

PESQUISAS SOCIAIS perguntar ou escutar?



Eleições 2022 | um estudo acadêmico

Trabalho realizado com o uso da plataforma



STILINGUE



INFORMAÇÃO, RESPOSTAS E DECISÕES

Dirigentes de organizações e, particularmente, **profissionais de Marketing e Comunicação**, precisam tomar decisões relevantes diariamente, e **informação** é um ingrediente fundamental para aumentar a probabilidade de **sucesso**.

Em 50 anos, saímos de um mundo, onde era necessário **ir a campo e perguntar** quando queríamos saber algo, para outro, onde boa parte do que precisamos saber pode ser **minerado na internet**, de forma automatizada, com **agilidade** compatível com as necessidades dos **contextos voláteis** e complexos da atualidade.

O objetivo da indústria da pesquisa sempre foi **encontrar as respostas** necessárias para apoiar **decisões relevantes**. Fazer perguntas, observar comportamentos (etnografia), promover discussões em grupo e analisar signos (semiótica), quaisquer que sejam as metodologias e processos aplicados, são algumas das formas tradicionalmente utilizadas para realizar essa tarefa.

UM NOVO UNIVERSO DE PESQUISA

A novidade é dispormos, hoje, de um imenso **repositório de manifestações humanas espontâneas**, permanentemente atualizado (genericamente, a “internet”).

Explorar esse Universo de opiniões, atitudes e interações, disponível para acesso a qualquer momento, cobrindo praticamente todas as áreas de interesse, é uma **oportunidade magnífica** para atender às mais diversas demandas dos decisores.

Mas será que podemos confiar nessas informações?

E até que ponto o que coletarmos nesse espaço será representativo da “sociedade real”?

E AS LIMITAÇÕES ESTATÍSTICAS?

Embora o acesso à internet esteja próximo de ser “universal”, a atividade dos que acessam difere por região, nível educacional, idade, poder aquisitivo e outras características sociodemográficas que podem ser importantes para o fenômeno a ser estudado, e as dimensões dessas divergências não são plenamente conhecidas.

Ou seja, o **grau de controle** oferecido pelas pesquisas quantitativas convencionais não existe nesse ambiente, o que requer uma ruptura com os modelos estatísticos tradicionais e uma lógica analítica mais próxima da “fuzzy” (lógica difusa, como contraposição à booleana).

A questão que se coloca é: **podemos substituir** (ou ao menos complementar) as **perguntas** dirigidas a uma amostra estatisticamente representativa da população **pela escuta** das manifestações caóticas da sociedade?



O DESAFIO DO REAL

Partindo do princípio de que **essas manifestações têm valor**, como utilizá-las com **razoável segurança** para subsidiar decisões?

Para ajudar a responder essa questão, compartilhamos com os associados da **ABA**, **em primeira mão**, um estudo recente, que exemplifica **as possibilidades analíticas** e o grau de confiança com que podemos utilizar a **escuta social** (social listening) para compreender fenômenos e prever resultados.

Trata-se do **acompanhamento do processo eleitoral nas redes sociais**, realizado pelo hub de pesquisa **SocialData** e a empresa de monitoramento social **Stilingue***, que oferece a **rara oportunidade** de compararmos dois caminhos distintos para obtenção de informações e **avaliar sua eficiência** a partir da aferição plena do comportamento da sociedade (o voto).

* A Stilingue não disponibiliza comercialmente sua plataforma para estudos relacionados com eleições. Participou do presente projeto considerando sua finalidade acadêmica e com a condição de que os resultados só fossem publicados após as eleições.

ESTUDO DE CASO | Eleições Presidenciais 2022

Dando sequência a um estudo iniciado durante as eleições de 2018, **SocialData** e **Stilingue** acompanharam as eleições presidenciais de 2022, com o objetivo de verificar se a métrica experimental, concebida quatro anos antes, seguiria tendo o mesmo **caráter preditivo** no novo processo eleitoral.

Eleições são uma “**prova de fogo**” para a instituição da pesquisa. É uma das raras oportunidades em que a pesquisa amostral **pode ser confirmada** (ou não) pela medição do fenômeno em todo o Universo pesquisado.





A DESCOBERTA DE 2018

Pesquisas Políticas costumam ser **contextualmente mais complexas** do que as de Consumo, porque temos mais variáveis a serem controladas e a dinâmica de escolha de um candidato é mais intrincada do que a de um produto de consumo, embora existam situações em que as pesquisas de consumo também enfrentam esse tipo de complexidade (lançamento de novas categorias ou segmentos de produtos mais sensíveis).

Porém, ao longo do estudo de 2018, identificamos uma **forte correlação** entre o **espaço ocupado pelas conversas** que mencionavam os candidatos nas redes sociais e o **resultado das eleições**. Essa dimensão demonstrou um grande poder de “explicação” para o voto. O sentido lógico dessa correlação é o de que seria impossível para o candidato convencer uma pessoa que esteja fora de sua “bolha” (espaço de conversa).

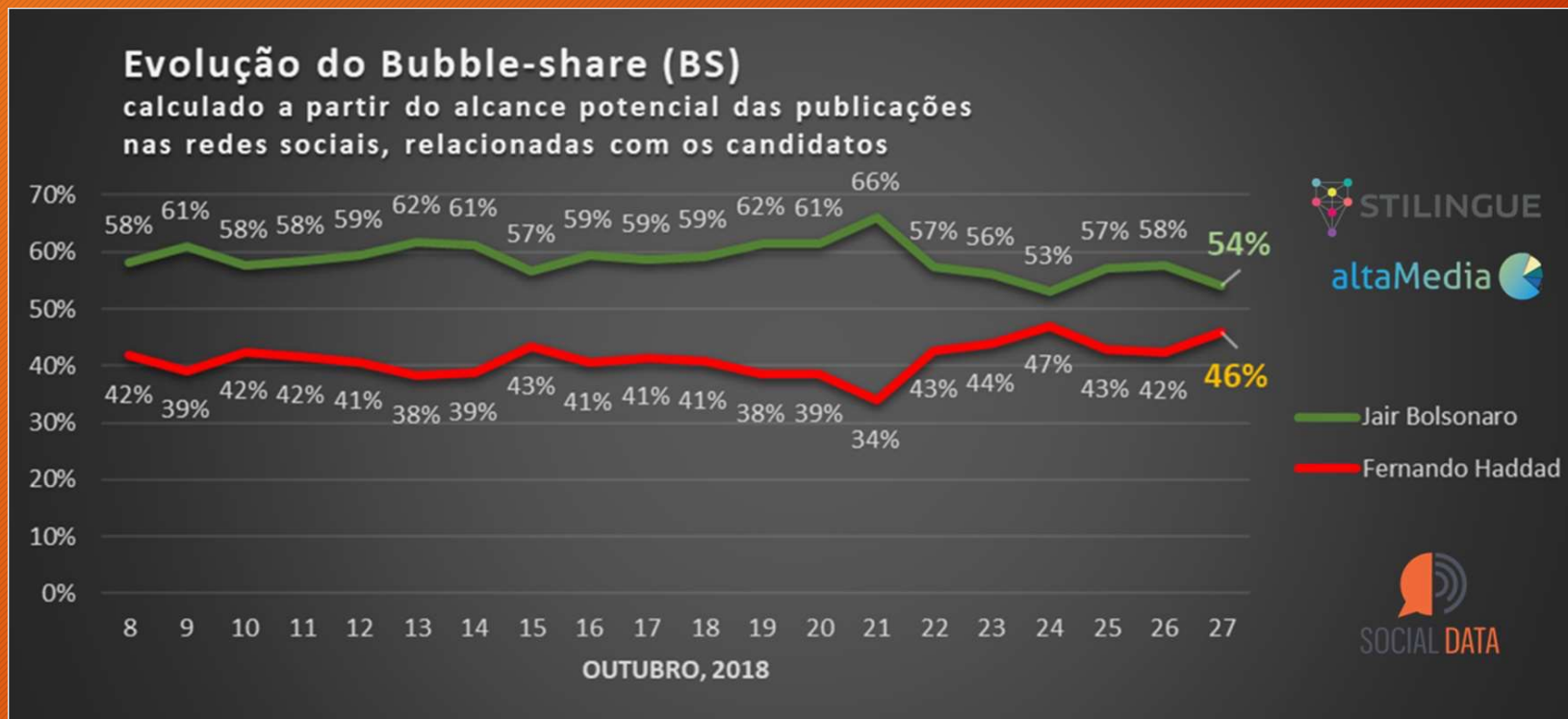
UMA NOVA MÉTRICA | o tamanho das “bolhas”

Partimos, então, de duas hipóteses:

- (1) Se as campanhas dos candidatos fossem igualmente eficientes, seus desempenhos na eleição seriam proporcionais ao tamanho de suas “bolhas”;
- (2) As bolhas das redes sociais poderiam ser consideradas uma “proxy” (representação) dos espaços de conversa no mundo “real”.

Criamos, então, uma **nova métrica**, relativamente simples, a partir dos dados disponíveis no dashboard da Stilingue, para estimar o **tamanho comparativo das bolhas** (não absoluto), batizada de **bubble-share | BS**, cuja configuração foi ajustada durante o primeiro turno e aplicada durante o segundo turno das eleições de 2018.

Vale observar que as características dos dados fornecidos pelas redes sociais e as limitações impostas pela LGPD inviabilizam o cálculo absoluto desse espaço.



Utilizamos como base para cálculo o Alcance Potencial da Stilingue, que resulta somatória do número de publicadores ponderados pela quantidade de seguidores. O BS é calculado dividindo-se o alcance potencial do candidato pela soma dos alcances potenciais de todos os candidatos (base 100). Para Presidente, o cálculo foi diário. Para Governador (SP), utilizamos uma média móvel de 7 dias, em função do menor volume de conversas.

PRESIDENTES 2018	10/out		15/out		18/out		23/out		25/out		27/out			28/out
candidatos	SD	DF	SD	IB	SD	DF	SD	IB	SD	DF	SD	IB	DF	Resultado
Bolsonaro	58	58	57	59	59	59	56	57	57	56	54	54	55	55
Haddad	42	42	43	41	41	41	44	43	43	44	46	46	45	45

Tabela 1: comparativo do bubble-share com os resultados das pesquisas de intenção de votos publicadas pelos institutos e o resultado final das eleições (SD – SocialData | DF – DataFolha | IB – Ibope)

Ao longo do período, comparamos o indicador diário da nova métrica (BS) com as pesquisas publicadas pelos dois principais institutos da ocasião, Ibope (hoje Ipec) e Datafolha, e com o resultado final da eleição.

Considerando as limitações inerentes à coleta de dados e a “simplicidade” da métrica BS, a similaridade dos números foi surpreendente.

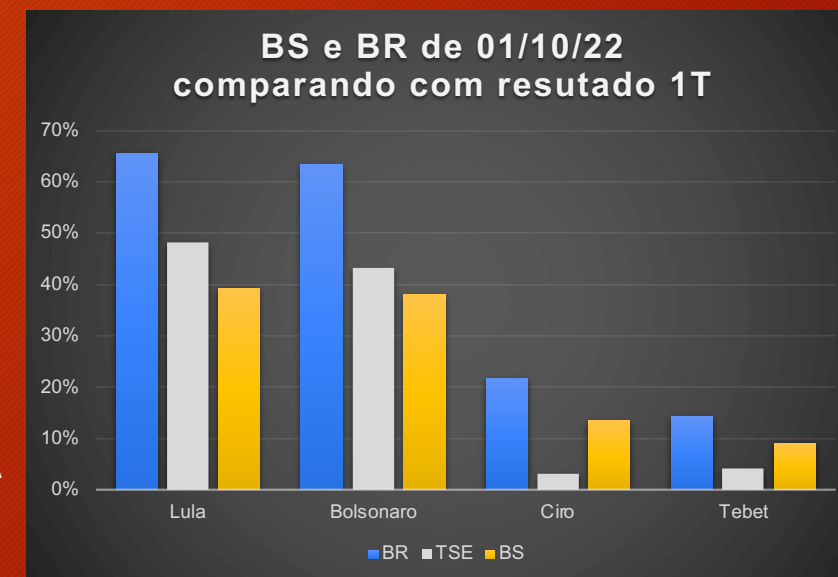
SEGUNDO ESTUDO | 2022 - primeiro turno

Objetivo: avaliar se a métrica (BS) seguiria preditiva, mesmo num contexto mais complexo, mantendo a comparação evolutiva com os mesmos institutos (Ipec e Datafolha)

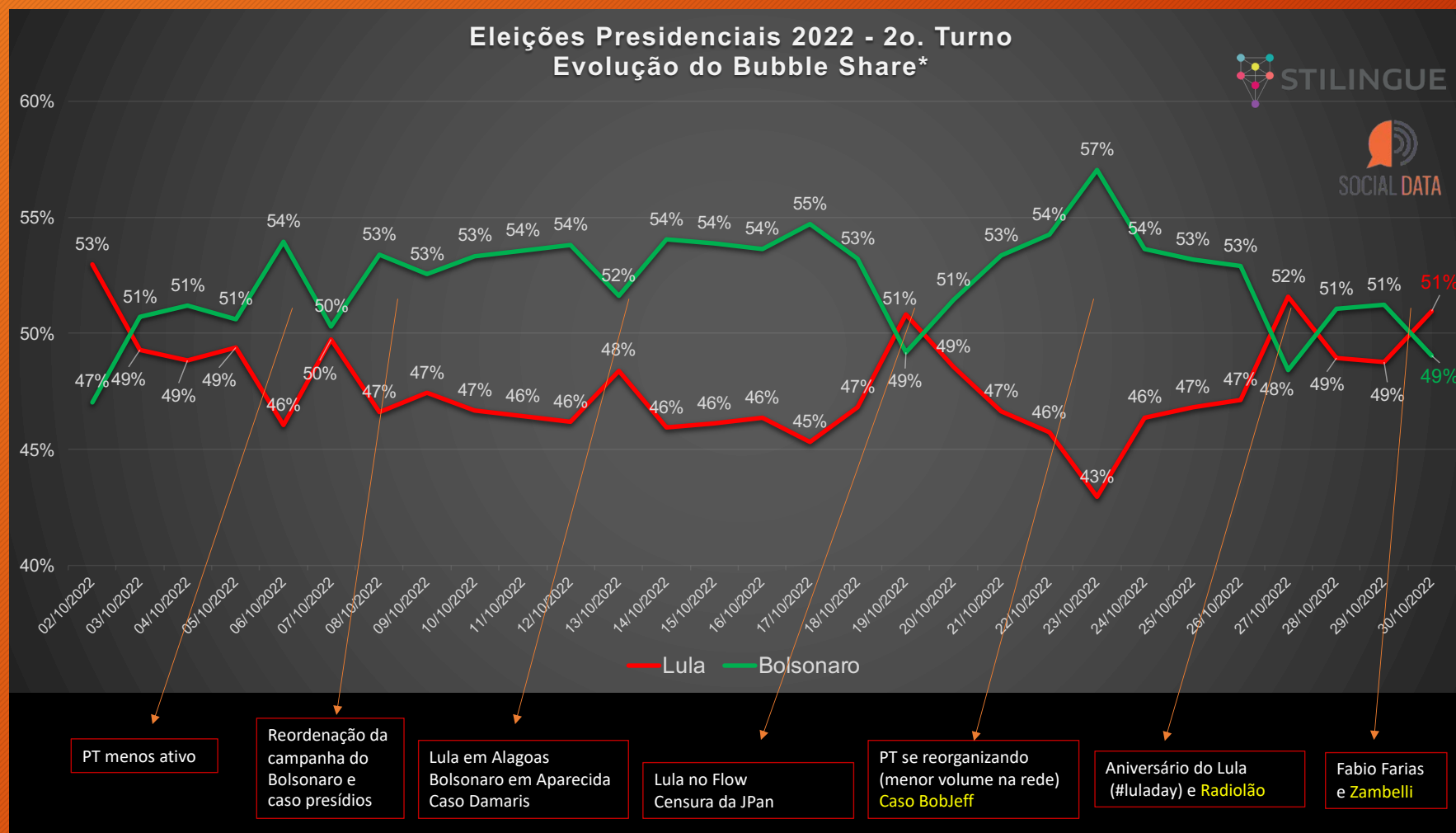
Parceiro:  STILINGUE

Primeiro turno | preocupação: a métrica (BS) indicava uma diferença muito menor entre os primeiros colocados do que as pesquisas de opinião - no dia anterior à eleição diferença de **apenas 2 pontos percentuais** a favor de Lula (15 pontos no caso do Ipec e 16 pontos no DataFolha).

Conceito adicional: no primeiro turno, calculamos os percentuais relativos dos alcances potenciais de cada candidato considerando o alcance potencial total das conversas sobre política (indicador que batizamos de “Bubble-Reach” | BR (com soma maior do que 100%). Esse é um indicador do “teto” de conversão orgânica, ou seja, o percentual que poderia ser alcançado se o candidato fosse capaz de convencer todos os participantes das conversas onde está inserido, sem levar em conta outras movimentações (como o voto útil).



SEGUNDO ESTUDO | 2022 - segundo turno



Bolsonaro à frente na maior parte do tempo e variações com causas identificáveis. Equilíbrio a partir de 27/10 com o aniversário do Lula e entrada dos artistas na campanha, e o caso do Radiolão (mudança de foco da campanha de Bolsonaro). Inversão na noite anterior à eleição com o caso Zambelli.

COMPARATIVO COM IPEC E DATAFOLHA

PRESIDENTE 2022	10/out		14/out		17/out		19/out		24/out		27/out		29/out			30/out	
candidatos	SD	IP	SD	DF	SD	IP	SD	DF	SD	IP	SD	DF	SD	IP	DF	SD	TSE
LULA	47%	55%	48%	53%	46%	54%	47%	52%	43%	54%	47%	53%	49%	54%	52%	51%	51%
BOLSONARO	53%	45%	52%	47%	54%	46%	53%	48%	57%	46%	53%	47%	51%	46%	48%	49%	49%

Tabela 3: Comparativo entre o BS, as pesquisas de intenção de voto e o resultado da eleição presidencial em 2022. SD – SocialData / IP – Ipec / DF - DataFolha

O aniversário de Lula (com a entrada de artistas na campanha) e os episódios de Roberto Jefferson, Radiolão e Carla Zambelli não impactam o resultado das pesquisas de declaração de intenção de voto. O BS indica o crescimento de Lula até o início da apuração.

O BS no dia da eleição foi calculado até as 17h00, para não incluir o período de apuração e as celebrações, e **correspondeu exatamente ao resultado da votação.**

CONCLUSÕES



- A maior **sensibilidade da métrica BS** a eventos relevantes, e o fato de corresponder ao próprio resultado no dia da votação, nos anima a credenciá-la como uma **boa métrica** para acompanhamento de cenários e complementação das pesquisas de opinião
- Os dois experimentos (2018 e 2022) demonstram o potencial do acompanhamento das redes sociais (Social Listening) para entendimento e previsão de fenômenos sociais de grande mobilização
- Também ficou claro o **benefício de contar com o dashboard** da Stilingue, utilizando Inteligência Artificial (IA), Processamento de Linguagem Natural (PLN) e Machine Learning, aplicados aos milhões de conteúdos coletados diariamente, mantendo as métricas atualizadas em tempo real e **classificando os conteúdos**.
- Vale ressaltar que a métrica que escolhemos (BS), considerada sua simplicidade, foi robusta em sua correlação com os resultados da eleição, inspirando-nos a considerar metodologias alternativas e intuitivas para a análise de dados difusos (lógicas “fuzzy”).

ESCUTA ATIVA



As **pesquisas “tradicionais”** são indispensáveis para aprofundar e detalhar investigações, superando muitas das limitações de coleta dos dados das redes, e **independem das limitações determinadas** pelos provedores (redes sociais) para as informações coletadas aos agregadores (empresas que integram e organizam as coletas).

Entretanto, considerando o **custo** do “social listening” (proporcionalmente menor), a **agilidade** na apresentação dos dados (em tempo real) e a **sensibilidade** a eventos críticos, nos parece um **complemento indispensável** para quem se dispõem a analisar **fenômenos sociais relevantes**, particularmente quando decisões precisam ser tomadas com **rapidez em cenários voláteis**.

Acreditamos que a “**escuta social**” (social listening) possa complementar e/ou **substituir as “perguntas”** de diversas pesquisas de mercado, oferecendo insights dinâmicos (escuta ativa), e planejamos explorar essas possibilidades em **novos projetos** para organizações e suas marcas.



Fundada em 2014, a **STILINGUE** é líder nacional de monitoramento de redes sociais, imprensa digital e atendimento ao cliente, utilizando Inteligência Artificial (IA), Processamento de Linguagem Natural (PLN) e Machine Learning, aplicados a 50 milhões de conteúdos coletados diariamente.

www.stilingue.com.br

Em tempo real, organiza as conversas derivadas das redes sociais, avaliações em sites e aplicativos, notícias, CRM e demais canais públicos e privados, para extração de relatórios e geração de insights, facilitando a condução das estratégias de social media e comunidades, atendimento ao consumidor e inteligência de mercado.



SocialData é o hub colaborativo que desenvolve e integra soluções para atender demandas de gestão de informação, particularmente as endereçadas ao conhecimento da sociedade, para entendimento (e previsibilidade) das ideias, motivações, atitudes e comportamentos das pessoas e seu impacto em situações específicas.

www.socialdata.com.br

Flavio Ferrari é professor da ESPM (módulos de Análise Estratégica de Cenários Futuros e Pesquisa de Mídia), criador do hub SocialData.

Dirigiu, como CEO ou MD, os três maiores institutos de pesquisa de mídia da América Latina