perquè realitzin una moderació més activa de les xarxes socials.

35. http://europa.eu/rapid/press-release_SPEECH-16-2197_en.htm

36. <https://edri.org/edri-access-now-withdraw-eu-commission-forum-discussions/>

37. <https://edri.org/guide-code-conduct-hate-speech/>

4. LES ESTRATÈGIES NO JURÍDIQUES PER COMBATRE EL DISCURS DE L'ODI A LES XARXES SOCIALS

En aquesta secció es presentaran, sense ànim exhaustiu, les principals estratègies “no jurídiques” que s'estan duent a terme per respondre al discurs de l'odi en general, i a les xarxes socials en particular.

Bona part d'aquestes estratègies estan directa o indirectament finançades per les institucions públiques, i per a alguns autors es plantegen dues qüestions importants. D'una banda, si són realment eficaces. I d'altra banda, la possibilitat que es produeixi una certa vulneració d'un suposat principi de neutralitat de l'estat (Kahn 2015). Malgrat això, la veritat és que fins i tot la majoria dels autors que s'oposen a les prohibicions del discurs de l'odi consideren que les institucions públiques no han de mantenir-se neutrals davant aquest tipus de discurs extrem. Molt pocs autors defensen un principi de neutralitat estatal estricta; aquests temen que una implementació equivocada o excessiva de certes estratègies pugui produir efectes no desitjats, com la marginació o l'expulsió del debat públic de certs discursos, que poden ser desagradables però legítims. Per a la gran majoria, però, no hi ha dubte que les institucions públiques han de dur a terme una important tasca en la creació d'un clima adequat per a un exercici de la llibertat d'expressió respectuós amb els principis d'igualtat i no discriminació (Article 19 2015).

Les administracions poden, a través de les declaracions dels polítics i els funcionaris públics, promoure l'empatia i el suport als grups víctimes del discurs de l'odi, difonent la idea que aquest tipus de manifestacions són nocives per a la convivència. Aquest tipus de declaracions són de vital importància en situacions de disputes entre comunitats ètniques en barris o ciutats o en contextos en els que les tensions identitàries empitjoren, com els períodes electorals. Altres iniciatives que les administracions públiques poden desenvolupar per garantir un clima adequat per a la igualtat i la no discriminació són la formació dels funcionaris públics en aquestes qüestions, la litigació estratègica, o la promoció del diàleg interreligiós (Article 19 2015).

En un pla més concret, i en referència a iniciatives dirigides específicament contra el discurs de l'odi *online*, i en particular a les xarxes socials, s'estan duent a terme estratègies no jurídiques de diversos tipus:

- Estratègies basades en el monitoreig i la recerca, que pretenen conèixer millor la magnitud del problema del discurs de l'odi i explorar la possibilitat d'utilitzar sistemes d'alerta precoç capaços d'identificar-lo automàticament;
- Estratègies de pressió a les empreses operadores de les xarxes socials, encaminades a que modifiquin les seves polítiques relatives al tipus de continguts que es poden compartir o a que retirin continguts específics;
- Estratègies dirigides als usuaris de la Xarxa i que s'han centrat a canviar les seves percepcions i actituds enfront del discurs de l'odi tant a Internet en general com a les xarxes socials en particular.

Entre aquest últim tipus d'estratègies, a les que aquest informe posarà especial atenció per la seva major aplicabilitat a nivell local, destaquen dos grans tipus: d'una banda, les que consisteixen en campanyes, que inclouen activitats d'índole diversa i l'objectiu de les quals sol ser atacar

els prejudicis i la intolerància com a arrel principal del problema del discurs de l'odi; i, d'altra banda, les estratègies que es basen en la formació i educació, que inclouen la realització de tallers, seminaris i altres tipus d'activitats destinades a dotar a les persones, especialment als joves, de les competències necessàries per identificar i combatre el discurs de l'odi a Internet.

4.1. Les estratègies de monitoreig i recerca

Aquest tipus d'estratègies són importants per entendre el fenomen del discurs de l'odi a les xarxes socials d'una manera integral: l'extensió del problema, els diferents tipus d'odi, com i en quines situacions es manifesta amb més virulència, quins són els perfils dels autors, quin tipus de llenguatge utilitzen, quines plataformes són més propícies per a la propagació del discurs extrem, quins efectes produeix en altres usuaris, etc. Aprofundir en la comprensió del discurs de l'odi és valuós pel propi coneixement, però sobretot, per aplicar-ho a solucions que disminueixin el problema, o almenys, a la prevenció de les seves conseqüències més nocives.

Les tasques de monitoreig tenen dificultats considerables, malgrat els indubtables progressos tècnics en l'automatització de la identificació del discurs de l'odi. En primer lloc, per una qüestió conceptual. Com s'assenyala en la primera part d'aquest informe, què és i què no és discurs de l'odi és objecte de controvèrsia, i aquestes diferències de concepte poden generar resultats poc homologables. En segon lloc, per la pròpia infinitud de les xarxes socials. En tercer lloc, pels diversos tipus de contingut (fotografies, vídeos, àudios, textos, etc.) a través dels quals el discurs de l'odi pot manifestar-se a les xarxes socials. I en últim lloc, perquè fins i tot quan el mitjà és textual, existeixen multitud de combinacions semàntiques i codis a través dels quals canalitzar el discurs de l'odi, sense necessitat d'utilitzar els insults i expressions que un programador pot preveure (Ruiz et al. 2010). Una prova fefaent de l'enorme dificultat d'aquesta tasca són els errors importants comesos pels algorismes de filtrat automàtic de les pròpies empreses operadores de xarxes socials, com Facebook, Twitter, Tumblr, etc, malgrat que compten amb els recursos tècnics i humans més potents.

Encara que amb prou feines existeixen recerques sistemàtiques sobre les connexions entre el discurs de l'odi i el desencadenament d'episodis de violència, en el pla més pràctic, les tasques de monitoreig poden exercir una tasca considerable com a sistemes d'alerta precoç en contextos inestables o en societats ètnicament dividides, per evitar successos com la tràgica onada de violència desencadenada després de les eleccions de Kenya de 2007 (Gagliardone et al 2015).

D'altra banda, en el camp de la recerca tenen especial interès els estudis i anàlisis, en l'àmbit de la psicologia social, sobre quines iniciatives són més efectives contra el discurs intolerant. Actualment són abundants els programes dirigits a combatre la intolerància i els prejudicis, però molt pocs estan basats en evidències sòlides de la seva eficàcia. Tenint en compte que bona part d'aquestes iniciatives són finançades directa o indirectament per les administracions públiques, aquest camp de recerca adquireix una importància cabdal perquè aquesta inversió de diners públics sigui socialment rendible, o almenys, no contraproductiu (Legault et al 2011).

Un dels centres de referència en el camp del monitoreig i la recerca, per la consistència dels seus estudis, és el *Centre for the Analysis of Social Media* del *think tank* britànic Demos que realitza freqüents estudis no només sobre l'extensió de diferents formes de discurs de l'odi (islamofòbia, sexisme, etc.), sinó també sobre com funcionen les contranarratives que s'enfronten a l'odi a les xarxes. En aquest camp d'estudi, Facebook ha encarregat a Demos la realització d'una sèrie de recerques sobre com es produeixen i comparteixen els continguts

que s'enfronten al discurs de l'odi, quins continguts tenen més èxit i quins són potencialment contraproductius³⁸.

4.2. Les estratègies de pressió a les xarxes socials

Conscients del rol clau de les empreses operadores de xarxes socials com a intermediàries dels missatges extremistes i intolerants, moltes organitzacions han optat per exercir pressió sobre aquestes empreses perquè actuïn de manera més contundent contra el discurs de l'odi i perquè siguin més transparents en relació amb els resultats dels seus mecanismes i polítiques de moderació i retirada de continguts. Aquest tipus d'estratègies poden enfocar directament la pressió cap a les xarxes socials, o fer-ho indirectament dirigint-se als seus anunciants.

Les estratègies de pressió directa a les xarxes socials s'han dut a terme tant des d'organitzacions de la societat civil, com des de les institucions públiques. Les organitzacions de la societat civil pressionen a les companyies a través de campanyes *online*, recollida de queixes, peticions *online*, etc. Per la seva banda, alguns estats i organitzacions supranacionals pressionen a les empreses perquè compleixin amb les lleis nacionals i internacionals i es comprometin en una moderació de contingut més activa, per temor a sancions i exigències de regulació més costoses (Article 19 2015). Com a exemple de resultat de la pressió de les institucions europees, al maig de 2016, Facebook, Twitter, Youtube i Microsoft van signar el Codi de Conducta amb la Unió Europea, al que s'ha fet referència en la secció precedent d'aquest informe.

Les campanyes de pressió als anunciants, per la seva banda, han aconseguit considerables èxits. Un exemple d'aquest tipus d'iniciatives va ser desenvolupat conjuntament a Gran Bretanya i Estats Units per WAM! (*Women, Action and the Media Group*) i *Everyday Sexism Project*, per exigir la retirada de contingut abusiu cap a les dones. Com a conseqüència d'una campanya de comunicació intensa, que va incloure la recollida de més 225.000 firmes a través de *change.org*, l'ús del *hashtag* #FBrape o la publicació d'un llistat de companyies que s'anunciaven en pàgines sexistes a Facebook, van aconseguir que 15 grans empreses retiressin la seva publicitat de Facebook, i que la companyia es posés en contacte amb les organitzadores de la campanya per buscar vies de cooperació, i es comprometés públicament a revisar les seves condicions de servei i els seus mecanismes de revisió de continguts (Gagliardone et al. 2015).

4.3. Les estratègies orientades a canviar percepcions i actituds

4.3.1. Les campanyes contra el discurs de l'odi a Internet

Atenent al seu contingut i orientació específica, aquestes campanyes poden categoritzar-se en tres tipus: de sensibilització, afirmatives o restrictives (Titley, Keen i Földi 2014). Abans d'examinar aquesta tipologia de campanyes que s'han dut a terme i alguns exemples específics de cadascun dels tipus, val la pena fixar-se en la campanya *Vivint junts online: acció i campanya contra el discurs de l'odi* (campanya *No Hate*), llançada pel Consell d'Europa (COE) al març de 2013, per ser una campanya integral, que inclou en si mateixa una àmplia varietat d'estratègies contra l'odi a Internet. Aquesta campanya ocupa un lloc privilegiat en l'àmbit de la lluita contra

³⁸<https://www.demos.co.uk/project/counter-speech-on-facebook-phase-2/>

el discurs de l'odi a Internet, ja que ha funcionat en els últims tres anys com un paraigua per a una multitud d'iniciatives en els nivells nacional, regional i local de l'entorn europeu.

a) La campanya *No Hate* del Consell d'Europa

La campanya *No Hate* s'emmarca dins de l'esforç del COE de promoure els drets humans a Internet i està dirigida especialment als joves europeus. La idea de realitzar aquesta campanya va sorgir dels joves representants del Consell Consultiu de la Joventut del COE després de la matança de Utøya a Noruega a l'estiu de 2012. Aquest atac en la que van morir 77 joves va posar de manifest els perills associats al discurs de l'odi a Internet, que en alguns casos pot traduir-se en violència contra determinats grups socials. Malgrat això, més que advocar per la simple repressió d'aquest discurs, aquesta campanya adopta un marc de drets humans i, per tant, reconeix també el dret a la lliure expressió, apostant per fomentar el desenvolupament de l'autoregulació per part dels usuaris (Keen i Georgescu 2016).

L'objectiu principal d'aquesta campanya és mobilitzar els joves europeus, creant un moviment social contra el discurs de l'odi a Internet (*No Hate Speech Movement*). Per això, aquesta campanya es va marcar quatre objectius específics: en primer lloc, reduir el nivell d'acceptació del discurs de l'odi *offline* i *online*; en segon lloc, prevenir i contrarestar aquest problema amb l'educació en drets humans; en tercer lloc, sensibilitzar sobre els riscos que comporta la propagació del discurs de l'odi per a la democràcia i el benestar dels joves; i, en últim lloc, desenvolupar i difondre eines i mecanismes per denunciar-lo, entre els que destaquen l'Observatori de Discurs d'Odi (*Hate Speech Watch*), que permet que els usuaris remetin casos de discurs d'odi que hagin identificat a les xarxes o en altres espais digitals.

L'abast d'aquesta campanya inclou Internet en el seu conjunt, però les xarxes socials tenen una especial importància, tant com a objecte de monitoreig de la campanya com pel seu ús com a eina en el seu combat. En aquest sentit, en la campanya s'han engegat activitats centrades en l'àmbit específic de les xarxes socials. Una de les iniciatives de la campanya que ha tingut major repercussió han estat els *Action Days*, en els quals els seus membres actuen de manera intensificada durant un dia per sensibilitzar sobre un tema específic, per exemple l'antisemitisme (9 de novembre de 2016). En concret, els *Action Days* preveuen accions específiques en les que els activistes vinculats a la campanya intervenen a les xarxes de quatre maneres diferents: en primer lloc, expressant solidaritat amb les persones víctimes de l'odi, a través de la difusió de la imatge del *Action Day* proporcionada per la campanya o afegint aquesta imatge a la foto del perfil de l'usuari a les xarxes socials; en segon lloc, denunciant el contingut d'odi a l'Observatori de Discurs d'Odi i afegint un contraargument, que es comparteix després a les xarxes; en tercer lloc, compartint el contingut amb contranarratives proporcionat per la campanya, en formats com MEMEs, imatges, infografies, etc.; i, per últim, organitzant activitats *offline*, amb l'objectiu de sensibilitzar i educar al públic jove sobre el problema, seguint alguns recursos proporcionats en el marc de la campanya, i que també es dinamitzen a les xarxes socials.

Aquesta campanya s'ha reproduït a nivell nacional, vehiculant-se a través dels Comitès Nacionals de Campanya. Aquests reuneixen a les principals parts interessades a cada país i segueixen les directrius proporcionades pel Consell d'Europa per mantenir la coherència general de la campanya. Els Comitès estan coordinats en la seva gran majoria pels departaments responsables de joventut a nivell de cada Estat, encara que en alguns casos estan formats solament per ONGs i altres parts interessades. Fins ara, s'han format més de 40 Comitès, tant als països membres del COE com en altres països com Estats Units, Mèxic o el Marroc. Els Comitès han engegat diferents campanyes nacionals, que en molts casos s'han fet ressò de les iniciatives

llançades des del COE. A Espanya, l'Institut de la Joventut (INJUVE) és l'organisme encarregat de posar-la en marxa i desenvolupar-la, amb la col·laboració de l'ONG Movimiento contra la Intolerancia³⁹. La presentació i el llançament de la campanya es van celebrar el 21 de març de 2014, amb motiu del Dia Internacional per a l'Eliminació de la Discriminació Racial.

Aquesta campanya també s'ha portat a l'àmbit local, a través d'una sèrie d'activitats presencials que han tingut lloc en diverses ciutats europees. A més, la ciutat d'Estrasburg ha col·laborat activament amb el COE, que té allí la seva seu, per engegar un pla d'acció local i un grup de suport de la campanya. Entre altres esdeveniments, aquest pla es va materialitzar en l'organització d'activitats com l'esdeveniment de dansa *No Hate Sounds*, que va tenir lloc el 22 de maig de 2016 a Estrasburg⁴⁰. No obstant això, fins al moment no hi ha hagut un moviment específic de ciutats o xarxa de ciutats que recolzés explícitament la campanya.

b) Una tipologia de campanyes

A més de la campanya *No Hate* del COE, al llarg del territori europeu s'han dut a terme altres campanyes d'abast més restringit contra el discurs de l'odi a Internet. Pel que fa al contingut i orientació específica d'aquestes campanyes, aquestes poden categoritzar-se, com ja s'ha assenyalat a l'inici d'aquesta secció, en tres tipus: de sensibilització, afirmatives o restrictives (Titley, Keen, i Földi 2014: 63).

En primer lloc, les **campanyes de sensibilització** pretenen conscienciar al públic sobre el discurs de l'odi i les seves conseqüències. En l'àmbit europeu, un exemple és la campanya *betterinternetforkids.eu*, engegada per INSAFE (*European Network of Awareness Centres*) i INHOPE (*Internet Association of Internet Hotlines*), dues xarxes actives en l'àmbit de la seguretat a Internet i la persecució de les modalitats de ciberdelinqüència que afecten a la infància. Encara que més centrada en l'àmbit de la seguretat a la Xarxa, aquesta campanya finançada per la Comissió Europea també va generar continguts vinculats al discurs de l'odi, sobretot a través de la celebració del "Dia per a la Seguretat a Internet" (*Safer Internet Day*)⁴¹. En la mateixa línia però en l'àmbit autonòmic, destaca la campanya "Internet sense riscos", engegada per l'Àrea de Joventut del Cabildo Insular de Tenerife i que té com a objectiu sensibilitzar contra la propagació del ciberodi i el ciberassetjament masculista⁴².

En segon lloc, les **campanyes afirmatives** tenen com a objectiu presentar les minories de manera positiva al públic general per prevenir comportaments discriminatoris. Aquestes campanyes solen tenir en compte també l'espai *offline*, i s'han engegat per canviar la percepció sobre diversos col·lectius que solen ser objecte de discurs de l'odi, com per exemple les persones LGTB, gitanes o musulmanes. En primer lloc, pel que fa al col·lectiu LGTB destaca la campanya d'àmbit global "Tots fora" (*All out*), que pretén reivindicar els drets d'aquest col·lectiu a través de la realització de micro campanyes davant esdeveniments com les Olimpíades de Sochi o accions en defensa d'activistes LGTB com la campanya #FreeTheFive, que exigia la llibertat per a cinc activistes LGTB xineses⁴³. En segon lloc, pel que fa a les persones gitanes, un altre exemple és la campanya internacional "Gitanos típics?" (*Typical Roma?*),

39. Més informació sobre la campanya a Espanya:

<http://www.injuve.es/convivencia/noticia/campana-contra-la-intolerancia-en-internet>.

40. Més informació a: <http://www.nohatespeechmovement.org/nohatesounds>

41. Més informació a: <https://www.saferinternetday.org/web/sid/resources/gallery>

42. Més informació a: <http://www.internetsinriesgos.com/wp-content/uploads/2013/11/folleto-ciberativista.pdf>

43. Més informació a: <https://allout.org/es/campanas-destacadas/>

engegada el 2010 a Macedònia, Albània i Bulgària, entre altres països, i que tenia com a objectiu desmuntar els prejudicis sobre a les persones d'aquest col·lectiu⁴⁴. En àmbit espanyol, ressalta també la campanya #YoNoSoyTrapacero, que protestava contra la imatge negativa dels gitanos que difon la Real Acadèmia Espanyola (RAE)⁴⁵. En tercer lloc, pel que fa a les persones musulmanes, destaca la campanya britànica “L'Islam és Paz” (*Islam is Peace*), centrada a proporcionar informació sobre l'Islam i les persones musulmanes, que desmunti els abundants prejudicis sobre aquesta religió i els seus practicants⁴⁶.

Una variant interessant de les campanyes afirmatives són les basades en contranarratives. Les contranarratives són missatges que ofereixen una alternativa positiva a la propaganda extremista o que pretenen deconstruir o deslegitimar les narratives extremistes (Silverman et al. 2016: 15). Desenvolupar contranarratives contra el discurs de l'odi a Internet és una de les prioritats de la campanya *No Hate* pel període 2016-2017. Així mateix, les campanyes basades en contranarratives contra l'odi s'estan potenciant des de la UE, a través de diversos instruments de finançament⁴⁷. En efecte, l'engegada d'aquest tipus de campanyes i iniciatives s'inclou com un dels objectius del Codi de Conducta signat entre la Comissió Europea i Facebook, Twitter, Youtube i Microsoft. En relació al potencial d'aquesta eina, el britànic *Institute for Strategic Dialogue* ha elaborat diversos estudis i materials de sensibilització sobre les contranarratives. Un d'ells és l'informe “L'impacte de les contranarratives” (“The Impact of Counter-narratives”), que afirmava després d'un estudi pilot d'un any de durada, que “l'exposició a punts de vista alternatius pot promoure el pensament crític o plantar una ‘llavor del dubte’ que més endavant maduri en un canvi d'alternatives o comportaments” (Silverman et al., 2016: 44). D'aquest institut van sorgir diverses iniciatives finançades per Facebook, com la “Iniciativa per al valor cívic en línia” (*Online Civil Courage Initiative*), en la qual s'anima a compartir històries personals amb l'objectiu de combatre l'extremisme i el discurs de l'odi a Internet⁴⁸, a més d'un Manual i una caixa d'eines sobre com elaborar contranarratives⁴⁹.

En tercer lloc, les **campanyes restrictives** tenen com a objectiu recollir informació sobre pàgines o accions *online* basades en contingut intolerant i prendre mesures per restringir aquesta activitat. En aquest àmbit destaca la fundació International Network Against Cyberhate (INACH), que està formada per organitzacions nacionals que recullen queixes i denúncies relatives a la discriminació a Internet, com per exemple el Movimiento contra la Intolerancia a Espanya. Especialment actius en l'àmbit de l'antisemitisme, aquesta xarxa va coorganitzar a l'abril de 2016 la primera “Conferència Internacional sobre l'antisemitisme *online*”, celebrada a Jerusalem⁵⁰. Així mateix, el 2016 va publicar l'informe “Retornar-los al mar” (*Kick them back into the sea*) sobre el discurs de l'odi contra els refugiats⁵¹. Les organitzacions nacionals membres d'aquesta entitat també han engegat campanyes que busquen augmentar les queixes i denúncies en aquest àmbit. Un exemple recent és la campanya internacional “*Get the Trolls Out*”, engegada per diverses

44. Diversos vídeos de la companya disponibles a: <https://www.youtube.com/user/typicalroma>

45. Aquesta i altres campanyes de sensibilització

46. Més informació a: <http://www.islamispeace.org.uk/>

47. En el col·loqui anual sobre drets fonamentals organitzat per la Comissió Europea el 2015, “Tolerance and respect: preventing and combating Antisemitic and anti-Muslim hatred in Europe”, es destacava la importància de promoure contranarratives provinents de la societat civil. Més informació sobre el Col·loqui a http://ec.europa.eu/justice/events/colloquium-fundamental-rights-2015/index_en.htm.

48. Més informació a: <https://www.facebook.com/onlinecivildcourage>

49. Més informació a: <http://www.counternarratives.org/about-us>

50. Més informació sobre la conferència i les seves recomanacions a http://www.inach.net/detail.html?tx_news_pi1%5Bnews%5D=10&cHash=3bf353f3d766835f68a466d2f8c5b170.

51. Informe disponible a: http://inach.zone35.net/fileadmin/user_upload/Refugee_Report20161.pdf

organitzacions i recolzada pel COE amb l'objectiu de restringir el discurs de l'odi antisemita a Internet⁵². En el marc d'aquesta campanya es van generar interessants recursos, com la guia "Com combatre el discurs de l'odi a Twitter" ("How to counter hate speech on Twitter")⁵³, així com iniciatives que pretenen estigmatitzar els abusadores com el "Trol del mes".

4.3.2. Les estratègies contra el discurs de l'odi a Internet basades en l'educació i la formació

Un altre grup d'estratègies per combatre l'odi a Internet són les basades en l'educació i la formació. Gran part d'aquestes estratègies entren dins de l'educació en drets humans, que inclou l'educació per a la diversitat i la interculturalitat, i estan orientades principalment als joves.

Fins al moment, s'han engegat diverses iniciatives educatives en els nivells europeu, nacional i local. D'una banda, en el marc de la campanya *No Hate* del COE destaca l'elaboració de la publicació "Bookmarks", un manual per combatre el discurs d'odi *online* a través de l'educació en drets humans. Aquest manual proposa una sèrie d'activitats educatives per contrarestar aquest tipus de discurs (Keen i Georgescu 2014: 21-133). D'altra banda, a més del nivell europeu, s'han engegat altres iniciatives d'aquest tipus en els nivells nacional i local, que en molts casos segueixen les recomanacions del manual Bookmarks. Sense ànim d'exhaustivitat, aquestes iniciatives poden classificar-se en dos grans grups: d'una banda, les que estan orientades a educar i formar a la població en general, per al desenvolupament del que s'ha conceptualitzat com a "ciutadania digital"; i, d'altra banda, les iniciatives destinades a dotar a persones ja conscienciades i activistes amb les aptituds necessàries per actuar més eficaçment en el mitjà *online*.

a) El pensament crític i la reflexió com a eines clau contra el discurs de l'odi a Internet

Des de la irrupció primer d'Internet i després de les xarxes socials al món de la comunicació, les persones han passat de ser només receptors de missatges a ser també emissors. Per tant, s'ha posat de manifest que les estratègies educatives han d'adaptar-se a aquesta nova realitat, promovent la recepció crítica dels missatges, però també l'empoderament en la creació de continguts (Hoechsmann i Poyntz 2012). Per aconseguir aquesta posada al dia, s'ha proposat la noció de "ciutadania digital", que pretén adaptar el concepte d'educació per a la ciutadania, incorporant a aquest el coneixement i aptituds necessàries per interaccionar en el mitjà digital (Gagliardone et al. 2015: 46). Aquestes aptituds fan referència a la necessitat tant d'un pensament crític que permeti filtrar els continguts que es reben, com d'una reflexió autocrítica abans de compartir continguts propis o aliens.

Algunes iniciatives contra el discurs de l'odi s'han centrat a treballar aquestes qüestions, principalment amb joves, ja que encara que tots els sectors de la societat generen, difonen i reben continguts amb discurs de l'odi, aquests són especialment vulnerables pel seu ús intensiu de les xarxes. En efecte, en una presentació del director de recerca sobre xarxes socials del *think tank* Demos el 2012 afirmava que un element clau en la lluita contra el discurs de l'odi és dotar d'eines als joves per desenvolupar un pensament crític adaptat a l'era digital⁵⁴. En aquest sentit, qüestio-

52. Més informació a <http://www.getthetrollsout.org/>

53. Disponible a: <http://stoppinghate.getthetrollsout.org/>

54. Carl Miller (2012), "Research perspective", discurs pronunciat a la conferència "Tackling hate speech. Living together online" el 26 i 27 de novembre de 2012 a Budapest. Disponible a http://hub.coe.int/c/document_library/get_file?uuid=8544adcf-f707-4bc1-8484-c6c55311cd1a&groupId=10227.

nar la informació que apareix a Internet i aprendre a buscar evidències que recolzin les opinions expressades és central per protegir als més joves contra el “populisme digital” o formes diverses de discurs de l’odi.

En línia amb el concepte de ciutadania digital, reforçar el pensament crític i l’alfabetització mediàtica s’estan promovent des de diversos fòrums, en els àmbits global i europeu. D’una banda, el programa d’educació de l’Organització de les Nacions Unides per a l’Educació, la Ciència i la Cultura (UNESCO) té com una de les seves àrees estratègiques de treball l’educació per a la ciutadania mundial, que pretén, entre altres coses, fomentar un “sentit de pertinença a una humanitat comuna, compartint valors i responsabilitats, empatia, solidaritat i respecte de les diferències i la diversitat” (Gagliardone et al. 2015). Entre els objectius específics d’aconseguir una ciutadania mundial, s’inclou el desenvolupament de capacitats i habilitats tècniques necessàries per usar les tecnologies digitals, així com el coneixement i habilitats per buscar, analitzar, avaluar i interpretar textos procedents de mitjans de comunicació, crear missatges en els mitjans i reconèixer la seva influència social i política (Hoechsmann i Poyntz 2012).

D’altra banda, els ministres d’educació de la UE van adoptar en una reunió informal al març de 2015 la denominada “Declaració de París sobre la promoció de la ciutadania i els valors comuns de llibertat, tolerància i no discriminació a través de l’educació”. En aquesta declaració s’identificaven una sèrie d’accions en l’àmbit educatiu per promoure la llibertat d’expressió, la inclusió social i el diàleg intercultural⁵⁵. A més, es marcava com a objectiu “reforçar el pensament crític i l’alfabetització mediàtica i informacional, en particular en l’ús d’Internet i les xarxes socials, de cara a desenvolupar resistències a tots els tipus de discriminació i adoctrinament”⁵⁶.

Pel que fa a les iniciatives específiques basades en millorar aquestes competències per prevenir i actuar contra el discurs de l’odi a Internet, es poden esmentar diversos exemples. En primer lloc, cal destacar el projecte “Prevenió del discurs de l’odi a Internet” (*Prevention of online hate speech*) del Centre Europeu Wergeland, situat a Noruega i especialitzat a desenvolupar programes d’educació per a la ciutadania democràtica. En el marc d’aquest projecte, es van dur a terme tallers i formacions que donaven als participants, educadors, estudiants adolescents i pares, eines per prevenir el discurs de l’odi i l’extremisme. Aquestes formacions es van basar en el manual Bookmarks del COE abans esmentat⁵⁷.

Una altra iniciativa interessant és el “Pack educatiu contra la discriminació digital”, elaborat en el marc del projecte ICUD⁵⁸. Està destinat a públics diversos (nens, joves, pares, educadors, professors, activistes, membres d’ONG’s, responsables polítics, acadèmics, entre altres), inclou eines pedagògiques resultat dels tallers realitzats amb joves i està disponible en 6 idiomes, entre ells castellà i català.

També val la pena esmentar el projecte pedagògic realitzat en el marc de la campanya *No Hate* per l’associació cultural belga Loupiote “NO HATE - Ce qui vous regarde”⁵⁹. Encara que més

55. https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/images/1/14/Leaflet_Paris_Declaration.pdf

56. *Ibid*, p. 3

57. <http://eng.theewc.org/Content/What-we-do/Completed-projects/Prevention-of-online-hate-speech>

58. Més informació sobre el projecte i accés al pack a l’enllaç: <http://digitaldiscrimination.eu/pack/>

59. http://www.culture-enseignement.cfwb.be/index.php?eID=tx_nawsecuredl&u=0&file=fileadmin/sites/cult_ens/upload/cult_ens_super_editor/cult_ens_editor/documents/News/Loupiote_NoHate.pdf&hash=ed1276c37b8dbfaa6864a9ee565d15fbc7a14b26

centrat en el tema de l'assetjament escolar a través de les xarxes, en el marc d'aquest projecte es van realitzar una sèrie de tallers i formacions amb joves i educadors de la regió de Brussel·les, que incloïen el desenvolupament d'eines pedagògiques destinades a combatre el ciberodi i el ciberassetjament entre els joves⁶⁰. L'eina principal resultat del projecte és un curtmetratge desenvolupat juntament amb un grup d'adolescents que mostra les conseqüències del discurs de l'odi a les xarxes socials, a través de l'experiència d'un jove que pateix els comentaris homòfobs i l'assetjament dels seus companys d'institut. A més, aquest curtmetratge està acompanyat de diversos vídeos de curta durada sobre temes vinculats com la llibertat d'expressió, l'anonimat a Internet i l'educació en els mitjans de comunicació, així com un vídeo musical sobre la propagació de l'odi *online*⁶¹.

b) Formant a ciberactivistes: algunes estratègies específiques dirigides a la formació de col·lectius ja conscienciats

A més de les estratègies esmentades a l'apartat anterior, també s'han dut a terme iniciatives destinades específicament a empoderar a les persones ja conscienciades o victimitzades amb aptituds específiques que els permetin actuar més eficaçment *online*. En particular, en els últims anys ha cobrat importància la necessitat d'empoderar a les pròpies víctimes. De fet, empoderar i recolzar a les víctimes del discurs i de crims d'odi a Internet és una de les prioritats de finançament del programa de Drets, Igualtat i Ciutadania de la UE per al període 2014-2020⁶².

Empoderar a activistes és l'objectiu principal del projecte Gent jove combatent el discurs de l'odi a Internet, que forma part de la campanya *No Hate* i que pretén formar a joves i associacions juvenils en les competències necessàries per reconèixer i actuar contra el discurs de l'odi i involucrar-los en el llançament de la campanya. L'objectiu de les formacions és compartir experiències i estratègies exitoses, per augmentar l'impacte de les activitats de bloguers i activistes. En el marc d'aquest projecte, es van dur a terme diversos tallers i seminaris sobre diferents temes, l'últim d'ells sobre com combatre el discurs de l'odi sexista⁶³.

Seguint la mateixa filosofia, i en l'àmbit espanyol, cal destacar, d'una banda, el taller "Com actuar davant el discurs de l'odi a Internet", i el curs "Drets Humans i Discurs de l'Odi", que incloïa una unitat específica sobre estratègies per combatre el discurs de l'odi a Internet, realitzats en el marc del Projecte Online contra la Xenofòbia i la Intolerància en mitjans digitals (PROXI), engegat per l'Institut de Drets Humans de Catalunya i United Explanations, en els que s'impartia formació sobre eines pràctiques per identificar comentaris que vulneren els drets humans i intervenir eficaçment en els fòrums dels mitjans digitals⁶⁴. D'altra banda, en el marc del projecte Xarxes Contra l'Odi, es va formar un grup de cibervoluntariat format per 15 joves activistes, encarregat de vigilar i denunciar l'existència de discurs d'odi a Internet, així com d'identificar situacions de ciberassetjament contra persones LGTB⁶⁵. En l'àmbit local, destaca la taula rodona "#Prourumors 2.0" sobre eines virtuals per desmuntar estereotips realitzada en el marc de l'Estratègia Antirumors de Barcelona i celebrada el juny de 2015 en aquesta ciutat, i on es van donar diferents eines per lluitar contra la rumorologia "virtual".

60.El video està disponible en francès al següent enllaç: <https://vimeo.com/111615933>

61.<https://loupioteasbl.wordpress.com/ce-qui-vous-regarde-no-hate/>

62.<http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52016SC0158&from=EN>

63.Més informació a: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680667bcd>

64.http://www.observatorioproxi.org/images/pdfs/Programa_taller_Proxi_003_VF.pdf

65.<http://www.felgtb.com/redescontraodio/>

CONCLUSIÓ

Encara que sigui complicat quantificar de manera integral el problema, existeix una sensació generalitzada a nivell internacional que el discurs de l'odi a les xarxes socials està preocupantment estès. L'odi, la ira i l'agressivitat s'han convertit en moneda comuna a les xarxes socials, provocant dany emocional als seus destinataris i contribuint a l'estigmatització i deshumanització de certs col·lectius. Se li atribueix fins i tot la possibilitat d'actuar com a desencadenant d'episodis de violència en contextos conflictius i societats ètnicament dividides. Les institucions públiques nacionals i internacionals porten anys intentant trobar solucions, o almenys remeis pal·liatius per a un problema que sembla descontrolat per les pròpies particularitats d'Internet i les xarxes socials: la sobreabundància comunicativa, l'anonimat, la transnacionalitat, etc.

Malgrat això, encara que existeix cert consens sobre l'extensió del problema, el concepte de discurs de l'odi segueix sent discutit i és objecte de debat a nivell internacional. És un concepte complex, perquè posa en joc principis fonamentals dels sistemes democràtics com són la igualtat, la dignitat humana, i la llibertat d'expressió, la concepció dels quals no és idèntica en tots els contextos sociopolítics. Per això, el concepte de discurs de l'odi pot ser implementat amb objectius més o menys legítims.

El terme 'discurs de l'odi' s'utilitza per referir-se a les expressions que inciten a la discriminació o la violència per motius d'odi racial, xenòfob, orientació sexual, o altres formes d'intolerància, però també per referir-se, en termes més amplis, a aquelles expressions que fomenten l'hostilitat a través de prejudicis o intolerància. En els mitjans de comunicació s'ha generalitzat l'ús del concepte per referir-se a una mescla heterogènia d'expressions que inclou manifestacions tan diverses, i que mereixen enfocaments molt diferents, com l'assetjament a famosos, les amenaces a individus o col·lectius per intolerància, el discurs ofensiu contra el poder, l'apologia del terrorisme, o la blasfèmia religiosa.

En qualsevol cas, l'objecte d'aquest informe, sense ànim de comprometre's amb una o una altra definició, són les expressions que en un sentit ampli fomenten i inciten a la violència, la discriminació o l'hostilitat contra individus o col·lectius per motius racistes, xenòfobs, religiosos, de gènere, orientació sexual, discapacitat i altres formes d'intolerància.

Per sistematitzar les diferents respostes i iniciatives que es duen a terme per combatre aquest fenomen, s'han classificat en estratègies jurídiques i estratègies no jurídiques. Entre aquestes últimes, s'ha parat atenció individualitzada al rol de les empreses operadores de les xarxes socials per la seva capacitat normativa i restrictiva com a intermediaris en l'àmbit privat.

Les estratègies jurídiques, que busquen la repressió i sanció del discurs de l'odi, suposen un important dilema perquè afecten un dret essencial per a les democràcies com és la llibertat d'expressió. Els arguments a favor i en contra de prohibir certes formes de discurs de l'odi poden ser igualment convinents. Per això, encara que la legislació penal emani de l'Estat, és important tenir en compte les implicacions d'aquest dilema com a punt de partida per establir una estratègia responsable enfront del discurs de l'odi.

L'última reforma del Codi Penal el 2015 va afectar als delictes de discurs de l'odi, tractant d'adaptar la legislació espanyola a la normativa europea i de corregir les deficiències que tenia: falta de proporcionalitat, massa marge per a la interpretació, i falta d'adequació al principi d'intervenció

mínima del dret penal. Com és habitual en assumptes d'aquest tipus, no hi ha consens entre els experts sobre si els objectius s'han aconseguit.

A més dels dubtes que planteja el recurs a les estratègies jurídiques pel conflicte que suposa amb l'exercici de la llibertat d'expressió, aquesta via es preveu insuficient per dues raons. En primer lloc, perquè la gravetat de les sancions de presó la converteix en una estratègia escassament aplicable, més enllà d'alguns casos extrems. I en segon lloc, perquè genera certa incertesa per l'escàs desenvolupament jurisprudencial de la matèria, i per la possibilitat que tingui efectes contraproductes.

Es innegable que el del discurs de l'odi és un camp d'actuació controvertit per a les autoritats. Per aquest motiu, i més enllà de les estratègies jurídiques, el rol de les empreses operadores de xarxes socials com a intermediaris podria ser crucial. Així ho ha considerat la Unió Europea, que ha signat el Codi de Conducta amb Facebook, Twitter, Youtube i Microsoft per aconseguir un compromís d'implicació en la moderació i la retirada de continguts que atemptin contra les normes d'ús d'aquestes empreses i les legislacions nacionals. Malgrat això, aquest tipus d'iniciatives ha rebut importants crítiques perquè traspassen la responsabilitat sobre una qüestió sensible per a la llibertat d'expressió a companyies privades, i els procediments de revisió i moderació de les quals no estan sotmesos a un marc de responsabilitat clar i transparent.

Existeix un consens considerable entre els experts en què les estratègies “no jurídiques” són les que plantegen menys objeccions. Potser l'assignatura pendent d'aquest tipus d'estratègies, que en gran mesura estan directa o indirectament finançades per les institucions públiques, sigui l'avaluació de la seva eficàcia i rendibilitat social.

S'han desenvolupat iniciatives i projectes molt diversos. Destaquen les estratègies de monitoreig i recerca, no només com a mètodes de coneixement, sinó també com a sistemes d'alerta precoç en situacions conflictives. També s'han desenvolupat estratègies de pressió a les empreses operadores de les xarxes socials, sol·licitant la modificació de les seves polítiques d'autoregulació o exigint la retirada de continguts específics. Així mateix, s'han engegat nombrosos projectes d'intervenció social amb l'objectiu de canviar les percepcions i actituds enfront del discurs de l'odi.

Entre aquest últim tipus d'estratègies destaquen les campanyes, que inclouen activitats diverses amb l'objectiu d'atacar els prejudicis i la intolerància com a arrel del problema del discurs de l'odi, i que poden categoritzar-se en tres tipus: de sensibilització, afirmatives o restrictives (Titley, Keen i Földi 2014). Esmert especial mereix la campanya *No Hate*, promoguda pel Consell d'Europa, per la repercussió aconseguida i per ser una campanya integral, que inclou una àmplia varietat d'estratègies, entre les quals destaca el desenvolupament de contranarratives. La contraargumentació a través d'un discurs alternatiu i deslegitimador del discurs de l'odi és una de les estratègies en les que s'han dipositat majors expectatives per part de les institucions públiques i de les pròpies xarxes socials.

Finalment, i també amb l'objectiu de canviar les percepcions i actituds enfront de l'odi, tenen especial importància les estratègies educatives i formatives, ja siguin dirigides al desenvolupament del que s'ha denominat “ciutadania digital”, o dirigides a la formació de persones ja conscienciades, que puguin exercir d'activistes a l'espai *online*. Com és habitual en el camp de la intervenció social, l'educació i la formació generen grans expectatives, pel seu potencial de combinar de manera òptima la responsabilitat amb la llibertat i la participació. El repte, per no quedar-se en el terreny de les bones intencions, està en com dotar-les de contingut concret que aconsegueixi els resultats desitjats.

BIBLIOGRAFIA

- Aguilar García, Miguel Ángel, ed. 2015. *Manual Práctico Para La Investigación Y Enjuiciamiento de Delitos de Odio y Discriminación*. Barcelona: Generalitat de Catalunya Centre d'Estudis Jurídics i Formació Especialitzada.
- Article 19. 2015. *Hate Speech' Explained A Toolkit*. Article 19. Londres. Disponible a https://www.article19.org/data/files/medialibrary/38231/Hate_speech_report-ID-files--final.pdf
- Bustos Gisbert, Rafael. 2015. "Libertad de expresión y discurso negacionista", a *Libertad de Expresión y discursos del odio*, Miguel Revenga Sánchez (dir.), *Cuadernos de la Cátedra de Democracia y Derechos Humanos* nº 12. Universidad de Alcalá.
- Castells, Manuel. 2012. *Redes de Indignación y Esperanza: Los Movimientos Sociales En La Era de Internet*. Alianza Editorial.
- Gagliardone, Iginio, Alisha Patel, i Matti Pohjonen. 2014. "Mapping and Analysing Hate Speech Online: Opportunities and Challenges for Ethiopia". University of Oxford and Addis Ababa University. <http://pcmlp.socleg.ox.ac.uk/sites/pcmlp.socleg.ox.ac.uk/files/Ethiopia%20hate%20speech.pdf>
- Gagliardone, Iginio, Danit Gal, Thiago Alves, i Gabriela Martinez. 2015. *Countering Online Hate Speech*. Programme in Comparative Media Law and Policy, University of Oxford.
- Gascón Cuenca, Andrés. 2015. "La Nueva Regulación Del Discurso Del Odio" a El Ordenamiento Jurídico Español: La Modificación Del Artículo 510 CP. *Cuadernos Electrónicos de Filosofía Del Derecho*, no. 32.
- Heinze, Eric. 2014. Nineteen arguments for hate speech bans- and against them. Disponible a: <http://freespeechdebate.com/en/discuss/nineteen-arguments-for-hate-speech-bans-and-against-them/> (Consultat el 23/11/2016)
- Hoehsmann, Michael., i Stuart R. Poyntz. 2012. *Media Literacies : A Critical Introduction*. Wiley-Blackwell.
- Jubany, Olga, i Malin Roiha. 2016. "Backgrounds, Experiences and Responses to Online Hate Speech: A Comparative Cross-Country Analysis"
- Jubany, Olga, Malin Roiha i Arlette Martínez. 2016. "Online Hate Speech in Spain. Fieldwork Report. PRISM. Preventing, Redressing & Inhibiting Hate Speech in New Media". Universitat de Barcelona. Fundamental Rights and Citizenship Programme of the European Union
- Kahn, Robert. 2015. Rethinking the Context of Hate Speech, a *First Amendment Law Review*; U of St. Thomas (Minnesota) *Legal Studies Research Paper* No. 15-20. https://works.bepress.com/robert_kahn/8/

Keen, Ellie, i Mara Georgescu. 2016. *Bookmarks - A Manual For Combating Hate Speech Online Through Human Rights Education*. Estrasburg: Council of Europe. <http://site.ebrary.com/lib/uacj/docDetail.action?docID=10961376>.

Legault, Lisa, Jennifer N. Gutsell, i Michael Inzlicht. 2011. "Ironic Effects of Antiprejudice Messages: How Motivational Interventions Can Reduce (but Also Increase) Prejudice". *Psychological Science*, 22, 1472-1477. DOI: 10.1177/0956797611427918

Ministeri d'Interior d'Espanya.

– 2015. "Protocolo de actuación de las fuerzas y cuerpos de seguridad para los delitos de odio y conductas que vulneran las normas legales sobre discriminación". Disponible a: <http://gestionpolicialdiversidad.org/PDFdocumentos/PROTOCOLO%20ODIO.pdf>

– 2016. "Informe 2015 Sobre Incidentes Relacionados Con Los Delitos de Odio En España". Disponible a: <http://datos.gob.es/catalogo/informe-2015-sobre-incidentes-relacionados-con-delitos-de-odio-espana>

Moretón Toquero i María Aranzazu. 2012. "El «ciberodio», La Nueva Cara Del Mensaje de Odio: Entre La Cibercriminalidad Y La Libertad de Expresión." *Revista Jurídica de Castilla y León*, no. 27.

Morozov, Evgeny. 2012. *El Desengaño de Internet: Los Mitos de La Libertad En La Red. Imago Mundi*. Barcelona: Ediciones Destino.

Phillipson, Gavin. 2015. "Hate Speech Laws: What they should and shouldn't try to do", *Revue générale du droit* (www.revuegeneraldudroit.eu), Etudes et reflexions 2015, no. 13.

Revenga Sánchez, Miguel. 2015. "Los discursos del odio y la democracia adjetivada: tolerante, intransigente, ¿militante?", a *Libertad de Expresión y discursos del odio (op.cit)*.

Rey Martínez, Fernando. 2015. "Discurso del odio y racismo líquido", a *Libertad de Expresión y discursos del odio (op. cit)*.

Rodríguez Izquierdo, Myriam. 2015 "El discurso del odio a través de Internet", a *Libertad de Expresión y discursos del odio (op. cit)*.

Ruiz, Carlos, Pere Masip, Josep Lluís Micó, Javier Díaz-Noci, i David Domingo 2010 "Conversación 2.0 y democracia. Análisis de los comentarios de los lectores en la prensa digital catalana". *Comunicación y Sociedad*. Vol. XXIII, No. 2, 2010, pp.7-39.

Silverman, Tanya, Christopher J Stewart, Jonathan Birdwell, i Zahed Amanullah. 2016. "The Impact of Counter-Narratives. Insights from a Year-Long Cross-Platform Pilot Study of Counter-Narrative Curation, Targeting, Evaluation and Impact."

Teruel Lozano, Germán M. 2015. "La libertad de expresión frente a los delitos de negacionismo y de provocación al odio y a la violencia : Sombras Sin Luces En La Reforma Del Código Penal." *Indret: Revista Para El Análisis Del Derecho*, no. 4.

Titley, Gavan, Ellie Keen, i László Földi. 2014. "Starting Points for Combating Hate Speech online. Three Studies about Online Hate Speech and Ways to Address It."

Van Spanje, J. i C. de Vreese. 2015. "The Good, the Bad and the Voter: The Impact of Hate Speech Prosecution of a Politician on Electoral Support for His Party." *Party Politics* 21 (1). SAGE Publications: 115–30. doi:10.1177/1354068812472553.

Vázquez Alonso, Víctor J. 2015. "Libertad de expresión y religión en la cultura liberal: de la moralidad cristiana al miedo postsecular", a *Libertad de Expresión y discursos del odio (op. cit.)*.

Vives Antón, Tomás. 2015. "Sobre la apología del terrorismo como 'discurso' del odio", a *Libertad de Expresión y discursos del odio (op. cit.)*.