INTERNET

Os desafios da identificação de conteúdo criminoso em redes sociais

Exame do crime de 'incitação', em qualquer das suas modalidades, possui enorme complexidade no contexto das redes

RAQUEL SCALCON

Crédito: Pexels

Nas linhas que seguem, compartilho algumas impressões incipientes sobre a relação entre liberdades constitucionais, direito penal e regulação das redes sociais. Considero que a moderação dos conteúdos que flanam nessa peculiar realidade virtual representa, sem dúvida, uma das mais densas indagações da atualidade.

A questão, que por certo é mundial, revelou-se em nosso país com invulgar intensidade. Perpassa desde o problema das fake news ou desinformação, com especial gravidade durante as eleições presidenciais de 2022, até culminar em uma possível incitação e organização dos atos antidemocráticos de 8 de janeiro de 2023.

A percepção de uma escalada no uso de plataformas digitais como instrumento e/ou meio (em sentido amplo) útil à prática de atos criminosos precipitou discussão intensa sobre *se* e *como* regulá-las. Não sendo possível falar sobre o tema puramente em abstrato, irei partir da

seguinte inquietação: quais seriam as dificuldades reais para avaliar se um determinado conteúdo postado em uma rede social configura crime? Como hipótese, considero que os desafios são robustos e decorrem, ao menos em parte, de uma peculiaridade que chamarei de "linguagem ao quadrado" ou "dupla indeterminação"[1]. Esse ponto e outros serão examinados a seguir.

Crimes e respectivos meios para sua consumação: o caso das postagens em redes sociais

De início, é fundamental pontuar que as normas incriminadoras descrevem uma *conduta* humana (preceito primário) e fixam uma *sanção* (preceito secundário). Somados os preceitos, temos o "crime e castigo" (Dostoiévski). A dificuldade que identifico quando em questão a pretensão de moderar conteúdo potencialmente criminoso está no preceito primário, ou seja, na descrição – pela via da linguagem escrita – de uma conduta reputada criminosa, que, em regra, *não é*, em si, também puramente linguagem escrita.

Como exemplo, pensemos no ato de "matar alguém" (tipificado no art. 121 do Código Penal) ou ainda no ato de "gerir fraudulentamente instituição financeira" (tipificado no art. 4º da Lei 7.492/86). Em ambos os casos, é

bastante difícil – senão impossível – imaginar uma forma de consumação por meio *apenas* da linguagem e/ou postagens em rede social. Logo, essa realidade virtual, isoladamente, ainda não seria apta a colocar certos bens jurídicos em risco.

De outro lado, as redes sociais, a depender do seu uso, podem ser um ambiente com potencialidade lesiva a outros bens jurídicos, como, ilustrativamente, a honra. A presença ou ausência desta aptidão reside na relação entre o bem jurídico protegido e os meios capazes de lesioná-lo ou colocá-lo em risco. Enquanto uma arma de fogo municiada é um meio potencialmente apto para a consumação do crime de homicídio, não o é para a consumação de uma calúnia. A lógica inversa ocorre relativamente às redes sociais.

O fato de se identificar um possível uso desvirtuado das redes e plataformas digitais precipita o debate sobre a necessidade de algum grau de regulação estatal. No entanto, a realidade virtual é pouco tangível e, logo, de mais arredio controle pela via das regras legais do que outros âmbitos da vida. Uma das razões está no problema da "linguagem ao quadrado" ou "dupla indeterminação", acima mencionado.

Regras jurídicas valem-se da linguagem escrita para qualificar comportamentos humanos como proibidos, permitidos ou mesmo obrigatórios. Por sua vez, no ambiente virtual boa parte dos conteúdos que se pretende regular também é linguagem escrita. A linguagem, contudo, é vaga, ambígua e polissêmica – ou seja, possui algum grau, maior ou menor, mas nunca nulo, de indeterminação – e tais características são potencializadas nas redes sociais[2]. A junção dessa dupla indeterminação traz, à toda evidência, caminhos tortuosos para sua regulação.

Pois bem, é fato que, no âmbito das condutas criminosas, é exceção o crime que possa ser consumado exclusivamente pela linguagem escrita, embora haja alguns bastante importantes, como o racismo (art. 20, Lei 7.716/89)[3], a ameaça (art. 147, Código Penal)[4], a calúnia (art. 138, Código Penal) etc. Já o homicídio (art. 121, Código Penal) e tantos outros (desde o estupro até a lavagem de capitais) simplesmente não estão

nesse rol. A linguagem escrita não é, pois, um *meio* capaz, por si mesmo, de consumar a grande maioria dos crimes do nosso ordenamento jurídico. Ao mesmo tempo, é bastante plausível reputá-la apta a *instigar* e/ou *incitar* a sua prática.

Portanto, o primeiro passo para bem refletir sobre a potencial relação entre (i) postagens em redes sociais e (ii) condutas criminosas é identificar os delitos passíveis de serem consumados *por meio* de redes sociais. Ou seja, esse ambiente pode ser desvirtuado e instrumentalizado para a consumação de *quais* crimes, e por quê? Como acima argumentado, postagens em redes sociais não possuem potencialidade para colocar em risco ou lesionar todo e qualquer bem jurídico, pelo contrário. É preciso, assim, identificar o espectro de maior risco.

Nesse sentido, a possibilidade de usurpação das redes, que é mais exceção do que regra, revela-se com grande intensidade em certas condutas criminosas, sendo tal *locus* um ambiente altamente propício à consumação de crimes contra a honra em geral (art. 138 a 140, Código Penal), contra a privacidade (por exemplo, art. 218-C, Código Penal)[5] e de racismo (Lei 7.716/89)[6]. Há ainda outros delitos que aqui denominarei de "suspeitos", isto é, potencialmente consumáveis por meio de postagens, a saber: (i) incitação ao crime e incitação à animosidade das Forças Armadas (modalidades do art. 286, *caput,* e parágrafo único, do Código Penal) e (ii) crimes contra as instituições democráticas (arts. 359-L e 359-M, Código Penal)[7].

Dentre eles, creio que o crime de instigação, em qualquer das suas modalidades, seja um dos que merece mais detido exame, por ser um dos mais "suspeitos" no sentido acima referido. Assim, como passo seguinte, proponho um singelo exercício prático: lançando mão de algumas postagens que circularam nas redes sociais em época próxima aos atos criminosos de 8 de janeiro de 2023, buscarei avaliar se elas configuram, ou não, o crime de incitação à animosidade das **Forças Armadas** (art. 286, P. Ú., Código Penal).

Crimes de incitação à animosidade das Forças Armadas e postagens com conteúdo 'suspeito'

Inicio com um conjunto de exemplos de postagens reais que circularam em plataformas digitais e que, aparentemente, se aproximam em seu conteúdo e sentido, tocando o tema da "intervenção militar" no Brasil. Veja-se:

- "COMPARTILHEM EM SUAS REDES, VAMOS PRO QUARTEL, BORA FAZER NOSSA HISTÓRIA!! BORA POVO, QUEM DEVE TEMER SÃO ELES !!!";
- "SOS Forças Armadas, Salvem o Brasil";
- "Intervenção militar e artigo 142 JA!";
- "O artigo 142 está na constituição. Cogitar usar não é crime. Não é crime algo que não aconteceu aliás o voto impresso pacificaria tudo";
- "INTERVENÇÃO MILITAR NOS TRÊS PODERES JÁ."

Diante do conteúdo dessas postagens, opto por indicar como hipótese de incriminação o tipo penal do parágrafo único do art. 286 do Código Penal, inserido pela Lei dos Crimes contra o Estado Democrático de Direito (Lei 14.197/2021)[8], que assim dispõe:

CP, Art. 286 – Incitar, publicamente, a prática de crime:

Pena – detenção, de três a seis meses, ou multa.

Parágrafo único. Incorre na mesma pena quem incita, publicamente, animosidade entre as Forças Armadas, ou delas contra os poderes constitucionais, as instituições civis ou a sociedade (grifado).

Examinando-se o tipo penal em questão, mas ainda na clássica modalidade do *caput*, é importante esclarecer que a sua consumação somente ocorre quando, constatada a incitação à prática de crime, esta *não* teve êxito. De tal maneira que, caso o crime incitado (um homicídio,

por exemplo) seja efetivamente praticado, o agente haveria de responder como partícipe por instigação do homicídio, não como autor do delito de incitação ao crime.

Essa questão, todavia, é mais complexa na modalidade especial do parágrafo único, porque não se trata de incitar *publicamente* a prática de um *crime,* mas de incitar *publicamente* uma *animosidade* (uma hostilidade), termo este que, contudo, é de difícil delimitação e subsunção. O seu significado, pois, enseja amplo e necessário debate. No caso das mensagens acima apresentadas, penso que a dificuldade não estará na avaliação da elementar *publicidade*, mas sim no exame da conduta de *incitar* em si, bem como do seu objeto – a animosidade –, que, no caso em análise, seria *das Forças Armadas contra os poderes constituídos*.

Veja-se que "incitar", verbo-núcleo *do caput,* é conduta complexa e já existente em nosso ordenamento jurídico desde o Código Penal de 1940, mas seguramente não pensada para a atual realidade virtual. A doutrina clássica, forjada em um mundo analógico, referia que "na incitação, há uma dupla substituição: à vontade do incitado substitui-se a do incitador; à ação do incitador substitui-se a ação do incitado. O incitador empolga a vontade do incitado".[9]

Quanto ao seu modo, afirmava-se que a incitação pode ser realizada "pela palavra só (a irradiação de um discurso incendiário)", mas também, por exemplo, pela "distribuição de escritos"[10]. Contudo, sustentava-se que, "uma vez instigado um número indeterminado de pessoas, será instigação (...), *per se,* se séria"[11], pois não há falar em instigação criminosa se ela não tiver, ao menos, "aparência de eficiente"[12] ou seja, se ela não se mostrar *provável*, *sedutora*, *convincente* ou a um passo de motivar o incitado.

Não penso que essas ideias doutrinárias precisam ser atualizadas, menos ainda descartadas. Contudo, há de se admitir, sem rodeios, que o exame do crime de "incitação", em qualquer das suas modalidades, possui enorme complexidade no contexto de redes sociais.

Feita essa ressalva, passo a examinar as postagens acima elencadas, no que assim constato, em análise preliminar: (i) elas cumpririam o requisito da publicidade; (ii) elas estavam, em alguma medida, dirigidas às Forças Armadas e (iii) havia uma aparente pretensão de *tensionar* a relação destas com os poderes constituídos. O item "i" seria pouco controverso, mas há algum espaço de discussão quanto aos itens "ii" e "iii". De todo modo, essas avaliações são insuficientes para sustentar uma conclusão no sentido de que as postagens são criminosas.

Isso porque há de se examinar, ainda, a presença de um elemento de difícil aferição, mas central, qual seja, (iv) se havia uma efetiva *aptidão* das postagens para incitar a pretendida animosidade. Como dito acima, a incitação há de ser "séria", "eficiente", apta a convencer o incitado. É preciso avaliar, pois, se as postagens são, ou não, dotadas de tal capacidade. A pergunta é de difícil resposta, mas, ao mesmo tempo, enfrentá-la é absolutamente essencial para tomadas de decisões no campo da moderação de conteúdo pelas plataformas.

Em uma avaliação ainda não amadurecida, admito que tenho grande dificuldade em reputar os conteúdos acima listados, *em postagem isolada*, como típicos para o crime do Parágrafo Único do art. 286 do Código Penal. E assim considero, porque não vejo, *prima facie,* a presença de "aparência de eficiência" para incitar a animosidade das Forças Armadas. Contudo, essa percepção parece mudar degrau a degrau, ganhando novos contornos, quando se considera a contínua *disseminação* destes mesmos conteúdos.

Em outras palavras, a exigida aptidão para "incitar a animosidade das Forças Armadas" parece revelar-se, pois, na exata conjugação entre (i) característica do conteúdo da postagem e (ii) intensidade da sua propagação. Ou seja, uma postagem isolada poderia até carecer dessa aptidão, mas tal qualidade ganharia mais e mais robustez quanto maior fosse a disseminação do conteúdo em si, em uma clara lógica de *risco na acumulação*. Essa conclusão não torna crime, em si, a postagem isolada, mas joga luz sobre *onde* reside o risco em questão e como mitigá-lo. O direito penal, nesse ponto, não deve ser o protagonista.

Caso as premissas anteriores sejam aceitas, então o papel central das plataformas e redes sociais, em caso de postagens altamente duvidosas e suspeitas, mas *não* evidentemente delituosas, seria menos a *moderação* do *conteúdo* em si, com sua imediata remoção, e mais a *moderação* da sua *disseminação*[13]. Seria relevante desenvolver, pois, instrumentos de autorregulação que possibilitem limitar o alcance, a propagação e o compartilhamento de conteúdo suspeito. Os algoritmos poderiam identificar conteúdos suspeitos e, ato contínuo, moderar a sua propagação. Esta singela ideia pode se prestar a mitigar, em algum grau, parcela dos dilemas decorrentes da imbricação entre liberdades constitucionais, direito penal e regulação de redes sociais.

---

[1] Cf. SCALCON, Raquel Lima. *Linguagem, interpretação e criminalização nas redes*. In: Jornal O Estado de São Paulo. Edição de 19 de abril de 2023. Disponível em: **https://www.estadao.com.br/opiniao/espaco-aberto/linguagem-interpretacao-e-criminalizacao-nas-redes/**. Acesso em 07 de maio de 2024.

[2] Evidentemente, uma regra legal orientadora de condutas deve evitar, ao máximo, ambiguidade, vagueza, imprecisão etc. Sobre o tema, cf. RAZ, Joseph. *The Authority of Law. Essays on Law and Morality.* Oxford: Oxford, 1979, p. 214 ss.

[3] Lei 7.716/89 – Art. 20. Praticar, induzir ou incitar a discriminação ou preconceito de raça, cor, etnia, religião ou procedência nacional. Pena: reclusão de um a três anos e multa.

[4] Código Penal – Art. 147 – Ameaçar alguém, por palavra, escrito ou gesto, ou qualquer outro meio simbólico, de causar-lhe mal injusto e grave: Pena – detenção, de um a seis meses, ou multa.

[5] Código Penal – Art. 138 – Caluniar alguém, imputando-lhe falsamente fato definido como crime: Pena – detenção, de seis meses a dois anos, e multa.

[6] Cf. nota 6.

[7] Código Penal – Art. 359-L. Tentar, com emprego de violência ou grave ameaça, abolir o Estado Democrático de Direito, impedindo ou restringindo o exercício dos poderes constitucionais: Pena – reclusão, de 4 (quatro) a 8 (oito) anos, além da pena correspondente à violência.

Código Penal – Art. 359-M. Tentar depor, por meio de violência ou grave ameaça, o governo legitimamente constituído: Pena – reclusão, de 4 (quatro) a 12 (doze) anos, além da pena correspondente à violência.

[8] Figura delitiva similar já existia na Lei de Segurança Nacional, mas que não tinha como pretensão a tutela do Estado Democrático de Direito. Veja-se: "Art. 23 – Incitar: (…) II – à animosidade entre as Forças Armadas ou entre estas e as classes sociais ou as instituições civis; (…) Pena: reclusão, de 1 a 4 anos".

[9] DRUMMOND, J. de Magalhães. *Comentários ao Código Penal*. Vol. IX. Arts. 250 a 361. Rio de Janeiro: Forense, 1944, p. 175-176.

[10] *Idem*, p. 176.

[11] *Idem*, p. 178-179.

[12] *Idem*, p. 179.

[13] Sobre moderação de conteúdo *versus* disseminação e legislação estrangeira, ver: COUTINHO, Diogo R.; KIRA, Beatriz. *PL das Fake News na regulação da internet*. **JOTA**, 2023. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/pl-das-fake-news-na-regulacao-da-internet-25042023>. Acesso em: 07 de maio de 2023 e KIRA, Beatriz; COUTINHO, Diogo R. *Desafios da regulação digital – parte 2*. **JOTA**, 2023. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/desafios-da-regulacao-digital-parte-2-27022023>. Acesso em: 07 de maio de 2023.

RAQUEL SCALCON – Consultora e parecerista. Professora da FGV Direito SP nos cursos de

graduação e pós-graduação. Doutora em Direito Penal pela UFRGS, com período de pesquisa na Georg-August-Universität Göttingen (Alemanha)

É estudante ?
Aproveite as condições especiais para quem está na graduação, mestrado ou doutorado.

ASSINE

Os artigos publicados pelo JOTA não refletem necessariamente a opinião do site. Os textos buscam estimular o debate sobre temas importantes para o País, sempre prestigiando a pluralidade de ideias.