

# PHÁT NGÔN THÙ GHÉT TRÊN MẠNG XÃ HỘI – NHỮNG VẤN ĐỀ PHÁP LÝ ĐẶT RA

Nguyễn Tiến Đức\*  
Trần Thị Thu Thủy\*\*

**Tóm tắt:** Sự phát triển của Internet và mạng xã hội không chỉ đem lại ích lợi, mà cũng tiềm tàng những mối đe dọa mới, trong đó có hiện tượng kích động thù ghét trên không gian ảo, hay còn gọi là thù ghét trực tuyến. Bài viết trình bày khái quát hai vấn đề pháp lý còn đang tranh luận: (i) Nhà nước có nên can thiệp để hạn chế phát ngôn thù ghét nói chung và thù ghét trực tuyến nói riêng; và (ii) Nếu có, khung pháp lý điều chỉnh vấn đề trên mạng xã hội nên như thế nào. Từ đó, bài viết đưa ra một số mô hình điều chỉnh vấn đề này và gợi mở cho những nghiên cứu tiếp theo.

**Abstract:** The emergence of Internet and social network sites not only bring about benefits, but also novel potential threats, including inciting hate speech in cyber space, also known as cyber hate. The article generalizes two controversial legal issues: (i) Should States intervene to tackle hate speech and cyber hate in particular; and (ii) If any, what should be a legal framework to regulate this phenomenon on social networks. Thereby, it theorizes models to regulate this phenomenon and draws implications for further research.

## 1. Mở đầu

Sự xuất hiện của Internet và các nền tảng mạng xã hội (*Social Networking Sites*, sau đây gọi tắt là MXH) đã tạo nên một cuộc cách mạng kết nối giữa con người với con người<sup>1</sup>. Nhờ tính nhanh chóng, tiện lợi, chi phí thấp của MXH, mọi thành viên trên đó đều có thể đăng tải thông tin gần như ngay lập tức; đồng thời, bất kỳ ai quan tâm cũng có thể nhanh chóng tiếp cận và chia sẻ.

Theo thống kê gần đây, có khoảng 2,45 tỷ người dùng Facebook<sup>2</sup>, 2 tỷ người dùng Youtube<sup>3</sup> và 330 triệu người dùng Twitter<sup>4</sup>. Những con số trên chỉ có thể đại diện phần nào tầm ảnh hưởng ngày càng lớn của MXH đối với đời sống thực của con người.

Dù vậy, dù một số lượng lớn các đối tượng đã lợi dụng Internet và MXH để đưa ra những phát ngôn kích động thù ghét (*Hate Speech*) nhắm tới những cá nhân khác, đặc biệt là những nhóm người yếu thế trong xã hội. Nhờ sự “trợ giúp” của MXH, những phát ngôn thù ghét đang lan rộng ở một quy mô và tần suất chưa từng thấy trên toàn cầu. Hiện tượng này còn được biết đến

\* NCS, Viện Nhà nước và Pháp luật.

\*\* ThS., Khoa Pháp luật quốc tế, Trường Đại học Luật Hà Nội.

Bài viết trong khuôn khổ đề tài cơ sở: “Quyền tự do ngôn luận trong quy định pháp luật quốc tế và pháp luật các quốc gia – Bài học kinh nghiệm cho Việt Nam” do TS. Mạc Thị Hoài Thương (Khoa Pháp luật quốc tế, Trường Đại học Luật Hà Nội) làm chủ nhiệm.

<sup>1</sup> Richard Stepe, *The Freedom of Speech on Social Networking Services – Do We Need Protection Against Our Expression?*, *Jura Falcones Jg.* 50(3) 2014-2015, <https://www.law.kuleuven.be/apps/jura/public/art/50n3/steppe.pdf>, truy cập ngày 29/7/2018, tr. 560-61.

<sup>2</sup> Facebook: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-world-wide/>, truy cập ngày 10/12/2019.

<sup>3</sup> Youtube: <https://www.omnicoreagency.com/youtube-statistics/>, truy cập ngày 10/12/2019.

<sup>4</sup> Twitter: <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>, truy cập ngày 12/12/2019.

với tên gọi: Thù ghét trực tuyến (*Cyber Hate*, sau đây gọi tắt là TGTT).

Lâu nay, tự do ngôn luận luôn được coi là “điều kiện không thể thiếu đối với sự phát triển toàn diện của con người”, của xã hội tự do và dân chủ<sup>5</sup>. Tuy nhiên, pháp luật quốc tế và quốc gia đều khẳng định việc thụ hưởng quyền này luôn đi kèm với những giới hạn nhất định, chẳng hạn ngôn luận có chứa nội dung khiêu dâm trẻ em, tục tĩu bị cấm tại tất cả các quốc gia<sup>6</sup>. Dù vậy, tranh cãi nảy sinh khi nhắc tới phát ngôn thù ghét ở cả thể giới thực hay ảo.

Thuật ngữ “phát ngôn thù ghét” đang ngày càng nhận được nhiều sự quan tâm, đặc biệt ở góc độ nghiên cứu. Tuy nhiên, cho tới nay vẫn chưa có cách định nghĩa cụ thể để kết luận một phát ngôn như thế nào là có tính chất “thù ghét”<sup>7</sup>. Một cách khái quát, thuật ngữ này thường được sử dụng để chỉ những phát ngôn rất tiêu cực, nhằm vào một số đặc điểm bản thân của “nạn nhân”, với mục đích kêu gọi, kích động lòng thù ghét. Cụ thể hơn, phát ngôn thù ghét là những phát ngôn tấn công, sỉ nhục một cá nhân hoặc một nhóm cá nhân cụ thể, với mục đích gieo rắc sự căm ghét hay kêu gọi bạo lực đối với những đối tượng này trên cơ sở xuất thân, tôn giáo, chủng tộc, giới tính,

quan điểm chính trị... của họ<sup>8</sup>. Cần lưu ý thêm rằng, “phát ngôn” không chỉ đơn thuần giới hạn ở lời nói, mà còn có thể mở rộng thông qua nhiều hành vi khác nhau.

Xuất phát từ sự khác biệt trong truyền thống và văn hóa pháp lý, các học giả trên thế giới vẫn đang thể hiện quan điểm trái chiều liên quan tới hai vấn đề: (i) Nhà nước có nên can thiệp để hạn chế phát ngôn thù ghét nói chung và TGTT nói riêng; và (ii) nếu có, khung pháp lý điều chỉnh vấn đề nên như thế nào? Câu hỏi thứ nhất liên quan tới mối quan hệ pháp lý giữa quyền tự do ngôn luận và phát ngôn thù ghét; trong khi đó, câu hỏi thứ hai đòi hỏi cần phải nhận diện rõ các yếu tố đặc trưng cũng như chủ thể liên quan tới TGTT. Vấn đề này đòi hỏi nhà lập pháp cần có cái nhìn bao quát và đầy đủ trước khi đưa ra bất kì một quyết định pháp lý quan trọng nào.

## 2. Phát ngôn thù ghét và quyền tự do ngôn luận

Thực tiễn cho thấy sự khác biệt sâu sắc trong tư duy pháp lý về nhu cầu và cách thức điều chỉnh phát ngôn thù ghét cũng như TGTT. Vụ *LICRA kiện Yahoo!* là một vụ kiện tiêu biểu cho sự khác biệt đó. Năm 2000, hai tổ chức xã hội ở Pháp đã đâm đơn kiện lên tòa án Pháp về việc Yahoo! đăng quảng cáo bán các hàng hóa cổ xúy tư tưởng phát xít, vi phạm vào Điều R645-1 Bộ luật Hình sự Pháp nghiêm cấm hành vi mặc hay trưng bày công cộng các trang phục, kí hiệu hay biểu tượng của tổ chức hoặc cá nhân phạm tội chống lại loài người<sup>9</sup>. Mặc dù nội

<sup>5</sup> Human Rights Committee, General Comment No. 34 on Article 19 Freedom of speech and expression, CCPR/C/GC/34, 12/9/2011, đoạn 1.

<sup>6</sup> Xem thêm Nguyễn Tiến Đức (2019), *Giới hạn quyền con người theo Công ước Nhân quyền châu Âu và gọi mở cho Việt Nam*, Tạp chí Nhà nước và Pháp luật số 4, tr. 60-69.

<sup>7</sup> Katherine Gelber, Luke McManama (2016), *Evidencing the Harms of Hate Speech, Social Identities* 22(3), tr. 324-325; UN Special Rapporteur (2019), *Report on Online Hate Speech*, UN Doc A/74/486, tr. 4. [https://www.ohchr.org/Documents/Issues/Opinion/A\\_74\\_486.pdf](https://www.ohchr.org/Documents/Issues/Opinion/A_74_486.pdf), truy cập ngày 20/12/2019.

<sup>8</sup> Lê Thị Thiên Hương, *Phát ngôn thù ghét trên mạng xã hội – Thực trạng và giải pháp của Cộng hòa Pháp*, <http://vpiis.edu.vn/1964-hate-speech-tren-mang-xa-hoi-thuc-trang-va-giai-phap-o-cong-boi-phap>, truy cập ngày 5/11/2019.

<sup>9</sup> *La Ligue Contre La Racisme et L'Antisemitisme (LICRA) and Union Des Etudiants Juifs De France (UEJF) v. Yahoo! Inc. and Yahoo! France*. Tại

dung quảng cáo này xuất phát từ lãnh thổ Hoa Kỳ, Tòa án Pháp vẫn xác định mình có thẩm quyền xét xử và phán quyết Yahoo! phải tìm cách ngăn ngừa công dân Pháp tiếp cận với những hàng hóa có tính chất như trên. Năm 2006, Yahoo! đưa vụ việc này ra Tòa án Quận Bắc California Hoa Kỳ và được phán quyết rằng, mặc dù phát ngôn thù ghét là rất đáng lên án, nhưng vẫn được bảo vệ trong phạm vi tự do ngôn luận trừ khi có bằng chứng cho thấy tồn tại yếu tố đe dọa tới tính mạng, sức khỏe của cá nhân hoặc tổ chức nhất định<sup>10</sup>.

Trường phái tư tưởng tự do cho rằng, tất cả các hình thức phát ngôn thù ghét cần được coi là hợp pháp bởi đây là tiền đề của tự do cá nhân và để hoàn thiện bản thân dù cho những phát ngôn này có tính chất thù hận mù quáng hay cố chấp<sup>11</sup>. Sự kiểm duyệt ngôn luận của chính quyền là rào cản đối với quá trình tìm kiếm sự thật khách quan, cản trở sự phát triển các ý tưởng sáng tạo và đột phá<sup>12</sup>. Thậm chí, Ronald Dworkin tranh luận rằng, việc cấm phát ngôn thù ghét còn có thể làm xói mòn nền dân chủ do nỗi sợ bị trừng phạt sẽ ngăn cản công dân lên tiếng, trình bày ý kiến trong đời sống chính trị. Theo Dworkin, thông qua các hình thức trình bày quan điểm (không chỉ qua lá phiếu) của mình, công dân nhận thức được

trách nhiệm chính trị của mình và từ đó hình thành nên ý thức tuân thủ pháp luật<sup>13</sup>.

Cho tới nay, Hoa Kỳ là quốc gia hiếm hoi vẫn còn bảo vệ phát ngôn thù ghét dù trong thế giới thực hay ảo, bằng chứng là những phát ngôn này được bảo vệ theo Tu chính án thứ nhất của Hiến pháp Hoa Kỳ về quyền tự do ngôn luận. Trong vụ *Virginia kiện Black* (2003), Tòa án Tối cao Hoa Kỳ tuyên quy định hình sự hóa hành vi đốt thánh giá tại nơi công cộng của bang Virginia là vi phạm quyền tự do ngôn luận. Tòa lưu ý rằng, quy định pháp luật không thể mặc nhiên cấm hình thức biểu đạt "đốt thánh giá", mà phải làm rõ được yếu tố chủ đích đe dọa, cố ý gây thương tích đối với người khác của hành vi này<sup>14</sup>. Năm 2017, trong vụ *Matal kiện Tam*, các thẩm phán Tòa án Tối cao Hoa Kỳ đồng thuận tuyên rằng, điều khoản chống miệt thị trong Đạo luật Lanham là vi hiến. Theo đó, Đạo luật này không cho phép đăng ký nhãn hiệu có tính chất "miệt thị chủng tộc". Coi nhãn hiệu là một hình thức ngôn luận, thẩm phán Samuel Alito viết: "Phát ngôn khinh miệt dựa trên chủng tộc, sắc tộc, giới tính, tôn giáo, độ tuổi, khuyết tật hay bất kì lý do nào, đều rất khó chịu; nhưng niềm tự hào lớn nhất trong văn hóa pháp lý về tự do ngôn luận của chúng ta là bảo vệ quyền tự do thể hiện "suy nghĩ [của người khác] mà chính chúng ta ghét bỏ"<sup>15</sup>. Có thể nói, tại Hoa Kỳ, quyền tự do ngôn luận luôn được đề cao kể cả khi nó gây ra cảm giác bất an và tổn thương đối với người khác<sup>16</sup>. Các nạn nhân không còn giải pháp nào ngoài việc

<http://www.juriscorn.net/txt/jurisfr/ct/yauctions2000522.htm>, truy cập ngày 5/11/2019.

<sup>10</sup> *Yahoo! Inc. v. La Ligue Contre Le Racisme et l'Antisemitisme*, 433 F.3d 1199 (9th Cir. 2006).

<sup>11</sup> Edwin Baker, *Human Liberty and Freedom of Speech*, Oxford: Oxford University Press, 1989, tr. 10.

<sup>12</sup> John S. Mill, *On Liberty*, reprinted in John Arthur, William Shaw (eds), *Social and Political Philosophy*, Prentice Hall, 1992, tr. 283-84; Caitlin Ring, *Hate Speech in Social Media: An Exploration of the Problem and Its Proposed Solutions*, University of Colorado, 2013, tr. 22.

<sup>13</sup> Ronald Dworkin (2006), *A New Map of Censorship*, *Index on Censorship* 35, tr. 131.

<sup>14</sup> *Virginia v. Black*, 538 U.S. 343 (2003).

<sup>15</sup> *Matal v. Tam*, 582 U.S. (2017).

<sup>16</sup> Michael McConnell, *You Can't Say That*, <https://www.nytimes.com/2012/06/24/books/review/you-can-say-that-by-jeremy-waldron.html?mtref=www.google.com&accfTune=PA&VW&T>

phải học cách làm quen và sống chung với những sự thù ghét quanh mình.

Ngược lại, Waldron cho rằng, phát ngôn thù ghét phải bị hạn chế bởi “mỗi cá nhân... cần được thụ hưởng quyền tự do của mình với bảo đảm không phải đối mặt với sự thù địch, bạo lực, bị phân biệt đối xử hay cần trợ giúp bởi người khác”<sup>17</sup>. Theo Waldron, kể cả khi không nhằm kích động bạo lực, phát ngôn thù ghét vẫn tác động rất tiêu cực tới sự hòa nhập xã hội của nhóm người dễ bị tổn thương. Hiện tượng này dần dần và gián tiếp xói mòn lợi ích công cộng (*Public good*), và sau đó có thể dẫn tới bùng phát bạo lực để giải quyết căng thẳng trong xã hội<sup>18</sup>. Nghiên cứu khảo sát thực nghiệm cho thấy, ngôn từ thù ghét có thể gây ra những tổn thương trực tiếp và gián tiếp đối với không chỉ nạn nhân hứng chịu nó và cả với những người xung quanh khi tiếp nhận thông điệp<sup>19</sup>.

Phải gánh chịu tổn thất do bạo lực xuất phát từ sự thù ghét, nhiều quốc gia phương Tây đã khởi động “cuộc chiến” chống lại hiện tượng này<sup>20</sup>. Tại Pháp, đạo luật áp dụng trong phần lớn các trường hợp phát ngôn thù ghét này lại là một luật ra đời từ rất lâu, từ những năm 1881<sup>21</sup>. Luật này liên quan đến tự do báo chí và đặt khuôn khổ

cho tự do ngôn luận, đồng thời xử phạt tội vu khống, làm nhục và tội phát ngôn kì thị kích động sự phân biệt đối xử, kích động thù hận hoặc kích động bạo lực. Ngoài ra, một số phát ngôn thù ghét khác còn được quy định trong luật hình sự Pháp. Tương tự, Anh nghiêm cấm các biểu hiện thù ghét dựa trên màu da, sắc tộc, quốc tịch, xuất thân, tôn giáo, giới tính, khuyết tật của một người<sup>22</sup>. Hình phạt đối với phát ngôn thù ghét bao gồm phạt tiền, phạt tù hoặc cả hai<sup>23</sup>. Đức là một trong những quốc gia thể hiện sự mạnh tay và quyết liệt nhất trong việc trừng phạt hành vi kích động thù ghét (*Volkshetze*) thông qua các quy định liên quan trong Bộ luật Hình sự của mình (Mục 185-189)<sup>24</sup>. Năm 2009, Tòa án Quận Regensburg ở Đức đã ban hành án phạt (*Strafbefehl*) đối với Richard Williamson (nhà báo Anh) vì phát ngôn chối bỏ tội ác diệt chủng ở Đức trong Thế chiến thứ hai trên kênh truyền hình Thụy Điển. Tại cấp phúc thẩm, Tòa nhận định ông Williamson phát biểu chối bỏ sự tồn tại của phòng khí ngạt, vốn được sử dụng vào việc hành quyết người Do Thái, cấu thành hành vi chối bỏ tội ác diệt chủng của Đức Quốc Xã, và vì vậy, Tòa vẫn giữ nguyên bản án trước đó. Nguyên đơn kháng cáo lên Tòa án Hiến pháp Liên bang và Tòa án Nhân quyền châu Âu nhưng bị từ chối xem xét lại vụ việc vì

<sup>17</sup> Jeremy Waldron, *The Harm in Hate Speech*, MA: Harvard University Press, 2014, tr. 10.

<sup>18</sup> Như trên, tr. 1-17.

<sup>19</sup> Tổn thương trực tiếp bao gồm: Làm cho nạn nhân bị lệ thuộc, cảm lạnh, sợ hãi hoặc biểu hiện cảm xúc tiêu cực khác, hạn chế tự do, tổn hại lòng tự trọng và phẩm giá con người. Tổn thương gián tiếp bao gồm: Thuyết phục người khác tin vào định kiến về một hoặc một số kiểu mẫu người, bình thường hóa hiệu tượng phân biệt chủng tộc, gây ảnh hưởng dẫn tới người nghe tiếp nhận sự thù ghét từ phía người phát ngôn. Xem thêm Katharine Gelber, Luke McNamara (2016), *Evidencing the Harms of Hate Speech, Social Identities, Journal for the Study of Race, Nation and Culture* 22(3), tr. 324-41.

<sup>20</sup> Michael McCornell, *ildd*.

<sup>21</sup> Lê Thị Thiên Hương, *ildd*.

<sup>22</sup> Ví dụ: Mục 29(A)(B) Đạo luật năm 2006 (Racial and Religious Hatred Act); Mục 74 Đạo luật Trật tự công cộng năm 1986 (Public Order Act) liên quan tới hành vi thù ghét dựa trên xu hướng tính dục.

<sup>23</sup> Tại: [http://www.legislation.gov.uk/ukxi/2003/2267/pdfs/ukxi\\_20032267\\_en.pdf](http://www.legislation.gov.uk/ukxi/2003/2267/pdfs/ukxi_20032267_en.pdf), truy cập ngày 5/11/2019.

<sup>24</sup> ARTICLE 19, Germany: Responding to Hate Speech, tại: <https://www.article19.org/wp-content/uploads/2018/07/Germany-Responding-to-%E2%80%98hate-speech%E2%80%99-v3-WEB.pdf>, tr. 21-48, truy cập ngày 5/11/2019.

không đủ căn cứ<sup>25</sup>. Tòa án Nhân quyền châu Âu nhất quán quan điểm trong các phán quyết của mình rằng: "... Sự bạo động và tôn trọng đối với phẩm giá bình đẳng của tất cả con người là nền móng của một xã hội dân chủ và đa nguyên. Do đó, về nguyên tắc, xã hội dân chủ có thể sẽ cần phải quy định hoặc thậm chí ngăn ngừa mọi hình thức biểu đạt nhằm truyền bá, thúc đẩy, kích động, bào chữa cho sự thù ghét thiếu khoan dung..., nhưng cần bảo đảm mọi "hình thức", "điều kiện", "giới hạn" hay "hình phạt" được áp dụng phải tương xứng với mục tiêu hướng tới"<sup>26</sup>.

Điều 20(2) Công ước quốc tế về các quyền dân sự và chính trị năm 1966 (ICCPR) quy định: "Mọi chủ trương gây thù hận dân tộc, chủng tộc hoặc tôn giáo để kích động sự phân biệt đối xử về chủng tộc, sự thù địch, hoặc bạo lực đều phải bị pháp luật nghiêm cấm". Quy định này có mối quan hệ chặt chẽ với Điều 19 ICCPR (quyền tự do ngôn luận) bởi nó được hiểu là một trong những giới hạn quyền tự do ngôn luận. Điều này cho thấy, quốc gia có thể có nghĩa vụ quốc tế nhằm ngăn ngừa phát ngôn thù hận, kể cả khi những phát ngôn này không nhắm tới mục tiêu kích động bạo lực, gây tổn thương cho người khác. Ngoài ra, Ủy ban giám sát Công ước quốc tế về xóa bỏ mọi hình thức phân biệt chủng tộc năm 1965 (ICERD) cho rằng quyền tự do ngôn luận không phải tuyệt đối<sup>27</sup>, đặc biệt trong vấn đề

phát ngôn kích động hận thù dựa trên chủng tộc: "Việc thể hiện quan điểm, ý tưởng trong quá trình thảo luận khoa học, đối thoại chính trị hay các hoạt động tương tự, cần phải được thực hiện một cách hợp pháp, không nhằm mục tiêu kích động thù hận, nổi loạn, bạo lực hay phân biệt đối xử"<sup>28</sup>. Đối phó với việc lạm dụng ngôn luận để tấn công sắc tộc, Ủy ban khuyến nghị các quốc gia khi xây dựng khung pháp lý điều chỉnh vấn đề này cần lưu ý tới những yếu tố như nội dung và hình thức phát ngôn; bối cảnh kinh tế, xã hội và chính trị; địa vị của người phát ngôn; khả năng lan tỏa của phát ngôn; mục tiêu của phát ngôn<sup>29</sup>. Đồng thời, Ủy ban cũng nhắc nhở quốc gia thành viên về nghĩa vụ theo Điều 7 ICERD trong hoạt động tuyên truyền, giáo dục nhằm loại bỏ gốc rễ tư tưởng phân biệt chủng tộc trong xã hội<sup>30</sup>.

### 3. Phát ngôn thù ghét trên nền tảng mạng xã hội – Nhận diện sự khác biệt

MXH, trong đó có thể kể tới Facebook, Youtube, Twitter, là một vài những cái tên đã quá quen thuộc với nhiều người. Những nền tảng này còn được gọi là Web 2.0, cung cấp dịch vụ miễn phí cho các cá nhân kết nối với nhau thông qua Internet<sup>31</sup>. Những công ty sở hữu nền tảng này còn được xếp vào nhóm các công ty trung gian Internet (*Internet intermediaries*) bởi tính chất trung

<sup>25</sup> Như trên, đoạn 25.

<sup>26</sup> Như trên, đoạn 15; Xem thêm Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence, đoạn 22, tại: [https://www.ohchr.org/Documents/Issues/Opinion/SEminarRabat/Rabat\\_draft\\_outcome.pdf](https://www.ohchr.org/Documents/Issues/Opinion/SEminarRabat/Rabat_draft_outcome.pdf), truy cập ngày 21/1/2020.

<sup>27</sup> Như chú thích 27, đoạn 30-44.

<sup>31</sup> Mark Davies, Barbara Lee (2008), *The Legal Implications of Student Use of Social Networking Sites in the UK and US: Current Concerns and Lessons for the Future*, Education and the Law 20(3), tr. 266.

<sup>25</sup> *Williamson v. Germany*, ECtHR App no. 64496/17 (2017), [https://www.echr.coe.int/Documents/FS\\_Hate\\_speech\\_ENG.pdf](https://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf), truy cập ngày 5/11/2019.

<sup>26</sup> *Erbakan v. Turkey*, ECtHR App no. 59405/00 (2006), đoạn 56, <https://hudoc.echr.coe.int/app/conversion/pdf/?library=ECHR&id=003-1728198-1812055&filename=003-1728198-1812055.pdf>, truy cập ngày 21/1/2020.

<sup>27</sup> General recommendation No 35, CERD/C/GC/35 (2012), đoạn 1-26, <https://undocs.org/en/CERD/C/GC/35>, truy cập ngày 21/1/2020.

gian tạo thuận lợi cho người dùng Internet<sup>32</sup>. Các công ty Web 1.0, ví dụ như Amazon, là những công ty tự xây dựng nội dung trên nền tảng của mình (*Company-driven content*) và ít có tính tương tác với người dùng. Trái lại, trên nền tảng Web 2.0, người dùng chính là “chủ nhân” xây dựng và phát triển nội dung (*User-driven content*). Trên đó, nhiều cộng đồng ảo được thành lập, cho phép người dùng chia sẻ tin nhắn, hình ảnh, video với nhau ở tốc độ chóng mặt và hoàn toàn miễn phí.

Cũng giống như phát ngôn thù ghét trong thế giới thực, TGTT cũng phản ánh những thông điệp có tính chất tương tự, như phân biệt chủng tộc, cực đoan hoặc khủng bố...<sup>33</sup> Tuy nhiên, điểm khác biệt rõ nét nhất giữa hai loại thù ghét này là phương tiện thực hiện. Theo ghi nhận của Tổng thư ký Liên hợp quốc, hiện tượng lợi dụng Internet, trong đó có MXH, để truyền bá thông tin, cảm xúc thù ghét đã và đang trở thành một trong những thách thức nhân quyền nghiêm trọng nhất trong thế giới hiện đại ngày nay<sup>34</sup>.

Ngày càng có nhiều tài liệu nghiên cứu về hình thức và phạm vi phát ngôn thù ghét trên MXH nhằm làm rõ sự khác biệt giữa phát ngôn thù ghét trong thế giới ảo và thực<sup>35</sup>. Đồng thời, các nghiên cứu này cũng

góp phần nhận diện những thách thức xã hội và pháp lý mà quốc gia phải đối mặt khi xây dựng khung pháp lý điều chỉnh vấn đề này.

Thứ nhất, tính ẩn danh (*Anonymity*) và tính bất khả kiến (*Invisibility*). Một trong những ưu điểm của Internet và MXH đó là không bắt buộc cá nhân phải đăng ký và công khai thông tin cá nhân thực. Trước đây, nhiều người tin rằng tính ẩn danh trên các nền tảng công nghệ sẽ mở đường cho quyền tự do ngôn luận phát triển mạnh mẽ, đặc biệt tại các quốc gia chuyên chế, toàn trị, bởi cá nhân có thể thoát khỏi nỗi sợ hãi bị trừng phạt vì bất đồng chính kiến<sup>36</sup>. Trên thực tế, điều này lại có thể trở thành con dao hai lưỡi bởi nó gián tiếp tạo ra xu hướng tâm lý phớt lờ những quy phạm đạo đức xã hội, còn gọi là hiện tượng *phi cá nhân hóa* (*Deindividuation*) trong tâm lý học đám đông<sup>37</sup>. Có những bằng chứng khoa học cho thấy, cá nhân có thể đưa ra phát ngôn hoặc cư xử trên không gian mạng mà họ sẽ không làm trong đời thực<sup>38</sup>. Nguyên nhân một

*Intermediaries and Hate Speech: Fostering Digital Citizenship for Our Information Age*, Boston University Law Review 91, tr. 1435–1484, Cohen-Almagor, *Confronting the Internet's Dark Side: Moral and Social Responsibility on the Free Highway*, Cambridge: Cambridge University Press, 2015; Richard Delgado, Jean Stefancic (2014), *Hate Speech in Cyberspace*, Wake Forest Law Review 49, tr. 319–343; Barbara Perry, Patrik Olsson (2009), *Cyberhate: The Globalization of Hate*, Information and Communication Technology Law 18, tr. 185–199; Alexander Tsesis (2001), *Hate in Cyberspace: Regulating Hate Speech on the Internet*, San Diego Law Review 38, tr. 817–874.

<sup>36</sup> Gordon Graham, *The Internet: A Philosophical Examination*, Routledge: London, 1999.

<sup>37</sup> Arnold Goldstein, *The Psychology of Group Aggression*, John Wiley and Sons, 2002, tr. 32; Brian Mullen (1983), *Operationalizing the Effect of the Group on the Individual: A Self-Attention Perspective*, Experimental Social Psychology Journal 19, tr. 295–322.

<sup>38</sup> John Suler (2004), *The Online Disinhibition Effect*, Cyber-Psychology and Behavior 7, tr. 321–326.

<sup>32</sup> Xem: <https://www.article19.org/resources/internet-intermediaries-dilemma-liability-q/>, truy cập ngày 31/7/2019.

<sup>33</sup> Xem thêm Anti-Defamation League (2010), *Responding to Cyber-hate: Toolkit for Action*, <https://www.adl.org/sites/default/files/documents/assets/pdf/combating-hate/ADL-Responding-to-Cyberhate-Toolkit.pdf>, truy cập ngày 31/7/2019.

<sup>34</sup> The Secretary General, *Preliminary Representation of the Secretary General on Globalization and Its Impact on the Full Enjoyment of All Human Rights*, UN Doc A/55/342 (31/8/2000), đoạn 26–28.

<sup>35</sup> Dianielle Citron, *Hate Crimes in Cyberspace*, Harvard, MA: Harvard University Press, 2014; Dianielle Citron, Helen Norton (2011),

phần là bởi khi ân danh, họ không lo sợ bị lên án hay trừng phạt bởi xã hội và nạn nhân của những lời nói trên mạng<sup>39</sup>. Đồng thời, do không phải chứng kiến những tổn thương mà nạn nhân phải chịu đựng nên họ không nảy sinh lòng thương xót hay sự đồng cảm<sup>40</sup>.

*Thứ hai, tính liên kết (Connectivity).* Nền tảng Internet và MXH có thể bị lợi dụng không chỉ để thể hiện những quan điểm thù ghét cực đoan, mà còn là nơi những người có cùng tư tưởng đó kết nối với nhau – mà trên thực tế, họ có thể bị cản trở bởi các yếu tố không gian và thời gian<sup>41</sup>. Ở đây, thù ghét được truyền bá, trao đổi, kích động giữa những người cùng tư tưởng, chứ không phải đối với nạn nhân của nó. Những người có tư tưởng cực đoan thường sử dụng Internet và MXH để tìm cách kết nối với nhau<sup>42</sup>. Ví dụ, trong vụ *R. v. Ahmad* (2011)<sup>43</sup>, công dân có hai quốc tịch Anh và Pakistan bị kết án từ 12 năm vì tội kích động thù hận tôn giáo, đã kết nối và viết bình luận trên một trang mạng Hồi giáo cực đoan kêu gọi người Hồi giáo “hãy giơ con dao thánh chiến... và giết tất cả các nghị sĩ Anh bỏ phiếu ủng hộ chiến tranh ở Iraq”.

*Thứ ba, tính tức thì (Instantaneousness).* Trên nền tảng MXH, chỉ cần một “cú nhấp

chuột” thì việc cung cấp, truyền bá thông điệp với bất kì nội dung nào có thể được thực hiện gần như ngay lập tức. Đây là điểm khác biệt lớn nhất giữa nền tảng Internet với các phương tiện truyền thông truyền thống khác; chẳng hạn như báo, đài, điện thoại... Brown cho rằng, tổng hợp tất cả các yếu tố: Nặc danh, bất khả kiến, thuận tiện và tức thì, tạo nên sự đặc thù của TGTT, dễ dẫn đến nguy cơ con người cư xử trên không gian một cách bộc phát, thiếu suy nghĩ và bất cần<sup>44</sup>. Đây là những yếu tố làm TGTT thêm phức tạp và trầm trọng.

#### 4. Khung pháp lý điều chỉnh thù ghét trực tuyến

Về lý luận, một số mô hình có thể được rút ra dựa trên phương pháp điều chỉnh của các quốc gia liên quan tới vấn đề này. Có thể kể tới các mô hình: Tự quản lý, cùng quản lý và bắt buộc quản lý, được sắp xếp tăng dần theo mức độ can thiệp của Nhà nước nhằm điều chỉnh hành vi của các chủ thể liên quan. Trong đó, hai mô hình đầu tiên nhận được nhiều sự chú ý hơn cả.

##### 4.1. Mô hình tự quản lý (Self-Regulation)

Hoa Kỳ là quốc gia tiêu biểu nhất cho mô hình tự điều chỉnh vấn đề thù hận trực tuyến. Dưới góc độ pháp luật Hoa Kỳ, phát ngôn thù ghét không nhằm mục đích kích động bạo lực thì vẫn được pháp luật bảo vệ vì đó là tự do ngôn luận<sup>45</sup>. Vì vậy, Nhà nước sẽ không can thiệp vào phát ngôn của công dân kể cả trên nền tảng MXH; và khu vực tư sẽ tự quản lý vấn đề này. Khu vực tư có thể bao gồm các hiệp hội, hội đồng chuyên ngành (báo chí, công nghệ, MXH...), tự đặt

<sup>39</sup> Citron (2014), *ltdđ*, tr. 58; Delgado, Stefancic (2014), *ltdđ*, tr. 322.

<sup>40</sup> Citron (2014), *ltdđ*, tr. 59.

<sup>41</sup> Richard Posner, *The Speech Market and the Legacy of Schenck* trong Lee Bollinger, Geoffrey Stone (eds) *Eternally Vigilant: Free Speech in the Modern Era*, Chicago, IL: University of Chicago Press, 2002, tr. 149-151.

<sup>42</sup> Citron (2014), *ltdđ*, tr. 61-62; Perry, Olsson (2009), *ltdđ*; Cass Sunstein, *Republic.com 2.0*, Princeton, NJ: Princeton University Press, 2007, tr. 57-58.

<sup>43</sup> The Guardian, *Radical Muslim Jailed for Calling for Jihad against MPs*, <https://www.theguardian.com/world/2011/jul/29/radical-muslim-bital-ahmad-jailed>, truy cập ngày 31/7/2019.

<sup>44</sup> Alexander Brown (2018), *What Is So Special about Online (As Compared to Offline) Hate Speech?*, *Ethnicities* 0(0), tr. 1-30.

<sup>45</sup> *Yahoo! Inc. v La Ligue Contre Le Racisme et L'Antisemitisme*, US 9<sup>th</sup> Cir. 2006, 169 F. Supp. 2d, 1192.

ra bộ quy tắc đạo đức, ứng xử chung cho các thành viên, hoặc mỗi công ty công nghệ có thể xây dựng và áp dụng bộ quy tắc của riêng mình.

Nhìn chung, vì thiếu sự quản lý thống nhất, các nền tảng MXH lớn đã trở thành “mảnh đất màu mỡ” góp phần làm trầm trọng hơn tình trạng thù ghét trong xã hội. Gần đây, các công ty công nghệ lớn của Hoa Kỳ đã dần thay đổi quan điểm và trở nên quyết liệt hơn trong đấu tranh với thù hận trực tuyến nhằm duy trì một môi trường MXH an toàn và văn minh hơn. Các công ty trung gian Internet như Youtube (Google)<sup>46</sup>, Facebook<sup>47</sup>, Twitter<sup>48</sup> đều đã xây dựng định nghĩa riêng của mình về “phát ngôn thù ghét” và buộc người dùng phải chấp nhận quy định khi sử dụng các ứng dụng của mình. Hình phạt áp dụng chủ yếu là khóa toàn bộ hoặc một vài tính năng trong một khoảng thời gian nhất định, mạnh tay hơn là từ chối cho phép người dùng tiếp tục tham gia nền tảng của những công ty công nghệ này<sup>49</sup>.

Mô hình tự quản lý này thường bị chỉ trích là “vừa đá bóng, vừa thổi còi” bởi các hiệp hội hay hội đồng chuyên ngành đặt ra quy tắc lại nhận tài trợ hoạt động từ chính các công ty công nghệ<sup>50</sup>. Ngoài ra, các công

ty công nghệ, truyền thông lớn đều cử đại diện của mình tham gia các tổ chức này. Điều này dễ dẫn tới tình trạng xung đột lợi ích khi nhóm này miễn cưỡng áp dụng chế tài đối với đồng nghiệp hay chính công ty của mình<sup>51</sup>. Vì vậy, khá là khó khăn để những tổ chức này có thể lấy được lòng tin từ phía công chúng<sup>52</sup>.

Năm 2018, David Kaye, báo cáo viên đặc biệt của Liên hợp quốc về thúc đẩy và bảo vệ quyền tự do ý kiến và biểu đạt, đã bày tỏ sự thất vọng về thực trạng quốc gia và các công ty trung gian Internet đều “thất bại” trong việc hạn chế TGTT<sup>53</sup>. Bất chấp sự thay đổi trong chính sách của các nền tảng MXH, TGTT vẫn phát triển “theo cấp số nhân” thông qua các công nghệ số. Báo cáo viên cũng nhận định thêm “các công ty [trung gian Internet] vẫn đang rất miễn cưỡng cam kết thực thi các chính sách và hành động phù hợp với những tiêu chuẩn của luật quyền con người”<sup>54</sup> do nhiều yếu tố, trong đó có chi phí và lợi nhuận của các công ty này.

#### 4.2. Mô hình cùng quản lý (Co-Regulation)

Mô hình này đôi khi còn được gọi là “quản lý việc tự quản lý” (*Regulated Self-Regulation*) bởi nó vẫn giữ lại một số đặc điểm của mô hình tự quản lý. Đặc điểm chung các quốc gia áp dụng mô hình này là phát ngôn thù ghét không được bảo vệ trong phạm vi quyền tự do ngôn luận. Trái lại,

<sup>46</sup> Xem: [https://support.google.com/youtube/answer/2801939?hl=en&ref\\_topic=2803176](https://support.google.com/youtube/answer/2801939?hl=en&ref_topic=2803176), truy cập ngày 31/7/2019.

<sup>47</sup> Xem: <https://www.facebook.com/communitystandards/>, truy cập ngày 31/7/2019.

<sup>48</sup> Xem: <https://www.theverge.com/2017/12/18/16789606/twitter-new-safety-policies-bate-groups>, truy cập ngày 31/7/2019.

<sup>49</sup> Brookings (2019), *How Should Social Media Platforms Combat Misinformation and Hate Speech*, tại: <https://www.brookings.edu/blog/techtank/2019/04/09/how-should-social-media-platforms-combat-misinformation-and-hate-speech/>, truy cập ngày 21/10/2019.

<sup>50</sup> ARTICLE 19 (2018), *Self-Regulation and Hate Speech on Social Media Platforms*, tại: [https://www.article19.org/wp-content/uploads/2018/03/Self-regulation-and-%E2%80%98hate-speech%20-%E2%80%99-on-social-media-platforms\\_March2018.pdf](https://www.article19.org/wp-content/uploads/2018/03/Self-regulation-and-%E2%80%98hate-speech%20-%E2%80%99-on-social-media-platforms_March2018.pdf), truy cập ngày 21/10/2019.

<sup>51</sup> Như trên.

<sup>52</sup> Ian Bartle, Peter Vass (2005), *Self-Regulation and the Regulatory State - A Survey of Policy and Practice*, Research Report No. 17, Centre for the Study of Regulated Industries, University of Bath, tr. 23.

<sup>53</sup> Xem: <https://news.un.org/en/story/2019/10/1049671>, truy cập ngày 21/10/2019.

<sup>54</sup> Như trên; xem thêm James Banks, *Regulating Hate Speech Online*, *International Reviews of Law, Computers and Technology* 24(3), tr. 233-239.



những phát ngôn này còn có thể bị xử lý hình sự hoặc các biện pháp khác. Với mô hình này, Nhà nước xây dựng khung pháp lý để các công ty công nghệ quản lý thù hận trực tuyến trên nền tảng của mình<sup>55</sup>.

Có thể nhận thấy, quá trình ra quyết định của mô hình này nhanh và hiệu quả hơn bởi các công ty trung gian Internet có *động lực* – dưới dạng nghĩa vụ pháp lý – nhận diện và loại bỏ TGTT một cách nhanh chóng. Cơ chế tìm kiếm và loại bỏ có thể được thực hiện thông qua chức năng báo cáo vi phạm của người dùng (*Report*) hoặc thuật toán nhận diện các từ khóa nhất định...

Hậu quả thảm khốc của chế độ phát xít đã biến Đức trở thành một trong những nước khá nhạy cảm với hiện tượng thù ghét<sup>56</sup>. Vì vậy, đây là một trong số ít các quốc gia phương Tây áp dụng các biện pháp rất quyết liệt đối với các công ty trung gian Internet thông qua Đạo luật Thực thi Mạng năm 2017 (*NetzDG*). Theo đó, Luật này yêu cầu các công ty truyền thông xã hội loại bỏ nội dung từ kích động thù địch nhanh chóng và kịp thời, nếu không sẽ phải đối mặt với mức phạt lên tới 50 triệu euro.

Dù vậy, theo một báo cáo của Chính phủ công bố năm 2017, Facebook và Twitter đều không đáp ứng được yêu cầu theo luật định là loại bỏ ít nhất 70% phát ngôn TGTT trong vòng 24 giờ sau khi được người dùng thông báo<sup>57</sup>. Bên cạnh đó, *NetzDG* bị cho là được ban hành quá vội

vàng mà thiếu đi một cơ quan độc lập xem xét và đánh giá kiến nghị của công chúng về phát ngôn chứa nội dung thù ghét; gián tiếp trao quyền quyết định quá lớn vào tay các công ty công nghệ<sup>58</sup>. Điều này dẫn đến khả năng các công ty này sẽ loại bỏ những phát ngôn nhạy cảm (dù chưa chắc đã trái luật) để tránh bị phạt nặng, và vì vậy, cản trở quyền tự do ngôn luận trực tuyến. Trước đó, Facebook ngầm ám chỉ rằng, pháp luật Đức sẽ đẩy trách nhiệm xác định tính hợp pháp của một phát ngôn từ phía tòa án sang cho các công ty tư nhân<sup>59</sup>. Điều này cho thấy, nguy cơ tiềm ẩn là “*các luật sư của Facebook, Google và Microsoft có quyền định đoạt tương lai của... tự do ngôn luận lớn hơn bất kỳ vị vua, tổng thống hay thẩm phán tòa án tối cao nào*”<sup>60</sup>.

### 4.3. Mô hình bắt buộc quản lý (Statutory Regulation)

Bên cạnh hai mô hình trên, còn có mô hình bắt buộc quản lý. Đây là mô hình thường áp dụng đối với các phương tiện phát thanh truyền thống. Trong đó, không chỉ gói gọn ở khung pháp lý, Nhà nước ban hành các quy định cụ thể quy định quyền và nghĩa vụ của các chủ thể liên quan, đặc biệt là các công ty truyền thông trung gian và người dùng cuối (*End-users*). Trường hợp vi phạm xảy ra, tùy vào mức độ, cả công ty truyền thông trung gian và người dùng cuối đều có thể bị xử phạt theo luật định. Có thể thấy, đây là mô hình quản lý chặt chẽ nhất,

<sup>55</sup> ARTICLE 19 (2018), *ibid.*

<sup>56</sup> Council on Foreign Relations (2019), *Hate Speech on Social Media: Global Comparisons*, tại: <https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons>, truy cập ngày 08/11/2019.

<sup>57</sup> Katy O'Donnell et al., *Germany's New Online Hate Speech Code Pushes Big Fines and Debate*, Link: <https://www.politico.eu/article/hate-speech-germany-twitter-facebook-google-fines/>, truy cập ngày 08/11/2019.

<sup>58</sup> Emma Thomasson, *Germany Looks to Revise Social Media Law as Europe Watches*, <https://www.reuters.com/article/us-germany-hatespeech/germany-looks-to-revise-social-media-law-as-europe-watches-idUSKCN1GK1BN>, truy cập ngày 01/08/2018.

<sup>59</sup> Như trên.

<sup>60</sup> Terry Gross, Jeffrey Rosen, *Interpreting the Constitution in the Digital Era*, <https://www.npr.org/2011/11/30/142714568/interpreting-the-constitution-in-the-c>

với mức độ can thiệp của Nhà nước lớn nhất. Tuy nhiên, dấu hỏi vẫn đặt ra đối với hiệu quả thực tế của nó khi áp dụng vào cấu trúc mạng Internet và nền tảng MXH, mà tại đó, khả năng kết nối và lan truyền tin tức diễn ra cực kì nhanh chóng.

### 5. Một số gợi mở nghiên cứu tiếp theo

Xây dựng khung pháp lý ngăn ngừa thù ghét trên MXH chưa bao giờ là dễ dàng do một loạt các rào cản. Trước tiên, cách hiểu về “thù ghét” vẫn còn gây tranh cãi bởi thuật ngữ này chịu nhiều ảnh hưởng bởi các yếu tố văn hóa và lịch sử của từng quốc gia, dẫn đến những trở ngại xây dựng khuôn khổ pháp lý thống nhất.

Bên cạnh đó, ngôn từ kích động thù ghét trực tuyến phản ánh sự thù hận trong thế giới thực<sup>61</sup>. Điều này cho thấy, việc phát hiện và giám sát phát ngôn thù ghét trong thế giới trực tuyến có thể là một bước tiến trong tổng thể giải pháp ngăn chặn sớm tội phạm thù hận trong thế giới thực. Dù vậy, vẫn cần tiếp tục làm rõ đặc trưng và khác biệt giữa phát ngôn thù ghét trong thế giới thực và thế giới ảo để tìm ra giải pháp hiệu quả. Ngoài ra, Nhà nước nên can thiệp ở mức độ như thế nào? Nên ban hành chính sách nâng cao nhận thức, xử phạt hành chính hay hình sự hóa? Và ai nên bị giám sát, các công ty công nghệ hay người dùng cuối? Nhiều người, đặc biệt là các công ty công nghệ, ủng hộ mô hình tự điều chỉnh bởi không chịu ràng buộc pháp lý, giảm thiểu chi phí giám sát, đồng thời tự do ngôn luận cũng có nhiều khoảng không phát triển. Tuy nhiên, phỏng vấn hơn 1700 nhà báo ở 14 quốc gia châu Âu, một nghiên cứu chỉ ra rằng, mặc dù các nhà báo châu Âu không muốn Nhà nước can thiệp do nguy cơ lạm dụng chính trị, họ nhận thấy mô hình tự

quản lý không đủ hiệu quả để ngăn chặn TGTT<sup>62</sup>.

Ở Việt Nam, phát ngôn thù ghét đang tồn tại tràn lan trên MXH. Tính tới năm 2017, có hơn 35 triệu người dùng MXH, chiếm 37% tổng dân số Việt Nam, dành trung bình hơn 2 giờ mỗi ngày trên MXH<sup>63</sup>. Theo khảo sát của chương trình do Chính phủ tài trợ, 61,7% người dùng MXH từng chứng kiến hoặc trở thành nạn nhân của trò nói xấu, phỉ báng, bôi nhọ danh dự và 46,6% từng bị vu khống, bịa đặt thông tin. Nạn nhân gần như bất lực, cách duy nhất có thể làm là yêu cầu gỡ thông tin sai trái đó<sup>64</sup>.

Dù vậy, thực tiễn cho thấy, pháp luật Việt Nam trong lĩnh vực này còn khá lỏng lẻo, trong khi khuyến khích chung của thế giới là yêu cầu MXH tích cực hơn nữa trong việc bảo vệ người sử dụng. Pháp luật Việt Nam hiện hành không có quy định cụ thể về phát ngôn thù ghét. Một số quy định rải rác trong Luật Phòng chống HIV/AIDS năm 2006, Luật Người khuyết tật năm 2010 không đủ hiệu quả để điều chỉnh thực trạng này<sup>65</sup>. Nghị định số 72/2013/NĐ-CP của Chính phủ về quản lý, cung cấp, sử dụng dịch vụ internet và thông tin trên mạng có một số quy định liên quan, nhưng lại không quy định chi tiết nghĩa vụ hỗ trợ người dùng

<sup>62</sup> Sussane Fengler *et al.* (2015), How Effective is Media Self-Regulation? Results from a Comparative Survey of European Journalists, *European Journal of Communication* 30(3), tr. 249-266.

<sup>63</sup> Tại: <https://vrv.vn/cong-nghe/phan-ngon-gay-thu-ghet-tren-mang-xa-hoi-van-nan-toan-cau-trong-ky-nguyen-so-20170412145326493.htm>, truy cập ngày 01/8/2019.

<sup>64</sup> Xem <https://vnexpress.net/thoi-su/gan-80-nguoi-dung-mang-xa-hoi-la-nan-nhan-cua-phan-ngon-gay-thu-ghet-3569175-p2.html>, truy cập ngày 01/8/2019.

<sup>65</sup> Xem thêm Viện Nghiên cứu chính sách và phát triển truyền thông, tại: <https://ips.org.vn/an-ninh-mang/luat-an-ninh-mang-viet-nam-cap-thiet-nhung-khong-duoc-voi-vang-110>, truy cập ngày 05/03/2020.

<sup>61</sup> Mainack Mondal *et al.* (2018), Characterizing Usage of Explicit Hate Expressions in Social Media, *New Review of Hypermedia and Multimedia*, tr. 110-130.

MXH hướng tới một môi trường lành mạnh, như nghĩa vụ của nhà mạng xã hội phải nhanh chóng gỡ bỏ những thông tin nhạy cảm, thông tin mang tính kì thị, thông tin giả mạo ảnh hưởng đến cá nhân khi nhận được “yêu cầu” của nạn nhân.

Tại Việt Nam, hiện nay, các MXH đang có ảnh hưởng lớn như Facebook, Youtube... nhưng các công ty này lại cho rằng họ không hoàn toàn phải chịu trách nhiệm về những nội dung được đăng tải<sup>66</sup>. Ở khía cạnh này, sự hợp tác giữa các công ty công nghệ thông tin và các tổ chức xã hội tại châu Âu là một điển hình tốt, trong đó khả năng lọc, kiểm duyệt nội dung khách quan nhất thuộc về các tổ chức xã hội. Năm 2016, Ủy ban châu Âu đã ban hành “Bộ luật ứng xử” với cam kết của các nhà cung cấp dịch vụ MXH trong việc chống lại phát ngôn thù ghét trên mạng tại châu lục này<sup>67</sup>. Bên cạnh đó, các trang mạng xã hội đều cài đặt chức năng “thông báo vi phạm” cho phép người dùng báo cáo những nội dung sai sự thật, kích động thù hận. Do người dùng không chú ý đến các chức năng nêu trên, thực tế khi phát ngôn được tung ra, vẫn dễ việc xóa bỏ còn chậm, không nhiều và chưa đáp ứng được so với yêu cầu thực tế<sup>68</sup>. Điều này cho thấy, việc xây dựng các giải pháp cũng cần dành sự chú ý tới nâng cao nhận thức người dùng.

Trong thời gian tới, Việt Nam nên xem xét rà soát, sửa đổi, bổ sung pháp luật để ngăn ngừa phát ngôn TGTT theo hướng:

*Thứ nhất*, cần xây dựng tiêu chuẩn rõ ràng về các khái niệm phát ngôn thù ghét, kích động, phi báng, bôi nhọ và các hình thức biểu đạt khác.

*Thứ hai*, việc xử phạt hành vi phạm nên tiếp cận dưới góc độ vi phạm hành chính, thay vì hình sự hóa hành vi này. Đây cũng là tiêu chuẩn chung trong pháp luật nhân quyền quốc tế và xu hướng chung tại các quốc gia trên thế giới<sup>69</sup>.

*Thứ ba*, đề cao mối quan hệ nhiều phía giữa Nhà nước, nhà cung cấp dịch vụ MXH, các tổ chức xã hội và người dùng. Sự hợp tác giữa Nhà nước và các công ty công nghệ là điểm then chốt. Nhà nước cần có cơ chế pháp lý thúc giục các công ty này phản ứng nhanh để loại bỏ phát ngôn thù ghét. Mô hình của Đức là một ví dụ tiêu biểu. Tuy nhiên, cũng cần phải nghiên cứu thêm để bảo đảm việc ngăn ngừa này không xâm phạm vào quyền tự do ngôn luận của các cá nhân khác trên MXH.

*Thứ tư*, tích cực phát động các phong trào, chiến dịch phổ biến quy tắc ứng xử trên MXH, đặc biệt hướng sự chú ý tới người trẻ, đề nâng cao tính lành mạnh trong môi trường này.

Trong quá khứ, điều chỉnh vấn đề tự do ngôn luận dễ dàng hơn khi lãnh thổ quốc gia có ranh giới xác định. Ngày nay, Internet và MXH đã mở đường cho quá trình toàn cầu hóa, tạo thành những liên kết xuyên quốc gia, xóa nhòa biên giới địa lý, nhưng đồng thời cũng làm nảy sinh những vấn đề toàn cầu, trong đó có TGTT. Những nền tảng công nghệ mới cởi mở hơn và đa dạng hơn như vậy chắc chắn sẽ khó quản lý hơn so với các phương tiện truyền thông truyền thống.

<sup>66</sup> Xem <http://tuyengiao.vn/cung-suy-ngam/phan-ngon-gay-thu-ghet-van-nan-cua-mang-xa-hoi-98852>, truy cập ngày 01/8/2019.

<sup>67</sup> Xem <https://ec.europa.eu/transparency/regcxpert/index.cfm?do=groupDetail.groupDetailDoc&id=29738&no=1>, truy cập ngày 01/12/2019.

<sup>68</sup> Xem <https://vnexpress.net/thoi-su/de-xuat-xay-dung-bo-quy-tac-loc-phat-ngon-xau-tren-mang-xa-hoi-3569454.html>, truy cập ngày 01/8/2019.

<sup>69</sup> Xem thêm Vũ Công Giao, Nguyễn Đình Đức (2018), *Chống phát ngôn thù ghét, phi báng trên Internet ở Hoa Kỳ, Liên minh châu Âu và những giá trị tham khảo cho Việt Nam*, Tạp chí Nghiên cứu lập pháp, số 17 (3-4).