

[Devemos confiar nos algoritmos? As tendências do Twitter, a questão da censura e nossa confiança nos algoritmos à nossa volta](#)

Por **Tarleton Gillespie**, Professor do Departamento de Comunicação da Universidade de Cornell



Data da publicação:

Dezembro de 2011

A questão a ser levantada neste texto não é se o Twitter está censurando sua lista de tendências do momento.¹ A questão a ser levantada é o que pensamos sobre a lista de tendências, o que ela representa e como funciona. Podemos presumir que exista algum tipo de responsabilização quando desconfiarmos que as tendências estão “erradas”? O que são estes algoritmos – e o que queremos que eles sejam?

Não é a primeira vez que esta questão vem à tona. Gilad Lotan², Vice Presidente de pesquisa e desenvolvimento da SocialFlow³ (que também já foi o profissional na Microsoft responsável pela área de design de experiências do usuário), faz esta mesma pergunta, provocado por questionamentos⁴ de participantes e apoiadores do movimento Occupy Wall Street: o Twitter está censurando sua lista de trends para excluir hashtags como #occupywallstreet e #occupyboston? Enquanto estes movimentos ganham força e atenção na mídia, participantes, observadores e críticos voltam-se para o Twitter para questioná-lo. Por que estas hashtags tão conhecidas não estão entrando para a lista das mais tuitadas? Por que elas não estão na lista das mais tuitadas nem mesmo nas cidades onde os protestos estão acontecendo, incluindo Nova York?

O que se presume - embora Gilad cautelosamente negue a possibilidade -, é que o Twitter esteja, por alguma razão, seguindo uma de duas alternativas: ou está retirando a #occupywallstreet da lista de tendências do momento, ou criou um algoritmo para dar prioridade a assuntos banais - como o casamento de Kim Kardashian - em detrimento de debates mais polêmicos e importantes. Acusações semelhantes emergiram em torno da ausência de #wikileaks⁵ das tendências do Twitter quando vários telegramas diplomáticos foram divulgados em dezembro de 2010, assim como aconteceu com os protestos de #demo2010⁶ no Reino Unido, a execução controversa de #TroyDavis⁷ no estado da Geórgia, a #flotill⁸ de Gaza, e até a morte de #SteveJobs.⁹ Por que, justamente quando estas questões parecem estar fervendo, elas não aparecem na lista dos temas mais tuitados?

A despeito de haver uma forte onda de ceticismo paranóico nas análises sobre este tema, e especialmente nas listas de comentários que seguem as análises, a maioria das pessoas que se debruçaram sobre a questão¹⁰ estão seguras de que o Twitter não está censurando estes tópicos. A ausência destes assuntos das listas de mais tuitados é na verdade produto de uma dinâmica particular do algoritmo que determina tendências - e a maioria dos usuários faz confusão por não saber exatamente como funciona este algoritmo, e o que ele deveria identificar. Eu não discordo desta análise e não é de meu interesse reabrir estas questões. Assim como Gilad Lotan, Angus Johnston tem uma série de posts¹¹, negando a acusação de censura em relação à hashtag #wikileaks. As tendências não foram criadas (e recriadas) pelo Twitter para simplesmente medir popularidade - por exemplo, a quantidade de posts utilizando uma certa palavra ou hashtag.

Na realidade, o Twitter desenvolveu os algoritmos de tendências para capturar tópicos que tornam-se populares repentinamente, crescendo de forma diferenciada, emergindo acima do nível normal de tweets, comentários e retweets. Para isso, seu algoritmo é feito com o objetivo de levar em consideração não somente o número de tweets, mas também fatores como: o aumento da velocidade com que um tema é tuitado; se este já foi parte da lista de trends ou não; se está sendo utilizado por várias redes de pessoas, e não somente por um conjunto de usuários intensamente interconectado; se os tweets são diferentes ou são somente retweets de um mesmo post. Como representantes do Twitter¹² já afirmaram, eles não querem ver destacadas na lista de tendências simplesmente as palavras mais tuitadas (o que faria a lista de tendências lembrar um exercício de gramática sobre pronomes e artigos indefinidos) ou os tópicos que são sempre populares e aparentemente estão destinados a continuar assim (como parece ser o caso do Justin Bieber¹³).

A acusação de censura é contra intuitiva. O Twitter, nos últimos anos, tem concordado com argumentos sobre seu papel catalizador em recentes movimentos políticos e civis, particularmente no mundo árabe. Neste caso em particular, o Twitter, sem modéstia, atribui a si mesmo grande importância política e alega empunhar a bandeira¹⁴ da coragem¹⁵ (ver Shepherd e Busch¹⁶). A ação de censurar estes tópicos tão populares nas listas de tendências iria contra seus objetivos autoproclamados, e, mais importante ainda, iria contra suas táticas de marketing. Como notou Johnston, se os próprios tweets estão disponíveis, então por quais motivos o Twitter removeria #wikileaks ou a #occupywallstreet da lista de trends, enquanto deixa a discussão mesma sobre estes tópicos correr solta?

Por outro lado, o vigor e a persistência das acusações de censura não surpreende. Os defensores de causas políticas como o caso Wikileaks e o Occupy Wall Street querem desesperadamente que seus temas ganhem visibilidade. As pessoas envolvidas nestas discussões provavelmente exageram a dimensão do quanto são importantes e de quão discutidos são seus tópicos. E, especialmente em relação a #wikileaks e #occupywallstreet, a possibilidade de o Twitter exercer censura sobre estes temas se encaixaria na perspectiva ideológica de seus apoiadores: o Twitter estaria agindo contra o Wikileaks assim como fez a Amazon, a Paypal, e a Mastercard; ou, no caso de #occupywallstreet, enquanto a rede de usuários do Twitter dá voz às pessoas, a empresa Twitter deve ter alianças firme com os “cachorros grandes” de Wall Street.

Mas o debate sobre ferramentas como as trends do Twitter é, acredito, um debate que vamos ter cada vez mais frequentemente. Enquanto a maior parte de nosso discurso público online ocorre em um conjunto seleto de plataformas de conteúdo e redes de comunicação privadas, e estes provedores adotam algoritmos complexos para gerenciar, curar e organizar enormes conjuntos de dados, há uma importante tensão que surge - entre o que esperamos que sejam estes algoritmos, e o que eles são de fato. Não somente devemos reconhecer que estes algoritmos não são neutros, que eles são o produto de escolhas políticas, mas também que eles enfocam informações de um modo particular. Também é importante entendermos o que significa o fato de que estamos começando a confiar neste algoritmos, que queremos que eles sejam neutros, queremos que sejam confiáveis - queremos que sejam eficientes meios de informação sobre os temas mais importantes.

A lista de trends do Twitter é apenas a mais visível deste tipo de ferramentas. A ferramenta de buscas em si é um algoritmo que promete um conjunto lógico de resultados em resposta a uma pergunta, mas na verdade é o resultado de um algoritmo feito para considerar uma vasta gama de critérios para fornecer resultados que satisfaçam não somente o usuário, mas também os objetivos dos provedores, sua visão de relevância, valor ou importância social, e as demandas particulares de seu modelo de negócios. Assim como James Grimmelman¹⁷ observou, “ferramentas de busca se vangloriam por serem automatizadas, mas de fato não são”. Quando a Amazon ou YouTube ou Facebook oferecem um relatório algorítmico e em tempo real do que é “mais popular” ou “mais comentado”, ou com mais “notas boas”, estes sítios estão elegendo e organizando uma lista cuja legitimidade é baseada na afirmativa de que ela não foi deliberadamente organizada. E nós queremos acreditar nesta legitimidade, a ponto de não nos interessarmos sequer em levantar questões sobre as escolhas e implicações dos algoritmos que usamos todos os dias.

Se analisarmos os algoritmos mais a fundo, suas implicações tornam-se bem aparentes. Sim, numa visita ao sítio do Twitter, vemos que a ferramenta de tendências é apresentada como uma lista de termos sem maiores implicações, o simples produto de um cálculo. Mas uma leitura casual da explicação sobre como esta lista de tendências funciona – nos termos e condições de uso¹⁸ do Twitter, em sua página de ajuda, no blog da empresa¹⁹, em tweets²⁰, em suas respostas a perguntas da imprensa²¹, até mesmo em comentários²² nas discussões sobre censura – mostra claramente a variedade de fatores que entram neste cálculo e atesta a ocasional e infeliz consequência destes algoritmos.

O Wikileaks pode não ter aparecido na lista de trends quando as pessoas assim esperavam devido ao fato de já ter aparecido antes; ou porque a discussão sobre a #wikileaks cresceu de forma lenta e consistente ao longo do tempo, sem despontar de forma aguda o suficiente para chamar a atenção do algoritmo; ou porque a maior parte das mensagens foram retweets; ou mesmo porque os usuários tuitando sobre o Wikileaks já estavam densamente interconectados. Quando o Twitter mudou significativamente seu sistema de algoritmos²³ em maio de 2010 (embora, na verdade, ele já tenha sofrido ajustes de formas menos visíveis antes e depois disso), a mudança foi anunciada no blog da empresa, ofereceu-se explicações sobre por que a mudança foi feita e até houve um pedido de desculpas diretamente ao Justin Bieber²⁴, cuja posição na lista de mais tuitados seria afetada por causa de tal mudança.

Em resposta às alegações de censura, os executivos do Twitter explicaram por que eles acreditavam que as listas deveriam privilegiar termos que crescem muito e repentinamente; que excedem grupos singulares de usuários interconectados; que trazem novo conteúdo ao invés de retweets, e novos termos ao invés dos que já estão sendo tuitados. Aqueles que são críticos à ferramenta reúnem evidências por vezes não confiáveis e fazem análises estatísticas utilizando ferramentas disponíveis online que rastreiam a popularidade de palavras de uma forma mais exaustiva e abrangente do que o Twitter - ou talvez mais abrangente do que os resultados que o Twitter está disposto a oferecer a seus usuários. Os algoritmos que definem o que está “tuitando” ou o que é “do momento” não são fórmulas triviais: são, sim, cautelosamente formuladas para capturar o que os provedores dos sítios Web querem capturar, e limpar os eventuais “enganos” que um cálculo simples faria.

Ao mesmo tempo, o Twitter certamente seleciona e organiza suas listas de tendências - e aplica a censura tradicional: por exemplo, um engenheiro do Twitter reconhece²⁵ que as listas excluem termos profanos, algo que fica óbvio para quem tenta incluir linguagem vulgar nas listas de tendências – para fazer isso, é preciso enveredar por um labirinto.²⁶ O Twitter remove tweets que constituem ameaças específicas de violência, violações de propriedade intelectual, uso indevido de identidade alheia, violação à privacidade, ou spam. O Twitter chegou até a ser criticado²⁷ por não remover alguns termos da lista de trends, como mostra a reclamação de um usuário²⁸ de que a hashtag #reasonstobeatyourgirlfriend (razões para bater em sua namorada) foi permitida.

O Twitter também se envolve com formas mais suaves de governança, desenvolvendo um algoritmo para privilegiar alguns tipos de conteúdo e excluir outros, assim como para privilegiar alguns usuários e não outros. Além disso, o Twitter oferece regras, guias e sugestões²⁹ para a tuitagem mais apropriada, na esperança de gentilmente orientar os usuários em direção a tipos de tópicos adequados a seu sítio Web e para longe de um tipo de conteúdo que, se entrasse na lista dos assuntos do momento, poderia significar um impacto negativo no sítio do Twitter. Com relação a algumas de suas regras referentes a conteúdo adequado de perfil, conteúdo de tweets e uso de hashtags, a punição imposta a usuários violadores das normas é que seus tweets não aparecerão na ferramenta de busca e na lista de trends, sacrificando, portanto, a própria lista de trends ao impedir que conteúdos sejam sequer considerados para estar nesta lista.

O Twitter inclui termos relacionados a seus parceiros promocionais em sua lista de trends, termos que não cresceriam em popularidade se não fosse esta preseleção. Esta lista automaticamente calculada é também o resultado de uma curadoria dos conteúdos, uma organização cuidadosa que decide o que deve ser representado, que tipo de conteúdo é “quente”. Ironicamente, termos como #wikileaks e #occupywallstreet são exatamente os tipos de termos que, a partir de uma perspectiva razoável, o Twitter deveria querer mostrar em sua lista de assuntos do momento. Se partirmos da lógica de que o Twitter está se beneficiando do papel que desempenhou em insurgências democráticas populares recentes; que está proclamando-se como uma ferramenta vital para discussões políticas importantes, e que quer dar destaque a termos que endossem esta visão, bem como atrair usuários para tópicos que parecem ser relevantes, então #occupywallstreet parece ter sido feita sob medida para isso. Se for assim, mesmo com a reformulação cautelosa dos algoritmos para desbancar a posição constante de Bieber nos trends e limpar as impurezas da linguagem vulgar, o Twitter ainda não consegue ser sempre bem sucedido em suscitar discussões populares relevantes, como parece almejar. Assim, o Twitter acaba indo ao

encontro de seus críticos; talvez #wikileaks devesse ter feito parte da lista de assuntos do momento, talvez devesse ter entrado para a lista de trends após o vazamento dos telegramas diplomáticos.

Estes algoritmos não são perfeitos; funcionam como facas cegas, quando seriam necessários bisturis. As listas de trends podem parecer por vezes um estudo sobre a insignificância - não apenas porque os interesses de uns poucos são irrelevantes para a maioria, mas também porque muito do que falamos no Twitter todos os dias são temas do cotidiano, mesmo com eventuais alegações heróicas sobre sua importância política. Entretanto, muitos usuários do Twitter tomam a lista de trends não somente como uma forma de mensurar visibilidade, mas como um meio para alcançar visibilidade, a despeito de não sabermos se a aparição de um termo ou hashtag aumenta ou não a sua audiência – isso ainda não está claro. O que a lista de tendências oferece é a possibilidade de alavancar um tópico a um nível de atenção maior, e também oferece provas da atenção que um determinado tópico já está recebendo – ou pelo menos parece estar recebendo.

Evidentemente, o Twitter tem em suas mãos o mais importante recurso para melhorar seu funcionamento: uma enorme e interessada base de usuários. É fácil imaginar resolver este problema através de crowdsourcing³⁰, pedindo a usuários para avaliar a qualidade das listas de trends e analisar estas respostas ao longo do tempo, utilizando um enorme número de pontos de dados. Mas aí o Twitter enfrenta um dilema, uma vez que revelar o funcionamento do algoritmo, “seja para responder a acusações de censura e manipulação, ou mesmo para compartilhar a tarefa de melhorar o sistema”, pode significar abrir a guarda para aqueles que colocariam o sistema do Twitter em jogo: de spammers a ativistas políticos, de fraudulentos a narcisistas – todos podem querer “otimizar” seus tweets e hashtags para aparecer nas tendências. Então o mecanismo subjacente a esta ferramenta, que se propõe a apresentar uma estimativa (quase) democrática sobre o que o público pensa ser importante num determinado momento, não pode revelar sua “receita secreta”.

Isso nos deixa, a nós e ao Twitter, em um beco sem saída. O verniz algorítmico das nossas práticas de agregação de dados sociais pode sempre ser lido - ou mal compreendido - como censura, se os resultados não forem equivalentes às expectativas. Se #occupywallstreet não está na lista de assuntos candentes, isso significaria que: (a) esta hashtag está sendo censurada de propósito? (b) que este assunto é popular, mas de forma consistente, e por isso não representa um pico de popularidade ou (c) ele é menos popular do que se pensava? Recortes amplos de grandes volumes de dados, como as listas de tendências do Twitter, servem para mostrar de alguma forma o que já sabemos ser verdade - e para mostrar que somos incapazes de perceber a verdade por causa de nosso escopo de visão limitado. Nunca podemos realmente distinguir o que o Twitter está mostrando-nos daquilo que não está. Seguimos presos em uma regressão algorítmica³¹, e nem mesmo o Twitter pode ajudar, pois não pode arriscar revelar os critérios que usa.

No entanto, o mais importante aqui não são as consequências dos algoritmos, mas nossa crescente e poderosa confiança neles. A ferramenta *Twitter Trends* mede “tendências”, um fenômeno que o Twitter tem o poder de definir e construir através de seu algoritmo. Nós somos levados a tratar as trends do Twitter como uma maneira aceitável e satisfatória de medir popularidade e importância - uma “tendência” no entendimento comum do termo. E o que queremos é que isso seja assim realmente. Queremos que as tendências sejam um árbitro imparcial do que é relevante... E queremos que nosso “assunto de estimação”, aquele que parece ser o foco da atenção de “todos” (ou que deveria sê-lo) seja corretamente reconhecido por esta medida objetiva e especificamente construída para isso. Queremos que o Twitter esteja “certo” sobre o que é importante. E às vezes também queremos que ele esteja errado, deliberadamente errado, porque isso também se encaixa em nossa visão de mundo: quando fatos não são bem representados é porque alguém fez isso deliberadamente, e não porque os fatos são, de muitas formas, o produto de como eles são criados.

Não temos vocabulário suficiente³² para avaliar a intervenção algorítmica em uma ferramenta como a *Twitter Trends*. Não somos bons em compreender a complexidade necessária para construir uma ferramenta como esta – que parece identificar com facilidade o que está acontecendo, que não é sufocada pelo mundano ou pelo que é irrelevante. Não temos uma linguagem para as inesperadas associações³³ que os algoritmos fazem, que vão além da intenção (ou até da compreensão) dos próprios engenheiros que os desenvolveram. Não temos uma ideia clara de como falar sobre esta política dos algoritmos. Se a ferramenta de trends, tal como está formulada, deixa #occupywallstreet fora da lista, mesmo quando seu uso está pipocando e quando as pessoas pensam que esta hashtag deveria estar lá... será que o algoritmo está analisando o que está acontecendo de verdade? Será que está buscando as coisas erradas? Será que o Twitter está obedecendo aos interesses de terceiros? Muito frequentemente - talvez na maioria das vezes que usamos estas plataformas -, não fazemos estas perguntas. Nós equiparamos os tópicos “quentes” ao que é popular, a lista de trends aos assuntos que importam. Mais importante ainda, podemos estar recusando-nos a reconhecer nossa crescente dependência destas ferramentas

algorítmicas, como um meio de navegar neste enorme corpo de dados porque temos que navegá-lo... porque queremos muito que estas ferramentas façam cálculos simples, neutros, sem viés, sem intervenções humanas, sem necessidade de aprimoramentos para “acertar” - e sem estar sendo moldadas pelos interesses de seus provedores.

Publicado originalmente em *Culture Digitally* - <http://culturedigitally.org/>

1. As tendências, ou trends, são as palavras mais usadas ou mais populares no momento no Twitter. Neste texto utilizamos ambos os termos para referir-nos a esta ferramenta.

2. <http://blog.socialflow.com/post/7120244374/data-reveals-that-occupying-t...>

3. Empresa especializada em desenvolvimento de tecnologias para otimização de mídias sociais – ver em <http://www.socialflow.com/>

4. Ver em <http://darinsullivan.blogspot.com/2011/10/scary-twitter-censorship-of.html>

5. <http://blogs.wsj.com/speakeasy/2010/12/08/twitter-denies-blocking-wikile...>

6. <http://studentactivism.net/2010/11/24/why-isnt-demo2010-trending-on-twit...>

7. <http://studentactivism.net/2011/09/21/is-twitter-blocking-troy-davis-fro...>

8. <http://modern-marketingjapan.blogspot.com/2010/06/twitter-censorship-not...>

9. <http://studentactivism.net/2011/10/05/why-isnt-your-cause-trending-no-it...>

10. Ver em <http://blog.bufferapp.com/five-twiter-secrets-about-censored-trending-to...>

11. Ver os posts no blog Student Activism: <http://studentactivism.net/2010/11/28/wikileaks>; <http://studentactivism.net/2010/12/05/twitter-wikileaks-trending-2>; <http://studentactivism.net/2010/12/05/wikileaks-twitter-3>; e <http://studentactivism.net/2010/12/11/twitter-wikileaks-why>

12. <http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html>

13. <http://new.music.yahoo.com/blogs/stopthepresses/271903/justin-bieber-acc...>

14. <http://blog.twitter.com/2011/01/tweets-must-flow.html>

15. <http://blog.twitter.com/2009/06/down-time-resche-duled.html>

16. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1924566

17. Ver em http://works.bepress.com/james_grimmelman/19/

18. Em <http://support.twitter.com/groups/31-twitter-basics/topics/111-features/...>

19. <http://blog.twitter.com/2010/12/to-trend-or-not-to-trend.html>

20. <http://twitter.com/#!/delbius/status/7547983465807872>

21. http://news.cnet.com/8301-13577_3-20025121-36.html?tag=mncol:2n

22. <http://studentactivism.net/2010/12/05/wikileaks-twitter-3/#comment-11619>

23. <http://mashable.com/2010/05/14/twitter-improves-trending-topic-algorithm...>

24. <http://twitter.com/#!/twitter/status/14080930416>

25. Ver em <http://blog.tweetsmarter.com/twitter-rules/twitter-admits-editing-offens...>

26. http://www.readwriteweb.com/archives/twitter_censoring_trending_topics.php

27. Ver em <http://www.policymic.com/article/show?id=1420> e em http://www.readwriteweb.com/archives/maybe_twitter_trends_shouldnt_be_en...

28. <http://twitter.com/#!/rachelklar/status/97744648621137920>

29. <http://support.twitter.com/articles/18311-the-twitter-rules>

30. De acordo com a Wikipedia, o crowdsourcing é um modelo de produção que utiliza a inteligência e os conhecimentos coletivos e voluntários espalhados pela Internet para resolver problemas, criar conteúdo e soluções ou desenvolver novas tecnologias.

31. http://en.wikipedia.org/wiki/Experimenter%27s_regress

32. <http://tarletongillespie.org/scrutiny/?p=121>

33. <http://www.theatlantic.com/technology/archive/2011/04/the-curious-conne...>

Categoria:

- [poliTICS 11](#)