
Discurso tóxico na internet está piorando? Estudo publicado na Nature analisa 30 anos de comentários e rejeita a hipótese

A internet é frequentemente descrita como um ambiente social nocivo e decadente, uma visão tão amplamente aceita que se tornou uma verdade inquestionável sobre a vida online. No entanto, estudo publicado em 20 de março na respeitada revista *Nature*, "[Persistent interaction patterns across social media platforms and over time](#)", oferece uma nova perspectiva que desafia essa premissa, sugerindo que o discurso online não está piorando.

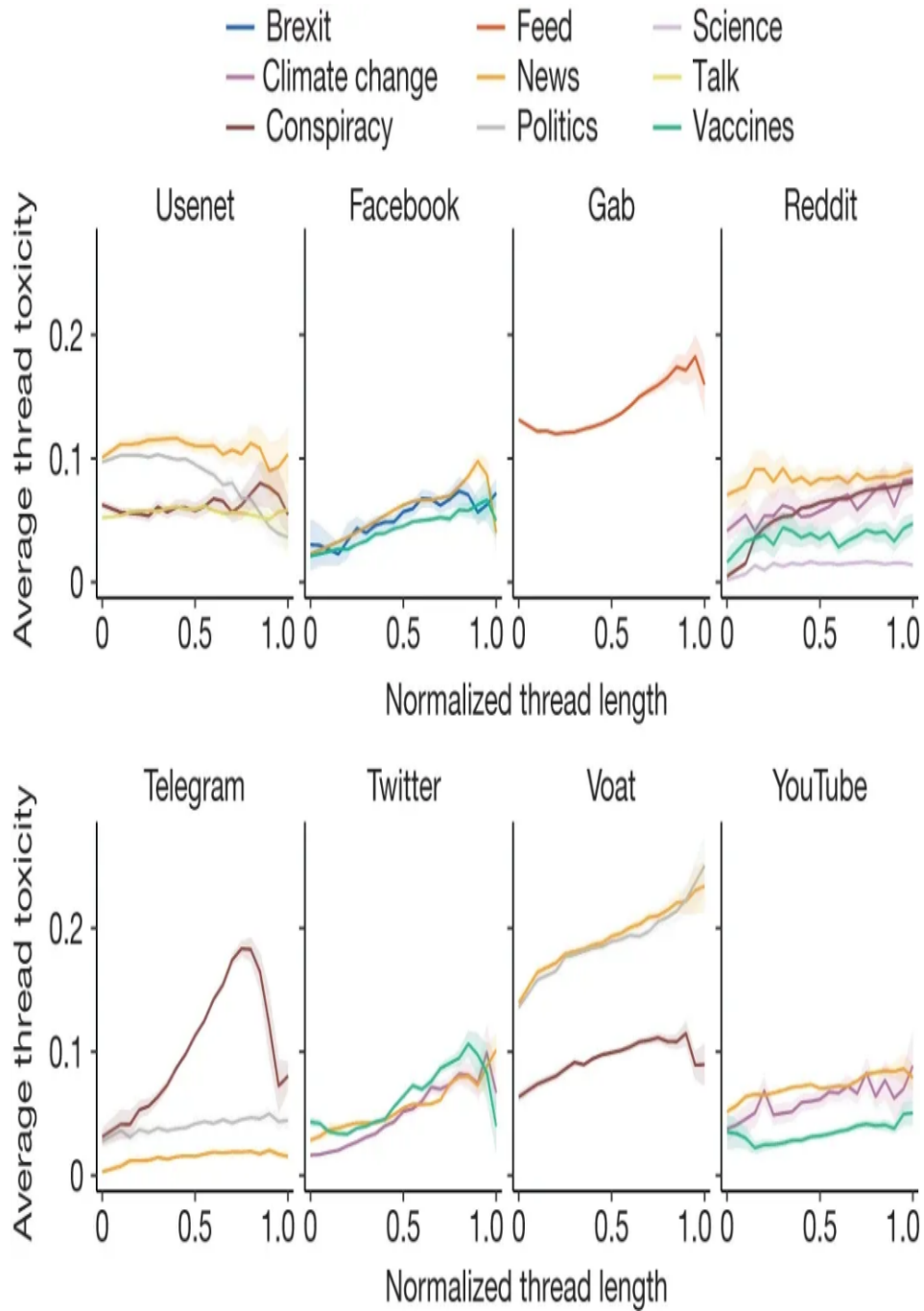
Walter Quattrociocchi, autor do estudo e líder do Data and Complexity for Society Lab na Universidade Sapienza de Roma, e sua equipe analisaram mais de meio bilhão de comentários ao longo de 30 anos. Os resultados indicam que o nível de toxicidade nas conversas online se manteve relativamente constante desde os anos 90. “

Apesar da evolução das plataformas e de como as utilizamos, os comportamentos humanos nesses espaços permaneceram surpreendentemente estáveis”, disse Quattrociocchi.

O estudo utilizou o Perspective AI para classificar a toxicidade dos comentários, avaliando discursos considerados “rudes, desrespeitosos ou irracionais” que provavelmente fariam alguém abandonar uma discussão. As análises mostraram que, embora a toxicidade e a participação dos usuários sejam independentes, discussões mais longas tendem a se tornar mais tóxicas, envolvendo menos participantes, mas mais ativos, à medida que progridem.

Esses padrões foram consistentemente observados em várias plataformas, de Usenet a YouTube e Twitter, com uma porcentagem de comentários tóxicos que raramente excede 10% mesmo em comunidades notoriamente pouco moderadas como Gab e Voot. No Facebook, por exemplo, apenas cerca de 4 a 6% dos comentários foram considerados tóxicos, enquanto no YouTube, a porcentagem variou de 4 a 7%.

Toxicidade aumenta com o tamanho da conversa



Esses padrões mostraram-se surpreendentemente consistentes ao longo do tempo e entre plataformas: No geral, o estudo descobriu que a prevalência tanto de discursos tóxicos quanto de usuários altamente tóxicos era extremamente baixa. No entanto, quanto mais uma conversa se prolonga, em praticamente qualquer plataforma, mais tóxica ela se torna. Ao mesmo tempo, as conversas tendem a envolver menos participantes, mas aqueles que permanecem se tornam mais ativos à medida que a discussão se estende.

Caitlin Dewey, jornalista especializada em cultura da internet publicou [análise sobre o estudo](#) e fez a seguinte analogia: **“[É] como um certo tipo de bar perto do horário de fechar: A maioria das pessoas já foi para casa, mas os retardatários são barulhentos, desinibidos e propensos a um certo nível de travessuras e/ou agressão.”** Na mesma linha, Quattrociochi, autor do estudo afirma: **“Conversas mais longas potencialmente oferecem mais oportunidades para que desacordos escalem para trocas tóxicas, especialmente à medida que mais participantes se juntam à conversa”**, disse.

A pesquisa também destaca a cultura de moderação em comunidades como a r/science no Reddit, onde a toxicidade é excepcionalmente baixa (apenas 1%), graças a uma moderação proativa e uma comunidade que valoriza altos padrões de conversação. Mais de mil moderadores voluntários, muitos dos quais cientistas, ajudam a manter a qualidade das discussões ao filtrar estudos não confiáveis, anedotas pessoais e comentários ofensivos ou fora do tópico.

Quattrociochi enfatiza que, embora as políticas de moderação e a dinâmica das plataformas influenciem a visibilidade e a propagação de conteúdo tóxico, as raízes do comportamento tóxico estão mais profundamente enraizadas na interação humana. Portanto, uma moderação eficaz pode exigir não apenas a remoção de conteúdo tóxico, mas também estratégias mais amplas para encorajar um discurso positivo.

O estudo enfrenta limitações, como não capturar completamente o impacto de campanhas de trolling coordenadas e concentrar-se apenas em mídia baseada em texto numa era dominada por vídeos. Ainda assim, **os resultados desafiam a narrativa de que as plataformas online deterioraram a qualidade do discurso público, sugerindo que a percepção de piora no discurso online pode ser ampliada pela maior visibilidade de do conteúdo tóxico nas redes sociais.**

Portanto, apesar das preocupações com o aumento da toxicidade online, este estudo reforça a ideia de que a qualidade geral do discurso pode não estar em declínio, mas sim revelando aspectos persistentes da natureza humana em novos ambientes digitais.

Material analisado

Dataset	Time range	Comments	Threads	Users	Toxicity
Facebook brexit	31 Dec 2015 to 29 Jul 2016	464,764	4,241	252,156	0.06
Facebook news	9 Sep 2009 to 18 Aug 2016	362,718,451	6,898,312	60,235,461	0.06
Facebook vaccines	2 Jan 2010 to 17 Jul 2017	2,064,980	153,137	387,084	0.04
Gab feed	10 Aug 2016 to 29 Oct 2018	14,641,433	3,764,443	166,833	0.13
Reddit climate change	1 Jan 2018 to 12 Dec 2022	70,648	5,057	26,521	0.07
Reddit conspiracy	1 Jan 2018 to 8 Dec 2022	777,393	35,092	92,678	0.07
Reddit news	1 Jan 2018 to 31 Dec 2018	389,582	7,798	109,86	0.09
Reddit science	1 Jan 2018 to 11 Dec 2022	549,543	28,33	211,546	0.01

Reddit vaccines	1 Jan 2018 to 6 Nov 2022	66,457	4,539	5,192	0.04
Telegram conspiracy	30 Aug 2019 to 20 Dec 2022	1,416,482	32,592	150,251	0.12
Telegram news	9 Apr 2018 to 20 Dec 2022	724,482	28,288	16,716	0.02
Telegram politics	4 Aug 2017 to 19 Dec 2022	491,294	27,749	6,132	0.04
Twitter climate change	1 Jan 2020 to 10 Jan 2023	9,709,855	130,136	3,577,890	0.07
Twitter news	1 Jan 2020 to 29 Nov 2022	9,487,587	97,797	1,710,213	0.07
Twitter vaccines	23 Jan 2010 to 25 Jan 2023	49,437,212	125,667	11,857,050	0.08
Usenet conspiracy	1 Sep 1994 to 30 Dec 2005	284,838	72,655	48,224	0.05
Usenet news	5 Dec 1992 to 31 Dec 2005	621,084	169,036	76,62	0.09

Usenet politics	29 Jun 1992 to 31 Dec 2005	2,657,772	625,945	209,905	0.08
Usenet talk	13 Feb 1989 to 31 Dec 2005	2,103,939	328,009	156,542	0.06
Voat conspiracy	9 Jan 2018 to 25 Dec 2020	1,024,812	99,953	27,641	0.10
Voat news	21 Nov 2013 to 25 Dec 2020	1,397,955	170,801	88,434	0.19
Voat politics	19 Jun 2014 to 25 Dec 2020	1,083,932	143,103	66,424	0.19
YouTube climate change	16 Mar 2014 to 28 Feb 2022	846,3	9,022	436,246	0.06
YouTube news	13 Feb 2006 to 8 Feb 2022	20,536,162	107,88	4,310,827	0.07
YouTube vaccines	31 Jan 2020 to 24 Oct 2021	2,648,909	14,147	902,34	0.04

Toxicity represents the fraction of toxic comments in the dataset, where a comment is considered toxic if its toxicity score is greater than 0.6.

Para saber mais do estudo:

Avalle, M., Di Marco, N., Etta, G. *et al.* Persistent interaction patterns across social media platforms and over time. *Nature* 628, 582–589 (2024). <https://doi.org/10.1038/s41586-024-07229-y>

Category

1. Notícias

Tags

1. discurso de ódio
2. liberdade de expressão
3. redes sociais

Date

08/09/2024

Date Created

17/04/2024