



Laboratório de
Estudos de Internet
e Redes Sociais



UFRJ
UNIVERSIDADE FEDERAL
DO RIO DE JANEIRO



NOVEMBRO 2024

ÍNDICE DE *TRANSPARÊNCIA* DE DADOS DAS *PLATAFORMAS DE* *REDES SOCIAIS*

RELATÓRIO DO OBSERVATÓRIO DA INDÚSTRIA DA
DESINFORMAÇÃO E SEU IMPACTO NAS RELAÇÕES
DE CONSUMO NO BRASIL

Equipe

Direção

R. Marie Santini

Coordenação de pesquisa

Débora Salles

Pesquisadores

Bruno Mattos
Marcela Canavarro
Carlos Eduardo Barros
Alékis Moreira

Pesquisadores Associados

Priscila Medeiros
Felipe Grael
Fernando Ferreira
Danielle Melo
Marcio Borges

Assistentes de Pesquisa

João Gabriel Haddad
Lucas Murakami
Daphne Silva

Equipe Técnica

Erick Dau
Felipe Loureiro

Revisão Técnica

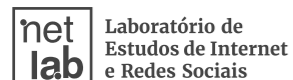
Adriano Belisário

Sumário

Apresentação	4
Sumário Executivo	5
A Importância da Pesquisa com Dados Digitais	9
Apagão de Dados Digitais	13
Por Que Medir a Transparência a Partir da Qualidade dos Dados?	18
Abordagem Metodológica	24
Resultados	33
YouTube	34
Facebook	35
Instagram	36
X/Twitter	37
Telegram	38
TikTok	39
Kwai	40
WhatsApp	41
Boas e Más Práticas na Disponibilização de Dados	42
Recomendações	45
Apêndice	47
Visão Geral da Avaliação	47
Detalhamento por Parâmetro de Avaliação	53
Referências Bibliográficas	97

Como citar

SANTINI, R. Marie; SALLES, Débora; MATTOS, Bruno; CANAVARRO, Marcela; BARROS, Carlos E.; MOREIRA, Alékis; MEDEIROS, Priscila; GRAEL, Felipe; FERREIRA, Fernando; MELO, Danielle; BORGES, Marcio; HADDAD, João G.; MURAKAMI, Lucas; SILVA, Daphne; DAU, Erick; LOUREIRO, Felipe. Índice de Transparência de Dados das Plataformas de Redes Sociais. Rio de Janeiro: NetLab – Laboratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). Publicado em 04 de novembro de 2024.



Escola de Comunicação | Universidade
Federal do Rio de Janeiro
Av. Pasteur, 250, Urca | Rio de Janeiro - RJ
CEP 21941-901

 WWW.NETLAB.ECO.UFRJ.BR

 NETLAB@ECO.UFRJ.BR

© NetLab UFRJ 2024

Apresentação

Observatório da **Indústria da Desinformação** e seu impacto nas **relações de consumo** no Brasil

Este relatório é um dos resultados das pesquisas desenvolvidas no âmbito do *Observatório da indústria da desinformação e seu impacto nas relações de consumo no Brasil*, projeto em parceria entre o **NetLab UFRJ** e a **Secretaria Nacional do Consumidor do Ministério da Justiça e Segurança Pública** (Senacon/MJSP).

O principal objetivo do Observatório é prover insumos que embasem **políticas públicas de proteção dos consumidores** a partir de análises sobre infraestrutura, economia política e as estratégias de manipulação das relações de consumo e da opinião pública das **operações de desinformação e influência que vêm se desenvolvendo nas plataformas de redes sociais**.

Diante da **escassez de informação qualificada** para pesquisas sociais aplicadas baseadas em dados digitais, apresentamos aqui o **Índice de Transparência de Dados das Plataformas de Redes Sociais (ITD)** no Brasil.

O **ITD** segue um roteiro estruturado, sistematizado e reproduzível, baseado em critérios de qualidade de dados, para avaliar **mecanismos de acesso a dados de interesse público** disponibilizados pelas plataformas.

Ele avalia o nível de transparência e de qualidade dos dados das principais redes sociais e aplicativos de mensageria que atuam no Brasil: YouTube, Facebook, Instagram, X/Twitter, Telegram, TikTok, Kwai e WhatsApp. A avaliação se baseia no acesso a dados de **conteúdos gerados por usuários, publicados sem impulso pago às plataformas**.

A partir das notas obtidas, as plataformas são classificadas em **cinco níveis**: Transparência **irrelevante ou nula** (0 a 20), Transparência **precaria** (21 a 40), Transparência **regular** (41 a 60), Transparência **satisfatória** (61 a 80) e Transparência **ideal** (81 a 100).

O **ITD** é parte de um estudo mais amplo sobre a transparência das plataformas de redes sociais no acesso a dados de interesse público. A pesquisa também inclui o **ITP — Índice de Transparência da Publicidade nas Plataformas de Redes Sociais**, que avalia diferentes parâmetros e dimensões de qualidade na disponibilização de dados sobre anúncios, ou seja, quaisquer publicações impulsionadas mediante pagamento às plataformas.

Sumário Executivo

Principais Resultados

Nenhuma plataforma avaliada atingiu o nível *ideal* na 1ª edição do Índice de Transparência de Dados das Plataformas de Redes Sociais (ITD).

O YouTube obteve a melhor pontuação, com 63,2 pontos, o que é considerado *satisfatório*. Com índice *regular*, Facebook e Instagram vêm em seguida (53,6 e 52,1 pontos, respectivamente), com boa parte da nota proveniente dos recursos da ferramenta CrowdTangle, que foi descontinuada pela Meta.

Universo inacessível: a maioria das plataformas não oferece API gratuita nem interface de coleta de dados.

Ao impedir o mapeamento e a recuperação de dados de forma sistemática e sem custos, as plataformas impõem barreiras técnicas e econômicas para a pesquisa, que resultam no enviesamento das análises realizadas com dados não representativos dos objetos de estudo, impactando também na reproduzibilidade das pesquisas.

Entre as más práticas recorrentes, destacam-se a limitação a um baixo volume diário de coleta e a entrega de resultados incoerentes em cada requisição.

Mesmo quando oferecem meios oficiais de coleta, persistem restrições que dificultam otimizar o trabalho de extração de dados de forma programática e em escala, com frequentes erros no processo de recuperação, número insuficiente de resultados e limitações para criar novos *tokens* de acesso à API, dificultando coletas simultâneas.

Prática discriminatória: plataformas como X/Twitter e TikTok oferecem bases de dados com diferentes níveis de qualidade e completude entre países.

Diferenças relevantes na transparência de dados de interesse público entre o Brasil e países da Europa e os EUA sugerem que é preciso estabelecer critérios mínimos satisfatórios de acesso e qualidade de dados, de forma justa e igualitária em todos os países onde tais empresas atuam.

Com o CrowdTangle, Facebook e Instagram ofereciam acesso *regular* aos dados de interesse público. Com a ferramenta descontinuada, as notas tendem a cair.

O CrowdTangle foi descontinuado pela Meta em agosto de 2024, mês em que começou a campanha eleitoral no Brasil. Mesmo com problemas de completude e consistência, sem a ferramenta, o cenário para a pesquisa tem piorado. Restrições da Meta à raspagem de dados agravam o problema da falta de transparência.

X/Twitter, Telegram, Kwai e WhatsApp não disponibilizam relatórios de transparência de ações de moderação e governança específicos sobre o contexto brasileiro.

Apesar do Kwai apresentar relatórios de transparência semestrais sobre ações de moderação, não disponibiliza dados específicos sobre o Brasil. Desde 2021, o X/Twitter não divulga relatórios de transparência para o contexto brasileiro, se juntando ao Telegram e WhatsApp que também não divulgam nenhum relatório desse tipo.

TikTok não disponibiliza ferramentas de coleta e acesso a dados no Brasil, mas se destaca na transparência de ações de moderação e governança.

O TikTok é a única plataforma analisada que disponibiliza informações sobre o volume e tipos de violações infringidas e moderadas pela plataforma, assim como os pedidos de moderação e de requisição de dados feitos pelo Estado brasileiro. No entanto, não oferece informações sobre metadados dos conteúdos e usuários moderados, impossibilitando a auditoria dessas ações.

A granularidade dos dados de relatórios de transparência é *insatisfatória*, dificultando a análise por país e a identificação dos tipos de violação de conteúdos removidos.

Uma vez que as plataformas não disponibilizam dados detalhados sobre suas ações de moderação, como metadados de casos concretos de publicações removidas, é impossível avaliar a aplicação de suas políticas de governança

Recomendações Prioritárias

Possibilidade de coleta do universo de dados públicos

- Recomenda-se a **disponibilização de API** oficial e gratuita para **acesso integral, programático e gratuito** ao universo de dados públicos para pesquisadores.
- Além disso, recomendamos a oferta de uma **interface** para coleta de dados, com fácil uso e navegação e sem exigência de conhecimentos técnicos em programação.
- Um outro fator a ser melhorado em todas as plataformas é a **permissão da raspagem de dados** para fins de **pesquisa**, oferecendo mais segurança jurídica para pesquisadores e facilitando o processo de coleta sistemática de dados.

Qualidade dos dados disponibilizados

- É preciso franquear acesso a dados de **conteúdos temporários** e a **dados históricos** aos pesquisadores, além de disponibilizar **links que não expirem** para possibilitar a análise de conteúdos permanentes em foto e vídeo.
- Também é importante que as plataformas tenham transparência sobre quando **publicações são removidas** e **usuários suspensos**, oferecendo acesso a seus **metadados**, ainda que o conteúdo das publicações fiquem restritos.
- Oferecer **busca customizada por palavras-chave** é fundamental para a coleta de dados relevantes para fins de pesquisa.

Clareza da documentação sobre API e termos de uso

- É importante que a **documentação da API** seja disponibilizada **publicamente**, sem necessidade de requisição individual, em **língua portuguesa**, e com descrição clara de seus **termos de uso**.
- A documentação oficial da API deve listar os **possíveis erros** de cada **endpoint** disponível e oferecer **exemplos claros e compreensíveis** de como fazer as requisições para obter os dados.



- Os **relatórios de transparência** relacionados à moderação de conteúdo devem ser publicados de forma periódica e com garantia da qualidade dos dados. Para possibilitar análises relevantes e aprofundadas das polí-

ticas de moderação e governança de todas as plataformas no país, os dados devem ser disponibilizados com maior detalhamento e com foco específico no Brasil.

A Importância da Pesquisa com Dados Digitais

Diferentes razões levam ao potencial dos dados digitais de não apenas identificar, descrever e compreender fenômenos e problemas sociais relevantes, mas também de ajudar a **lidar com eles ou até resolvê-los**. Aspectos como **volume** de dados, **granularidade** da amostra, **atualidade** e **não intrusividade** na coleta desses dados permitem às Ciências Sociais Computacionais — assim como suas correlatas em áreas como Biomedicina ([Correia et al., 2020](#)), Economia ([Tumasjan, 2023](#)), Saúde Pública ([Terry et al., 2023](#)), Biodiversidade ([Chowdhury, 2024](#)) e Meio Ambiente ([Ghermandi et al., 2023](#)), entre outras — realizar pesquisas com suficiente relevância e conexão com a realidade social. Assim, são capazes de **informar políticas públicas e de governança** que terão impacto no desenvolvimento das próprias tecnologias e no tipo de espaços de sociabilidade que oferecem.

Um aspecto-chave para compreender a relevância dos dados de redes sociais para a pesquisa é sua crescente influência na vida pública, caracterizando-se, hoje, como uma das principais arenas de debate público ([Staab; Thiel, 2022](#); [Yasseri, 2023](#)). Isso envolve não apenas temáticas políticas e sociais que circulam organicamente, mas também campanhas planejadas para influenciar gostos, hábitos de consumo, modos de vida, opiniões e comportamentos.

No entanto, cidadãos comuns estão sujeitos ao uso de seus dados para uma infinidade de ações comerciais que atualmente são difíceis de auditar, embora impactem tanto na vida individual e coletiva. Quando o usuário aceita um contrato de adesão para utilizar a plataforma, concordando com seus termos de uso, cria-se um vínculo comercial e uma **relação de consumo** que, como todas as outras relações desse tipo, deve oferecer meios para a **proteção do consumidor**. A pesqui-

sa acadêmica tem o papel de reduzir a **assimetria de informações** que tende a ser prejudicial ao consumidor: a plataforma sabe muito sobre o usuário, mas o usuário pouco sabe sobre o que é feito com os seus dados extraídos e minerados pela plataforma.

Este é, portanto, o ganho mais abrangente da pesquisa com dados digitais: a possibilidade de **mensurar o impacto das próprias tecnologias** na vida social, já que boa parte do recente desenvolvimento dos métodos computacionais “é voltado para analisar a estrutura e dinâmica da comunicação humana” ([Van Atteveldt; Peng, 2018, p. 81](#)). Dados nativamente digitais não são produzidos especificamente para a pesquisa, mas sim como um resultado de interações e processos sociais que acontecem online ([Edwards et al., 2013](#); [Lee et al., 2008](#); [Marres, 2017](#)). Isto é, os dados são produzidos e coletados sem a interferência dos pesquisadores sobre o objeto de estudo, a partir de uma **abordagem não obstrutiva** ([Rogers, 2009](#)).

Para analisar esses dados, pesquisadores devem conhecer a fundo o **funcionamento** do ambiente digital de onde foram extraídos, já que os dados “tendem a ser fortemente marcados pelos **efeitos da plataforma**, como a busca por termos sugeridos em funções de autocompletar ou a utilização de hashtags que estão em alta na plataforma em questão” ([Shaw, 2015, p. 2](#)). Essa é uma das razões pelas quais os dados digitais trouxeram, junto aos novos tipos de informações a serem analisadas, também novos desafios de pesquisa. Emergiram **novos métodos** para explorá-los e para compreender diferentes características das plataformas digitais e, em particular, das plataformas de redes sociais, que moldam as dinâmicas sociais nesses ambientes ([Rogers, 2009](#)). Esses novos métodos trouxeram

também desafios para os pesquisadores, que devem aplicá-los com foco no interesse público e nos problemas sociais contemporâneos.

Os métodos digitais, por serem voltados para o processamento e a descoberta de informações relevantes em meio a quantidades massivas de dados, também têm sido amplamente utilizados em diversas áreas do conhecimento. Esse é o aspecto que coloca as **Ciências Sociais Computacionais** como uma disciplina científica que também é **instrumental** para outras áreas

([Cioffi-Revilla, 2018](#)). Estudos epidemiológicos e de farmacovigilância, por exemplo, utilizam métodos digitais e computacionais há anos ([El-Sayed et al., 2012](#); [Pappa; Stergioulas, 2019](#)), abordagens que foram úteis durante a pandemia de Covid-19, no início dos anos 2020, em aspectos como controle de aglomerações ([Cecilia et al., 2020](#)), avaliação de sofrimentos psicológicos ([Dhelim et al., 2023](#)), predição da evolução da doença, visualização de dados e definição de prognósticos ([Senthilraja, 2021](#)).

soais, de acordo com a LGPD, parte da natureza do órgão que realiza a pesquisa — compreensão que também é adotada aqui para definirmos **pesquisas de interesse público**. Assim, entendemos que devem ter acesso a dados públicos das plataformas de redes sociais o “órgão ou entidade da administração pública direta ou indireta ou pessoa jurídica de direito privado sem fins lucrativos legalmente constituída sob as leis brasileiras, com sede e foro no País, que **inclua em sua missão institucional ou em seu objeto social ou estatutário a pesquisa básica ou aplicada de caráter histórico, científico, tecnológico ou estatístico**” ([ANPD, \[S.d.\], p. 26](#)). Nesse conceito, são abrangidas as universidades públicas e privadas sem fins lucrativos, fundações e institutos dedicados à pesquisa e órgãos públicos como Ipea, IBGE e Fiocruz, por exemplo.

Dados de conteúdos públicos, abertos e acessíveis por qualquer usuário de plataformas de redes sociais devem estar disponíveis para fins de **pesquisa, análise e auditoria de interesse público**. Além disso, a LGPD estipula que dados tornados manifestamente públicos pelo seu titular **não requerem consentimento para uso**, definição que viabiliza, no âmbito da finalidade deste índice, a pesquisa com dados originados em espaços de opinião pública com potencial de impactar inúmeros cidadãos. Assim, podem ter acesso a dados públicos das plataformas de redes sociais o órgão ou entidade da administração pública sem fins lucrativos ou pessoa jurídica legalmente constituída e com sede no Brasil, que tenham sua **missão na pesquisa “de caráter histórico, científico, tecnológico ou estatístico”** ([Brasil, 2018](#)).

Também são abarcados nesta definição os grupos e canais públicos no WhatsApp e no Telegram, quando **divulgados na internet**, para captar novos membros e nos quais qualquer usuário pode entrar ([Evangelista; Bruno, 2019](#); [Garimella; Tyson, 2018](#); [Resende et al., 2019](#)). Visto que qualquer pessoa com o link de acesso ao grupo pode se afiliar, pesquisadores estabeleceram critérios éticos sobre o que pode ser considerado público nesses aplicativos. Listas online que elencam os grupos públicos e ferramentas de busca que possibilitam encontrá-los também têm sido úteis para a identificação desse tipo de grupo por pesquisadores ([Melo, 2022](#); [Garimella, Tyson, 2018](#); [Resende et al., 2019](#)).

Na primeira década dos anos 2000, a pesquisa com aplicativos de mensageria se limitava, em grande parte, a amostras reduzidas e metodologias qualitativas, como entrevistas e *surveys* com participantes de grupos, por exemplo ([Garimella; Tyson, 2018](#); [Rosenfeld et al., 2018](#)). Porém, a ampla utilização de aplicativos de mensageria na organização e no fluxo de mobilizações políticas tem aumentado o interesse no estudo de grandes volumes de dados de grupos públicos do WhatsApp e do Telegram ([Moura; Michelson, 2014](#); [Treré, 2020](#)). Por serem importantes vias para a propagação de desinformação política, esses aplicativos ganham cada vez mais relevância para a pesquisa social ([Evangelista; Bruno, 2019](#); [Ozawa et al., 2023](#); [Melo et al., 2019](#); [Wendratama; Yusuf, 2023](#); [Calvo-Gutiérrez; Marín-Lladó, 2023](#); [Smith et al., 2023](#)).

Por que a pesquisa com dados digitais é importante?

- Possibilita mensurar o impacto das próprias tecnologias na vida social;
- Permite analisar comportamentos de forma não-observada, em complemento a métodos tradicionais que dependem da autodeclaração dos participantes;
- Facilita o trabalho colaborativo, já que torna mais fácil compartilhar e reutilizar dados;
- Fundamenta a formulação e atualização de políticas públicas com base em um novo tipo de evidência;
- Aumenta a escala de um universo amostral relativamente pequeno para possibilidades de grandes amostras de dados, facilitando o estudo de relações ou efeitos sutis e de nicho;
- Possibilita construir estudos longitudinais de forma mais refinada, a partir da variedade e granularidade das amostras;
- Facilita a observação do comportamento das pessoas em ambientes reais, ao invés da simulação em laboratório.

FONTES: Van Atteveldt e Peng (2018) e Marres (2017)

Um dos objetivos que emergem como mais urgentes, nesse contexto, é a pesquisa que ajude a produzir diagnósticos e panoramas, coletar evidências e indicar soluções para fenômenos atuais relevantes, como a desinformação, o uso de tecnologias por crianças e adolescentes, os crimes cibernéticos, o impacto das redes sociais na saúde mental e os novos hábitos de consumo, entre outros. Estes e outros aspectos afetam a sociedade brasileira de formas diversas e demandam compreensão profunda, que pode ser oferecida por pesquisas de qualidade.

O ITD tem como objetivo contribuir para a pesquisa com dados digitais, em especial nas Ciências Sociais Aplicadas. Sua missão é melhorar o acesso gratuito e universal a quaisquer dados públicos nas plataformas de redes sociais, que sejam úteis para **pesquisas acadêmicas de interesse público**.

Embora o conceito de “interesse público” seja abrangente, o entendimento da Autoridade Nacional de Proteção de Dados (ANPD) para aplicar condições especiais no tratamento de dados pes-

Dados digitais públicos

- Conteúdos com visibilidade definida como pública pelos seus autores;
- Conteúdos temporários com visibilidade pública enquanto estiverem no ar;
- Mensagens enviadas em grupos e canais do WhatsApp e Telegram divulgados em espaços abertos da internet.

Dados digitais não públicos

- Publicações e vídeos com visibilidade definida como *privada* pelos seus autores;
- Vídeos não listados ou privados no YouTube;
- Mensagens de bate-papo do Messenger, Instagram e X/Twitter;
- Conteúdos feitos por contas trancadas;
- Conversas privadas entre dois ou mais usuários, em mensagens um a um ou em grupos privados, no WhatsApp e no Telegram.

Apagão de Dados Digitais

Embora plataformas de redes sociais venham **ganhando relevância** como espaços de formação de **opinião pública**, a transparência e o conhecimento público sobre seu funcionamento não são proporcionais ao seu potencial de impacto na vida social. Além da **opacidade** dos algoritmos (Lu, 2021) e da dificuldade de acesso aos dados (Bruns, 2019), as plataformas de redes sociais também **mudam constantemente** as regras de funcionamento **embutidas nos algoritmos**, nem sempre de forma transparente, gerando importantes preocupações éticas e políticas (Selinger; Hartzog, 2016; Guess et al., 2023).

Os dados digitais são vitais para compreender o modo como as plataformas online impactam dinâmicas políticas e sociais, mas o acesso a dados **completos, atualizados, consistentes, disponíveis no tempo necessário, em conformidade legal e relevantes** para a pesquisa é cada vez mais **limitado, dificultado** ou **mesmo extinto pelas big tech** (Greene; Martens; Schmueli, 2022). Ao analisar o cenário do Facebook, Instagram e X/Twitter, Bruns (2019) considera que as *big tech* procuram limitar o desenvolvimento de pesquisas críticas baseadas em dados sobre as plataformas. Ao mesmo tempo, lançam ferramentas para acesso a dados incompletos e inconsistentes para gerar publicidade positiva baseada em falsa imagem de transparência ativa e capacidade de autorregulação

Entre os exemplos mais emblemáticos, está o **encerramento do CrowdTangle** pela Meta (Soares, 2024), em agosto de 2024; a ferramenta garantia acesso a dados do **Facebook** e **Instagram**. Esse movimento acompanha a mudança do **X/Twitter** que, desde março de 2023, cobra um **alto valor** pelo acesso à sua API, que até então era gratuita (Mozelli, 2023). A tendência, entretanto, não é restrita aos últimos anos; logo após o escândalo da Cambridge Analytica, em

2018, o então **Facebook** modificou e restringiu funcionalidades de sua API sem aviso prévio, prejudicando o desenvolvimento de pesquisas acadêmicas (Bruns, 2019).

Diante desse cenário, pesquisadores vinculados a diversos institutos e universidades no mundo tecem críticas frequentes às novas restrições impostas ao acesso a APIs de plataformas de redes sociais, alegando que as decisões podem **inviabilizar o desenvolvimento de pesquisas** (Coalition for Independent Technology research, 2023; Mozilla Foundation, 2024). No entanto, a falta de critérios vinculativos sobre o tema motiva as *big tech* a ignorar esse tipo de solicitação recorrentemente.

O domínio das plataformas digitais na distribuição de conteúdo permitiu que elas estabelecessem políticas e discursos contundentes na defesa de sua posição dominante no mercado e influenciassem ativamente na forma como são vistas e entendidas pela opinião pública. Ao mesmo tempo que as plataformas de redes sociais buscam ser percebidas como facilitadoras da produção e distribuição de conteúdos gerados por usuários, elas tentam se afastar das responsabilidades sobre o que esses usuários publicam e da percepção de que realizam uma curadoria algorítmica de conteúdo (Gillespie, 2010). Frequentemente, diminuir o acesso aos dados é parte desses esforços. Bossetta (2020) lembra que, apesar de às vezes adotarem medidas que podem contribuir para o incremento da transparência na esfera online, as operações das plataformas são motivadas por interesses econômicos. Elas também têm as condições materiais para rapidamente alterar suas instâncias organizacionais e arquitetura do *software* diante de escândalos, remediando crises reputacionais em que se envolvem.

“A informação que as plataformas decidem incluir — e, mais importante, excluir — [...] podem ser vistas como um posicionamento estratégico imposto por meio do design”

(Bossetta, 2020, p. 2).

Apesar da grande influência das plataformas de redes sociais sobre a vida social e política desde os anos 2010 (Zuboff, 2021; van Dijck; Nieborg; Poell, 2019; Gerbaudo, 2021), elas ainda não têm obrigação legal sobre como seus dados devem ser disponibilizados às diferentes partes interessadas, como pesquisadores, por exemplo. A dificuldade no acesso e/ou utilização de dados de plataformas de redes sociais é, atualmente, um grande empecilho para pesquisadores, especialmente no Sul Global, que recebem dados com menos qualidade e ainda mais incompletos que suas contrapartes do Norte Global (Orembo; Berger; Simon, 2023; Lurie, 2023).

A prática discriminatória de oferecer bases de dados com diferentes níveis de qualidade e completude de acordo com a localização geográfica sugere que não basta criar mecanismos regulatórios para obrigar as *big tech* a oferecer acesso a seus dados para escrutínio público de pesquisadores — é preciso vincular critérios satisfatórios de acesso e qualidade a essas bases de dados. Por exemplo, estudos sobre a aplicação da Lei de Acesso à Informação (LAI) no Brasil (2011) apontam que a obrigação legal de fornecer informação de interesse público ou individual nem sempre é suficiente (Transparência Brasil, 2018; Braga; Cunha, 2022; Santos, 2023), se não for acompanhada de critérios objetivos de qualidade de dados. Padronizar critérios aceitáveis de qualidade, transparência e portabilidade é importante para garantir previsibilidade a pesquisadores sobre a longevidade da qualidade dos dados.

Diante desse cenário, o índice aqui proposto parte do princípio de que é urgente reduzir a assimetria entre países no acesso aos dados que

são produzidos nas plataformas de redes sociais. Diferenças nacionais entre o acesso e a qualidade dos dados disponibilizados no Brasil e em países como EUA e as nações do bloco europeu foram consideradas nas análises dos resultados, além de ajudarem a embasar o desenvolvimento do formulário de avaliação. Detectamos que parte considerável dos parâmetros que não são atendidos pelas plataformas na operação brasileira são disponibilizados no cenário europeu por estas mesmas plataformas, evidenciando que a baixa qualidade das bases disponibilizadas aqui é uma decisão política e não técnica.

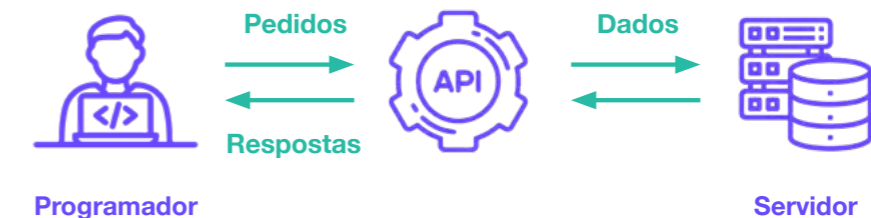
Assim, em paralelo ao esforço para garantir e ampliar o acesso a dados em plataformas de redes sociais para pesquisa acadêmica, a definição de indicadores de qualidade é fundamental para mapear e avaliar as medidas de transparência oferecidas por elas. Dessa forma, é possível fornecer evidências sobre dados incompletos, inconsistentes e de baixa qualidade que impedem ou dificultam a análise sistemática, o diagnóstico de riscos e a responsabilização sobre possíveis danos. A padronização de critérios de qualidade e transparência também estabelece parâmetros comuns que possibilitam comparação e acompanhamento sistemáticos de diferentes plataformas.

Glossário

API: acesso oficial a dados de forma programática

API é a sigla em inglês para Interface de Programação de Aplicações. Sua funcionalidade básica é estabelecer a comunicação entre dois componentes de software (Goodwin, 2024) — como um banco de dados e o dispositivo de um pesquisador —, mediante pedidos de dados que devem cumprir definições e protocolos próprios. Assim, as APIs possibilitam o compartilhamento de dados (PostMan, [S.d.]) entre aplicações, sistemas, dispositivos e plataformas de diferentes naturezas, facilitando a interoperabilidade entre sistemas diversos.

Para acessar uma API, é preciso gerar *tokens*, que são pequenos códigos entregues a cada programador ou pesquisador cadastrado e que funcionam como senhas para autenticar e validar as requisições. Outros componentes específicos da API incluem os *endpoints*, com os quais o programador realiza um pedido ao servidor e recebe a resposta (CloudFlare, [S.d.]), de acordo com as instruções indicadas em sua documentação.



Interface de coleta de dados: acesso oficial a dados

Refere-se à aplicação que permite a visualização, consulta, exploração e o download das publicações geradas por usuários, disponibilizando dados de maneira acessível e com soluções específicas. A interface de coleta de dados deve permitir a busca de publicações a partir de termos de busca específicos ou publicadas por páginas de seu interesse, gerar sumarizações e visualizações, e exportá-las em formato adequado para análises mais profundas.

Diferente da API, a interface de coleta de dados facilita o acesso ao ecossistema de publicações por não exigir conhecimento em programação, mas, em contrapartida, não permite a automatização dos processos de coleta de dados e monitoramento sistemático de posts.



Interface de usuário da plataforma X Interface de coleta de dados

Uma interface é um ambiente online que permite a interação entre um usuário e um banco de dados por meio de elementos gráficos, como ícones, janelas, menus e outros indicadores visuais. No ITD, nos referimos a dois tipos de interface.

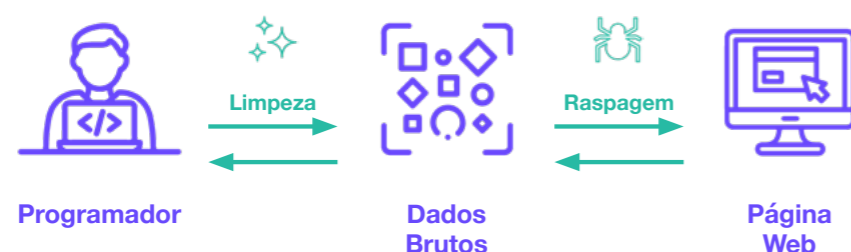
- **Interface de usuário da plataforma:** ambiente da plataforma acessado por usuários para visualizar conteúdos, interagir com outros usuários e manter seus respectivos perfis, seja através de um site ou de um aplicativo.

- **Interface de coleta de dados:** ambiente destinado a pesquisadores e outras partes interessadas em explorar e monitorar consistentemente discussões em torno de temáticas relevantes em plataformas de redes sociais com o objetivo de entender fenômenos sociais e tendências de mercado. Refere-se à aplicação que permite a visualização, consulta, exploração e o download dos dados das publicações geradas por usuários, disponibilizando-os de maneira acessível e com soluções específicas. Idealmente, a interface de coleta de dados deve permitir a busca de publicações a partir de termos de busca específicos.

Técnicas de raspagem: acesso extraoficial a dados de forma programática

A raspagem de dados (*web scraping*) é um processo de extração e combinação de conteúdo online (Bar-Ilan, 2001; Mooney; Westreich; El-Sayed, 2015). Por conta das limitações encontradas em meios oficiais para coleta de dados, como APIs, não é incomum que pesquisadores precisem recorrer a soluções baseadas em *web scraping* para desenvolver estudos. Comumente, as técnicas de raspagem não são permitidas nas políticas de uso das plataformas, que criam obstáculos técnicos para impedi-las, exigindo esforço crescente e contínuo de pesquisadores para coleta e análise de dados.

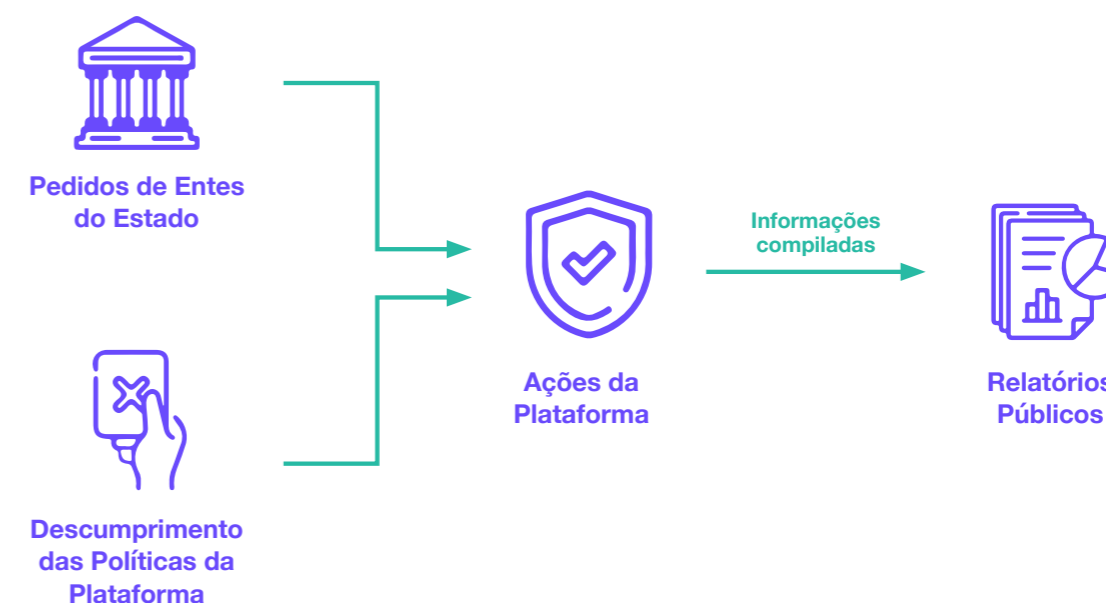
Além disso, diferentes casos de conflitos judiciais impostos por plataformas a pessoas que realizaram raspagem de dados (Roth, 2023; Conger, 2016) apontam para o problema da insegurança jurídica que esse tipo de metodologia impõe. As plataformas alegam que dificultam a prática de raspagem para impedir a ação de malfeitores e garantir a integridade dos usuários. Os motivos listados incluem proteger dados de indivíduos (Meta, [S.d.]), garantir a autenticidade da sua base de usuários, impedir a ação de bots (X/Twitter, [S.d.]) e, possivelmente, evitar que modelos de Inteligência Artificial (IA) de outras empresas sejam alimentados com dados públicos de usuários (Mehta, 2023).



O que são os relatórios de transparência?

Tradicionalmente, os relatórios de transparência são documentos voluntários das plataformas. Eles reúnem as ações motivadas por pedidos de autoridades e pelo descumprimento dos termos de uso da plataforma. As decisões por trás destas ações podem incluir, por exemplo, o fornecimento (ou não) de informações não disponíveis publicamente a entes estatais e a moderação ou a remoção de um conteúdo.

Algumas das críticas frequentes aos relatórios estão relacionadas à disponibilização de dados agregados com granularidade insatisfatória e à ausência de exemplos concretos que permitam o detalhamento das políticas de moderação. O foco excessivo em ações governamentais, a subnotificação das ações das plataformas e a falta de padronização entre diferentes relatórios também são pontos negativos.



Por Que Medir a Transparência a Partir da Qualidade dos Dados?

“A capacidade de medir a qualidade [dos dados] é dependente da capacidade de medir o grau com que os dados atendem aos requisitos estipulados”

(Benson, 2019)

Ideais de transparência são, há muito tempo, parte de modelos de compreensão e responsabilização de diferentes sistemas e envolvem, por um lado, saber **como eles funcionam** e, de outro, **como é possível governá-los** (Crawford, 2021). De modo amplo, a governança de dados envolve tanto a **governança de recursos de tecnologia da informação** de empresas e organizações quanto a **governança corporativa** e “foca em **princípios de organização e controle** sobre esses insumos [dados]” (Barbieri, 2019, p. 36) de modo a possibilitar o melhor uso da informação e do conhecimento que eles podem gerar.

A **governança de dados** não é uma área nova, já que organizações de todo porte sempre tiveram que lidar com dados. Porém, a revolução digital passou a envolver o entrecruzamento de novas disciplinas, sempre “com **foco central em qualidade de dados** no sentido mais amplo”, o que envolve melhorias na produção dos dados, monitoramento de seu uso e diferentes aspectos críticos de segurança, privacidade, ética e *compliance* (Barbieri, 2019, p. 36).

As plataformas de redes sociais já **desenvolvem ativamente processos e políticas de governança de dados**, porque são fundamentais para extrair informações e conhecimentos úteis para seus negócios. No entanto, é necessário que, em meio a governança corporativa e tecnológica, as plataformas de redes sociais considerem de forma

central o interesse público na **governança de dados**. Para isso, é fundamental definir **mecanismos** que permitam aos pesquisadores analisar os processos que **fornecem ou produzem dados** nesses espaços.

Uma das dificuldades da gestão de dados é que eles tendem a parecer “abstratos” e “imateriais” e, por isso, sistemas baseados em dados podem “facilmente ficar de fora (...) de responsabilidades de cuidado, consentimento e risco” (Crawford, 2021, p. 113). Iniciativas legislativas recentes buscam contornar essas dificuldades e mitigar os **efeitos sociais negativos do alto poder de influência das plataformas de redes sociais na vida social**, por meio da ampliação de sua responsabilidade como intermediárias, da regulação de seus serviços e de obrigações legais vinculativas no que diz respeito à transparência de dados para a pesquisa. Entre os exemplos estão o **PL 2.630/2020**, que tramitou no Brasil por quatro anos e foi retirado de pauta em abril de 2024, o **Digital Services Act (DSA)** em vigor na União Europeia (European Commission, [S.d.]) e a **Lei de Responsabilidade e Transparência das Plataformas**, em tramitação no Senado americano, após ser reintroduzida em junho de 2023 (Perrino, 2023).

A transparência, aspecto central no atual debate sobre a governança das plataformas digitais, é entendida nesse contexto como a prática das plataformas de disponibilizar informações internas que sejam de interesse público (Urman; Makhortykh, 2023). Acessar informações sobre as funcionalidades e a aplicação de políticas das plataformas é fundamental para permitir que a sociedade, os governos e outros *stakeholders* avaliem a atuação dessas empresas (Urman; Makhortykh, 2023).

Experiências Anteriores com Índices de Transparência de Dados

Uma fonte de importantes aprendizados no manejo de dados de interesse público no Brasil é a **Lei de Acesso à Informação (LAI)** (Brasil, 2011), em vigor desde 2012. A LAI resultou no avanço da transparência de dados públicos não sigilosos e levou ao acúmulo de **boas referências na padronização e na avaliação de critérios de qualidade em dados públicos**, que podem **informar iniciativas semelhantes para dados de plataformas de redes sociais**.

No **setor público**, alguns dos problemas de acesso à informação têm sido identificados por meio de diversos índices de transparência que avaliam portais do Executivo, Legislativo e Judiciário municipais, estaduais e federais, assim como Ministérios Públicos e Tribunais de Conta. São exemplos de índices de transparência e qualidade:

- **Índice de Transparência de Portais Legislativos (ITpL)**: criado pelo Congresso

Brasileiro (2022), é aplicável a todas as casas legislativas do Brasil e tem cinco critérios de avaliação, com diferentes pesos: disponibilidade da informação, atualidade da informação, uso de técnicas de Linguagem Simples nos textos; existência de série histórica; e disponibilidade da informação em dados abertos.

- **Selo de Qualidade de Transparência Pública**: criado pela Atricon (Associação de Membros do Tribunal de Contas do Brasil) (2023), o selo se baseia em um índice de “transparência ativa” em portais dos Poderes Executivo, Legislativo e Judiciário, assim como dos Tribunais de Contas, Ministérios Públicos e Defensorias Públicas, nos níveis federal, estadual/distrital e municipal.

- **Índice de Transparência Ativa (ITA)**: criado pelo Governo de Brasília (2023),

mensura o grau de cumprimento da LAI por órgãos e entidades do Poder Executivo do Distrito Federal. Em 2023, o ITA também avaliou a transparência passiva.

Além disso, a **sociedade civil organizada** também desenvolve índices para avaliar a transparência de dados de interesse público, mobilizando a ação de entes de Estado e fomentando o debate público sobre temáticas sensíveis. Um exemplo bem sucedido foi o Índice de Transparência da Covid-19:

- **Índice de Transparência da Covid-19**: criado pela Open Knowledge Brasil ([S.d.]), avalia a transparência dos dados divulgados sobre contágio e infraestrutura de saúde para enfrentamento da pandemia de Covid-19. O índice atribui uma nota e compara as iniciativas de transparência de estados e capitais.

Sistemas baseados em dados, como são as plataformas digitais e, particularmente, as plataformas de redes sociais, são altamente dependentes não apenas da qualidade dos dados que os alimentam, mas também de **como esses dados são utilizados durante todo o seu ciclo de vida**. Em linhas gerais, as medidas de qualidade de dados estabelecem critérios objetivos que ajudam

a identificar o quão adequados para o propósito almejado estão os dados de uma organização (Mahanti, 2018). Essas medidas ajudam a avaliar se os dados estão de acordo com dimensões definidas, reconhecidas e empregadas internacionalmente, como as da **ISO 8000** ([S.d.]).

ISO 8000

A ISO 8000 é um conjunto de normas internacionais que especificam os requisitos para qualidade e integridade dos dados em situações de portabilidade entre diferentes softwares ou aplicativos. Desenvolvidas pela **Organização Internacional de Padronização (ISO)**, essas normas cobrem aspectos do gerenciamento da qualidade de dados, como, por exemplo, métodos de medição, processos de validação e práticas de manutenção de registros.

A ISO 8000 é importante por estabelecer uma padronização comum a diferentes países para a portabilidade de dados e informações, com garantias de qualidade. Do ponto de vista técnico, a norma considera como passíveis de portabilidade os “dados que podem ser movidos de uma aplicação de *software* para outra sem perder significado” (Benson, 2019). A norma trata de aspectos como a possibilidade dos dados serem lidos em qualquer aplicativo

sem perder informações e sem que seja preciso pagar taxas ou *royalties*. Para isso, define padrões para o *encoding* dos dados, a apresentação dos metadados e os tipos de dados que devem ser incluídos em casos de portabilidade.

Garantir a qualidade dos dados coletados e analisados é uma forma de assegurar a **confiabilidade** e a **reprodutibilidade** de estudos, além de possibilitar **generalizações** importantes para a pesquisa social (Srivastava; Mishra, 2021). Em última instância, aumenta o **nível de transparência sobre o funcionamento do sistema como um todo**. Além disso, a qualidade dos dados permite à pesquisa acadêmica assegurar se as medidas de transparência anunciadas pelas plataformas de redes sociais estão, de fato, sendo implementadas, se não em todo, mas em boa parte do ciclo de vida dos dados.

Embora muitas padronizações, como a ISO 8000, sejam de **adoção opcional**, elas oferecem uma **referência internacional** de qualidade e transparência que traz mais confiança para ações de interesse público e para a economia baseada em dados, além de possibilitar a **manutenção da utilidade e da qualidade** dos dados no longo prazo.

Segundo a Electronic Code Management Association (ECCMA)¹, “dados de baixa qualidade são

(...) a principal causa de problemas de transparência que elevam os custos do *compliance* regulatório” ([S.d.]).

Os altos custos do *compliance* regulatório a que a ECCMA se refere tendem a ficar ainda maiores, caso o problema de transparência não seja resolvido em paralelo às crescentes exigências legais em relação à proteção de dados pessoais, após a chegada das chamadas leis de proteção de dados de quarta geração, como o **Regulamento Geral de Proteção de Dados (RGPD)**, na União Europeia (2016), e a **Lei Geral de Proteção de Dados (LGPD)** no Brasil (2018), assim como as crescentes demandas relacionadas à IA generativa.

Uma das medidas de *compliance* mais utilizadas pelas plataformas, atualmente, são os **relatórios de transparência**, que são regularmente disponibilizados pelas empresas com informações sobre a remoção de conteúdos, incluindo

¹A ECCMA é uma entidade sem fins lucrativos e administradora do grupo de aconselhamento técnico da Organização Internacional de Padronização (ISO, na sigla em inglês) nos EUA.

aqueles retirados por requisição de governos (Urman; Makhortykh, 2023). Tradicionalmente, os relatórios de transparência sobre moderação são documentos voluntários e sem um formato obrigatório. Em muitos casos, promovem apenas uma “visibilidade moderada” (Wagner et al., 2020), já que, sem meios para uma **auditoria independente**, não é possível saber se há liberação apenas de certas informações e enquadramentos que promovam uma **perspectiva favorável às plataformas**. Recentemente, o DSA passou a obrigar que grandes plataformas online e sites de busca divulguem esse tipo de relatório sobre suas atividades em países da União Europeia. Os primeiros relatórios após a nova regra foram disponibilizados em novembro de 2023.

A história dos relatórios de transparência das plataformas de redes sociais ainda é recente. O primeiro foi publicado pelo Google, em 2010 (Google, [S.d.]), seguido pelo Twitter (X/Twitter, [S.d.]), em 2012. As plataformas se viram pressionadas, por diferentes empresas e por governos locais, a retirar conteúdos protegidos por direito autoral e a derrubar conteúdos ilegais. Mas as requisições também envolviam restrições à liberdade de expressão de cidadãos quanto a questões políticas sensíveis, especialmente (mas não apenas) em países com governos autoritários. Isso motivou o Google, e depois outras plataformas, a informar publicamente casos em que a retirada do conteúdo estava relacionada a ordens judiciais ou requisições governamentais.

Uma das críticas comuns aos relatórios de transparência das plataformas é o foco em **dados agregados, com granularidade insatisfatória**, e sem a inclusão de casos concretos que permitam um exame mais detalhado sobre as políticas aplicadas na moderação de conteúdo (Kosta; Brewczyńska, 2019; Suzor et al., 2019). Ou seja, normalmente os conceitos e as justificativas da moderação dos conteúdos não são transparentes. Uma outra crítica de pesquisadores é que as plataformas tendem a tornar público e dar bastante visibilidade às requisições governamentais de remoção de conteúdo e perfis, porém

há uma subnotificação proposital das ações de moderação decididas e realizadas diariamente por elas mesmas (Hovyadinov, 2019; Kosta; Brewczyńska, 2019). Essas ações levantam a suspeita de que haja “*transparency-washing*” (Zalnieriute, 2021), termo utilizado para descrever a prática corporativa de distrair o público com determinadas questões para minimizar a responsabilização de plataformas digitais, incluindo as de redes sociais, sobre as decisões de moderação de conteúdo que tomam arbitrária e proativamente, afastando-as do compromisso com a transparência (Urman; Makhortykh, 2023).

Os relatórios de transparência sobre moderação de conteúdo na Europa mais recentes, após a vinculação legal do DSA, também têm recebido críticas. Uma delas é que falta **padronização** entre diferentes relatórios, impossibilitando comparações analíticas entre as plataformas. Entre os aspectos mais notados estão as **diferenças de qualidade e no nível de granularidade** dos dados disponibilizados (Miller, 2023).

Assim, para avaliar diferentes perspectivas das políticas e das práticas de transparência das plataformas, é importante uma análise abrangente dos recursos disponíveis, como:

- a **disponibilidade e a qualidade dos dados, por meio de métodos oficiais de coleta** (como a API e a interface de coleta de dados);
- a possibilidade de utilizar **métodos extra-oficiais** para este fim (como a raspagem de dados);
- o detalhamento da **documentação oficial** para guiar a utilização desses recursos; e
- os relatórios de transparência.

Todos esses aspectos são avaliados neste **Índice de Transparência de Dados (ITD)**, que procura revelar o nível de transparência das plataformas de redes sociais no Brasil, por meio de critérios objetivos e aplicáveis a todas as plataformas

avaliadas. O ITD parte da premissa de que a padronização de critérios para avaliar sistematicamente a transparência e a qualidade dos dados das plataformas de redes sociais e de interesse para pesquisa tende a impactar positivamente na transparência desses espaços, que tanto afetam a vida pública.

Para isso, o ITD toma por base dimensões de qualidade preconizadas pela literatura científica, independentemente das particularidades de cada plataforma. Entre as dimensões endógenas aos dados, são avaliadas a completude, a atualidade, a consistência e a acessibilidade (Batini;

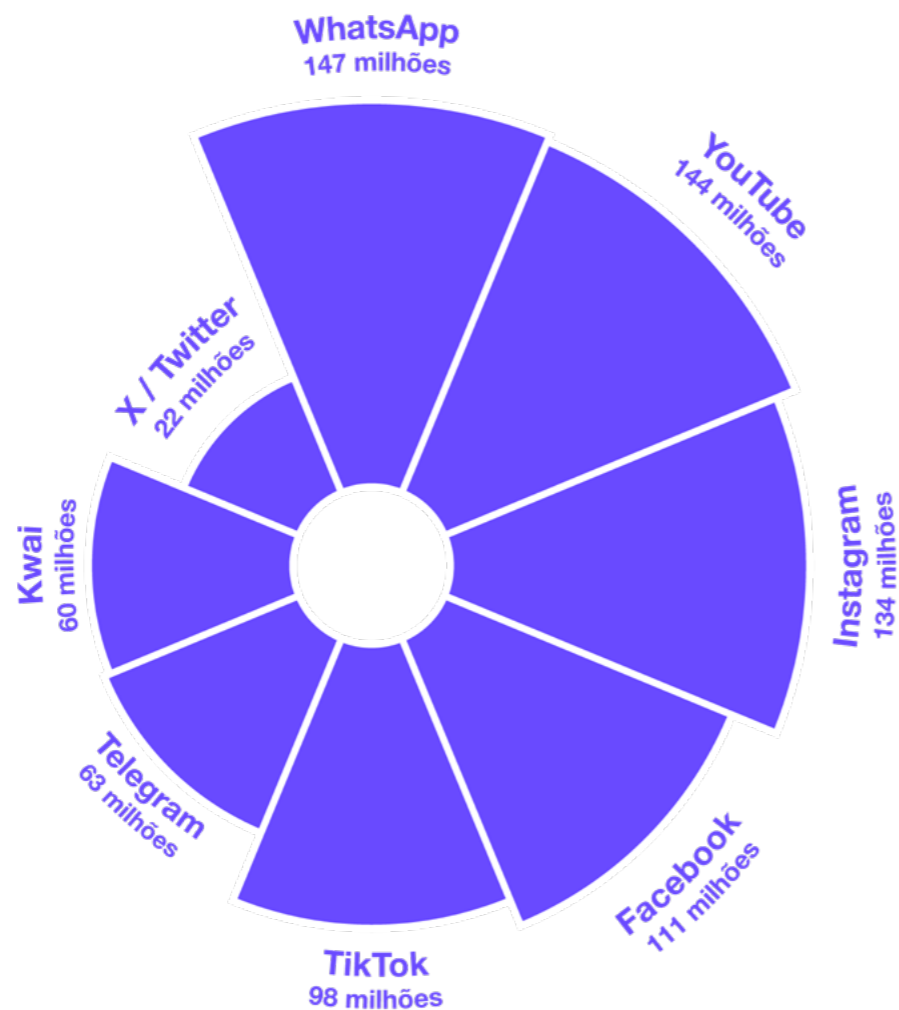
Scannapieca, 2006; McGilvray, 2008; Loshin, 2008; Barbieri, 2019). Outras dimensões, como conformidade e relevância (Barbieri, 2019), dependem de fatores exógenos e, por isso, podem variar de acordo com as normas legais em vigor em cada país ou com os objetivos específicos da pesquisa. Vistas de forma conjunta, dimensões endógenas e exógenas de qualidade de dados podem indicar o nível de transparência das plataformas.

Escopo do ITD

O Índice de Transparência de Dados das Plataformas de Redes Sociais avalia o acesso e a qualidade dos dados de publicações orgânicas e públicas nas principais plataformas de redes sociais e mensageria: YouTube, Facebook, Instagram, X/Twitter, Telegram, TikTok, Kwai e WhatsApp.

Fontes: Global AD (2024); Bianchi (2024); Shewale (2024); Data Reportal (2024); Opinion Box (2024)

Número de usuários, no Brasil, em cada plataforma analisada



O ITD Pretende Responder:

Quais são as medidas de transparência e acesso a dados das principais plataformas de redes sociais no Brasil?

Qual é o nível de qualidade dos dados disponibilizados por essas plataformas para atividades de pesquisa?

Objetivos

Padronizar

Definir parâmetros de avaliação sobre o nível de acesso e qualidade dos dados de interesse público provenientes de plataformas de redes sociais.

Comparar

Aferir comparativamente a performance de cada plataforma a partir de critérios comuns e metodologia padronizada.

Avaliar

Identificar, de forma sistemática e transparente, os pontos fortes e fracos no acesso e na qualidade dos dados.

Aprimorar

Indicar, pública e objetivamente, o que precisa ser melhorado na disponibilização de dados para a pesquisa.

Justificativa: Que Plataformas Analisamos?

O ITD abrange as plataformas de redes sociais que têm grande popularidade no Brasil para a difusão de conteúdos gerados por usuários (dos Santos, 2021). Consideramos como plataformas de redes sociais os espaços em que usuários produzem e consomem conteúdos, interagindo e se conectando com outros usuários (Ellison; Boyd, 2013).

Esse conceito inclui os chamados aplicativos de mensageria, como Telegram e WhatsApp. Esses apps também rompem a barreira entre a comunicação interpessoal e massiva, desempenhando papel relevante na disseminação e no compartilhamento de conteúdos produzidos por usuários em grupos e canais públicos. Permitem ainda a criação de redes de interação com base em afinidades em comum (Júnior et al., 2021; Rogers, 2020), nas quais os usuários tornam-se simultaneamente, produtores empoderados de conteúdo e alvo para a exploração por parte das empresas.

Para o recorte do estudo, nos baseamos na premissa do Digital Services Act (DSA), que

estabelece medidas de responsabilidade e transparência para plataformas digitais que alcançam mais de 10% da União Europeia.

Este critério foi utilizado para a seleção das plataformas avaliadas no índice, considerando o contexto brasileiro, incluindo as plataformas de maior impacto social. Por serem as maiores e mais relevantes, devem garantir investimento em infraestrutura robusta de transparência e seguir as melhores práticas do mercado por terem os recursos necessários para tanto.

Nessa versão do Índice, não estão incluídas avaliações da transparência de Pinterest e LinkedIn. Ainda que tenham uma base significativa de usuários, essas plataformas se fazem pouco presentes em investigações acadêmicas e científicas no Brasil e no mundo (Kapoor et al., 2017; Zuckerman, 2021). Essas e outras plataformas poderão ser incluídas nas próximas atualizações do ITD.

Abordagem Metodológica

Para mensurar a transparência das principais plataformas disponíveis no Brasil, um **roteiro de avaliação** foi elaborado em um processo iterativo e deliberativo que estabeleceu os parâmetros e critérios de avaliação e suas definições conceituais.

Os parâmetros foram avaliados e justificados por oito pesquisadores do **NetLab UFRJ**, divididos em duplas que incluíam um especialista em coleta, infraestrutura e processamento de dados e outro com experiência em análise de dados e desenho de pesquisa em Ciências Sociais Computacionais. As duplas também ficaram responsáveis

pela revisão de respostas feitas por outros pares, conforme a divisão apresentada na **Tabela 1**. A distribuição de plataformas entre os pesquisadores levou em consideração o conhecimento prévio e a elaboração de pesquisas envolvendo dados da plataforma avaliada.

Ao longo do processo de elaboração do índice, a adequação dos parâmetros e a pertinência de suas justificativas foram continuamente deliberadas em conjunto pelos avaliadores e outros pesquisadores envolvidos no estudo. A avaliação foi realizada e revisada ao longo do primeiro semestre de 2024.

Plataforma	Pesquisadores Responsáveis pela Resposta	Pesquisadores Responsáveis pela Revisão
 YouTube	E1 e E2	E7 e E5
 Facebook	E3 e E4	E6 e E2
 Instagram	E3 e E4	E6 e E2
 X/Twitter	E3 e E5	E6 e E8
 Telegram	E6 e E2	E3 e E4
 Kwai	E3 e E2	As respostas foram deliberadas em conjunto
 TikTok	E7 e E5	E1
 WhatsApp	E6 e E8	E3 e E4

Tabela 1: Divisão das respostas dos parâmetros por duplas de especialistas (E_n)

O roteiro é composto por **40 parâmetros** que analisam seis dimensões de qualidade de dados: **acessibilidade, conformidade, completude, consistência, relevância e atualidade**. As avaliações foram realizadas e justificadas com base em cinco diferentes fontes de informação: a documentação oficial da plataforma, a literatura acadêmica, a realização de testes de acesso e coleta de dados, a experiência acumulada do **NetLab UFRJ** e os relatórios de transparência.

Destas, consideramos principalmente a **documentação da API**. A documentação de um *software* relata, detalha e explica o seu funcionamento, indicando aos usuários como utilizá-lo. Plataformas que disponibilizam APIs comumente incluem documentações para que desenvolvedores possam entendê-las durante as requisições. Além da documentação, consultamos as **políticas** e os **termos de uso** da plataforma e das APIs para responder aos parâmetros avaliados. Em casos excepcionais, contatamos diretamente o suporte da plataforma para esclarecimentos.

Para as avaliações, também levamos em consideração a produção acadêmica nacional e internacional **publicada em periódicos de impacto**, com metodologias desenvolvidas, testadas e aprovadas por pares. A literatura acadêmica foi usada, principalmente, para identificar e conferir o uso de métodos não oficiais de coleta. Conduzimos, ainda, **experimentos controlados** realizados entre dezembro de 2023 e janeiro de 2024, simulando situações reais de uso e coleta de dados, para **testar** e **verificar** a consistência

das respostas da API, a persistência de conteúdos removidos, a viabilidade da raspagem via navegador e a ocorrência de bloqueios à raspagem de dados.

Também levamos em consideração os obstáculos enfrentados e as soluções desenvolvidas pelo **NetLab UFRJ** na construção de sua infraestrutura própria e customizada para garantir o monitoramento de diferentes plataformas ao longo dos anos. Desde 2020, o laboratório desenvolve e mantém uma coleta contínua e ininterrupta para o **monitoramento constante** de diferentes plataformas de redes sociais. A construção dessa infraestrutura de coleta depende do manuseio de diferentes sistemas, assim como do entendimento pleno dos tipos de dados disponíveis nas plataformas e de como coletá-los.

Por fim, também avaliamos a disponibilização de **relatórios de transparência** por parte das plataformas e, nos casos em que são disponibilizados, a periodicidade de sua publicação e o detalhamento de suas informações.

Critérios de Avaliação: As Dimensões da Qualidade de Dados

Acessibilidade (16 parâmetros):

Dimensão mais importante analisada, já que os parâmetros de avaliação referentes às outras

dimensões são dependentes do acesso aos dados. A acessibilidade se refere à disponibilidade de dados e a facilidade de localizá-los, acessá-los, obtê-los e explorá-los para um determinado fim

(Mahanti, 2018). Portanto, não basta torná-los acessíveis. É preciso, também, haver condições para que sejam facilmente compreendidos e analisados por pesquisadores com variados graus de conhecimento técnico, principalmente no que diz respeito à necessidade de programação. Nesta dimensão, foram analisados fatores como a disponibilização gratuita ou paga de API e se ela permite a extração total ou parcial de dados de interesse público.

Cinco parâmetros desta dimensão se desdobram em dois critérios especiais de avaliação, que compõem 50% da nota, sendo eles:

Critério Especial 1: É possível acessar pela API da plataforma o universo de dados públicos, de forma gratuita, para fins de pesquisa?

O universo a ser monitorado é composto por todas as postagens públicas imediatamente localizáveis, acessíveis e recuperáveis por qualquer usuário. Assim, entendemos que, caso a pessoa interessada assim deseje, as plataformas devem disponibilizar todo o conjunto de dados das publicações públicas correspondentes às requisições feitas, ao invés de apenas disponibilizar conjuntos recortados e definidos arbitrariamente por elas, de forma a permitir a reprodutibilidade das coletas, das análises e o avanço das pesquisas, evitando o possível enviesamento.

A nota máxima é dada à plataforma que garanta acesso total ao universo a ser monitorado por meio de API gratuita. No caso de não haver uma API de coleta de dados disponível gratuitamente para o público geral, avaliamos se ao menos pesquisadores dispõem de meios específicos para acessar os dados de forma programática.

Com base no acesso aos dados, os parâmetros que avaliam o acesso universal de forma gratuita por pesquisadores e/ou outros interessados compõem 25% da nota final de cada plataforma.

Parâmetros de avaliação que compõem o Critério Especial 1

P1: A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários?

P2: O universo a ser monitorado é recuperável pela API da plataforma?

P3: O acesso à API da plataforma é gratuito?

P4: A plataforma oferece a pesquisadores acesso gratuito e específico à API?

Critério Especial 2: A plataforma oferece uma interface para coletar os dados por meio de busca customizável?

É esperado que a plataforma disponibilize uma interface gráfica para que interessados com pouco ou nenhum conhecimento técnico em programação possam coletar os mesmos dados retornados pela API.

Esse tipo de ferramenta colabora com a democratização da pesquisa e a transparência dos dados para o cidadão. No entanto, a existência de uma interface do tipo não dispensa a necessidade de existência de APIs, já que estas garantem uma maior customização dos processos de coleta de dados e permitem que estes processos sejam automatizados e ganhem escala. Esse parâmetro compõe 25% da nota final das plataformas.

Parâmetro de avaliação que compõe o Critério Especial 2

P5: A plataforma oferece uma interface para coletar os dados por meio de busca customizável?

Outros parâmetros que compõem a dimensão Acessibilidade

P6: É possível extrair os dados requisitados diretamente da resposta da API da plataforma?

P7: A API da plataforma provê uma forma de autenticação que permite a renovação automática, sem bloqueios à aquisição de dados?

P8: A criação de tokens de acesso à API da plataforma pode ser feita de forma gratuita?

P9: É possível criar novos tokens de acesso à API da plataforma sem limitações de quantidade?

P10: O processo para que pesquisadores tenham acesso à API da plataforma é claro, descomplicado e com prazo bem definido?

P11: A API da plataforma disponibiliza um endpoint para recuperar dados de uma publicação específica?

P12: A API da plataforma disponibiliza um endpoint para recuperar dados de um autor específico?

P13: A API da plataforma disponibiliza um endpoint para recuperar dados por meio de termos de busca?

P14: É possível realizar aquisição de dados por raspagem, sem necessidade de autenticação, por meio da interface de usuário da plataforma?

P15: É possível realizar aquisição de dados por raspagem sem necessidade de outros dispositivos?

P16: É possível recuperar os dados por meio de raspagem, sem necessidade de contornar ferramentas e técnicas que visam impedir o acesso programático aos dados?

Conformidade (11 parâmetros):

Avalia se a documentação oficial e os dados recuperados estão adequados quanto aos formatos adotados e às normas legais vigentes no país (Mahanti, 2018). Trata-se de uma dimensão exógena, ou seja, relacionada mais ao “entorno” dos dados do que com eles próprios” e, portanto, mais atrelada “à sua governança e gerência do que ao seu próprio conteúdo” (Barbieri, 2019). Nesta dimensão, é avaliado, por exemplo, se dados específicos, como datas e URLs, são entregues em acordo com padrões internacionais. Também analisamos as documentações disponibilizadas por cada plataforma referentes aos meios de coleta de dados disponibilizados, a fim de entender se são facilmente acessíveis e compreensíveis, se apresentam descrições e exemplos claros de utilização e quais são suas condições e termos de uso. Além disso, também avaliamos a divulgação e o detalhamento de relatórios de transparência sobre ações de moderação por parte das plataformas analisadas.

Parâmetros que compõem a dimensão Conformidade

P17: A estrutura dos dados disponibilizados pela API da plataforma é estável?

P18: Os dados retornados pela API da plataforma estão em formato padronizado?

P19: A documentação da API da plataforma é publicada em acesso aberto?

P20: A documentação da API da plataforma está escrita de forma clara e exemplificada?

P21: A documentação da API da plataforma descreve seus termos de uso?

Parâmetros que compõem a dimensão Conformidade (continuação)

P22: A documentação descreve o formato da resposta dos *endpoints* da API da plataforma?

P23: A documentação da API da plataforma é disponibilizada nativamente em português?

P24: A plataforma permite raspagem e outros tipos de acesso automático nos seus termos de uso?

P25: A plataforma produz relatórios de transparência periódicos sobre a moderação de conteúdos no Brasil e os disponibiliza publicamente, sem necessidade de requisição?

P26: Em seus relatórios de transparência, a plataforma indica o volume de cada tipo de violação identificada no Brasil de acordo com as políticas de moderação vigentes?

P27: Os relatórios de transparência especificam informações sobre a quantidade e o tipo de requisições feitas por entes do Estado brasileiro à plataforma, além da quantidade e do tipo de solicitações acatadas?

Completeness (6 parâmetros):

Indica se os dados recuperados apresentam os atributos indispensáveis para sua compreensão e se é possível realizar um **monitoramento íntegro em cada uma das plataformas analisadas** ao coletá-los, considerando, principalmente, os critérios exigidos para a pesquisa acadêmica e científica (Mahanti, 2018). Nesta dimensão, um dado é considerado completo quando **pode ser utilizado e aplicado em situações diversas de pesquisa**, ainda que alguns campos de dados opcionais não tenham sido preenchidos. Assim, é avaliado se a frequência permitida para coleta de dados por vias oficiais possibilita o monitoramento consistente e se é possível recuperar dados referentes a comentários e conteúdos temporários, por exemplo.

Parâmetros que compõem a dimensão Completeness

P28: É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?

P29: É possível recuperar dados de conteúdos temporários por meio da API da plataforma?

P30: É possível recuperar dados históricos por meio da API da plataforma?

P31: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 1 milhão de publicações em 24 horas?

P32: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 100 mil publicações em 24 horas?

P33: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 10 mil publicações em 24 horas?

Consistency (4 parâmetros):

Avalia se o formato e a apresentação dos dados são consistentes e idênticos em todas as bases extraídas e, em especial, em requisições idênticas entre si (Mahanti, 2018). Também verifica se os termos de busca e os filtros usados recuperam dados **coerentes e sem contradições, duplicações e discrepâncias**. Esta dimensão analisa, por exemplo, se os mesmos dados são recuperados quando requisitados em momentos diferentes, evitando **inconsistências** no monitoramento sistemático. A consistência é imprescindível para produzir relatórios **precisos e ágeis**, pois evita a necessidade de conferência e/ou correção constante dos dados e permite maior **auditabilidade**.

Parâmetros que compõem a dimensão Consistency

P34: Os dados retornados pela API da plataforma são persistentes?

P35: Os dados recuperados pela API da plataforma refletem o que é exibido em sua interface de usuário?

P36: A resposta retornada pela API da plataforma é sempre a esperada?

P37: A resposta retornada pela API da plataforma é coerente com os parâmetros e filtros utilizados na requisição?

Relevance (2 parâmetros):

Avalia se os dados são **pertinentes para a finalidade** à qual se destinam (Mahanti, 2018), ou seja, se estão de acordo com os **objetivos da pesquisa e da requisição**. Os dados recuperados também devem ser suficientes para embasar uma análise robusta. Esta dimensão analisa, por exemplo, a possibilidade de **filtrar dados por meio de termos de busca** no processo de coleta.

Parâmetros que compõem a dimensão Relevância

P38: As entidades retornadas pela API da plataforma são suficientes para compreender os dados em todos os seus níveis de detalhes?

P39: A API da plataforma permite a utilização de filtros para refinar a requisição de dados?

Atualidade (1 parâmetro):

Qualifica o **impacto da passagem de tempo** na disponibilização dos dados, de modo a avaliar a **agilidade** com que o processo de **atualização** dos dados ocorre (Mahanti, 2018).

Parâmetros que compõem a dimensão Atualidade

P40: É possível recuperar dados recém-publicados, quase em tempo real à publicação, por meio da API da plataforma?

Composição da Nota

Para o cálculo das avaliações do índice, cinco parâmetros relativos à dimensão de acessibilidade foram agrupados em dois critérios especiais, que compõem 50% da nota. Os outros parâmetros correspondem aos 50% restantes da nota. Cada um dos 35 parâmetros restantes possui o mesmo peso na composição da nota final, de modo que dimensões com mais parâmetros têm maiores pesos.

Dessa forma, realizamos um cálculo ponderado em que:

1. 25% da nota corresponde ao Critério Especial 1: “É possível acessar pela API da plataforma o universo de dados públicos, de forma gratuita, para fins de pesquisa?”.


Só pontuam nesse campo plataformas que disponibilizam acesso a todo o conteúdo público e permitem a recuperação total, sistematizada e sem custos, evitando vieses e garantindo a reprodutibilidade de coletas e análises. Ou seja, a plataforma precisa ser avaliada positivamente nos parâmetros P1 (A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários?) e P2 (O universo a ser monitorado é recuperável pela API da plataforma?), além de responder positivamente a pelo menos um dos parâmetros P3 (O acesso à API da plataforma é gratuito?) ou P4 (A plataforma oferece a pesquisadores acesso gratuito e específico à API?).

2. 25% da nota corresponde ao Critério Especial 2: “A plataforma oferece uma interface para coletar os dados por meio de busca customizável?”.

Para pontuar neste critério a plataforma deve responder positivamente ao parâmetro P5 (A plataforma oferece uma interface para coletar os

dados por meio de busca customizável?). O fator determinante para pontuar neste parâmetro é oferecer meios para coleta de fácil compreensão e navegação por pessoas com pouco conhecimento técnico.

3. 50% da nota é composta por 35 parâmetros com o mesmo peso, de modo que dimensões com mais parâmetros têm maior peso na composição da nota final. A pontuação de cada plataforma corresponde à porcentagem de respostas SIM ☑ obtidas em cada dimensão em relação ao total de parâmetros daquela dimensão. A visualização na próxima página detalha a distribuição de pesos e parâmetros que compõem essa parcela da nota.

 O cálculo da nota final de cada plataforma é representado por:

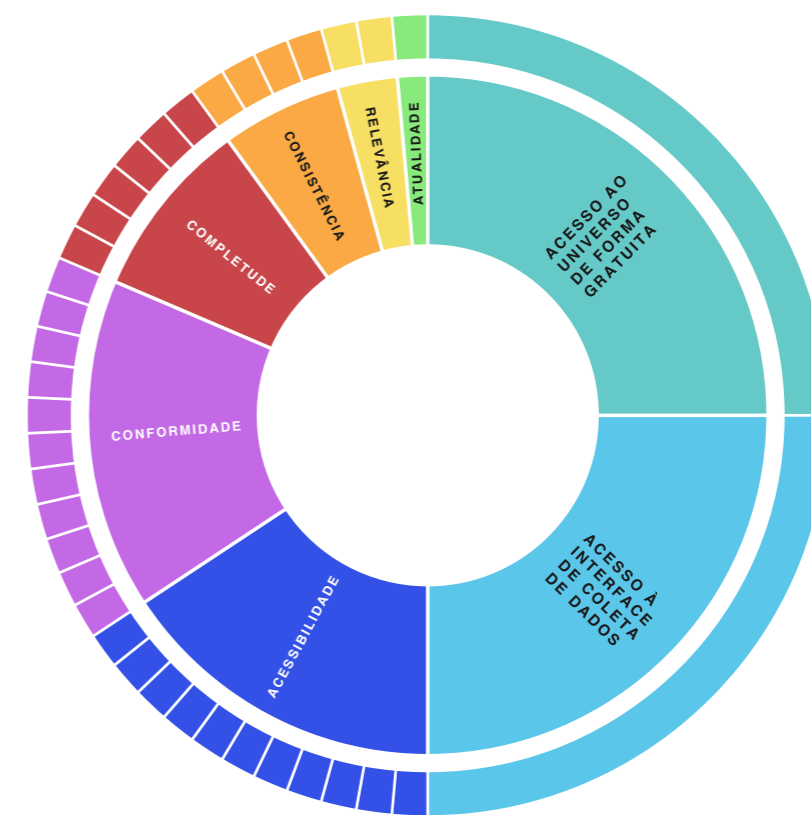
$$(xi * 25) + (xii * 25) + \left(\frac{\text{Positivas}}{\text{Parâmetros}} * 50 \right)$$

Em que:

- xi** é a resposta binária negativa ou positiva (0 ou 1) para o Critério Especial 1;
- xii** é a resposta binária negativa ou positiva (0 ou 1) para o Critério Especial 2;
- Positivas** é o número de respostas positivas de cada plataforma;
- Parâmetros** é o número de parâmetros de avaliação aplicáveis² no questionário.

²Nas avaliações de WhatsApp e YouTube, desconsideramos um dos parâmetros de avaliação e readequamos o restante dos cálculos em torno desta decisão.

Representação Gráfica da Nota



<p>ACESSO AO UNIVERSO DE FORMA GRATUITA (CRITÉRIO ESPECIAL 1)</p> <p>PESO: 25 PTS Nº DE PARÂMETROS: 4</p>	<p>ACESSO À INTERFACE DE COLETA DE DADOS (CRITÉRIO ESPECIAL 2)</p> <p>PESO: 25 PTS Nº DE PARÂMETROS: 1</p>	<p>ACESSIBILIDADE</p> <p>PESO: 15,71 PTS Nº DE PARÂMETROS: 11</p>	<p>CONFORMIDADE</p> <p>PESO: 15,71 PTS Nº DE PARÂMETROS: 11</p>
<p>COMPLETUDE</p> <p>PESO: 8,57 PTS Nº DE PARÂMETROS: 6</p>	<p>CONSISTÊNCIA</p> <p>PESO: 5,71 PTS Nº DE PARÂMETROS: 4</p>	<p>RELEVÂNCIA</p> <p>PESO: 2,86 PTS Nº DE PARÂMETROS: 2</p>	<p>ATUALIDADE</p> <p>PESO: 1,43 PT Nº DE PARÂMETROS: 1</p>

Níveis de Transparência de Dados

Para facilitar a interpretação das pontuações obtidas, o índice enquadra as plataformas analisadas em cinco categorias:

Transparência ideal (81 a 100 pontos)

Plataformas com soluções oficiais eficientes para coleta de dados, que incluem APIs e interface de coleta de dados, com documentações bem exemplificadas e sem impedimento à raspagem. Costumam divulgar relatórios de transparência periódicos, com detalhes sobre violações e remoções a pedido do Estado, no Brasil.

Transparência satisfatória (61 a 80 pontos)

Plataformas que disponibilizam dados sem restrições financeiras, porém com limitações quanto ao volume de dados que pode ser requisitado e/ou apresentando problemas de qualidade, especialmente de consistência. Divulgam relatórios de transparência sobre suas ações de moderação no Brasil com alguma periodicidade.

Transparência regular (41 a 60 pontos)

Plataformas que apresentam algumas medidas de transparência e acesso a dados, mas com diversas limitações relacionadas ao tipo de conteúdo que pode ser acessado e à amostra do universo de dados públicos passível de coleta. Em geral, divulgam relatórios de transparência com ações de moderação, no Brasil, mas sem o detalhamento esperado.

Transparência precária (21 a 40 pontos)

Plataformas que impõem barreiras técnicas, operacionais e/ou financeiras significativas às suas medidas de acesso a dados, inviabilizando o monitoramento para a maioria dos pesquisadores e pessoas interessadas. Também não têm o costume de publicar relatórios de transparência periódicos sobre suas ações de moderação de conteúdo no Brasil.

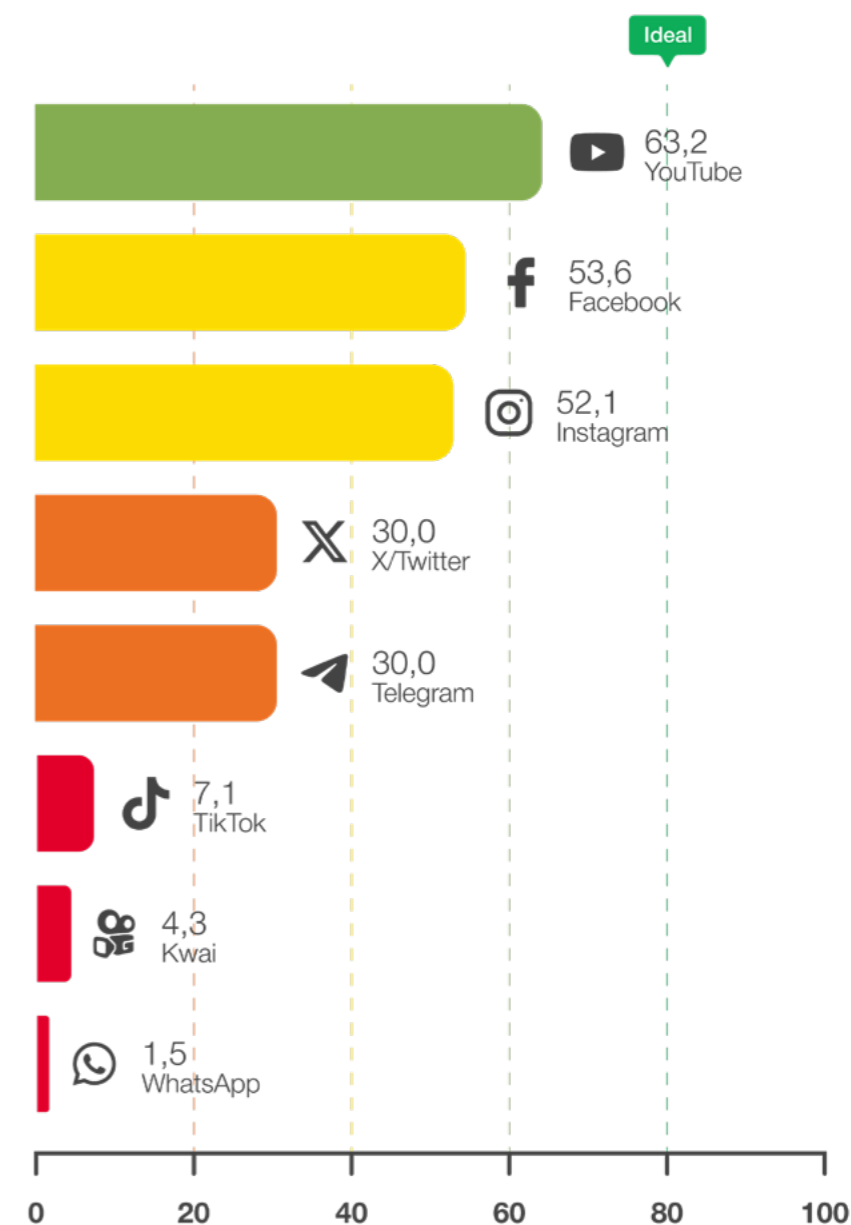
Transparência irrelevante ou nula (0 a 20 pontos)

Plataformas que não investem em quaisquer medidas de transparência e acesso a dados. Recebem poucos pontos graças às possibilidades de raspagem de dados, que, em geral, não são oficialmente permitidas. Não costumam publicar relatórios de transparência periódicos sobre suas ações de moderação de conteúdo no Brasil.

Resultados

Os resultados do ITD apontam que nenhuma plataforma avaliada alcança pontuação ideal no que diz respeito às medidas de transparência e de acesso a dados e à qualidade dos dados retornados.

A seguir, apresentaremos uma visão geral do que foi observado em cada plataforma analisada. A visão geral de cada plataforma, dividida por dimensão, assim como as respostas e justificativas específicas a cada parâmetro de avaliação estão disponíveis no [Apêndice](#).

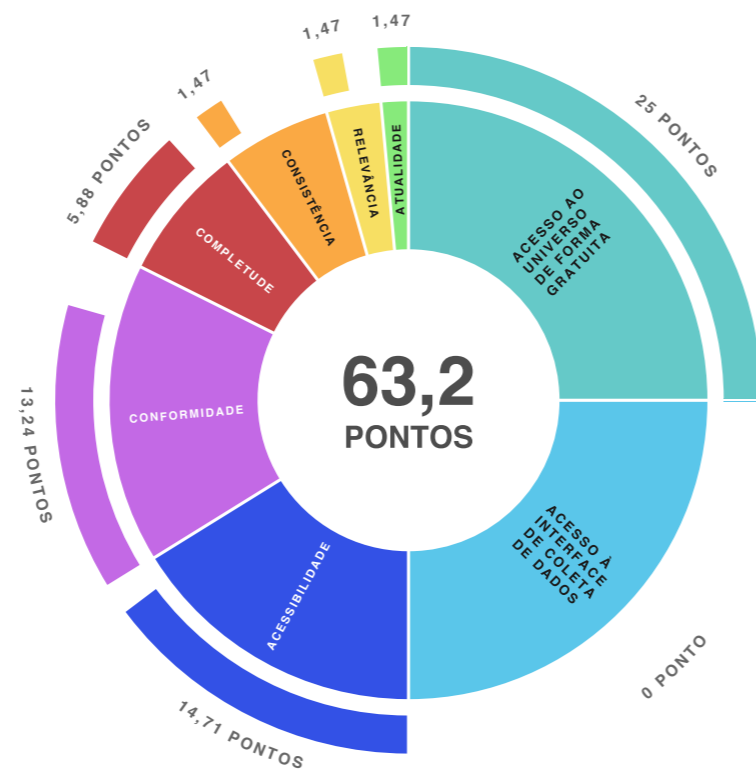


YouTube

Transparência de dados: **Satisfatório**

Legenda:

■ ACESSO AO UNIVERSO DE FORMA GRATUITA 25 DE 25 PONTOS	■ COMPLETUDE 5,88 DE 7,35 PONTOS
■ ACESSO À INTERFACE DE COLETA DE DADOS 0 DE 25 PONTOS	■ CONSISTÊNCIA 1,47 DE 5,88 PONTOS
■ ACESSIBILIDADE 14,71 DE 16,18 PONTOS	■ RELEVÂNCIA 1,47 DE 2,94 PONTOS
■ CONFORMIDADE 13,24 DE 16,18 PONTOS	■ ATUALIDADE 1,47 DE 1,47 PONTO



Dentre as plataformas pesquisadas, o YouTube obteve a melhor pontuação (63,2 pontos), sendo sua transparência de dados considerada satisfatória³. A plataforma disponibiliza uma API oficial (YouTube, [S.d.]) de acesso gratuito para qualquer usuário (P1 e P3), permitindo a busca em todo o universo de vídeos considerados públicos. A plataforma também disponibiliza uma API específica para pesquisadores (P4), cujo acesso pode ser solicitado facilmente (P10). Outros ponto positivo é que o YouTube também permite a recuperação ágil de dados recém-publicados (P40) e de dados históricos (P30).

Por outro lado, a plataforma não pontua no Critério Especial 2 em virtude da ausência de uma interface de coleta de dados (P5). Parâmetros relativos às dimensões de consistência e relevância também prejudicam o desempenho do YouTube, uma vez que há diversos problemas relacionados a respostas inconsistentes (P36) e com dados incoerentes em relação aos parâmetros e filtros aplicados nos momentos da coleta (P37). Além disso, os dados recuperados pela API da plataforma não são persistentes (P34): ela não disponibiliza metadados de publicações removidas, ao contrário da interface da platafor-

ma, que exibe uma mensagem sobre a remoção e seus motivos, caso o usuário tente assistir a um vídeo que não está mais no ar.

No Relatório sobre Cumprimento das diretrizes da comunidade (Google, [S.d.]), o YouTube disponibiliza trimestralmente o volume de vídeos removidos que foram publicados por usuários no Brasil e que violaram as diretrizes da plataforma (P25). Entretanto, não são disponibilizadas informações como o número de canais removidos, o número de remoções com base em denúncias e agregações por tipo de violações no país (P26), apesar desse tipo de informação ser cedida a outros países sobre os quais o YouTube publica relatórios específicos. A plataforma também aparece em um relatório semestral junto a outros serviços associados ao Google, no qual constam o volume de solicitações governamentais de remoção de conteúdo, por tipo de violação, no Brasil (P27).

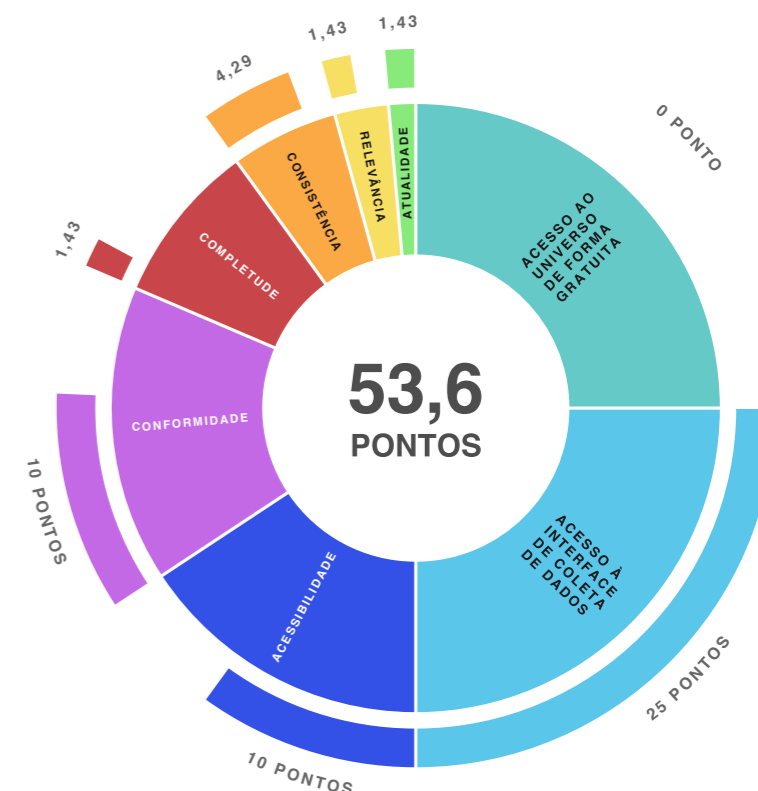
³Para o cálculo final da nota do YouTube, desconsideramos o parâmetro "É possível recuperar dados de conteúdos temporários por meio da API da plataforma?", referente à dimensão de completude, uma vez que a plataforma não permite que usuários publiquem conteúdos temporários. Portanto, além dos dois critérios especiais, consideramos 34 dos 35 parâmetros aplicáveis e readequamos todos os demais cálculos.

Facebook

Transparência de dados: **Regular**

Legenda:

■ ACESSO AO UNIVERSO DE FORMA GRATUITA 0 DE 25 PONTOS	■ COMPLETUDE 1,43 DE 8,57 PONTOS
■ ACESSO À INTERFACE DE COLETA DE DADOS 25 DE 25 PONTOS	■ CONSISTÊNCIA 4,29 DE 5,71 PONTOS
■ ACESSIBILIDADE 10 DE 15,71 PONTOS	■ RELEVÂNCIA 1,43 DE 2,86 PONTOS
■ CONFORMIDADE 10 DE 15,71 PONTOS	■ ATUALIDADE 1,43 DE 1,43 PONTO



A pontuação da transparência de dados do Facebook de 53,6 pontos é considerada regular. Quase metade da nota da plataforma depende dos 25 pontos recebidos pelo CrowdTangle, interface e API para coleta de dados (P5), que a Meta disponibilizava para pesquisadores e jornalistas coletarem amostras de dados públicos do Instagram e do Facebook (CrowdTangle, [S.d.]).

A API do CrowdTangle apresentava várias limitações e não permitia acesso a dados de comentários vinculados às publicações originais (P28) ou a conteúdos temporários (P29), como stories, o que prejudicava a completude dos dados retornados. Além disso, a ferramenta também não indicava quando conteúdos eram removidos da plataforma (P34): publicações deletadas eram tratadas pelo CrowdTangle como se nunca tivessem existido.

A API do CrowdTangle também não permitia a recuperação de todo o universo de dados públicos da plataforma (P2), uma vez que só podiam ser recuperadas publicações feitas: (i) por páginas com mais de 25 mil seguidores ou curtidas, (ii) por perfis verificados, e (iii) em grupos públicos com mais de 95 mil membros

(CrowdTangle, [S.d.]). O Facebook também perde pontos em parâmetros relacionados a métodos alternativos de coleta por colocar limitações para impedir a raspagem de dados (P14 e P16).

Na Central de Transparência da Meta (Meta, [S.d.]), a empresa disponibiliza semestralmente informações sobre publicações, perfis e comentários moderados em suas plataformas conforme as leis locais de cada país em que atua, incluindo o Brasil (P25). A Meta também divulga dados semestrais de requisições feitas por entes do Estado brasileiro para a moderação de conteúdo em suas plataformas (P27). Porém, dados sobre o total de ações de moderação, incluindo remoções por violação dos termos de uso da Meta, são disponibilizados apenas em nível global ou para alguns países com relatórios específicos — o que não é o caso do Brasil (P26).

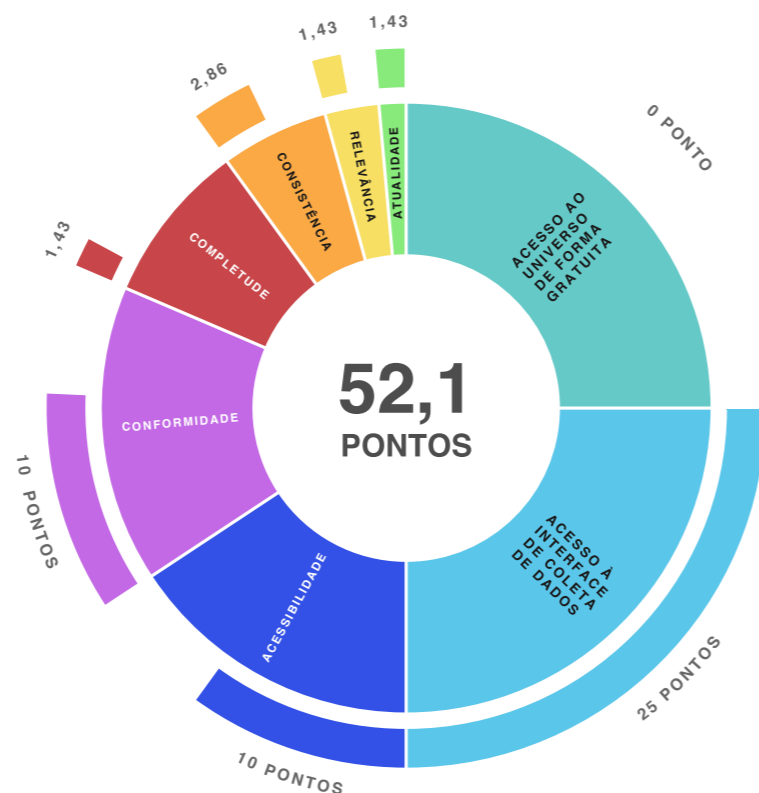
Durante a elaboração deste índice, a Meta anunciou que o serviço seria oficialmente descontinuado a partir de agosto de 2024 (CrowdTangle, 2024). Se ferramentas equivalentes de acesso a dados não forem implementadas, a pontuação do Facebook tende a ser ainda menor em avaliações futuras.

Instagram

Transparência de dados: **Regular**

Legenda:

■ ACESSO AO UNIVERSO DE FORMA GRATUITA 0 DE 25 PONTOS	■ COMPLETUDE 1,43 DE 8,57 PONTOS
■ ACESSO À INTERFACE DE COLETA DE DADOS 25 DE 25 PONTOS	■ CONSISTÊNCIA 2,86 DE 5,71 PONTOS
■ ACESSIBILIDADE 10 DE 15,71 PONTOS	■ RELEVÂNCIA 1,43 DE 2,86 PONTOS
■ CONFORMIDADE 10 DE 15,71 PONTOS	■ ATUALIDADE 1,43 DE 1,43 PONTO



A pontuação da transparência de dados do Instagram de **52,1 pontos** é considerada **regular**. Assim como no Facebook, nos baseamos na ferramenta CrowdTangle para responder aos parâmetros de avaliação da plataforma. A ferramenta oferecia uma API oficial (CrowdTangle, [S.d.]) para extração de dados (P1), além de uma interface (CrowdTangle, [S.d.]) dedicada à coleta de dados (P5).

No entanto, o CrowdTangle não permitia a recuperação de todo o universo de dados públicos da plataforma (P2), uma vez que apenas dados de publicações de perfis com mais de 50 mil seguidores e/ou verificados podiam ser recuperadas e pelo fato de dados de *reels* não poderem ser requisitados. O Instagram também foi penalizado pela falta de completude dos dados retornados pelo CrowdTangle, que não permitia a coleta de comentários e *stories* (P28 e P29), por exemplo.

Como ocorria com o Facebook, o CrowdTangle não indicava quando uma publicação havia sido removida do Instagram (P34). Ainda como faz para o Facebook, a Meta disponibiliza semestralmente informações (Meta, [S.d.]) discriminadas em nível de país sobre publicações, perfis

e comentários moderados no Instagram (P25) e um breve relatório sobre requisições de moderação e/ou acesso a dados da plataforma feitas por entes do Estado Brasileiro (P27).

A principal diferença entre as avaliações do Facebook e do Instagram no ITD reside na dimensão de consistência. No caso do Instagram, o CrowdTangle retornava links de redirecionamento para acessar as imagens presentes nas publicações identificadas e recuperadas, mas estes expiravam quase instantaneamente e não podiam ser analisados de forma sistemática, fazendo com que os dados coletados não refletissem o que era exibido na interface de usuário da plataforma (P35). Isso é especialmente problemático no caso do Instagram, já que os conteúdos publicados são majoritariamente visuais, em formatos de foto e/ou vídeo.

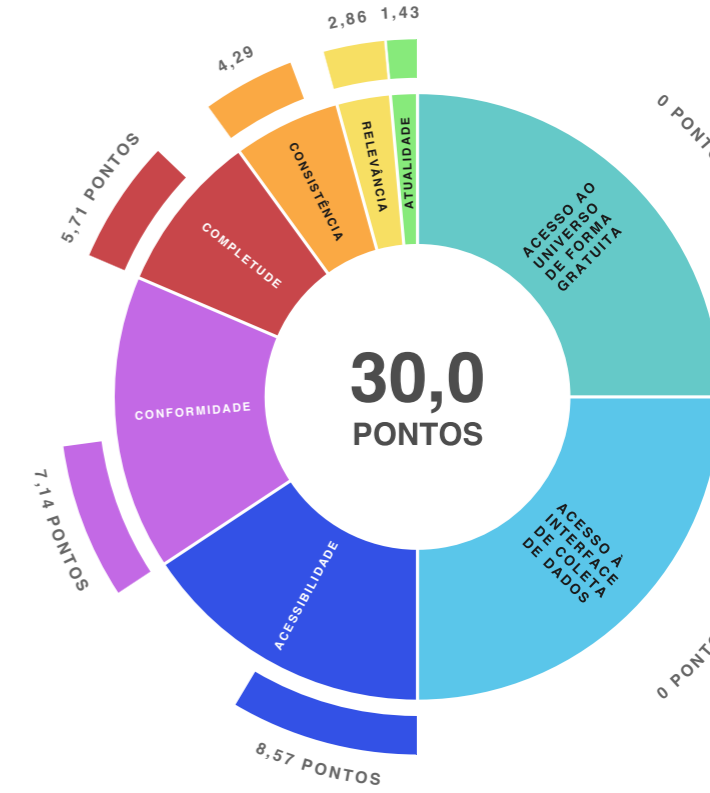
Durante a elaboração deste índice, a Meta anunciou que o CrowdTangle seria oficialmente descontinuado a partir de agosto de 2024 (CrowdTangle, 2024). Assim como observado no Facebook, caso ferramentas equivalentes de acesso a dados não sejam implementadas, a pontuação do Instagram tende a ser ainda menor em avaliações futuras.

X / Twitter

Transparência de dados: **Precária**

Legenda:

■ ACESSO AO UNIVERSO DE FORMA GRATUITA 0 DE 25 PONTOS	■ COMPLETUDE 5,71 DE 8,57 PONTOS
■ ACESSO À INTERFACE DE COLETA DE DADOS 0 DE 25 PONTOS	■ CONSISTÊNCIA 4,29 DE 5,71 PONTOS
■ ACESSIBILIDADE 8,57 DE 15,71 PONTOS	■ RELEVÂNCIA 2,86 DE 2,86 PONTOS
■ CONFORMIDADE 7,14 DE 15,71 PONTOS	■ ATUALIDADE 1,43 DE 1,43 PONTO



Marcando apenas **30 pontos**, a transparência de dados da plataforma é considerada **precária**. No passado, o X/Twitter havia se consolidado como uma das plataformas mais acessíveis para coleta de dados públicos (Zuckerman, 2021). No entanto, desde 2023, a API do X/Twitter passou a condicionar o acesso aos dados à adesão a planos pagos (X/Twitter, [S.d.]), com taxas proibitivas, sendo eles o plano *Basic*, que custa US\$100 (cerca de R\$543,00, segundo cotação de setembro de 2024) mensais, o *Pro*, que custa US\$5.000 (cerca de R\$27,5 mil, segundo cotação de setembro de 2024) mensais, e o plano *Enterprise*, que custa a partir de US\$42.000 (cerca de R\$231,4 mil, segundo cotação de setembro de 2024).

Para a avaliação da plataforma, consideramos as funcionalidades dos planos *Basic* e *Pro*, já que o X/Twitter comercializa o plano *Enterprise* apenas para empresas aprovadas mediante solicitação (X/Twitter, [S.d.]). Com a implementação destes planos pagos, o X/Twitter deixou de garantir uma forma gratuita de acesso aos dados para qualquer usuário (P3), e não garante mais qualquer forma de acesso para pesquisadores (P4). Embora alegue disponibilizar dados para pesquisadores que investigam riscos sistêmicos

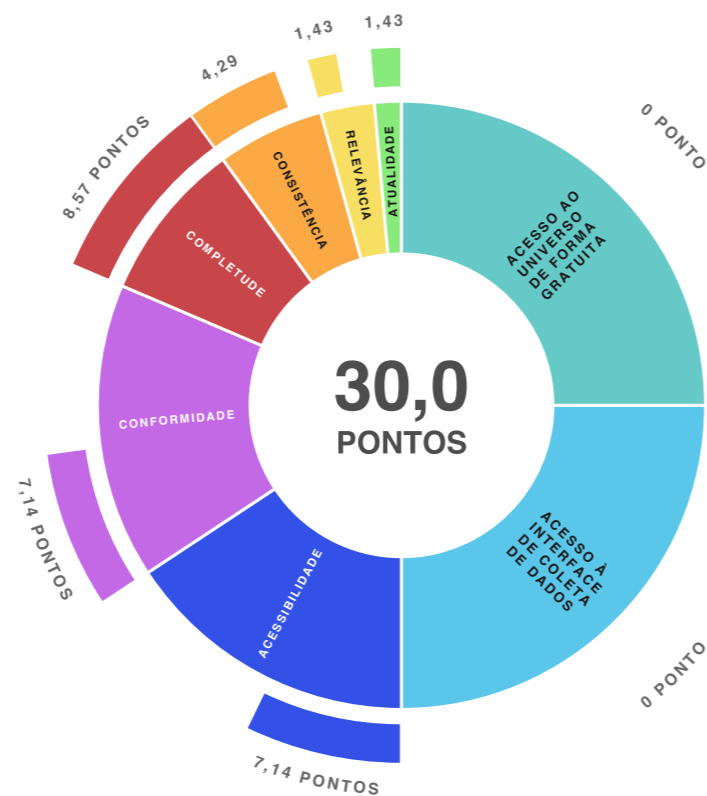
na União Europeia, a plataforma é acusada de não cumprir essa obrigação do DSA (European Commission, 2023).

Devido à alta qualidade dos dados retornados por meio da API, ainda que paga, a plataforma garante pontos nas dimensões de completude e consistência. Entre eles, o X/Twitter é a única plataforma que sinaliza um conteúdo removido adequadamente (P34). Porém, a plataforma não disponibiliza nenhum relatório com as informações sobre conteúdos removidos no Brasil ou submetidos a outras formas de moderação, como faz para outros países (P25) (X/Twitter, [S.d.]). Na Central de Transparência do X/Twitter, os últimos dados referentes ao Brasil foram publicados em 2021 e estão indisponíveis atualmente (X/Twitter, [S.d.]).

Telegram

Transparência de dados: **Precária**

Legenda:



O Telegram somou **30 pontos**, com sua transparência de dados sendo considerada **precária**. A plataforma disponibiliza uma API ([Telegram, \[S.d.\]](#)) oficial gratuita (**P1 e P3**), o que possibilita o acesso programático a dados de maneira confiável, consistente e, em grande parte, em conformidade com boas práticas. No entanto, a plataforma não permite recuperar dados de mensagens sobre um determinado tema sem que os grupos ou canais em que elas circularam sejam conhecidos previamente. Como não há maneira de buscar na API por grupos e canais públicos, o acesso ao universo de dados de interesse é limitado (**P2**), o que contribui para sua **baixa pontuação**. A plataforma também não oferece uma interface dedicada à coleta de dados (**P5**).

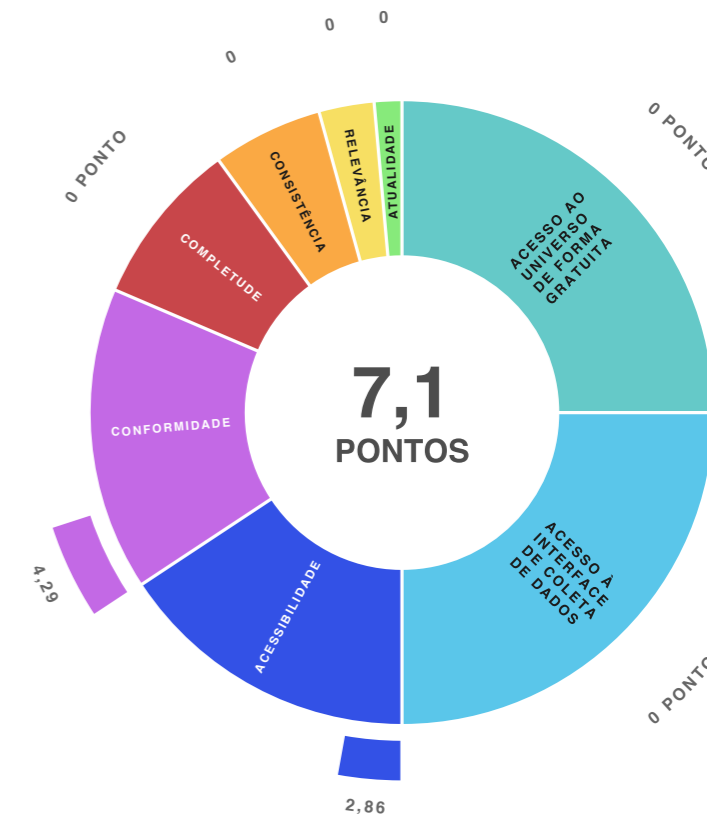
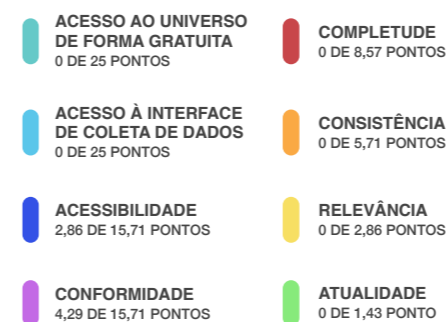
Por um lado, o Telegram é a única plataforma analisada que atende a todos os critérios de completude por permitir a coleta de dados de comentários (**P28**), dados de conteúdos temporários (**P29**) e dados históricos (**P30**), além de garantir a recuperação de um grande volume de dados sem dificuldades (**P31, P32 e P33**).

Por outro, a plataforma é muito penalizada na dimensão de conformidade. O Telegram não publica relatórios de transparência públicos e periódicos, o que impossibilita a identificação de conteúdos moderados no Brasil (**P25, P26 e P27**). A documentação de sua API também é deficiente por não definir diretamente como chamar *endpoints* (**P22**), não descrever os formatos dos dados retornados (**P18**) e não ser disponibilizada em português (**P23**).

TikTok

Transparência de dados: **Irrelevante**

Legenda:



O TikTok não oferece acesso a uma API gratuita de coleta de dados oficial para interessados no Brasil, apenas para pesquisadores registrados nos Estados Unidos e na Europa ([TikTok, \[S.d.\]](#)). Em grande parte por conta disso, a transparência de dados da plataforma alcança uma pontuação considerada **irrelevante**, de apenas **7,1 pontos**.

Além da indisponibilidade de API no país (**P1**) e da consequente impossibilidade de se acessar programaticamente todo o universo de dados públicos (**P2**), a plataforma tampouco oferece interface dedicada à coleta de dados (**P5**).

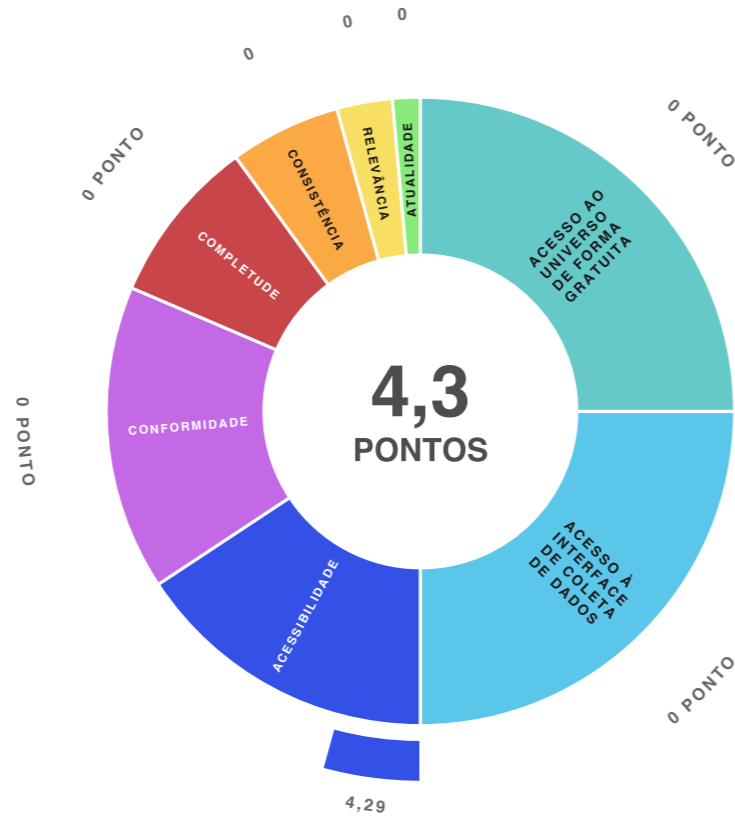
Os poucos pontos alcançados na dimensão de acessibilidade decorrem da possibilidade de raspagem de dados a partir da interface da plataforma (**P14 e P15**). No entanto, a necessidade de contornar processos de verificação por *Captcha*, empregados pela plataforma para evitar o uso automatizado, limitam substancialmente o processo de coleta de dados por meios extraoficiais (**P16**).

Por outro lado, o TikTok é a única a pontuar nos três parâmetros referentes a relatórios públicos de transparência (**P25, P26 e P27**) por disponibilizar informações sobre o volume e os tipos de violações moderadas pela plataforma ([TikTok, 2024](#)) e por apontar, por meio de relatórios semestrais, os pedidos de moderação e de requisições de dados feitos pelo Estado brasileiro, desde 2019 ([TikTok, 2023](#)).

Transparência de dados: **Irrelevante**

Legenda:

- ACESSO AO UNIVERSO DE FORMA GRATUITA 0 DE 25 PONTOS
- ACESSO À INTERFACE DE COLETA DE DADOS 0 DE 25 PONTOS
- ACESSIBILIDADE 4,29 DE 15,71 PONTOS
- CONFORMIDADE 0 DE 15,71 PONTOS
- COMPLETUDE 0 DE 8,57 PONTOS
- CONSISTÊNCIA 0 DE 5,71 PONTOS
- RELEVÂNCIA 0 DE 2,86 PONTOS
- ATUALIDADE 0 DE 1,43 PONTO



O Kwai não disponibiliza nenhuma API para coleta de dados no Brasil ou em qualquer outro lugar no mundo (P1), inviabilizando a auditabilidade do universo de postagens públicas da plataforma (P2). O Kwai tampouco disponibiliza uma interface dedicada à coleta de dados (P5). Estes fatores contribuem amplamente para os **4,3 pontos** alcançados pela plataforma na avaliação de sua transparência de dados, considerada **irrelevante**.

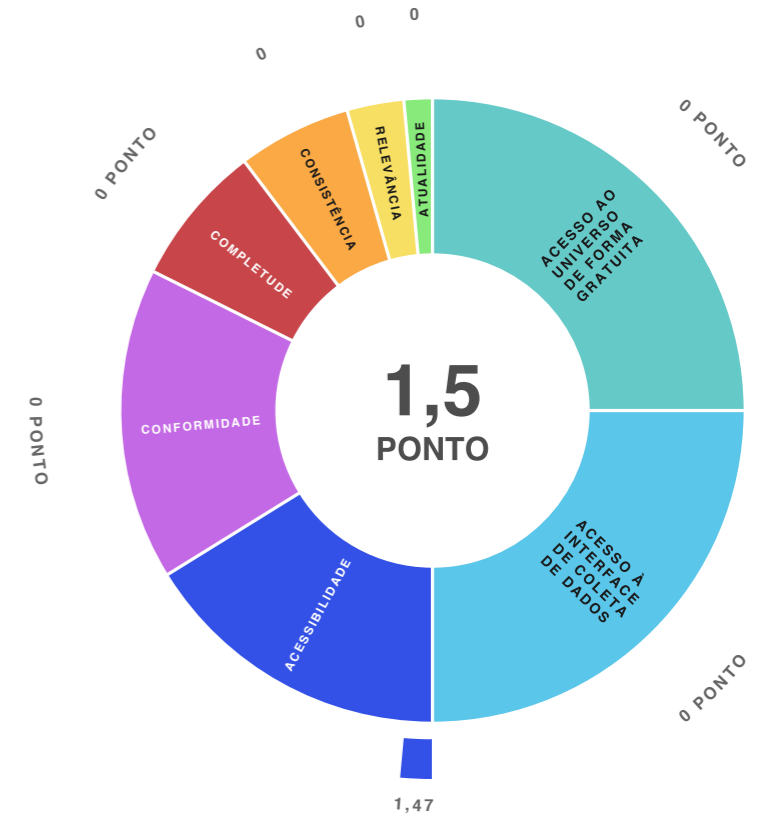
Os poucos pontos obtidos pelo Kwai na avaliação do ITD decorrem de três parâmetros da dimensão de acessibilidade, devido à possibilidade de raspagem de dados (P14, P15 e P16) sem necessidade de autenticação na interface de usuário da plataforma. No entanto, os dados extraídos são limitados e enviesados pelo baixo volume de publicações retornadas a cada busca na interface web da plataforma.

Outro ponto negativo é a ausência de relatórios de transparência dedicados a publicações removidas e usuários suspensos no Brasil (P25, P26 e P27), uma vez que o Kwai disponibiliza informações de transparência apenas no nível do continente (Kwai, [S.d.]).

Transparência de dados: **Irrelevante**

Legenda:

- ACESSO AO UNIVERSO DE FORMA GRATUITA 0 DE 25 PONTOS
- ACESSO À INTERFACE DE COLETA DE DADOS 0 DE 25 PONTOS
- ACESSIBILIDADE 1,47 DE 16,8 PONTOS
- CONFORMIDADE 0 DE 16,18 PONTOS
- COMPLETUDE 0 DE 7,35 PONTOS
- CONSISTÊNCIA 0 DE 5,88 PONTOS
- RELEVÂNCIA 0 DE 2,94 PONTOS
- ATUALIDADE 0 DE 1,47 PONTO



O WhatsApp aparece em último lugar na avaliação do ITD, apresentando uma transparência de dados considerada **irrelevante** e apenas 1,5 ponto⁴ conquistado na dimensão de acessibilidade. A plataforma não disponibiliza API oficial (P1), ao contrário de seu concorrente Telegram, e tampouco interface dedicada à coleta de dados (P5).

O WhatsApp teve resposta positiva em apenas um dos parâmetros de avaliação, relativo ao processo de raspagem de dados (P16). De todo modo, ainda que seja possível obter dados pela raspagem, é impossível ter acesso total ao universo de mensagens de interesse no WhatsApp, visto que a coleta se limita a uma amostra de grupos públicos previamente selecionados pelos pesquisadores.

A falta de transparência, influenciada também pela ausência de relatórios públicos que apresentem as ações de moderação empreendidas pela plataforma (P25), acende um sinal de alerta, uma vez que o WhatsApp é visto como uma das principais plataformas para a disseminação

de desinformação no Sul Global, especialmente no Brasil (Kalogeropoulos; Rossini, 2023).

⁴Para o cálculo final da nota do WhatsApp, desconsideramos o parâmetro “É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?”, referente à dimensão de completude, uma vez que a plataforma não apresenta suporte para comentários em suas mensagens. Portanto, além dos dois critérios especiais, consideramos 34 dos 35 parâmetros aplicáveis restantes e readequamos todos os demais cálculos.

Boas e Más Práticas na Disponibilização de Dados

A partir das evidências observadas na análise de cada plataforma, apresentamos um panorama de medidas, **dividido por dimensões de qualidade de dados**, que devem ser amplamente adotadas ou evitadas por elas a fim de garantir um **nível ideal de transparência e disponibilização de dados para pesquisa de interesse público**.

Boas Práticas Que Podem Ser Replicadas

Acessibilidade:

O **YouTube** é a única plataforma analisada que garante o **acesso integral ao universo de dados públicos por meio de uma API gratuita**. Qualquer usuário da API pode resgatar todos os dados principais dos vídeos, incluindo identificador, título, descrição, autor e comentários, consagrando a plataforma como a mais acessível entre as avaliadas.

Uma boa prática adotada por **YouTube, X/Twitter, Facebook e Instagram** é a possibilidade de recuperação programática de dados por meio de termos de busca específicos.

É possível coletar dados do **YouTube, WhatsApp e Telegram** por meio de técnicas de raspagem via interface de usuário da plataforma, de forma relativamente descomplicada.

Conformidade:

O **YouTube** disponibiliza a documentação da API traduzida para português e descreve clara e acessivelmente seus termos de uso, sem redirecionamentos para outras páginas.

Embora não disponibilize meios oficiais e programáticos para coleta de dados, o **TikTok** divulga relatórios de transparência sobre ações de moderação empreendidas no Brasil, com grande nível de detalhe.

Completude:

YouTube e **Telegram** oferecem API oficial para coleta de dados de diferentes tipos, como **comentários, publicações temporárias e dados históricos**.

Essas plataformas também possibilitam uma coleta massiva e célere de dados, com destaque para o **Telegram**, a única que atende a todos os parâmetros dessa dimensão.

Consistência:

As APIs do **Telegram, X/Twitter e Facebook** seguem os filtros aplicados por pesquisadores nas requisições. Requisições feitas às APIs em momentos distintos ou por usuários diferentes recuperaram dados praticamente idênticos, como esperado.

Relevância:

Além de oferecer filtros eficientes de coleta em sua API, como idioma e intervalo de tempo, o **X/Twitter** retorna dados que permitem mapear as **relações entre diferentes conteúdos**, incluindo compartilhamentos e respostas. Ela é a única plataforma que atende aos dois parâmetros dessa dimensão.

Más Práticas Que Devem Ser Evitadas

Acessibilidade:

TikTok, Kwai e WhatsApp não disponibilizam API oficial para coleta de dados públicos de suas publicações no Brasil, embora o **TikTok** o faça nos Estados Unidos e na Europa.

X/Twitter, Facebook, Instagram e Telegram limitaram a quantidade de *tokens* que podem ser criados, dificultando processos paralelos de coleta. Com exceção do **Telegram**, essas plataformas também atuam ativamente para restringir a raspagem de dados.

Conformidade:

X/Twitter, Facebook, Instagram e Telegram não disponibilizaram as documentações de suas APIs em acesso aberto e completamente traduzidas para o português. O **Telegram** se destaca negativamente pela ausência de exemplos claros para sua utilização e pela má redação da documentação.

X/Twitter, Telegram, WhatsApp e Kwai não divulgam relatórios de transparência específicos para o Brasil.

Atualidade:

YouTube, Telegram, X/Twitter, Facebook e Instagram, as plataformas que, no momento da análise, ofereciam APIs oficiais no Brasil para coleta de dados, disponibilizaram dados em tempo (quase) real para coleta e análise.

Completude:

As APIs de coleta de dados do **Facebook** e do **Instagram** não permitiam a exportação de mais de 10 mil publicações por vez, prejudicando a coleta de grandes volumes de dados.

Além disso, **dados de comentários e conteúdos temporários** eram inacessíveis nas duas plataformas.

Consistência:

A API do **YouTube** apresenta **resultados inconsistentes** para coletas com parâmetros idênticos entre si. Os conjuntos de dados recuperados em buscas idênticas, realizadas em um curto intervalo de tempo durante os testes, foram consideravelmente diferentes entre si e não obedeceram aos critérios definidos nas requisições. Nenhuma das diferenças percebidas nos testes ocorreu pela inclusão de dados mais recentes.

Relevância:

Apesar de disponibilizarem APIs oficiais de coleta, **Telegram** e **YouTube** possuem limitações relacionadas à aplicação de filtros. O **YouTube** apresenta uma possibilidade de filtrar vídeos por idiomas mas que, na prática, não funciona. Inclusive, a documentação da API reconhece que ela **pode retornar vídeos em qualquer idioma** caso os julgue relevantes para uma busca, segundo critérios que não ficam claros aos usuários. Já o **Telegram** não permite a filtragem de resultados por idioma.

Atualidade:

TikTok, Kwai e WhatsApp não disponibilizam API oficial para coleta de dados públicos de suas publicações no Brasil.

Recomendações

Os principais pontos de melhoria na transparência das plataformas se enquadram em **quatro eixos**:

01

Possibilidade de Coleta do Universo de Dados Públicos

Quanto à **possibilidade de coleta dos dados**, os **casos mais problemáticos são do TikTok, Kwai e WhatsApp**. Recomenda-se a **disponibilização de uma API oficial e gratuita** para o acesso ao universo de dados públicos das plataformas para pesquisadores. Além disso, é recomendada a oferta de uma **interface dedicada à coleta de dados**, com fácil uso e navegação e sem exigência de conhecimentos técnicos em programação. A maior parte dos parâmetros do ITD sequer são aplicáveis às três plataformas porque a impossibilidade de coletar dados impede a avaliação de outros parâmetros de qualidade e transparência.

Embora oferecessem, à época da análise, API gratuita e interface para a coleta, **Facebook** e **Instagram** também

precisavam, mesmo antes do encerramento do CrowdTangle, reduzir as restrições, disponibilizando **acesso integral e gratuito** ao seu universo de dados públicos. No caso do **YouTube** e do **Telegram**, ainda que haja amplo acesso aos dados via API, seria importante oferecer uma **interface de coleta de dados** de fácil utilização. Quanto ao **X/Twitter**, é fundamental que volte a oferecer uma **API gratuita**, ao menos para pesquisadores brasileiros, para aprimorar suas práticas de transparência.

Um outro fator a ser melhorado em todas as plataformas é a **permissão oficial da raspagem de dados** para fins de **pesquisa**, oferecendo mais segurança jurídica e facilitando o processo para a análise sistemática para quem utiliza esse método.

02

Qualidade dos Dados Disponibilizados

Quanto à **qualidade dos dados disponibilizados**, todas as APIs gratuitas apresentam problemas em diferentes dimensões avaliadas. No caso do **YouTube**, é preciso corrigir diversas **inconsistências** nos resultados entregues, assim como **inco-**

erências entre os resultados e os parâmetros de busca. Além disso, quase todas as plataformas precisam melhorar o acesso a dados de **conteúdos temporários**, como os *stories* do Instagram e do Facebook, por exemplo. No caso do

Instagram, a situação é ainda mais crítica, já que os *stories* são frequentemente priorizados pelos usuários ao publicar. É preciso dar acesso a esses dados, além de disponibilizar **links que não expirem** para possibilitar a análise de conteúdos permanentes em foto e vídeo.

Também é importante que as plataformas sinalizem quando **publicações são removidas e usuários suspensos**, oferecendo acesso aos **metadados**, ainda que o conteúdo da publicação fique restrito. Esse tipo de dado é essencial para pesquisas sobre a circulação de **desinformação e conteúdos ilegais ou**

abusivos nas plataformas de redes sociais, assim como para compreender as práticas de **moderação** das plataformas.

Vale destacar, também, que a **busca customizada por palavra-chave** é fundamental para a coleta de dados relevantes para os fins da pesquisa, recurso que não é oferecido pelo Telegram para todo o universo de dados públicos, apenas para grupos previamente conhecidos.

observados por Facebook, Instagram, X/Twitter e Telegram. No caso do Telegram, também é preciso aprimorar as instruções sobre o uso da API oficial, que não são **claras e exemplificadas**, nem descrevem o formato das respostas esperadas com o uso de cada *endpoint*.

cífico no Brasil, impossibilitando análises mais relevantes e aprofundadas das políticas de moderação e governança das plataformas. Algumas plataformas, como **X/Twitter, Telegram, Kwai e WhatsApp** sequer publicam esse tipo de documento, agravando ainda mais o problema.

Quando à **documentação da API**, é importante que seja disponibilizada **publicamente**, sem necessidade de requisição individual, em **língua portuguesa**, e com descrição clara de seus **termos de uso**. Alguns destes aspectos, que ajudariam a **democratizar** o acesso aos recursos existentes, não foram

Quando aos **relatórios de transparência**, o ponto mais crítico reside na disponibilização do documento de forma **periódica**. Mesmo quando os relatórios são disponibilizados com alguma regularidade, seus dados **apresentam baixa qualidade** e são entregues com granularidade insatisfatória, normalmente sem foco espe-

Apêndice I: Visão Geral da Avaliação

✓ Indica uma avaliação positiva do parâmetro

✗ Indica uma avaliação negativa do parâmetro

— Indica não aplicabilidade do parâmetro à plataforma avaliada

Acessibilidade		Critério Especial 1							
Parâmetros de avaliação		YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
P1	A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários?	✓	✓	✓	✓	✓	✗	✗	✗
P2	O universo a ser monitorado é recuperável pela API da plataforma?	✓	✗	✗	✓	✗	✗	✗	✗
P3	O acesso à API da plataforma é gratuito?	✓	✓	✓	✗	✓	✗	✗	✗
P4	A plataforma oferece a pesquisadores acesso gratuito e específico à API?	✓	✓	✓	✗	✗	✗	✗	✗

Acessibilidade		Critério Especial 2							
Parâmetros de avaliação		YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
P5	A plataforma oferece uma interface para coletar os dados por meio de busca customizável?	✗	✓	✓	✗	✗	✗	✗	✗

03 Clareza da Documentação Sobre API e Termos de Uso

04 Disponibilização de Relatórios de Transparência Detalhados Sobre o Brasil

Outros Parâmetros de Acessibilidade

Parâmetros de avaliação								
P6 É possível extrair os dados requisitados diretamente da resposta da API da plataforma?								
P7 A API da plataforma provê uma forma de autenticação que permite a renovação automática, sem bloqueios à aquisição de dados?								
P8 A criação de <i>tokens</i> de acesso à API da plataforma pode ser feita de forma gratuita?								
P9 É possível criar novos <i>tokens</i> de acesso à API da plataforma sem limitações de quantidade?								
P10 O processo para que pesquisadores tenham acesso à API da plataforma é claro, descomplicado e com prazo bem definido?								
P11 A API da plataforma disponibiliza um <i>endpoint</i> para recuperar dados de uma publicação específica?								
P12 A API da plataforma disponibiliza um <i>endpoint</i> para recuperar dados de um autor específico?								
P13 A API da plataforma disponibiliza um <i>endpoint</i> para recuperar dados por meio de termos de busca?								
P14 É possível realizar aquisição de dados por raspagem, sem necessidade de autenticação, por meio da interface de usuário da plataforma?								

Outros Parâmetros de Acessibilidade (continuação)

Parâmetros de avaliação								
P15 É possível realizar aquisição de dados por raspagem sem necessidade de outros dispositivos?								
P16 É possível recuperar os dados por meio de raspagem, sem necessidade de contornar ferramentas e técnicas que visam impedir o acesso programático aos dados?								
P17 A estrutura dos dados disponibilizados pela API da plataforma é estável?								
P18 Os dados retornados pela API da plataforma estão em formato padronizado?								
P19 A documentação da API da plataforma é publicada em acesso aberto?								
P20 A documentação da API da plataforma está escrita de forma clara e exemplificada?								
P21 A documentação da API da plataforma descreve seus termos de uso?								
P22 A documentação descreve o formato da resposta dos <i>endpoints</i> da API da plataforma?								

Conformidade

Parâmetros de avaliação								
P17 A estrutura dos dados disponibilizados pela API da plataforma é estável?								
P18 Os dados retornados pela API da plataforma estão em formato padronizado?								
P19 A documentação da API da plataforma é publicada em acesso aberto?								
P20 A documentação da API da plataforma está escrita de forma clara e exemplificada?								
P21 A documentação da API da plataforma descreve seus termos de uso?								
P22 A documentação descreve o formato da resposta dos <i>endpoints</i> da API da plataforma?								

Conformidade (continuação)

Parâmetros de avaliação								
P23 A documentação da API da plataforma é disponibilizada nativamente em português?								
P24 A plataforma permite raspagem e outros tipos de acesso automático nos seus termos de uso?								
P25 A plataforma produz relatórios de transparência periódicos sobre a moderação de conteúdos no Brasil e os disponibiliza publicamente, sem necessidade de requisição?								
P26 Em seus relatórios de transparência, a plataforma indica o volume de cada tipo de violação identificada no Brasil de acordo com as políticas de moderação vigentes?								
P27 Os relatórios de transparência especificam informações sobre a quantidade e o tipo de requisições feitas por entes do Estado brasileiro à plataforma, além da quantidade e do tipo de solicitações acatadas?								

Completude

Parâmetros de avaliação								
P28 É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?								

Completude (continuação)

Parâmetros de avaliação								
P29 É possível recuperar dados de conteúdos temporários por meio da API da plataforma?								
P30 É possível recuperar dados históricos por meio da API da plataforma?								
P31 A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 1 milhão de publicações em 24 horas?								
P32 A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 100 mil publicações em 24 horas?								
P33 A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 10 mil publicações em 24 horas?								

Consistência

Parâmetros de avaliação								
P34 Os dados retornados pela API da plataforma são persistentes?								
P35 Os dados recuperados pela API da plataforma refletem o que é exibido em sua interface de usuário?								
P36 A resposta retornada pela API da plataforma é sempre a esperada?								
P37 A resposta retornada pela API da plataforma é coerente com os parâmetros e filtros utilizados na requisição?								

Relevância								
Parâmetros de avaliação								
P38 As entidades retornadas pela API da plataforma são suficientes para compreender os dados em todos os seus níveis de detalhes?								
P39 A API da plataforma permite a utilização de filtros para refinar a requisição de dados?								

Atualidade								
Parâmetros de avaliação								
P40 É possível recuperar dados dados recém-publicados, quase em tempo real à publicação, por meio da API da plataforma?								

Apêndice II: Detalhamento por Parâmetro de Avaliação



P1

A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários? (Critério especial 1)

Aqui, observamos se a plataforma oferece uma API com ao menos um *endpoint* para acesso aos dados de conteúdos gerados por usuários.

YouTube

Desde 2014, a plataforma opera com a API YouTube Data v3. Toda a documentação e instruções para uso da API oficial do YouTube podem ser acessadas [aqui](#).

Facebook

Para o acesso e a extração de dados públicos do Facebook e do Instagram, a Meta oferecia, no momento da análise, a ferramenta CrowdTangle, que, além de apresentar uma interface de usuário para este fim, também **disponibilizava uma API**. No entanto, a Meta **encerrou a ferramenta em agosto de 2024** e a substituiu pela *Meta Content Library*, que não oferece recursos equivalentes.

Instagram

Para o acesso e a extração de dados públicos do Facebook e do Instagram, a Meta oferecia, no momento da análise, a ferramenta CrowdTangle, que, além de apresentar uma interface de usuário para este fim, também **disponibilizava uma API**. No entanto, a Meta **encerrou a ferramenta em agosto de 2024** e a substituiu pela *Meta Content Library*, que não oferece recursos equivalentes.

X / Twitter

A **v2 da API do X/Twitter** apresenta planos de acesso pago com disponibilização de *endpoints* para a recuperação de dados de publicações públicas, ainda que com taxas mensais proibitivas para a coleta de dados de publicações.

Telegram

A plataforma **oferece uma API** que permite a coleta de dados públicos em conversas do Telegram.

TikTok

O TikTok possui **API para pesquisa acadêmica**, mas o acesso aos dados públicos dos usuários só está disponível para pesquisadores dos Estados Unidos e da Europa.

Kwai

O Kwai não disponibiliza nenhuma API oficial para coleta de dados em nenhum lugar do mundo.

WhatsApp

O WhatsApp não disponibiliza uma API para acesso a dados para pesquisa em nenhum lugar do mundo.



P2

O universo a ser monitorado é recuperável pela API da plataforma? (Critério especial 1)

Neste campo, busca-se verificar se a plataforma permite a descoberta e a coleta de dados, de forma programática, de todo o conjunto de publicações públicas de interesse.

YouTube

O YouTube permite buscar todo o universo de vídeos públicos disponíveis via API. Dessa forma, entendemos que é possível, por meio da API oficial, estruturar a coleta programática de vídeos, buscando por quaisquer canais ou vídeos públicos.

Facebook

Embora disponibilizasse um volume considerável de dados, o CrowdTangle limitava o universo monitorável do Facebook às publicações públicas de perfis verificados, de páginas com mais de 25 mil seguidores ou curtidas e de grupos públicos com mais de 95 mil membros. Segundo **dados mais recentes**, isso engloba cerca de 7 milhões de páginas, perfis e grupos no mundo todo.

Instagram

Embora disponibilizasse um volume considerável de dados, o CrowdTangle limitava o universo monitorável do Instagram às publicações públicas de perfis com mais de 50 mil seguidores e de todos os perfis verificados pela plataforma. Segundo **dados mais recentes**, isso engloba cerca de 2 milhões de páginas e perfis no mundo todo. O CrowdTangle também **não permitia a coleta de dados gerados a partir de reels**, vídeos curtos introduzidos pela plataforma para concorrer com o TikTok.

X / Twitter

Apesar de paga, a API do X/Twitter permite acompanhar todo o universo de publicações públicas em torno de um determinado tema de interesse, inclusive em tempo (quase) real. A plataforma só exclui as atividades de perfis privados que não são seguidos pelo usuário que está fazendo a busca, conforme estabelece sua **Política de Privacidade**.

Telegram

A coleta é possível apenas dentro de determinado grupo ou canal, sendo necessário o conhecimento prévio daqueles a serem coletados. Assim, é impossível localizar e recuperar todo o conteúdo considerado público disponível na plataforma.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P3

O acesso à API da plataforma é gratuito? (Critério especial 1)

Verificamos se a plataforma oferece acesso geral à sua API de forma gratuita.

YouTube

O uso da API do YouTube é limitado por **cotas diárias de dados** para cada projeto criado no Google Cloud, mas a obtenção de **autorização de credenciais para o acesso** é totalmente gratuita.

Facebook

Até seu encerramento, em agosto de 2024, o CrowdTangle disponibilizava acesso gratuito à sua API aos parceiros credenciados, que incluíam jornalistas, pesquisadores e *fact-checkers*.

Instagram

Até seu encerramento, em agosto de 2024, o CrowdTangle disponibilizava acesso gratuito à sua API aos parceiros credenciados, que incluíam jornalistas, pesquisadores e *fact-checkers*.

X / Twitter

Atualmente, o X/Twitter apenas oferece o acesso pago à sua API de coleta de dados, dividido em diferentes faixas de preço.

Telegram

Não é preciso pagar para utilizar a API, como esclarece a **página inicial da documentação**.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P4

A plataforma oferece a pesquisadores acesso gratuito e específico à API? (Critério especial 1)

Neste campo, é verificado se a API da plataforma disponibiliza algum tipo de acesso a pesquisadores, seja na forma de um *token* específico ou de *endpoints* exclusivos.

YouTube

O **YouTube Researcher Program** é uma opção de registro para acesso à API YouTube Data v3 direcionada especificamente para pesquisadores.

Facebook

O CrowdTangle permitia em seus termos o acesso gratuito para **pesquisadores e estudantes** vinculados a instituições acadêmicas.

Instagram

O CrowdTangle permitia em seus termos o acesso gratuito para **pesquisadores e estudantes** vinculados a instituições acadêmicas.

X / Twitter

Há apenas acesso privilegiado e gratuito para pesquisadores vinculados a instituições acadêmicas da União Europeia, por conta de imposições legais do DSA — ainda que a norma **não esteja sendo seguida** de forma adequada por lá. Até o momento, a plataforma não anunciou planos de expandir o acesso ao resto do mundo.

Telegram

O Telegram não garante acesso específico a pesquisadores a seus dados públicos. No entanto, o acesso à API é garantido a qualquer interessado, o que é levado em consideração para a resposta do critério especial.

TikTok

A API do TikTok está disponível apenas para pesquisadores dos Estados Unidos e da Europa.

Kwai

O Kwai não oferece nenhuma API oficial para coleta de dados, tampouco para uso de pesquisadores.

WhatsApp

O WhatsApp não oferece nenhuma API oficial para coleta de dados, tampouco para uso de pesquisadores.



P5

A plataforma oferece uma interface para coletar os dados por meio de busca customizável? (Critério especial 2)

Neste campo, é verificado se a plataforma oferece uma interface de observação e coleta de dados direcionada a pessoas sem conhecimento técnico de programação.

YouTube

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface gráfica para usuários não especializados.

Facebook

Além da API, o CrowdTangle oferecia uma interface composta de *dashboards*, por meio dos quais era possível testar e elaborar buscas específicas, acompanhar a evolução de discussões na plataforma, analisar o engajamento de páginas de interesse e solicitar a recuperação de dados de publicações.

Instagram

Além da API, o CrowdTangle oferecia uma interface composta de *dashboards*, por meio dos quais era possível testar e elaborar buscas específicas, acompanhar a evolução de discussões na plataforma, analisar o engajamento de páginas de interesse e solicitar a recuperação de dados de publicações.

X / Twitter

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface para usuários não especializados.

Telegram

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface para usuários não especializados.

TikTok

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface para usuários não especializados.

Kwai

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface para usuários não especializados.

WhatsApp

A plataforma não oferece nenhum método oficial de coleta de dados por meio de uma interface para usuários não especializados.



P6

É possível extrair os dados requisitados diretamente da resposta da API da plataforma?

Neste campo, verificamos se a API retorna dados estruturados como resposta ao pedido, ao invés de entregar um link que redireciona para os dados. Dados de mídias audiovisuais, como arquivos de imagens, vídeos e áudio, não foram considerados para a avaliação deste parâmetro.

YouTube

É possível coletar diretamente **todos os dados principais**, incluindo título, descrição, ID, autor, comentários e link para o conteúdo audiovisual, disponíveis na API.

Facebook

As respostas das requisições à API do CrowdTangle traziam os dados esperados em formato adequado, sem necessidade de redirecionamento.

Instagram

As respostas das requisições à API do CrowdTangle traziam os dados esperados em formato adequado, sem necessidade de redirecionamento.

X / Twitter

Nos *endpoints* de recuperação de publicações, os dados são retornados na resposta da requisição.

Telegram

As respostas das requisições à API do Telegram trazem os dados esperados em formato adequado, sem necessidade de redirecionamento.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P7

A API da plataforma provê uma forma de autenticação que permite a renovação automática, sem bloqueios à aquisição de dados?

Neste campo, é avaliado se os *tokens* disponibilizados para o uso da API expiram e se a renovação dos mesmos pode ser feita de forma automática.

YouTube

A coleta não exige renovação de *tokens*, apenas se limita às cotas diárias para cada projeto credenciado pelo usuário.



Instagram

Os *tokens* gerados pelo CrowdTangle não expiravam e a renovação ficava a cargo dos usuários, que podiam gerar novos *tokens* a qualquer momento.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Facebook

Os *tokens* gerados pelo CrowdTangle não expiravam e a renovação ficava a cargo dos usuários, que podiam gerar novos *tokens* a qualquer momento.



X / Twitter

Salvo em situações de falta de pagamento, o X/Twitter não bloqueia *tokens* já existentes e o usuário pode renová-los sempre que desejar.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

Apesar do Telegram prever a atualização automática de *tokens*, a experiência nos testes realizados foi diferente. Houve inúmeros bloqueios à conta e, portanto, ao *token*, sem que o Telegram esclarecesse os motivos que levaram a isso. Ademais, até a data de fechamento deste relatório, a plataforma não havia respondido às solicitações de contato para esclarecimentos via e-mail enviadas ao longo de 2022 e 2023.



P8

A criação de *tokens* de acesso à API da plataforma pode ser feita de forma gratuita?

Neste campo, foi avaliado se a plataforma permite a criação de novos *tokens* de forma gratuita, sem a necessidade de cadastramento de mais de uma conta para a utilização da API.

YouTube

O registro para acesso à API pode ser realizado por qualquer usuário de forma gratuita.



Instagram

O CrowdTangle não apresentava quaisquer funcionalidades pagas, inclusive a geração de *tokens* de acesso à API.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Facebook

O CrowdTangle não apresentava quaisquer funcionalidades pagas, inclusive a geração de *tokens* de acesso à API.



X / Twitter

Com os novos planos pagos da API do X/Twitter, não é mais possível criar novos *tokens* de acesso a dados da plataforma gratuitamente. É apenas possível criar *tokens* gratuitos para publicação, mas não para a coleta de publicações de forma programática.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

O Telegram prevê a atualização automática de *tokens* de forma gratuita.





P9

É possível criar novos *tokens* de acesso à API da plataforma sem limitações de quantidade?

Aqui, é verificado se a plataforma limita a quantidade de *tokens* de acesso à API que podem ser criados por um mesmo usuário.

YouTube

Apesar do YouTube não utilizar um sistema baseado em renovação de *tokens*, o acesso aos dados é feito por uma via única e limitada à cota padrão de 10.000 unidades por dia para cada projeto no Google Cloud. Diferentes desenvolvedores devem compartilhar a cota de um mesmo projeto, cadastrando suas contas Google individuais.



Instagram

O CrowdTangle permitia a utilização de apenas um *token* de acesso à API para cada equipe de trabalho. Assim, era possível renovar o *token* mas cada equipe só podia usar um, e seus membros deviam compartilhá-lo, sendo impossível a utilização simultânea por mais de um usuário.



Telegram

O Telegram só permite um *token* de coleta de dados por conta, por isso, é necessário ter mais de um número de celular caso a pessoa necessite expandir o volume da coleta. Em nossos testes, passamos por experiências de banimentos não justificados e não previstos nos termos de uso da API, o que justifica a necessidade de diversificação da coleta em diferentes contas por parte de pesquisadores e outros interessados.



Facebook

O CrowdTangle permitia a utilização de apenas um *token* de acesso à API para cada equipe de trabalho. Assim, era possível renovar o *token* mas cada equipe só podia usar um, e seus membros deviam compartilhá-lo, sendo impossível a utilização simultânea por mais de um usuário.



X / Twitter

Com as novas regras da API do X/Twitter, é possível ter apenas um *token* por aplicativo. Enquanto o plano *Free* (gratuito) permite a criação de um aplicativo, que não pode ser destinado à coleta de dados, o plano *Pro* (pago) permite a criação de três aplicativos.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P10

O processo para que pesquisadores tenham acesso à API da plataforma é claro, descomplicado e com prazo bem definido?

Aqui, se verifica se o processo para solicitação de acesso à API exclusivo a pesquisadores é bem descrito pela plataforma. Avaliamos se a plataforma deixa clara a documentação necessária para comprovar vínculo institucional e o tempo para que o acesso seja concedido, em caso de aprovação da solicitação do pesquisador.

YouTube

O acesso a pesquisadores é descrito [aqui](#). Para solicitá-lo, é necessário preencher um formulário detalhando as atribuições do pesquisador na universidade e as informações sobre o departamento ao qual está vinculado.

Muitas informações são solicitadas, e com muitos detalhes, mas há links para esclarecimentos nos termos de serviço sempre que necessário.



Instagram

O CrowdTangle não liberava acesso à API a **novos solicitantes** desde 2022 e a Meta encerrou a ferramenta em agosto de 2024.



X / Twitter

O X/Twitter não oferece acesso a dados voltado especificamente a pesquisadores brasileiros.



Facebook

O CrowdTangle não liberava acesso à API a **novos solicitantes** desde 2022 e a Meta encerrou a ferramenta em agosto de 2024.



Telegram

O Telegram não oferece acesso a dados voltado especificamente a pesquisadores.



TikTok

O TikTok não oferece acesso a dados voltado especificamente a pesquisadores brasileiros.



Kwai

O Kwai não oferece nenhuma API oficial para coleta de dados, tampouco para uso de pesquisadores.



WhatsApp

O WhatsApp não oferece nenhuma API oficial para coleta de dados, tampouco para uso de pesquisadores.





P11

A API da plataforma disponibiliza um *endpoint* para recuperar dados de uma publicação específica?

Neste campo, verifica-se a possibilidade de recuperar dados de uma publicação pública específica na plataforma em questão, por meio de um identificador único e não necessariamente por termos de busca ou outros parâmetros e filtros.

YouTube

É possível coletar dados de um único vídeo a partir do seu **registro de identificação**.

Facebook

A API do CrowdTangle tinha o *endpoint* **GET /post/:id**, que permitia recuperar publicações específicas, quando disponibilizadas na API, a partir de seus identificadores únicos ou dos identificadores de seus autores.

Instagram

A API do CrowdTangle tinha o *endpoint* **GET /post/:id**, que permitia recuperar publicações específicas, quando disponibilizadas na API, a partir de seus identificadores únicos ou dos identificadores de seus autores. Diferentemente do Facebook, que disponibilizava estes identificadores a qualquer usuário da plataforma, o Instagram exigia um cadastro em sua **Graph API**, separada da API do CrowdTangle, para que as informações fossem fornecidas, o que complexificava, desnecessariamente, a tarefa de recuperação de dados.

X / Twitter

É possível fornecer uma lista de IDs de publicações específicas a serem coletadas por meio do *endpoint* **Tweets Lookup** da API da plataforma.

Telegram

É possível, desde que se conheça o identificador de um grupo ou canal público, recuperar os dados de uma mensagem específica por meio de seu identificador único.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P12

A API da plataforma disponibiliza um *endpoint* para recuperar dados de um autor específico?

Neste campo, é verificado se a API permite a recuperação de dados de publicações públicas feitas na plataforma em questão por um autor específico, por meio de seu nome de usuário ou identificador único.

YouTube

A API da plataforma permite coletar dados dos vídeos publicados por **um único canal** sem grandes dificuldades.

Facebook

Identificadores de usuários específicos, quando com mais de 25 mil seguidores, figuravam entre os parâmetros possíveis para recuperar publicações por meio dos *endpoints* **GET /posts** e **GET /posts/search** da API do CrowdTangle.

Instagram

Os identificadores de usuários específicos, quando com mais de 50 mil seguidores, figuravam entre os parâmetros possíveis para recuperar publicações por meio dos *endpoints* **GET /posts** e **GET /posts/search** da API do CrowdTangle.

X / Twitter

A expressão de busca sobre um determinado tema pode ser construída com a indicação do(s) autor(es) das publicações a serem coletadas. Também é possível recuperar publicações e informações de uma conta específica por meio do *endpoint* **Users Lookup**. Além disso, o *endpoint* **Timeline** permite recuperar toda a linha do tempo de publicações de um usuário específico.

Telegram

O Telegram não oferece um *endpoint* dedicado à coleta de mensagens enviadas por um autor específico. Além disso, embora seja possível identificar as mensagens de um autor em um determinado grupo ou canal, é impossível rastrear todas as mensagens públicas enviadas por ele na plataforma.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P13

A API da plataforma disponibiliza um endpoint para recuperar dados por meio de termos de busca?

Este campo verifica se é possível recuperar dados de publicações públicas da plataforma em questão por meio de termos de busca, ou seja, montar uma base de dados com publicações mencionando esses termos.

YouTube

A API YouTube Data v3 disponibiliza a opção *Search* para recuperar dados de vídeos que correspondam a termos ou expressões de busca específicos.

Facebook

A busca por termos era um dos parâmetros possíveis para a recuperação de dados de publicações públicas disponíveis na API, com os endpoints *GET /posts* e *GET /posts/search* da API do CrowdTangle.

Instagram

A busca por termos era um dos parâmetros possíveis para a recuperação de dados de publicações públicas disponíveis na API, com os endpoints *GET /posts* e *GET /posts/search* da API do CrowdTangle.

X / Twitter

A API do X/Twitter permite recuperar dados disponíveis na API por meio de termos de busca, com o endpoint *Search Tweets*.

Telegram

A busca por mensagens que mencionam termos de interesse no Telegram é permitida apenas em grupos e canais públicos **já conhecidos anteriormente** pela pessoa interessada. Como a busca é limitada, a plataforma não atinge o mínimo esperado no parâmetro de avaliação.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P14

É possível realizar aquisição de dados por raspagem, sem necessidade de autenticação, por meio da interface de usuário da plataforma?

Neste campo, é avaliado se é possível fazer a coleta dos dados por meio de técnicas de raspagem sem necessidade de criação de contas e logins.

YouTube

Diversas publicações científicas indicam métodos simples para a aquisição de dados do YouTube por raspagem, incluindo a transcrição automática dos vídeos, que não é disponibilizada pela API da plataforma.

Facebook

É impossível navegar de forma adequada e consistente pelo conteúdo da plataforma sem autenticação via login.

Instagram

É impossível navegar de forma adequada e consistente pelo conteúdo da plataforma sem autenticação via login.

X / Twitter

É impossível navegar de forma adequada e consistente pelo conteúdo da plataforma sem autenticação via login.

Telegram

Sem concluir a autenticação via login, é possível navegar por uma prévia limitada do conteúdo de canais públicos da plataforma, mas não de grupos públicos, o que apenas permite uma coleta parcial dos dados de interesse.

TikTok

É possível fazer a raspagem de dados sem necessidade de login. No entanto, os métodos alternativos de aquisição limitam-se à identificação de, no máximo, 1.000 vídeos por busca, cujos resultados são determinados pelo algoritmo de recomendação da plataforma.

Kwai

É possível fazer a raspagem de dados sem autenticação. No entanto, a quantidade de publicações retornadas pelo Kwai é consideravelmente reduzida. A aquisição de dados limita-se a, no máximo, algumas centenas de vídeos que podem ser descobertos pela busca na interface de usuário da plataforma, o que dificulta um monitoramento consistente.

WhatsApp

Não é possível fazer raspagem de dados sem ter uma conta no WhatsApp e estar inscrito no grupo a ser coletado.



P15

É possível realizar aquisição de dados por raspagem sem necessidade de outros dispositivos?

Esse campo verifica se há a necessidade de outros dispositivos eletrônicos (como celulares) para a realização da coleta por meios alternativos, não sendo possível realizá-la por meio da interface de usuário da plataforma em um computador.

YouTube

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos.

Instagram

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos.

TikTok

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos, ainda que o volume de dados retornados seja limitado.

Facebook

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos.

X / Twitter

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos.

Kwai

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos, ainda que o volume de dados retornados seja limitado.

Telegram

É possível realizar raspagem de dados por meio da interface de usuário da plataforma em um computador, sem necessidade de outros dispositivos.

WhatsApp

É possível realizar coleta de dados por técnicas de raspagem, mas é necessário utilizar outros dispositivos que não o computador.



P16

É possível recuperar os dados por meio de raspagem, sem necessidade de contornar ferramentas e técnicas que visam impedir o acesso programático aos dados?

Aqui, é verificado se a plataforma utiliza ferramentas como Cloudflare para impedir o acesso aos dados por meio de *paywall*, taxa de tráfego limitada e/ou bloqueio por detecção de comportamento automatizado.

YouTube

Não identificamos quaisquer bloqueios em coletas de dados por métodos extraoficiais.

Facebook

O Facebook bloqueia a atividade de usuários com base na detecção de comportamento automatizado para raspagem de dados.

Instagram

O Instagram bloqueia a atividade de usuários com base na detecção de comportamento automatizado para raspagem de dados.

X / Twitter

O X/Twitter bloqueia a atividade de usuários com base na detecção de comportamento automatizado para raspagem de dados.

Telegram

O Telegram não apresenta impedimentos para usuários realizarem a raspagem do conteúdo disponível em grupos e canais públicos.

TikTok

A plataforma solicita, periodicamente, que o usuário resolva um *Captcha* para verificar se há utilização automatizada, impedindo uma raspagem satisfatória.

Kwai

Não identificamos quaisquer bloqueios em coletas de dados por métodos extraoficiais.

WhatsApp

É possível raspar dados de grupos públicos da plataforma, sem bloqueios, desde que o aplicativo esteja funcionando em um celular.

Conformidade



P17

A estrutura dos dados disponibilizados pela API da plataforma é estável?

Neste campo, é verificada a estabilidade da estrutura das respostas retornadas pela API. Ela é considerada estável se não mudar constantemente e sem aviso prévio de, pelo menos, 30 dias, com ampla divulgação. Também foi avaliado se as mudanças na API impactam o funcionamento de aplicações integradas a elas.

YouTube

A API se mantém na versão 3 desde 2014 e, desde então, não apresentou grandes alterações sem aviso prévio. Além disso, a plataforma disponibiliza um **histórico de revisões bem documentado**.

Facebook

Como o CrowdTangle não passava por grandes atualizações desde 2021, a estrutura dos dados retornados nos pedidos à API da ferramenta vinha sendo estável até a ferramenta ser descontinuada, em agosto de 2024.

Instagram

Como o CrowdTangle não passava por grandes atualizações desde 2021, a estrutura dos dados retornados nos pedidos à API da ferramenta vinha sendo estável até a ferramenta ser descontinuada, em agosto de 2024.

X / Twitter

O X/Twitter apresentou um longo planejamento para garantir a migração da v1 para a v2 de sua API e há **documentação dedicada aos processos de migração**. Ainda que a estrutura dos dados tenha mudado consideravelmente, todos os processos foram bem documentados e exemplificados.

Telegram

Não há relatos de usuários sobre o Telegram alterar a estrutura de sua API de coleta regularmente sem aviso prévio, o que pôde ser confirmado em nossa experiência de uso.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Conformidade



P18

Os dados retornados pela API da plataforma estão em formato padronizado?

Aqui, é avaliado se os tipos de dados disponibilizados pela API correspondem ao consenso e/ou aos padrões da área como, por exemplo, o formato de datas segundo a norma ISO 8601, já que a estruturação padronizada facilita o armazenamento e a utilização dos dados.

YouTube

As datas das publicações retornadas pela API do YouTube seguem o formato ISO 8601 e as URLs são completas.

Facebook

As datas das publicações retornadas pelo CrowdTangle seguem o formato ISO 8601 e as URLs eram completas.

Instagram

As datas das publicações retornadas pelo CrowdTangle seguem o formato ISO 8601 e as URLs eram completas.

X / Twitter

Atualmente, os resultados retornados pelo X/Twitter apresentam datas segundo a norma ISO 8601 e a plataforma disponibiliza URLs em formato completo ou utilizando seu encurtador próprio (t.co).

Telegram

As datas das publicações retornadas pela API do Telegram seguem o formato ISO 8601 e as URLs são completas.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P19

A documentação da API da plataforma é publicada em acesso aberto?

Aqui, é verificado se a plataforma publica na internet a documentação para o uso de sua API, de forma aberta, sem necessidade de cadastro e login.

YouTube

A plataforma possui extensa documentação sobre a versão mais atual da API, assim como sobre as versões anteriores.

Facebook

A documentação da API do CrowdTangle pode ser acessada por qualquer usuário no GitHub, sem necessidade de autenticação, ainda que tenha sido descontinuada.

Instagram

A documentação da API do CrowdTangle pode ser acessada por qualquer usuário no GitHub, sem necessidade de autenticação, ainda que tenha sido descontinuada.

X / Twitter

Toda a documentação da API do X/Twitter pode ser acessada sem necessidade de autenticação, com exceção de detalhes sobre as permissões de seus planos pagos, disponíveis apenas para usuários pagantes.

Telegram

A documentação da API do Telegram pode ser acessada por qualquer usuário, sem necessidade de autenticação.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Instagram

Na documentação de sua API, o CrowdTangle trazia exemplos claros sobre como requisições deveriam ser feitas e as respostas esperadas para cada tipo de requisição.

X / Twitter

A documentação da API do X/Twitter apresenta exemplos claros para sua utilização, referentes aos diferentes tipos de objetos retornados e às respostas esperadas.

Telegram

A documentação da API do Telegram não deixa clara a forma de chamar os *endpoints* e os padrões adotados não são usualmente utilizados em outras APIs, tornando sua utilização dependente de soluções desenvolvidas por terceiros, como o Telethon.



P20

A documentação da API da plataforma está escrita de forma clara e exemplificada?

Aqui, é avaliado se a documentação para uso da API da plataforma está escrita de forma clara, é completa e tem exemplos de implementação.

YouTube

A documentação da API do YouTube apresenta exemplos claros para sua utilização, referentes aos diferentes tipos de objetos retornados e às respostas esperadas.

Facebook

Na documentação de sua API, o CrowdTangle trazia exemplos claros sobre como requisições deveriam ser feitas e as respostas esperadas para cada tipo de requisição.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P21

A documentação da API da plataforma descreve seus termos de uso?

Neste campo, é verificado se a documentação da API apresenta, de forma clara e sem ambiguidades, os termos para sua utilização e seus aspectos legais.

YouTube

A documentação da API do YouTube detalha bem seus termos de uso, inclusive aqueles **específicos para certas regiões do mundo**.



Instagram

Os termos de uso da API do CrowdTangle não eram claramente descritos na documentação oficial, ainda que pudessem ser encontrados em outras páginas referenciadas nela.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

A página dos **termos de uso da API** do Telegram é disponibilizada na página inicial da API da plataforma.



Facebook

Os termos de uso da API do CrowdTangle não eram claramente descritos na documentação oficial, ainda que pudessem ser encontrados em outras páginas referenciadas nela.



X / Twitter

Os termos de uso da API do X/ Twitter não estão claramente descritos na documentação oficial, ainda que possam ser encontrados em outras páginas referenciadas na documentação e no portal de desenvolvedores da plataforma.



P22

A documentação descreve o formato da resposta dos endpoints da API da plataforma?

Neste campo, é verificado se a documentação da API descreve o formato de cada resposta, incluindo exemplos e os erros possíveis.

YouTube

Há descrições e exemplos das respostas da API em diversas seções da documentação, a exemplo da **Video List**.



Instagram

A documentação da API do CrowdTangle trazia exemplos claros para cada um dos endpoints de coleta, além de ter **uma página dedicada aos erros possíveis** de cada tipo de requisição.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Facebook

A documentação da API do CrowdTangle trazia exemplos claros para cada um dos endpoints de coleta, além de ter **uma página dedicada aos erros possíveis** de cada tipo de requisição.



X / Twitter

A documentação da API disponibiliza um dicionário de dados, no qual constam exemplos das respostas esperadas para endpoints que lidam com diferentes tipos de objetos.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

A documentação oficial não descreve os tipos de dados de forma convencional e facilmente compreensível, e utiliza termos de objetos próprios.





P23

A documentação da API da plataforma é disponibilizada nativamente em português?

Aqui, é verificado se a plataforma disponibiliza a documentação de sua API em língua portuguesa, para melhor compreensão de usuários brasileiros.

YouTube

Sim, todo o site em que se encontra a documentação da API da plataforma está disponível em diversos idiomas, incluindo o português.



Instagram

Toda a documentação da API do CrowdTangle, bem como a maior parte das páginas de ajuda ao usuário, era apenas disponibilizada em inglês.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

O Telegram apenas disponibiliza a documentação de sua API em inglês.



Facebook

Toda a documentação da API do CrowdTangle, bem como a maior parte das páginas de ajuda ao usuário, era apenas disponibilizada em inglês.



X / Twitter

O X/Twitter apenas disponibiliza a documentação de sua API em inglês.



P24

A plataforma permite raspagem e outros tipos de acesso automático nos seus termos de uso?

Aqui, é verificado se a plataforma expressamente proíbe técnicas de raspagem para obtenção de dados em seus termos de uso.

YouTube

Nos **termos de uso**, o YouTube afirma não ser permitido “1) acessar, reproduzir, fazer download, distribuir, transmitir, exibir, vender, licenciar, alterar, modificar ou usar de outra forma qualquer parte do Serviço ou qualquer Conteúdo, exceto: (a) se autorizado de forma expressa pelo Serviço; ou (b) mediante uma permissão prévia por escrito do YouTube e, se aplicável, dos respectivos detentores dos direitos; [...] 3) Acessar o Serviço usando qualquer meio automatizado (como robôs, botnets ou scrapers), exceto (a) no caso de mecanismos de pesquisa públicos compatíveis com o arquivo robots.txt do YouTube; ou (b) mediante permissão prévia por escrito do YouTube”.



Instagram

Segundo a **Central de Ajuda do Instagram**, a plataforma tem equipes dedicadas à detecção e bloqueio de contas que apresentam padrões de comportamento automatizados associados à extração indevida de dados.



X / Twitter

Nas **regras sobre automação de contas**, o Twitter alega que o uso de *scripts* externos para automatizar a navegação em sua interface de usuário pode acarretar em eliminação permanente da conta de usuário.



Telegram

Na **lista de proibições e utilizações não permitidas** do Telegram, em seus termos de serviço (tanto de usuários quanto da API), a plataforma não menciona raspagem ou outras práticas que inviabilizariam a raspagem de dados.



Facebook

Segundo a **Central de Ajuda do Facebook**, a plataforma age para dificultar que atores externos coletem dados de suas plataformas via raspagem, sob a justificativa de impedir que estes lucrem indevidamente com dados de seus usuários.



TikTok

Segundo os **termos de uso da plataforma**, não é permitido o uso de *scripts* automatizados para coletar informações dos serviços disponíveis ou para interagir com eles.



Kwai

A utilização de *web crawlers* e ferramentas de mineração de dados viola os **termos de uso** do Kwai.



WhatsApp

Em seus **termos de serviço**, o WhatsApp inclui proibições que resultam em restrições ao uso de aplicações de coleta de dados, tais como fazer engenharia reversa de partes da aplicação e ter um sistema paralelo com dados extraídos da aplicação sem a autorização prévia da plataforma.





P25

A plataforma produz relatórios de transparência periódicos sobre a moderação de conteúdos no Brasil e os disponibiliza publicamente, sem necessidade de requisição?

Neste campo, é verificado se a plataforma produz e disponibiliza, publicamente e sem necessidade de requisição pelas partes interessadas, relatórios de transparência com periodicidade mínima semestral. Nestes relatórios, devem ser detalhadas informações sobre a aplicação de suas políticas de governança e as ações de moderação no Brasil, como a quantidade de publicações removidas ou restritas e/ou de usuários suspensos no país.

YouTube

O Relatório sobre Cumprimento das diretrizes da comunidade do YouTube disponibiliza, trimestralmente, o número de vídeos removidos e feitos por IPs no Brasil. O mesmo relatório disponibiliza outros dados em nível global, como o número de canais removidos, o número de remoções decorrentes de denúncias, detecção automatizada ou outras fontes. Vale destacar, porém, que este documento é referente apenas a remoções decorrentes da violação das diretrizes da plataforma: a Central de Relatórios de Transparência do Google não apresenta informações sobre todas as remoções realizadas pelo YouTube no Brasil, apesar de este tipo de informação ser disponibilizada em relatórios específicos para certos países, como Coreia do Sul e Turquia.

Facebook

Na página do Relatório de Restrições de conteúdo baseadas na legislação local, é possível baixar uma base de dados com os números gerais do Facebook para publicações, perfis, páginas e comentários com acesso restringido pela Meta no Brasil (e em outros países onde ela atua) com base na legislação do país. Vale destacar, porém, que estes dados não incluem as remoções aplicadas em decorrência de violações das diretrizes da Meta, e sua Central de Transparência não apresenta informações sobre todas as remoções e/ou outras medidas de moderação realizadas pelo Facebook no Brasil. Este tipo de informação é disponibilizado em relatórios específicos para certos países, como Coreia do Sul, Áustria, Turquia e Índia.

Instagram

Na página do Relatório de Restrições de conteúdo baseadas na legislação local, é possível baixar uma base de dados com os números gerais do Instagram para contas e publicações com acesso restringido pela Meta no Brasil (e em outros países onde ela atua) com base na legislação do país. Vale destacar, porém, que estes dados não incluem as remoções aplicadas em decorrência de violações das diretrizes da Meta, e sua Central de Transparência não apresenta informações sobre todas as remoções e/ou outras medidas de moderação realizadas pelo Instagram no Brasil. Este tipo de informação é disponibilizado em relatórios específicos para certos países, como Coreia do Sul, Áustria, Turquia e Índia.



P25

A plataforma produz relatórios de transparência periódicos sobre a moderação de conteúdos no Brasil e os disponibiliza publicamente, sem necessidade de requisição?

Neste campo, é verificado se a plataforma produz e disponibiliza, publicamente e sem necessidade de requisição pelas partes interessadas, relatórios de transparência com periodicidade mínima semestral. Nestes relatórios, devem ser detalhadas informações sobre a aplicação de suas políticas de governança e as ações de moderação no Brasil, como a quantidade de publicações removidas ou restritas e/ou de usuários suspensos no país.

X / Twitter

O X/Twitter não disponibiliza informações de transparência sobre ações de moderação de conteúdos no Brasil. Em sua Central de Transparência, a seção sobre Brasil foi atualizada pela última vez em 2022, e apenas retorna uma mensagem alertando que não há dados disponíveis.

Telegram

O Telegram não disponibiliza qualquer relatório de transparência para o Brasil. Em suas Políticas de Privacidade, a plataforma afirma que, caso compartilhe dados de usuários com autoridades mediante confirmação de suspeita de terrorismo, isso será informado em relatório de transparência semestral, a ser disponibilizado no canal próprio @transparency. De qualquer forma, o Telegram alega que isso nunca aconteceu.

TikTok

O TikTok disponibiliza um Community Guidelines Enforcement Report atualizado trimestralmente, com informações sobre o volume e o tipo de ações de moderação realizadas em todos os países. Também é possível exportar uma base de dados e filtrá-la de acordo com o país de origem dos conteúdos com violações identificadas, incluindo o Brasil.

Kwai

O Safety Center do Kwai apresenta relatórios de transparência semestrais nos quais segmenta ações de moderação referentes à América Latina e outras regiões do mundo, não permitindo identificar dados específicos sobre o Brasil.

WhatsApp

O WhatsApp não disponibiliza qualquer relatório de transparência sobre iniciativas de moderação de conteúdos e usuários no Brasil, nem em outros países.



P26

Em seus relatórios de transparência, a plataforma indica o volume de cada tipo de violação identificada no Brasil de acordo com as políticas de moderação vigentes?

Aqui, é verificado se os relatórios de transparência da plataforma apresentam informações sobre o volume de violações identificadas, separadas por tipo de violação, durante a aplicação de suas políticas de governança e ações de moderação no Brasil. Os tipos de violação podem incluir, por exemplo, a disseminação de conteúdo ilegal, discurso de ódio e informações falsas.

YouTube

A agregação por tipo de violação identificada só está disponível em nível global no **Relatório sobre Cumprimento das diretrizes da comunidade do YouTube**. Em nível nacional, apenas as **solicitações governamentais de remoção de conteúdo** são agrupadas por tipo de violação, e não o número de remoções de fato realizadas — apesar de este tipo de informação ser disponibilizada em relatórios específicos para certos países, como **Coreia do Sul** e **Turquia**.

Facebook

Nenhum documento da **Central de Transparência** apresenta o número de ações de moderação no Brasil agrupado por tipo de violação — como é disponibilizado em nível global e em relatórios específicos para certos países, como **Coreia do Sul**, **Áustria** e **Índia**.

Instagram

Nenhum documento da **Central de Transparência** apresenta o número de ações de moderação no Brasil agrupado por tipo de violação — como é disponibilizado em nível global e em relatórios específicos para certos países, como **Coreia do Sul**, **Áustria** e **Índia**.

X / Twitter

O X/Twitter não disponibiliza informações de transparência sobre ações de moderação de conteúdos no Brasil. Em **sua Central de Transparência**, a seção sobre Brasil foi atualizada pela última vez em 2022 e apenas retorna uma mensagem alertando que não há dados disponíveis.

Telegram

O Telegram não disponibiliza qualquer relatório de transparência para o Brasil. Em suas **Políticas de Privacidade**, a plataforma afirma que, caso compartilhe dados de usuários com autoridades mediante confirmação de suspeita de terrorismo, isso será informado em um relatório de transparência semestral, a ser disponibilizado no canal próprio @transparency. De qualquer forma, o Telegram alega que isso nunca aconteceu.



P26

Em seus relatórios de transparência, a plataforma indica o volume de cada tipo de violação identificada no Brasil de acordo com as políticas de moderação vigentes?

Aqui, é verificado se os relatórios de transparência da plataforma apresentam informações sobre o volume de violações identificadas, separadas por tipo de violação, durante a aplicação de suas políticas de governança e ações de moderação no Brasil. Os tipos de violação podem incluir, por exemplo, a disseminação de conteúdo ilegal, discurso de ódio e informações falsas.

TikTok

Nas bases de dados globais disponibilizadas na página **Community Guidelines Enforcement Report** do TikTok, é possível exportar uma base de dados que apresenta o número de violações identificadas para cada país onde a plataforma atua, incluindo o Brasil. Nesta base de dados, são indicados os tipos de violações cometidas, com categorias como “*Misinformation*”, “*Civic & Election Integrity*”, “*Hate Speech & Hateful Behavior*”, “*Youth Exploitation & Abuse*”, entre outras.

Kwai

O **Safety Center do Kwai** apresenta relatórios de transparência semestrais nos quais segmenta ações de moderação referentes à América Latina e outras regiões do mundo, não permitindo identificar dados específicos sobre o Brasil.

WhatsApp

O WhatsApp não disponibiliza qualquer relatório de transparência sobre iniciativas de moderação de conteúdos e usuários no Brasil, nem em outros países.



P27

Os relatórios de transparência especificam informações sobre a quantidade e o tipo de requisições feitas por entes do Estado brasileiro à plataforma, além da quantidade e do tipo de solicitações acatadas?

Este campo verifica se os relatórios de transparência produzidos pela plataforma elencam os pedidos de moderação e de entrega de dados realizados por entes do Estado brasileiro, detalhando a natureza do pedido, o total de requisições e o volume de solicitações deferidas e indeferidas.

YouTube



O YouTube integra o relatório de **solicitações governamentais de remoção de conteúdo** do Google, com outros serviços do mesmo grupo empresarial. São disponibilizados uma descrição geral e um conjunto de bases de dados que incluem o volume de solicitações por tipo, a cada semestre e para cada plataforma e país, incluindo o Brasil.

Instagram



A Meta apresenta um relatório sobre *Government Requests for User Data* que inclui requisições feitas por entes do Estado brasileiro e de outros países para a moderação e/ou acesso a dados do Instagram.

Telegram



O Telegram não disponibiliza qualquer relatório de transparência para o Brasil. Em suas **Políticas de Privacidade**, a plataforma afirma que, caso compartilhe dados de usuários com autoridades mediante confirmação de suspeita de terrorismo, isso será informado em um relatório de transparência semestral a ser disponibilizado no canal próprio @transparency. De qualquer forma, o Telegram alega que isso nunca aconteceu.

X / Twitter



O X/Twitter não disponibiliza informações de transparência sobre ações de moderação de conteúdos no Brasil. Em sua **Central de Transparência**, a seção sobre Brasil foi atualizada pela última vez em 2022 e apenas retorna uma mensagem alertando que não há dados disponíveis.

Facebook



A Meta apresenta um relatório sobre *Government Requests for User Data* que inclui requisições feitas por entes do Estado brasileiro e de outros países para a moderação e/ou acesso a dados do Facebook.



P27

Os relatórios de transparência especificam informações sobre a quantidade e o tipo de requisições feitas por entes do Estado brasileiro à plataforma, além da quantidade e do tipo de solicitações acatadas?

Este campo verifica se os relatórios de transparência produzidos pela plataforma elencam os pedidos de moderação e de entrega de dados realizados por entes do Estado brasileiro, detalhando a natureza do pedido, o total de requisições e o volume de solicitações deferidas e indeferidas.

TikTok



O *Government Removal Requests Report* do TikTok apresenta informações sobre o volume e os tipos de requisições de dados feitas por entes do Estado brasileiro e de outros países desde 2019, segmentadas por semestre.

Kwai



O *Safety Center do Kwai* apresenta relatórios de transparência semestrais nos quais segmenta ações de moderação referentes à América Latina e outras regiões do mundo, não permitindo identificar dados específicos sobre o Brasil.

WhatsApp



O WhatsApp não disponibiliza qualquer relatório de transparência sobre iniciativas de moderação de conteúdos e usuários no Brasil, nem em outros países.



P28

É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?

Aqui, verifica-se a possibilidade da recuperação de dados de comentários, incluindo o conteúdo destes, quando disponíveis na plataforma, seja junto aos dados da publicação, seja por meio de um *endpoint* específico.

YouTube

Os comentários podem ser coletados por endpoints específicos da API oficial da plataforma.

Facebook

O CrowdTangle retornava apenas a informação referente ao volume de comentários feitos em uma publicação, mas não era possível acessar o conteúdo dos comentários.

Instagram

O CrowdTangle retornava apenas a informação referente ao volume de comentários feitos em uma publicação, mas não era possível acessar o conteúdo dos comentários

X / Twitter

Há operadores específicos para a construção de expressões de busca que permitem recuperar respostas a uma publicação. Além disso, a API disponibiliza endpoints para a captura de quote tweets feitos sobre uma determinada publicação.

Telegram

Todas as mensagens enviadas em canais públicos no Telegram são retornadas pela API da plataforma, sem especificar se a mensagem seria um comentário.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

Todas as publicações em grupos de WhatsApp são tratadas como “mensagens”, sem diferenciação entre a mensagem original e a resposta a outra publicação. Assim, não há “comentários” em conversas em grupos de WhatsApp.



P29

É possível recuperar dados de conteúdos temporários por meio da API da plataforma?

Neste campo, verificamos se a API da plataforma fornece ao menos um *endpoint* para a recuperação de dados de publicações temporárias, como *stories* e mensagens temporárias.

YouTube

Não é possível publicar conteúdos temporários no YouTube.

Facebook

Como declarado na documentação oficial, o CrowdTangle não arquivava nenhum tipo de conteúdo temporário, ainda que seja possível publicar conteúdos dessa natureza no Facebook, como *stories*.

Instagram

Como declarado na documentação oficial, o CrowdTangle não arquivava nenhum tipo de conteúdo temporário, ainda que seja possível publicar conteúdos dessa natureza no Instagram, como *stories*.

X / Twitter

A v2 da API do X/Twitter disponibiliza um *endpoint* específico para a captura de metadados de spaces, ainda que não seja possível recuperar seu conteúdo em áudio.

Telegram

A API do Telegram permite coletar dados sobre stories (fotos temporárias), enquanto o conteúdo ainda está no ar.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



P30

É possível recuperar dados históricos por meio da API da plataforma?

Aqui, é avaliado se a API oferece *endpoints* que permitam indicar um período de tempo específico que abranja, ao menos, os últimos 365 dias para a coleta de dados, contados a partir do momento da requisição.

YouTube

Sim, a API permite coletar dados de quaisquer vídeos que estejam disponíveis na plataforma, a partir de períodos indicados pelo usuário.



Instagram

O CrowdTangle permitia, tanto na interface de coleta quanto na API, a busca e a extração de dados históricos de publicações disponibilizadas, a partir de períodos indicados pelo usuário.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Facebook

O CrowdTangle permitia, tanto na interface de coleta quanto na API, a busca e a extração de dados históricos de publicações disponibilizadas, a partir de períodos indicados pelo usuário.



X / Twitter

Os planos pagos analisados da API do X/Twitter apenas permitem acessar dados publicados ao longo dos sete dias anteriores à data em que eles forem requisitados.



Telegram

O Telegram não estipula limite de tempo para busca e coleta de mensagens, contanto que sejam públicas no canal ou grupo de interesse a ser coletado.



P31

A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 1 milhão publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 1 milhão de publicações em 24 horas, por meio da API da plataforma.

YouTube

De acordo com a calculadora de cotas disponibilizada na documentação da API da plataforma, não é possível realizar uma coleta nesse patamar, sejam dados de vídeos ou de comentários.



Instagram

Como afirmava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



Facebook

Como afirmava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



X / Twitter

Como o teto de publicações no plano *Pro* é de 1 milhão por mês, não seria possível fazer um monitoramento dessa escala, considerando os planos de acesso analisados.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

Com base nos testes realizados e na documentação da plataforma, não foram identificadas limitações a respeito da taxa de requisição de mensagens para a quantidade especificada.





P32

A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 100 mil publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 100 mil publicações em 24 horas, por meio da API da plataforma.

YouTube

De acordo com a calculadora de cotas disponibilizada na documentação da API da plataforma, é possível realizar a coleta desde volume de dados, sejam de vídeos ou de comentários, contanto que os IDs dos vídeos de interesse sejam conhecidos previamente.



Instagram

Como informava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



Facebook

Como informava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



X / Twitter

Como o teto de publicações no plano *Pro* é de 1 milhão por mês, seria possível fazer um monitoramento dessa escala, considerando as faixas de acesso analisadas.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

Com base nos testes realizados e na documentação da plataforma, não foram identificadas limitações a respeito da taxa de requisição de mensagens para a quantidade especificada.



P33

A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 10 mil publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 10 mil publicações em 24 horas, por meio da API da plataforma.

YouTube

De acordo com a calculadora de cotas disponibilizada na documentação da API da plataforma, é possível realizar a coleta deste volume de publicações, sejam vídeos ou comentários, contanto que os IDs dos vídeos de interesse sejam conhecidos previamente.



Instagram

Como informava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.



Kwai

O Kwai não disponibiliza API para acesso a dados públicos.



Facebook

Como informava a documentação do CrowdTangle, a API não era recomendada para a extração de mais de 10 mil publicações de uma vez. Para casos entre 10 mil e 100 mil publicações, a interface de coleta da ferramenta era mais recomendada.



X / Twitter

Como o teto de publicações no plano *Pro* é de 1 milhão por mês, seria possível fazer um monitoramento dessa escala, considerando as faixas de acesso analisadas.



WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.



Telegram

Com base nos testes realizados e na documentação da plataforma, não foram identificadas limitações a respeito da taxa de requisição de mensagens para a quantidade especificada.



Consistência



P34

Os dados retornados pela API da plataforma são persistentes?

Este campo verifica se os dados recuperados pela API da plataforma estão imunes à expiração, ainda que sejam links. Espera-se que alguns metadados das publicações removidas da plataforma não sejam excluídas da resposta da API, mas que haja uma sinalização da remoção do conteúdo.

YouTube

A API não retorna quaisquer metadados de vídeos removidos ou tornados privados, ao contrário da interface de usuário da plataforma, que exibe mensagem sobre a remoção e seus motivos, caso o usuário tente assistir a um vídeo que não está mais no ar.

Facebook

O CrowdTangle tratava as publicações deletadas do Facebook como se nunca tivessem existido, mesmo quando era feita uma busca específica por elas. Além disso, nossos testes apontaram que a API retornava publicações aleatórias quando havia o *input* de um ID de publicação inexistente ou grafado de forma errada. Como ponto positivo, links de outros conteúdos não expiram.

Instagram

Campos de links de imagens retornados pelo CrowdTangle expiravam tão rapidamente (erro *URL Signature Expired*) que era impossível manipulá-los ou analisá-los, mesmo no caso de publicações ativas.

X / Twitter

Os links retornados pela requisição da API não expiram e seus *endpoints de compliance* garantem maior persistência no retorno de metadados, mesmo quando as publicações não estão mais disponíveis na plataforma; o status e a disponibilidade são exibidos, tanto em relação a perfis, quanto em relação a publicações.

Telegram

A API do Telegram trata as mensagens deletadas como se nunca tivessem existido.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Consistência



P35

Os dados recuperados pela API da plataforma refletem o que é exibido em sua interface de usuário?

Neste campo, é verificado se os dados retornados pela API correspondem às informações exibidas na interface de usuário das plataformas. Deve ser possível identificar na resposta da API, por exemplo, informações como autoria, conteúdo completo e principais interações.

YouTube

Sim, a API permite a recuperação de todos os principais dados disponíveis na interface de usuário da plataforma, com exceção do próprio conteúdo do vídeo. Há, no entanto, um identificador para acessar o vídeo na plataforma.

Facebook

Uma vez que o CrowdTangle retornava as principais informações das publicações coletadas, que incluíam, entre outras, seu conteúdo, autoria, data e hora da publicação, além do volume das principais interações, era possível dizer que os dados refletiam o que era exibido na plataforma. Não era possível, no entanto, identificar os usuários responsáveis pelas curtidas, comentários ou compartilhamentos.

Instagram

Não era possível usar os links de imagens de forma sistemática porque eles expiravam rapidamente. Devido à importância do conteúdo visual na plataforma, entendemos que a API do CrowdTangle deveria disponibilizar links funcionais para as mídias publicadas. Por isso, consideramos que os dados não refletiam o que era mostrado na plataforma, ainda que a API disponibilizasse dados como a autoria, data e hora de publicação além do volume das principais interações.

X / Twitter

Não é possível coletar todas as informações porque o X/Twitter não disponibiliza o número de impressões recebidas pelas publicações por meio de sua API, embora mostre essa informação na interface de usuário. Esse dado só é disponibilizado via API mediante autorização de seus autores.

Telegram

Não há inconsistências entre os dados retornados pela plataforma na coleta e o que é exibido na interface de usuário, inclusive em relação às mídias (*imagens, vídeos, áudios*), que podem ser baixadas e analisadas de forma descomplicada.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Consistência



P36

A resposta retornada pela API da plataforma é sempre a esperada?

Neste campo, é avaliado se os dados recuperados por meio da API são sempre os mesmos, de acordo com os parâmetros e filtros utilizados, ou ao menos coerentes com outras coletas feitas de forma idêntica, excetuando-se publicações apagadas ou realizadas entre elas.

YouTube

Há inconsistências frequentes, mesmo em coletas por parâmetros fixos, ou seja, sem o envolvimento da recomendação algorítmica de conteúdos. Testes de coleta feitos com poucos minutos de intervalo recuperam dados consideravelmente diferentes entre si, e essa diferença não ocorre necessariamente pela inclusão de dados mais recentes.

Facebook

Com base nos testes realizados, requisições à API do CrowdTangle, em momentos diferentes e por pessoas diferentes, retornavam dados consistentes, sem variações substanciais, com exceção de publicações removidas ou tornadas privadas.

Instagram

Com base nos testes realizados, requisições à API do CrowdTangle, em momentos diferentes e por pessoas diferentes, retornavam dados consistentes, sem variações substanciais, com exceção de publicações removidas ou tornadas privadas.

X / Twitter

Com base nos testes realizados, requisições à API do X/Twitter, em momentos diferentes e por pessoas diferentes, retornam dados consistentes, sem variações substanciais, com exceção de publicações removidas ou tornadas privadas.

Telegram

Com base nos testes realizados, requisições à API do Telegram, em momentos diferentes e por pessoas diferentes, retornam dados consistentes, sem variações substanciais, com exceção de publicações removidas ou tornadas privadas.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Consistência



P37

A resposta retornada pela API da plataforma é coerente com os parâmetros e filtros utilizados na requisição?

Neste campo, é verificado se os dados recuperados pela API da plataforma refletem as escolhas de parâmetros e filtros determinadas no momento da requisição.

YouTube

Há incoerências frequentes, mesmo em coletas por parâmetros fixos, sem o envolvimento da recomendação algorítmica de conteúdos. Em nossos testes, verificamos que a API do YouTube não obedece, necessariamente, a filtros determinados pelos usuários (mais notadamente os de data e de idioma).

Facebook

As publicações retornadas pelo CrowdTangle eram coerentes com os parâmetros e filtros definidos. Na interface de coleta de dados, inclusive, era possível conferir quais publicações resultavam de uma busca específica e confirmar o motivo de cada uma ter sido retornada.

Instagram

As publicações retornadas pelo CrowdTangle eram coerentes com os parâmetros e filtros definidos. Na interface de coleta de dados, inclusive, era possível conferir quais publicações resultavam de uma busca específica e confirmar o motivo de cada uma ter sido retornada.

X / Twitter

É possível identificar na resposta da API a aplicação apropriada dos filtros e parâmetros indicados pelo usuário em relação ao intervalo de tempo, ao idioma das publicações e aos termos de busca de interesse, por exemplo.

Telegram

Com base nos testes realizados, os dados retornados sempre estão de acordo com os filtros selecionados na requisição à API.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Relevância



P38

As entidades retornadas pela API da plataforma são suficientes para compreender os dados em todos os seus níveis de detalhes?

Aqui, é avaliado se a resposta da API entrega todos os dados necessários para compreender o ciclo de vida completo de uma publicação específica, incluindo comentários, compartilhamentos, respostas e outros possíveis relacionamentos, assim como seus autores e conteúdos referenciados, como nos casos de compartilhamentos e menções.

YouTube

Sim, a API retorna **todas as principais entidades observadas** em um vídeo na interface de usuário da plataforma. Embora o conteúdo do vídeo não seja retornado pela API, ela fornece um identificador para que ele seja acessado na plataforma.

Facebook

Como o CrowdTangle **não retornava respostas e comentários de publicações**, não era possível avaliar os dados com todos os níveis de detalhe esperados.

Instagram

Como o CrowdTangle **não retornava respostas e comentários de publicações**, não era possível avaliar os dados com todos os níveis de detalhe esperados.

X / Twitter

Além do ID da publicação em questão, a v2 da API do X/ Twitter **disponibiliza os identificadores de publicações relacionadas** e, caso sejam identificadas na requisição, os dados de publicações relacionadas.

Telegram

Os dados são tecnicamente compreensíveis. Há uma exceção: a API do Telegram não permite, por padrão, a coleta programática de dados relativos aos autores das mensagens, já que eles não são representados por entidades próprias. Ainda assim, consideramos que o Telegram atende ao critério por ser possível identificar autores e cruzar informações de autoria entre mensagens enviadas em diferentes grupos.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Relevância



P39

A API da plataforma permite a utilização de filtros para refinar a requisição de dados?

Este campo verifica se a API da plataforma permite a utilização de filtros de busca, como a localização do publicador, idioma ou período específico, entre outros.

YouTube

A API permite aplicar diversos filtros de busca. No entanto, o filtro de idioma não garante que os resultados se limitarão ao idioma escolhido, pois, de acordo com a **documentação da plataforma**, se conteúdos em outras línguas forem considerados relevantes, serão incluídos na resposta. Por isso, as opções apresentadas não atendem ao mínimo esperado neste parâmetro.

Facebook

O idioma das publicações, o país declarado por seus autores e o intervalo de tempo em que foram feitas figuravam entre os **filtros possíveis para recuperar publicações** por meio da API do CrowdTangle.

Instagram

O idioma das publicações, o país declarado por seus autores e o intervalo de tempo em que foram feitas figuravam entre os **filtros possíveis para recuperar publicações** por meio da API do CrowdTangle.

X / Twitter

Certos operadores funcionam apenas em **determinados planos de acesso à API**, mas, de modo geral, é possível aplicar filtros de acordo com os objetivos do usuário.

Telegram

Na documentação da API da plataforma, não localizamos quaisquer menções a filtros de localização ou de idioma para buscar por publicações.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

P40

É possível recuperar dados recém-publicados, quase em tempo real à publicação, por meio da API da plataforma?

Aqui, é avaliado se a API da plataforma permite a recuperação de dados de um conjunto de publicações específicas em até uma hora após sua publicação.

YouTube

Sim, testes com vídeos recém-publicados apontam que, uma vez aplicado o ID do vídeo, a coleta ocorre normalmente. A coleta de vídeos de transmissões ao vivo (*lives*) também ocorre de forma similar, com a entrega dos mesmos tipos de dados.

Facebook

Os testes com o CrowdTangle indicaram que publicações muito recentes se tornavam rapidamente disponíveis para coleta por meio da ferramenta.

Instagram

Os testes com o CrowdTangle indicaram que publicações muito recentes se tornavam rapidamente disponíveis para coleta por meio da ferramenta.

X / Twitter

Os testes indicam que os dados são disponibilizados na API logo após uma publicação ser feita, sendo possível recuperar publicações inseridas em uma janela de poucos minutos.

Telegram

Os testes indicam que a API disponibiliza a mensagem praticamente de forma quase instantânea.

TikTok

O TikTok não disponibiliza API para acesso a dados públicos no Brasil.

Kwai

O Kwai não disponibiliza API para acesso a dados públicos.

WhatsApp

O WhatsApp não disponibiliza API para acesso a dados públicos.

Referências Bibliográficas

ANPD. Autoridade Nacional de Proteção de Dados. Glossário de proteção de dados pessoais e privacidade: 2.0. **Governo Federal**, [S.d.]. Disponível em: <<https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/glossario-anpd-protecao-de-dados-pessoais-e-privacidade.pdf>>. Acesso em: 3 abr. 2024.

ATRICON. Acesso à informação na prática: Orientações para cidadãos, gestores públicos e Tribunais de Contas. **Atricon**, 2023. Disponível em: <https://atrimon.org.br/wp-content/uploads/2023/06/Cartilha-com-orientacoes-Ciclo-2023.pdf>. Acesso em: 11 mar. 2024.

BAR-ILAN, J. Data collection methods on the Web for infometric purposes — A review and analysis. **Scientometrics**, [S.l.], v. 96, n. 3, p. 7–32, 2001. Disponível em: <<https://link.springer.com/article/10.1023/A:1005682102768>>. Acesso em: 9 abr. 2024.

BARBIERI, C. **Governança de Dados: Práticas, conceitos e novos caminhos**. Rio de Janeiro: Alta Books, 2019.

BATINI, C.; SCANNAPIECO, M. **Data Quality Concepts, Methodologies and Techniques**. Nova Iorque: Springer Berlin Heidelberg, 2006.

BENSON, P. R. The Guide to ISO 8000. **ECCMA**, [S.l.], 13 fev. 2019. Disponível em: <https://eccma.org/forms/Publications/download_library_data/VGhIE1aWRIHRvIElTtyA4MDAw/V2hp-dGVwYXBldGcg==>. Acesso em: 9 abr. 2024.

BIANCHI, T. Topic: WhatsApp in Brazil. **Statista**, [S.l.], 10 jan. 2024. Disponível em: <<https://www.statista.com/topics/7731/whatsapp-in-brazil/>>. Acesso em: 3 abr. 2024.

BOSSETTA, M. Scandalous Design: How Social Media Platforms' Responses to Scandal Impacts Campaigns and Elections. **Social Media + Society**, [S.l.], v. 6, n. 2, p. 1–4, jun. 2020. Disponível em: <<https://journals.sagepub.com/doi/10.1177/2056305120924777>>. Acesso em: 9 abr. 2024.

BRAGA, L. CUNHA, B. Violações dos direitos de povos tradicionais e as barreiras de acesso à informação ambiental: Uma análise em transparência ativa. In: Fórum de Direito de Acesso a Informações Públicas (org.). **A LAI É 10: O Brasil após uma década da Lei de Acesso à Informação**. São Paulo: Abraji, 2022. Disponível em: <https://www.transparencia.org.br/downloads/publicacoes/a_lai_e_10_ebook.pdf>. Acesso em: 9 abr. 2024.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Lei de Acesso à Informação (LAI). **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 18 nov. 2011. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 26 abr. 2024.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial [da] República Federativa do Brasil**, Poder Executivo, Brasília, DF, 15 ago. 2018. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13709.htm>. Acesso em: 10 abr. 2024.

BRUNS, A. After the 'APIcalypse': social media platforms and their fight against critical scholarly research. **Information, Communication & Society**, [S.l.], v. 22, n. 11, p. 1544–1566, 19 set. 2019. Disponível em: <<https://www>>.

[tandfonline.com/doi/full/10.1080/1369118X.2019.1637447](https://doi.org/10.1080/1369118X.2019.1637447)>. Acesso em: 9 abr. 2024.

CALVO-GUTIÉRREZ, E.; MARÍN-LLADÓ, C. Combatting Fake News: A Global Priority Post COVID-19. *Societies*, [S.l.], v. 13, n. 7, p. 1-13, 2023. Disponível em: <<https://www.mdpi.com/2075-4698/13/7/160>>. Acesso em: 15 abr. 2024.

CECILIA, J. M.; CANO, J.-C.; HERNÁNDEZ-ORALLO, E.; CALAFATE, C. T.; MANZONI, P. Mobile crowdsensing approaches to address the COVID-19 pandemic in Spain. *IET Smart Cities*, [S.l.], v. 2, n. 2, p. 58-63, 2020. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1049/iet-smc.2020.0037>>. Acesso em: 9 abr. 2024.

CHOWDHURY, S.; FULLER, R. A.; AHMED, S.; ALAM, S.; CALLAGHAN, C. T.; DAS, PRIYANKA; CORREIA, R. A.; DI MARCO, M.; DI MININ, E.; JARIĆ, I.; LABI, M. M.; LADLE, R. J.; ROKONUZZAMAN, M.; ROLL, U.; SBRAGAGLIA, V.; SIDDIKA, A.; BONN, A. Using social media records to inform conservation planning. *Conservation Biology*, [S.l.], v. 38, n. 1, p. 1-11, 2024. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/cobi.14161>>. Acesso em: 9 abr. 2024.

CIOFFI-REVILLA, C. *Introduction to Computational Social Science: Principles and Applications*. Cham: Springer International Publishing, 2018.

CLOUDFLARE. O que é um endpoint de API? *CloudFlare*, [S.d.]. Disponível em: <<https://www.cloudflare.com/pt-br/learning/security/api/what-is-api-endpoint/>>. Acesso em: 3 abr. 2024.

COALITION FOR INDEPENDENT TECHNOLOGY RESEARCH. Letter: Imposing Fees to Access the Twitter API Threatens Public-Interest Research. *Coalition for Independent Technology Research*, 2023. Disponível em: <[\[-threatens-public-interest-research/\]\(https://independent-techresearch.org/letter-twitter-api-access-\)>. Acesso em: 4 abr. 2024.](https://independent-techresearch.org/letter-twitter-api-access-</p></div><div data-bbox=)

CODE MANAGEMENT ASSOCIATION (ECCMA). The Code Management Association. What is ISO 8000?. *ECCMA*, [S.d.]. Disponível em: <<https://eccma.org/what-is-iso-8000/>>. Acesso em: 3 abr. 2024.

CONGER, K. LinkedIn sues anonymous data scrapers. *TechCrunch*, [S.l.], 15 ago. 2016. Disponível em: <<https://techcrunch.com/2016/08/15/linkedin-sues-scrapers/>>. Acesso em: 15 abr. 2024.

CONGRESSO NACIONAL. Índice de Transparência dos Portais Legislativos. *Congresso Nacional*, 2022. Disponível em: https://www12.senado.leg.br/transparencia/indice-de-transparencia-dos-portais-legislativos/manual_itpl/view. Acesso em: 12 mar. 2024.

CORREIA, R. B.; WOOD, I. B.; BOLLEN, J.; ROCHA, M. R.. Mining Social Media Data for Biomedical Signals and Health-Related Behavior. *Annual Review of Biomedical Data Science*, [S.l.], v. 3, n. 1, 2020, p. 433-458, 20 jul. 2020. Disponível em: <<https://www.annualreviews.org/content/journals/10.1146/annurev-biodata-sci-030320-040844>>. Acesso em: 4 abr. 2024.

CRAWFORD, K. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Nova Iorque: Yale University Press, 2021.

CROWDTANGLE. CrowdTangle/API: API Documentation. *CrowdTangle*, [S.d.]. Disponível em: <<https://github.com/CrowdTangle/API>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. CrowdTangle API. *CrowdTangle*, [S.d.]. Disponível em: <<https://help.crowdtangle.com/en/articles/1189612-crowdtangle-api>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. Errors. *CrowdTangle*, [S.d.]. Disponível em: <<https://github.com/CrowdTangle/API/wiki/Errors>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. GET /posts. *CrowdTangle*, [S.d.]. Disponível em: <<https://github.com/CrowdTangle/API/wiki/Posts>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. Important Update to Crowd-Tangle | March 2024. *CrowdTangle*, 2024. Disponível em: <<https://help.crowdtangle.com/en/articles/9014544-important-update-to-crowdtangle-march-2024>>. Acesso em: 3 abr. 2024.

CROWDTANGLE. Página inicial. *CrowdTangle*, [S.d.]. Disponível em: <<https://www.crowdtangle.com/>>. Acesso em: 29 abr. 2024.

CROWDTANGLE. Request Access. *CrowdTangle*, [S.d.]. Disponível em: <<https://www.crowdtangle.com/request/>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. Search | GET /posts/search. *CrowdTangle*, [S.d.]. Disponível em: <<https://github.com/CrowdTangle/API/wiki/Search>>. Acesso em: 4 abr. 2024.

CROWDTANGLE. What data is CrowdTangle tracking? *CrowdTangle*, [S.d.]. Disponível em: <<https://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking>>. Acesso em: 4 abr. 2024.

DATA REPORTAL. Digital 2024: Brazil. *Data Reportal*, 23 fev. 2024. Disponível em: <<https://datareportal.com/reports/digital-2024-brazil>>. Acesso em: 3 abr. 2024.

DHELIM, S.; CHEN, L.; DAS, S. J.; NING, H.; LEAVEY, G.; PESCH, D.; BANTRY-WHITE, E.; BURNS, D. Detecting Mental Distresses Using Social Behavior Analysis in the Context of COVID-19: A Survey. *ACM Computing Surveys*, [S.l.], v. 55, n. 14, p. 1-30, 17 jul. 2023. Disponível em: <<https://dl.acm.org/doi/10.1145/3589784>>. Acesso em: 9 abr. 2024.

EDWARDS, A.; HOUSLEY, W.; WILLIAMS, M.; SLOAN, L.; WILLIAMS, M. Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology*, [S.l.], v. 16, maio 2013. Disponível em: <<https://www.researchgate.net/publication/263038400>>. Acesso em: 9 abr. 2024.

EL-SAYED, A. M.; SCARBOROUGH, P.; SEEMANN, L.; GALEA, S. Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives & Innovations*, [S.l.], v. 9, n. 1, p. 1-9, fev. 2012. Disponível em: <<https://epi-perspectives.biomedcentral.com/articles/10.1186/1742-5573-9-1>>. Acesso em: 9 abr. 2024.

EUROPEAN COMMISSION. Commission opens formal proceedings against X under the DSA. *European Commission*, Bruxelas, 18 dez. 2023. Disponível em: <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6709>. Acesso em: 3 abr. 2024.

EUROPEAN COMMISSION. The EU's Digital Services Act. *European Commission*, [S.d.]. Disponível em: <https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/eu-rome-fit-digital-age/digital-services-act_en>. Acesso em: 3 abr. 2024.

EVANGELISTA, R.; BRUNO, F. WhatsApp and political instability in Brazil: Targeted messages and political radicalisation. *Internet Policy Review*, [S.l.], v. 8, n. 4, 31 dez. 2019. Disponível em: <<https://policyreview.info/node/1434>>. Acesso em: 9 abr. 2024.

FACEBOOK. KoPI-G Transparency Report. *Facebook*, 2023. Disponível em: <<https://transparency.fb.com/sr/kopi-report-english-jul-23>>. Acesso em: 16 abr. 2024.

FACEBOOK. O que é extração de dados e o que posso fazer para proteger as minhas infor-

mações no Facebook? **Facebook**, [S.d.]. Disponível em: <<https://www.facebook.com/help/463983701520800>>. Acesso em: 4 abr. 2024.

FACEBOOK. Türkiye Transparency Report for Content Removal Requests under Law No. 5651. **Facebook**, [S. d.]. Disponível em: <<https://transparency.fb.com/sr/turkey-report-english-fb-jan-24>>. Acesso em: 16 abr. 2024.

FAN, C. API Cheat Sheet. **CrowdTangle**, [S.d.]. Disponível em: <<http://help.crowdtangle.com/en/articles/3443476-api-cheat-sheet>>. Acesso em: 4 abr. 2024.

GARIMELLA, K.; TYSON, G. WhatsApp, doc? A first look at WhatsApp public group data. In: Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018), 12., jun. 2018, Califórnia. **Anais [...]**. [S.l.]: Association for the Advancement of Artificial Intelligence, 2018. Disponível em: <<https://arxiv.org/abs/1804.01473>>. Acesso em: 9 abr. 2024.

GERBAUDO, P. **Redes e Ruas**. São Paulo: Editora Funilaria, 2021.

GHERMANDI, A.; LANGEMEYER, J.; VAN BERKEL, D.; CALCAGNI, F.; DEPIETRI, Y.; VIGL, L. E.; FOX, N.; HAVINGA, I.; JÄGER, H.; KAISER, N.; KARASOV, O.; MCPHEARSON, T.; PODSCHUN, S.; RUIZ-FRAU, A.; SINCLAIR, M.; VENOHR, M.; WOOD, S. A. Social media data for environmental sustainability: A critical review of opportunities, threats, and ethical use. **One Earth**, [S.l.], v. 6, n. 3, p. 236–250, 17 mar. 2023. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S259033222300088X>>. Acesso em: 9 abr. 2024.

GILLESPIE, T. The politics of ‘platforms’. **New Media & Society**, [S.l.], v. 12, n. 3, p. 347–364, maio 2010. Disponível em: <<https://journals.sagepub.com/doi/10.1177/1461444809342738>>. Acesso em: 9 abr. 2024.

GLOBAL AD. Kwai em 2024: Um panorama detalhado da plataforma. **Global AD**, [S.l.], 22 fev. 2024. Disponível em: <<https://globalad.com.br/blog/explorando-o-kwai-em-2024-um-panorama-detalhado/>>. Acesso em: 3 abr. 2024.

GOODWIN, M. What Is an API (Application Programming Interface)?. **IBM**, [S.l.], 9 abr. 2024. Disponível em: <<https://www.ibm.com/topics/api>>. Acesso em: 26 abr. 2024.

GOOGLE. Cumprimento das diretrizes da comunidade do YouTube – Google Relatório de Transparência. **Google**, [S.d.]. Disponível em: <<https://transparencyreport.google.com/youtube-policy/removals>>. Acesso em: 16 abr. 2024.

GOOGLE. Google Relatório de Transparência. **Google**, [S.d.]. Disponível em: <https://transparencyreport.google.com/?hl=pt_BR>. Acesso em: 16 abr. 2024.

GOOGLE. Removals under the Law No. 5651 (Turkish Internet Law). **Google**, [S.d.]. Disponível em: <https://storage.googleapis.com/transparencyreport/report-downloads/turkish-internet-law_2023-7-1_2023-12-31_en_v1.pdf>. Acesso em: 16 abr. 2024.

GOOGLE. Solicitações governamentais de remoção de conteúdo – Google Relatório de Transparência. **Google**, [S.d.]. Disponível em: <<https://transparencyreport.google.com/government-removals/overview>>. Acesso em: 16 abr. 2024.

GOOGLE. South Korea Network Act and the Telecommunications Business Act Transparency Report. **Google**, [S.d.]. Disponível em: <https://storage.googleapis.com/transparencyreport/report-downloads/south-korea-network-act_2023-1-1_2023-12-31_en_v1.pdf>. Acesso em: 16 abr. 2024.

GOVERNO DE BRASÍLIA. Controladoria Geral: Índice de Transparência Ativa. **Governo de Brasília**, 2016. Disponível em: <https://www.cg.df.gov.br/wp-content/uploads/2018/02/>

[Crit%C3%A9rios-Avalia%C3%A7%C3%A3o-i-TA-2016.pdf](https://www.instagram.com/help/463983701520800). Acesso em: 12 mar. 2024.

GREENE, T.; MARTENS, D.; SHMUELI, G. Barriers to academic data science research in the new realm of algorithmic behaviour modification by digital platforms. **Nature Machine Intelligence**, [S.l.], v. 4, n. 4, p. 323–330, abr. 2022. Disponível em: <<https://www.nature.com/articles/s42256-022-00475-7>>. Acesso em: 9 abr. 2024.

GUESS, A. M.; MALHOTRA, N.; PAN, J.; BARBERÁ, P.; ALLCOTT, H.; BROWN, T.; CRESPO-TENORIO, A.; DIMMERY, D.; FREELON, D.; GENTZKOW, M.; GONZÁLEZ-BAILÓN, S.; KENNEDY, E.; KIM, Y. M.; LAZER, D.; MOEHLER, D.; NYHAN, B.; RIVERA, C. V.; SETTLE, J.; THOMAS, D. R.; THORSON, E.; TROMBLE, R.; WILKINS, A.; WOJCIESZAK, M.; XIONG, B.; DE JONGE, C. K.; FRANCO, A.; MASON, W.; STROUD, N. J.; TUCKER, J. A. How do social media feed algorithms affect attitudes and behavior in an election campaign? **Science**, [S.l.], v. 381, n. 6656, p. 398–404, 28 jul. 2023. Disponível em: <<https://www.science.org/doi/10.1126/science.abp9364>>. Acesso em: 9 abr. 2024.

HOVYADINOV, S. Toward a More Meaningful Transparency: Examining Twitter, Google, and Facebook’s Transparency Reporting and Removal Practices in Russia. **SSRN Scholarly Paper**, 2019. [Manuscrito em pré-publicação] Disponível em: <<https://papers.ssrn.com/abstract=3535671>>. Acesso em: 15 abr. 2024.

INSTAGRAM. KoPl-G Transparency Report. **Instagram**, 2023. Disponível em: <<https://transparency.fb.com/sr/kopi-report-english-IG-jul-23>>. Acesso em: 16 abr. 2024.

INSTAGRAM. Por que minha conta foi restringida por extração de dados e o que posso fazer? **Instagram**, [S.d.]. Disponível em: <<https://help.instagram.com/740480200552298>>. Acesso em: 4 abr. 2024.

INSTAGRAM. Türkiye Transparency Report for Content Removal Requests under Law No. 5651. **Instagram**, 2024. Disponível em: <<https://transparency.fb.com/sr/turkey-report-english-ig-jan-24>>. Acesso em: 16 abr. 2024.

ISO. ISO 8000. Data quality — Part 1: Overview. **International Organization for Standardization**, [S.l.], [S.d.]. Disponível em: <<https://www.iso.org/obp/ui/#iso:std:iso:8000:-1:ed-1:v1:en>>. Acesso em: 3 abr. 2024.

KALOGEROPOULOS, A.; ROSSINI, P. Unraveling WhatsApp group dynamics to understand the threat of misinformation in messaging apps. **New Media & Society**, [S.l.], v. 0, n. 0, n.p., 2023. DOI: 14614448231199247. Disponível em: <<https://doi.org/10.1177/14614448231199247>>. Acesso em: 16 abr. 2024.

KOSTA, E.; BREWCZYŃSKA, M. Government Access to User Data: Towards More Meaningful Transparency Reports. In: BALLARDINI, R.; KUOPPAMÄKI, P.; PITKÄNEN, O. (Eds.). **Regulating industrial internet through IPR, data protection and competition law**. Países Baixos: Kluwer Law International, 2019. Disponível em: <<https://papers.ssrn.com/abstract=3601661>>. Acesso em: 15 abr. 2024.

KWAI. Safety Center. **Kwai**, [S.d.]. Disponível em: <<https://www.kwai.com/safety>>. Acesso em: 16 abr. 2024.

KWAI. Terms of Service. **Kwai**, [S.d.]. Disponível em: <<https://www.kwai.com/legal>>. Acesso em: 4 abr. 2024.

LEE, R. M.; FIELDING, N.; BLANK, G. The Internet as a research medium: An editorial introduction to the sage handbook of online research methods. In: FIELDING, N.; LEE, R. M.; BLANK, G. **The Sage handbook of online research methods**. Oxford: SAGE Research Methods, 2008.

LOSHIN, D. **Master Data Management**. Burlington: Morgan Kaufmann, 2008.

LU, S. Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence. *Vanderbilt Journal of Entertainment & Technology Law* 99, [S.l.], v. 23, n. 1, p. 99–159, 20 out. 2021. Disponível em: <<https://papers.ssrn.com/abstract=3582222>>. Acesso em: 3 abr. 2024.

LURIE, E. Comparing Platform Research API Requirements. *TechPolicy.Press*, [S.l.], 22 mar. 2023. Disponível em: <<https://techpolicy.press/comparing-platform-research-api-requirements>>. Acesso em: 4 abr. 2024.

MAHANTI, R. *Data Quality: Dimensions, Measurement, Strategy, Management, and Governance*. Milwaukee: ASQ Quality Press, 2018.

MARRES, N. *Digital Sociology: The Reinvention of Social Research*. Cambridge: Wiley, 2017.

MCGILVRAY, D. *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information (TM)*. São Francisco: Morgan Kaufmann Publishers, 2008.

MEHTA, I. X updates its terms to ban crawling and scraping. *TechCrunch*, [S.l.], 8 set. 2023. Disponível em: <<https://techcrunch.com/2023/09/08/x-updates-its-terms-to-ban-crawling-and-scraping/>>. Acesso em: 4 abr. 2024.

MELO, P.; VIEIRA, C.; GARIMELLA, K.; MELO, P. V.; BENEVENUTO, F. Can WhatsApp Counter Misinformation by Limiting Message Forwarding?. In: *International Conference on Complex Networks and their Applications*, 8., dez. 2019, Lisboa. *Anais [...]*. Suíça: Springer, 2019. Disponível em: <<https://arxiv.org/abs/1909.08740>>. Acesso em: 16 abr. 2024.

MELO, P. de F. *Activism and Misinformation on WhatsApp*. 218 p. 2022. *Tese [Doutorado em Ciência da Informação]* – Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Minas Gerais, 2022. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/49506/4/Tese_PhilipeMelo_ActivismAndMisinformationOnWhatsApp.pdf>. Acesso em: 15 abr. 2024.

repositorio.ufmg.br/bitstream/1843/49506/4/Tese_PhilipeMelo_ActivismAndMisinformationOnWhatsApp.pdf>. Acesso em: 15 abr. 2024.

META. Content Restrictions Based on Local Law | Transparency Center. *Meta*, [S.d.]. Disponível em: <<https://transparency.fb.com/reports/content-restrictions/>>. Acesso em: 16 abr. 2024.

META. Government Requests for User Data | Transparency Center. *Meta*, [S.d.]. Disponível em: <<https://transparency.fb.com/reports/government-data-requests/country/br/>>. Acesso em: 16 abr. 2024.

META. Graph API do Instagram. *Meta for Developers*, [S.d.]. Disponível em: <<https://developers.facebook.com/docs/instagram-api/>>. Acesso em: 4 abr. 2024.

META. India Monthly Report under the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. *Meta*, 2024. Disponível em: <<https://transparency.fb.com/sr/india-monthly-report-feb29-2024/>>. Acesso em: 16 abr. 2024.

META. Korea Transparency Report. *Meta*, 2024. Disponível em: <https://z-p3-s-content.fsdu2-2.fna.fbcdn.net/v/t39.6/425023426_406911275040892_142225897538078248_n.pdf?_nc_cat=108&ccb=1-7&_nc_sid=b8d81d&_nc_ohc=PR_Ve-EqU2usAb7cb6fv&_nc_ht=z-p3-scontent.fsdu2-2.fna&oh=00_AfAtew8y7ejiaiUAz-bhY6GJ8oFpnuyFFR2_iRGVHskUoOQ&oe=6623459A>. Acesso em: 16 abr. 2024.

META. Meta Transparency Center. *Meta*, [S.d.]. Disponível em: <<https://transparency.meta.com/pt-br/>>. Acesso em: 29 abr. 2024.

META. Progresso na privacidade. *Meta*, [S.d.]. Disponível em: <<https://about.meta.com/br/privacy-progress/>>. Acesso em: 4 abr. 2024.

META. Transparency reports | Transparency Center. *Meta*, [S.d.]. Disponível em: <<https://transparency.fb.com/reports/>>. Acesso em: 16 abr. 2024.

MILLER, G. First Transparency Reports Under Digital Services Act Are Difficult to Compare. *TechPolicy.Press*, [S.l.], 22 nov. 2023. Disponível em: <<https://techpolicy.press/first-transparency-reports-under-digital-services-act-are-difficult-to-compare>>. Acesso em: 15 abr. 2024.

MOONEY, S. J.; WESTREICH, D. J.; EL-SAYED, A. M. Commentary: Epidemiology in the Era of Big Data. *Epidemiology*, [S.l.], v. 26, n. 3, p. 390–394, maio 2015. Disponível em: <<https://journals.lww.com/epidem/fulltext/2015/05000/>>. Acesso em: 4 abr. 2024.

MOURA, M.; MICHELSON, M. R. WhatsApp in Brazil: Mobilising voters through door-to-door and personal messages. *Internet Policy Review*, [S.l.], v. 6, n. 4, 2017. Disponível em: <<https://policyreview.info/articles/analysis/whatsapp-brazil-mobilising-voters-through-door-door-and-personal-messages>>. Acesso em: 15 abr. 2024.

MOZELLI, R. Twitter: API cara impede uso para pesquisas acadêmicas. *Olhar Digital*, [S.l.], 01 jun. 2023. Disponível em: <<https://olhardigital.com.br/2023/06/01/internet-e-redes-sociais/twitter-api-cara-impede-uso-para-pesquisas-academicas/>>. Acesso em: 3 abr. 2024.

MOZILLA FOUNDATION. Open Letter To Meta: Support CrowdTangle Through 2024 and Maintain CrowdTangle Approach. *Mozilla Foundation*, 2024. Disponível em: <<https://foundation.mozilla.org/en/campaigns/open-letter-to-meta-support-crowdtangle-through-2024-and-maintain-crowdtangle-approach/>>. Acesso em: 4 abr. 2024.

OPEN KNOWLEDGE BRASIL. Transparência COVID-19 3.0. *Open Knowledge Brasil*, 2021.

Disponível em: <<https://transparenciacovid19.ok.org.br/>>. Acesso em: 15 abr. 2024.

OPINION BOX. Relatório Kwai no Brasil 2024. *Opinion Box*, 2024. Disponível em: <<https://content.app-us1.com/JY8yY/2024/05/24/f32d-1193-1919-4170-8727-db73a82d84a1.pdf>>. Acesso em: 20 set. 2024.

OREMBO, L.; BERGER, G.; SIMON, N. Need for Data Access to Tackle AI-powered Disinformation in the Global South. *TechPolicy.Press*, [S.l.], 20 dez. 2023. Disponível em: <<https://techpolicy.press/need-for-data-access-to-tackle-ai-powered-disinformation-in-the-global-south>>. Acesso em: 4 abr. 2024.

OZAWA, J. V. S.; WOOLEY, S. C.; STRAUBHAR, J.; RIEDL, M.; JOSEFF, K.; GURSKY, J. How Disinformation on WhatsApp Went From Campaign Weapon to Governmental Propaganda in Brazil. *Social Media + Society*, [S.l.], v. 9, n. 1, n.p., 2023. Disponível em: <<https://journals.sagepub.com/doi/10.1177/20563051231160632>>. Acesso em: 15 abr. 2024.

PAPPA, D.; STERGIOULAS, L. K. Harnessing social media data for pharmacovigilance: A review of current state of the art, challenges and future directions. *International Journal of Data Science and Analytics*, [S.l.], v. 8, n. 2, p. 113–135, set. 2019. Disponível em: <<https://link.springer.com/article/10.1007/s41060-019-00175-3>>. Acesso em: 9 abr. 2024.

PERRINO, J. Platform Accountability and Transparency Act Reintroduced in Senate. *TechPolicy.Press*, [S.l.], 8 jun. 2023. Disponível em: <<https://techpolicy.press/platform-accountability-and-transparency-act-reintroduced-in-senate>>. Acesso em: 3 abr. 2024.

POSTMAN. What is an API? A Beginner's Guide to APIs. *PostMan*, [S.d.]. Disponível em: <<https://www.postman.com/what-is-an-api/>>. Acesso em: 3 abr. 2024.

RESENDE, G.; MELO, P.; SOUSA, H.; MESSIAS, J.; VASCONCELOS, M.; ALMEIDA, J.; BENEVENUTO, F. (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. In: The World Wide Web Conference, 28., maio 2019, Califórnia. **Anais [...]**. Nova Iorque: Association for Computing Machinery (ACM), 13 maio 2019. Disponível em: <<https://dl.acm.org/doi/10.1145/3308558.3313688>>. Acesso em: 3 abr. 2024.

ROGERS, R. **The end of the virtual**: Digital methods. Amsterdã: Vossiuspers UvA, 2009.

ROSENFELD, A.; SINA, S.; SARNE, D.; AVIDOV, O.; KRAUS, S. WhatsApp usage patterns and prediction of demographic characteristics without access to message content. **Demographic Research**, [S.l.], v. 39, n. 22, p. 647-670, 2018. Disponível em: <<https://www.demographic-research.org/volumes/vol39/22/>>. Acesso em: 15 abr. 2024.

ROTH, E. Elon Musk's X sues anti-hate researchers for allegedly scraping data from Twitter. **The Verge**, [S.l.], 2023. Disponível em: <<https://www.theverge.com/2023/8/1/23815515/twitter-ccdh-anti-hate-research-group-lawsuit>>. Acesso em: 15 abr. 2024.

RUDNITZKI, E. Como funciona e por que ainda usar o CrowdTangle. **Aos Fatos**, [S.l.], 11 dez. 2023. Disponível em: <<https://aosfatos.org/noticias/como-funciona-e-porque-ainda-usar-o-crowdtangle/>>. Acesso em: 3 abr. 2024.

SANTOS, E. Após 11 anos de LAI, governos locais ainda dificultam pedidos de acesso à informação. **Open Knowledge Brasil**, [S.l.], 18 maio 2023. Disponível em: <<https://ok.org.br/noticia/apos-11-anos-de-lai-governos-locais-ainda-dificultam-pedidos-de-acesso-a-informacao/>>. Acesso em: 3 abr. 2024.

SELINGER, E.; HARTZOG, W. Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments

where users lack control. **Research Ethics**, [S.l.], v. 12, n. 1, p. 35-43, jan. 2016. Disponível em: <<https://journals.sagepub.com/doi/10.1177/1747016115579531>>. Acesso em: 9 abr. 2024.

SENTHILRAJA, M. Application of Artificial Intelligence to Address Issues Related to the COVID-19 Virus. **SLAS**, [S.l.], v. 26, n. 2, p. 123-126, abr. 2021. Disponível em: <[https://slas-technology.org/article/S2472-6303\(22\)01102-5/full-text](https://slas-technology.org/article/S2472-6303(22)01102-5/full-text)>. Acesso em: 9 abr. 2024.

SHAW, R. Big Data and reality. **Big Data & Society**, [S.l.], v. 2, n. 2, p. 1-4, dez. 2015. Disponível em: <<https://journals.sagepub.com/doi/10.1177/2053951715608877>>. Acesso em 9 abr. 2024.

SHEWALE, R. Telegram Statistics In 2024 (Usage By Country & Financials). **DemandSage**, [S.l.], 18 jan. 2024. Disponível em: <<https://www.demandsage.com/telegram-statistics/>>. Acesso em: 3 abr. 2024.

SMITH, R.; SMITH, R.; CHEN, K.; WINNER, D.; FRIEDHOFF, S.; WARDLE, C. A Systematic Review Of COVID-19 Misinformation Interventions: Lessons Learned. **Health Affairs**, [S.l.], v. 42, n. 12, p. 1738-1746, 2023. Disponível em: <<https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2023.00717>>. Acesso em: 15 abr. 2024.

SOARES, M. Meta encerra CrowdTangle e impõe alternativa limitada e burocrática. **Desinformante**, [S.l.], 14 ago. 2024. Disponível em: <<https://desinformante.com.br/meta-crowdtangle-alternativa/#:~:text=Fim%20do%20CrowdTangle%20amea%C3%A7a%20continuidade%20de%20projetos>>. Acesso em: 12 set. 2024.

SRIVASTAVA, A. K.; MISHRA, R. Analyzing Social Media Research: A Data Quality and Research Reproducibility Perspective. **IIM Ko-**

zhikode Society & Management Review, Calicut, v. 12, n. 1, p. 39-49, mai. 2021. Disponível em: <<https://journals.sagepub.com/doi/10.1177/22779752211011810>>. Acesso em: 9 abr. 2024.

STAAB, P.; THIEL, T. Social Media and the Digital Structural Transformation of the Public Sphere. **Theory, Culture & Society**, [S.l.], v. 39, n. 4, p. 129-143, 5 set. 2022. Disponível em: <<https://journals.sagepub.com/doi/10.1177/02632764221103527>>. Acesso em: 26 abr. 2024.

STOKEL-WALKER, C. X nega a pesquisadores que estudam desinformação acesso a sua API. **Fast Company Brasil**, [S.l.], 1 mar. 2024. Disponível em: <<https://fastcompanybrasil.com/tech/x-nega-a-pesquisadores-que-estudam-desinformacao-acesso-a-sua-api/>>. Acesso em: 4 abr. 2024.

SUZOR, N. P.; ESTE, S. M.; QUODLING, A.; YORK, J. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. **International Journal of Communication**, [S.l.], v. 13, n. 0, p. 1526-1543, 2019. Disponível em: <<https://ijoc.org/index.php/ijoc/article/view/9736>>. Acesso em: 15 abr. 2024.

TELEGRAM. messages.search. **Telegram**, [S.d.]. Disponível em: <<https://core.telegram.org/method/messages.search>>. Acesso em: 4 abr. 2024.

TELEGRAM. Methods. **Telegram**, [S.d.]. Disponível em: <<https://core.telegram.org/methods>>. Acesso em: 4 abr. 2024.

TELEGRAM. Política de Privacidade do Telegram. **Telegram**, [S.d.]. Disponível em: <<https://telegram.org/privacy/br>>. Acesso em: 16 abr. 2024.

TELEGRAM. Telegram APIs. **Telegram**, [S.d.]. Disponível em: <<https://core.telegram.org/>>. Acesso em: 4 abr. 2024.

TELEGRAM. Telegram API Terms of Service. **Telegram**, [S.d.]. Disponível em: <<https://core.telegram.org/api/terms>>. Acesso em: 4 abr. 2024.

TELEGRAM. User Authorization. **Telegram**, [S.d.]. Disponível em: <<https://core.telegram.org/api/auth>>. Acesso em: 4 abr. 2024.

TELETHON. Methods/Stories. **Telethon**, [S.d.]. Disponível em: <<https://tl.telethon.dev/methods/stories/index.html>>. Acesso em: 4 abr. 2024.

TELETHON. Telethon's Documentation. **Telethon**, [S.d.]. Disponível em: <<https://docs.telethon.dev/en/stable/>>. Acesso em: 4 abr. 2024.

TERRY, K.; YANG, F.; YAO, Q.; LIU, C. The role of social media in public health crises caused by infectious disease: a scoping review. **BMJ Global Health**, [S.l.], v. 8, n. e013515, p. 1-13, dez. 2023. Disponível em: <<https://gh.bmj.com/content/8/12/e013515>>. Acesso em 9 abr. 2024.

TIKTOK. Community Guidelines Enforcement Report. **TikTok**, 2024. Disponível em: <<https://www.tiktok.com/transparency/en/community-guidelines-enforcement-2023-4/>>. Acesso em: 16 abr. 2024.

TIKTOK. Content Posting API. **TikTok for Developers**, [S.d.]. Disponível em: <<https://developers.tiktok.com/products/content-posting-api/>>. Acesso em: 4 abr. 2024.

TIKTOK. Government Removal Requests Report. **TikTok**, 2023. Disponível em: <<https://www.tiktok.com/transparency/en/government-removal-requests-2023-1/>>. Acesso em: 15 abr. 2024.

TIKTOK. Research API. **TikTok**, [S.d.]. Disponível em: <<https://developers.tiktok.com/products/research-api/>>. Acesso em: 3 abr. 2024.

TIKTOK. Termos de Serviço. **TikTok**, [S.d.]. Disponível em: <https://www.tiktok.com/legal/page/row/terms-of-service/pt-BR?enter_method=bottom_navigation>. Acesso em: 4 abr. 2024.

TRANSPARÊNCIA BRASIL. O que a população quer saber do poder público? Uma análise de respostas a pedidos de acesso à informação de órgãos de todos os poderes e níveis federativos. **Transparência Brasil**, 2018. Disponível em: <https://www.transparencia.org.br/downloads/publicacoes/RelatorioLAI_TransparenciaBrasil_2018_vf.pdf>. Acesso em: 3 abr. 2024.

TRERÉ, E. The banality of WhatsApp: On the everyday politics of backstage activism in Mexico and Spain. **First Monday**, [S.l.], v. 25, n. 1, 2020. Disponível em: <<https://journals.uic.edu/ojs/index.php/fm/article/view/10404>>. Acesso em: 15 abr. 2024.

TUMASJAN, A. The many faces of social media in business and economics research: Taking stock of the literature and looking into the future. **Journal of Economic Surveys**, [S.l.], v. 38, n. 2, p. 389–426, 23 jun. 2023. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12570>>. Acesso em 9 abr. 2024.

UNIÃO EUROPEIA. Regulation (EU) 2016/679 of the European Parliament and of the Council. General Data Protection Regulation. **Official Journal of the European Union**, Bruxelas, 4 maio 2016. Disponível em: <<https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=OJ:L:2016:119:FULL>>. Acesso em: 17 jun. 2023.

URMAN, A.; MAKHORTYKH, M. How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms. **Telecommunications Policy**, [S.l.], v. 47, n. 3, p. 1–15, 2023. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0308596122001793>>. Acesso em: 15 abr. 2024.

VAN ATTEVELDT, W.; PENG, T.-Q. When Communication Meets Computation: Opportunities, Challenges, and Pitfalls in Computational Communication Science. **Communication Methods and Measures**, [S.l.], v. 12, n. 2–3, p. 81–92, 20 abr. 2018. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/19312458.2018.1458084>>. Acesso em 9 abr. 2024.

VAN DIJCK, J.; NIEBORG, D.; POELL, T.. Reframing platform power. **Internet Policy Review**, [S.l.], v. 8, n. 2, 30 jun. 2019. Disponível em: <<https://policyreview.info/articles/analysis/reframing-platform-power>>. Acesso em: 9 abr. 2024.

WAGNER, B.; ROZGONYI, K.; SEKWENZ, M-T.; COBBE, J.; SINGH, J. Regulating Transparency? Facebook, Twitter and the German Network Enforcement Act. In: Conference on Fairness, Accountability, and Transparency, 20., 2020, Barcelona. **Anais [...]**. [S.l.]: Association for Computing Machinery (ACM), 2020. Disponível em: <<https://dl.acm.org/doi/10.1145/3351095.3372834>>. Acesso em 15 abr. 2024.

WENDRATAMA, E.; YUSUF, I. A. COVID-19 Falsehoods on WhatsApp: Challenges and Opportunities in Indonesia. In: SOON, C. (org.). **Mobile Communication and Online Falsehoods in Asia: Trends, Impact and Practice**. Países Baixos: Springer, 2023. Disponível em: <https://link.springer.com/chapter/10.1007/978-94-024-2225-2_2>. Acesso em: 15 abr. 2024.

WHATSAPP. Termos de Serviço do WhatsApp. **WhatsApp**, [S.d.]. Disponível em: <<https://www.whatsapp.com/legal/terms-of-service>>. Acesso em: 4 abr. 2024.

X/TWITTER. Brazil. **X/Twitter Transparency Center**, [S.d.]. Disponível em: <<https://transparency.x.com/en/reports/countries/br.html>>. Acesso em: 16 abr. 2024.

X/TWITTER. Compliance streams | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/compliance/streams/introduction>>. Acesso em: 4 abr. 2024.

X/TWITTER. Conversation ID | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/conversation-id>>. Acesso em: 4 abr. 2024.

X/TWITTER. Customer directory Enterprise data customers. **X/Twitter**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/products/twitter-api/enterprise/directory>>. Acesso em: 3 abr. 2024.

X/TWITTER. DSA Transparency Report. **X/Twitter Transparency Center**, [S.d.]. Disponível em: <<https://transparency.twitter.com/dsa-transparency-report.html>>. Acesso em: 16 abr. 2024.

X/TWITTER. GET /2/spaces | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/spaces/lookup/api-reference/get-spaces>>. Acesso em: 4 abr. 2024.

X/TWITTER. GET /2/tweets/:id/quote_tweets | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <https://developer.twitter.com/en/docs/twitter-api/tweets/quote-tweets/api-reference/get-tweets-id-quote_tweets>. Acesso em: 4 abr. 2024.

X/TWITTER. Getting Started with the Twitter API | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api>>. Acesso em: 4 abr. 2024.

X/TWITTER. GET users/lookup | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/v1/accounts-and-users/follow-search-get-users/api-reference/get-users-lookup>>. Acesso em: 4 abr. 2024.

X/TWITTER. Migrating to the New Twitter API v2 | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/migrate/overview>>. Acesso em: 4 abr. 2024.

X/TWITTER. Relatórios de Transparência, **X/Twitter**, [S.d.]. Disponível em: <<https://transparency.x.com/pt/reports.html>>. Acesso em: 16 abr. 2024.

X/TWITTER. Search Tweets introduction | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/tweets/search/introduction>>. Acesso em: 4 abr. 2024.

X/TWITTER. Timelines introduction | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/tweets/timelines/introduction>>. Acesso em: 4 abr. 2024.

X/TWITTER. Transparency Report: Brazil, **X/Twitter**, [S.d.]. Disponível em: <<https://transparency.x.com/en/reports/countries/br.html>>. Acesso em: 16 abr. 2024.

X/TWITTER. Tweet object | Data dictionary: Standard v1.1. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/v1/data-dictionary/object-model/tweet>>. Acesso em: 4 abr. 2024.

X/TWITTER. Tweets lookup quick start guide | Docs. **X/Twitter Developer Platform**, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/tweets/lookup/quick-start>>. Acesso em: 4 abr. 2024.

X/TWITTER. Update on X's Rate Limits. **X/Twitter**, [S.d.]. Disponível em: <<https://business.x.com/en/blog/update-on-tweets-limited-usage.html>>. Acesso em: 4 abr. 2024.

X/TWITTER. X API Documentation | Docs. X/ Twitter Developer Platform, [S.d.]. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api>>. Acesso em: 4 abr. 2024.

X/TWITTER. X Privacy Policy. X/Twitter, [S.d.]. Disponível em: <<https://twitter.com/en/privacy>>. Acesso em: 4 abr. 2024.

X/TWITTER. X's automation development rules. X/Twitter Help Center, 3 nov. 2017. Disponível em: <<https://help.twitter.com/en/rules-and-policies/x-automation>>. Acesso em: 4 abr. 2024.

YASSERI, T. From Print to Pixels: The Changing Landscape of the Public Sphere in the Digital Age. SSRN Scholarly Paper, 2019. [Manuscrito em pré-publicação] Disponível em: <<https://papers.ssrn.com/abstract=4543907>>. Acesso em 26 abr. 2024.

YOUTUBE. API YouTube Data (v3) – Calculadora de cotas | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <https://developers.google.com/youtube/v3/determine_quota_cost?hl=pt-br>. Acesso em: 4 abr. 2024.

YOUTUBE. Channels: list | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3/docs/channels/list?hl=pt-br#try-it>>. Acesso em: 4 abr. 2024.

YOUTUBE. Conseguir credenciais de autorização | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <https://developers.google.com/youtube/registering_an_application?hl=pt-br>. Acesso em: 4 abr. 2024.

YOUTUBE. Histórico de revisões | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <https://developers.google.com/youtube/v3/revision_history?hl=pt-br>. Acesso em: 4 abr. 2024.

YOUTUBE. Implementação: comentários | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3/guides/implementation/comments?hl=pt-br>>. Acesso em: 4 abr. 2024.

YOUTUBE. Search | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3/docs/search?hl=pt-br>>. Acesso em: 4 abr. 2024.

YOUTUBE. Termos de Serviço. Google for Developers, [S.d.]. Disponível em: <<https://www.youtube.com/t/terms>>. Acesso em: 4 abr. 2024.

YOUTUBE. Videos: list | YouTube Data API. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3/docs/videos/list?hl=pt-br>>. Acesso em: 4 abr. 2024.

YOUTUBE. Visão geral da API YouTube Data. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3/getting-started?hl=pt-br>>. Acesso em: 4 abr. 2024.

YOUTUBE. YouTube API Services Terms of Service. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/terms/api-services-terms-of-service>>. Acesso em: 4 abr. 2024.

YOUTUBE. YouTube Data API. Google for Developers, [S.d.]. Disponível em: <<https://developers.google.com/youtube/v3?hl=pt-br>>. Acesso em: 4 abr. 2024.

YOUTUBE. YouTube Research – How It Works. YouTube, [S.d.]. Disponível em: <<https://research.youtube/how-it-works/>>. Acesso em: 4 abr. 2024.

ZALNIERIUTE, M. “Transparency-Washing” in the Digital Age: A Corporate Agenda of Procedural Fetishism. UNSW Law & Justice Legal

Studies Research Paper Series, Sydney, v. 8, n. 1, p. 39–53, 2021. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3805492>. Acesso em: 16 abr. 2024.

ZUBOFF, S. A era do capitalismo de vigilância. Rio de Janeiro: Editora Intrínseca, 2021.

ZUCKERMAN, E. Why study media ecosystems? Information, Communication & Society, [S.l.], v. 24, n. 10, p. 1495–1513, 28 jun. 2021. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1942513>>. Acesso em 9 abr. 2024.



 WWW.NETLAB.ECO.UFRJ.BR

 NETLAB@ECO.UFRJ.BR