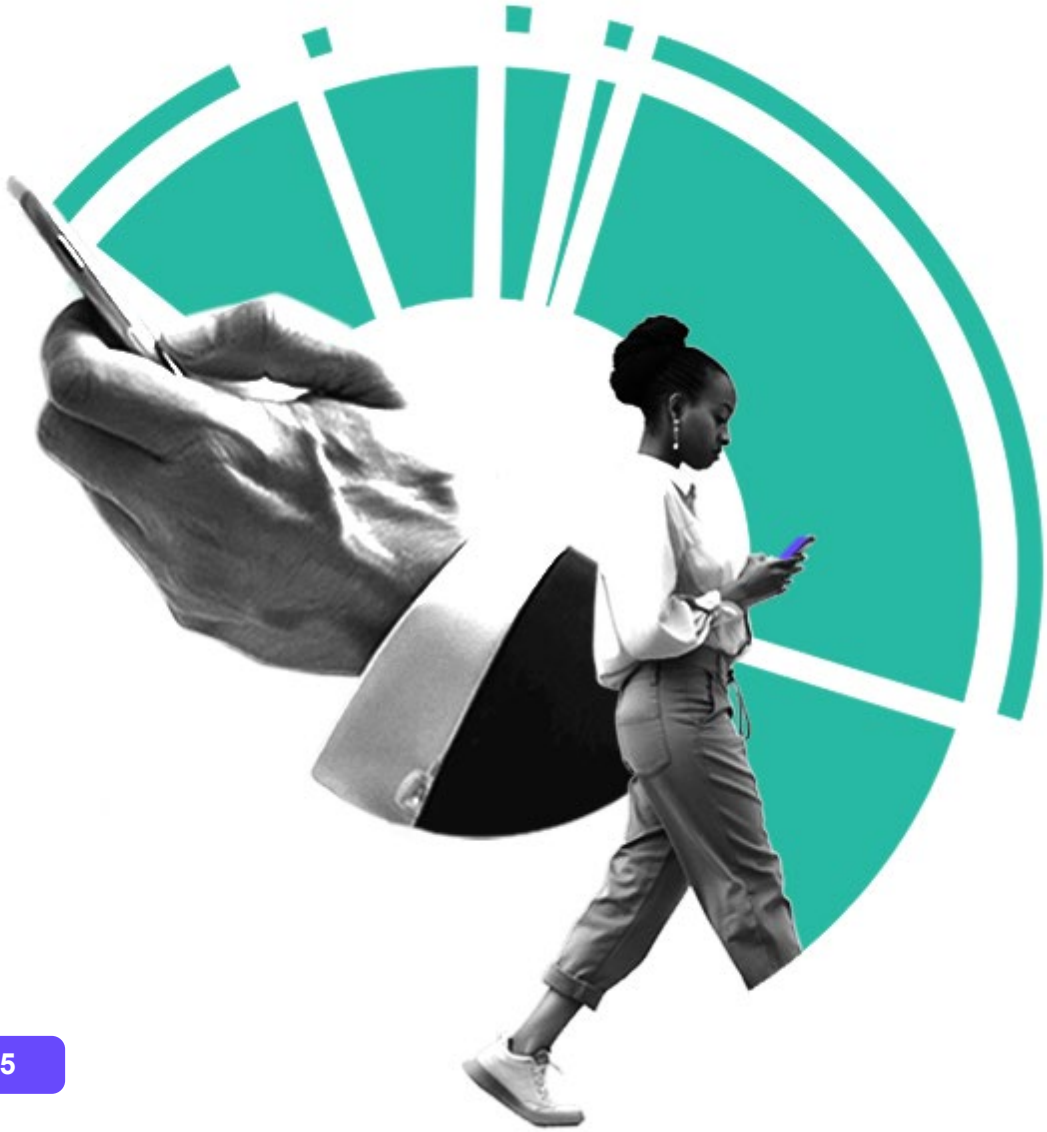
Laboratório de Estudos de Internet e Redes Sociais

UFRJ UNIVERSIDADE FEDERAL DO RIO DE JANEIRO



JULY 2025

DATA TRANSPARENCY INDEX FOR SOCIAL MEDIA PLATFORMS

REPORT BY THE OBSERVATORY ON THE DISINFORMATION INDUSTRY AND ITS IMPACT ON CONSUMER RELATIONS IN BRAZIL

Team

Director

R. Marie Santini

Research coordination

Débora Salles

Researchers

Bruno Mattos
Marcela Canavarro
Carlos Eduardo Barros
Alékis Moreira

Associate Researchers

Priscila Medeiros
Felipe Grael
Fernando Ferreira
Danielle Melo
Marcio Borges

Research Assistants

João Gabriel Haddad
Lucas Murakami
Daphne Silva

Technical Team

Erick Dau
Felipe Loureiro
Lena Benzecry

Technical Review

Adriano Belisário

How to cite

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, Eri- ck; LOUREIRO, Felipe. Data Transparency Index for Social Media Platforms. Rio de Janeiro: NetLab - Internet and Social Media Research Lab, Federal University of Rio de Janeiro (UFRJ). Published on July 9, 2025.

Laboratório de Estudos de Internet e Redes Sociais

ECO

UFRJ Universidade Federal do Rio de Janeiro

Support

Ministério da Justiça e Segurança Pública

Governo Federal Brasil União e Reconstrução

School of Communication| University Federal do Rio de Janeiro
Av. Pasteur, 250, Urca| Rio de Janeiro - RJ
ZIP CODE 21941-901

WWW.NETLAB.ECO.UFRJ.BR

NETLAB@ECO.UFRJ.BR

Summary

Presentation

Observatory of the **Disinformation Industry** and its impact on consumer relations in Brazil

This report is one of the outcomes of the research carried out as part of the *Observatory on the disinformation industry and its impact on consumer relations in Brazil*, a partnership between **NetLab UFRJ** and the **National Consumer Secretariat of the Ministry of Justice and Public Security** (Senacon/MJSP).

The Observatory's main objective is to provide inputs that support **public policies for consumer protection** by analysing infrastructure, the political economy and strategies for manipulating consumer relations, as well as analysing public opinion on disinformation and influence operations that have been gaining space on social media platforms.

In view of the scarcity of qualified information for applied social research based on digital data, we present in this report the **Data Transparency Index for Social Media Platforms** (**DTI**) in Brazil.

The **DTI** follows a structured, systematized and reproducible roadmap, based on data quality criteria, to evaluate mechanisms for accessing data of public interest made available by platforms.

It assesses the level of transparency and data quality of the main social networks and messaging apps operating in Brazil: YouTube, Facebook, Instagram, X/Twitter, Telegram, TikTok, Kwai and WhatsApp. The evaluation is based on access to **user-generated content data, published without paid boosts to the platforms**.

Based on the scores obtained, the platforms are classified into **five levels**: **irrelevant transparency** (0 to 20), **precarious transparency** (21 to 40), **regular transparency** (41 to 60), **satisfactory transparency** (61 to 80) and **ideal transparency** (81 to 100).

The **DTI** is part of a broader study on the transparency of social media platforms in accessingdata of public interest. The research also includes the **ATI - Social Media Platforms Advertising Transparency Index**, which assesses different parameters and dimensions of quality in the provision of data on advertisements, i.e. any driven by payment to the platforms.

Summary Executive

Main *Results*

No social media platform evaluated reached an ideal or satisfactory level in the 1st edition of the Advertising Transparency Index (ATI).

Meta obtained the best score, with 49.8 points, which is considered a regular level of transparency. Telegram, LinkedIn and Google come next, with 22.8, 18.3 and 8.2 points respectively, with transparency levels classified as poor.

Recurring bad practices include limiting collection to a low daily volume and delivering inconsistent results with each request.

Even when they offer official means of collection, there are still restrictions that make it difficult to optimize data extraction work programmatically and at scale, with frequent errors in the retrieval process, an insufficient number of results and limitations on creating new API access , making simultaneous collection difficult.

Inaccessible universe: most platforms do not offer a free API or data interface.

By preventing the systematic, cost-free mapping and retrieval of data, the platforms impose technical and economic barriers to research, which result in the biasing of analyses carried out with data that not representative of the objects of study, also impacting on the reproducibility of research.

Recurring bad practices include limiting collection to a low daily volume and delivering inconsistent results with each request.

Significant differences in how data of public interest is transfered comparing Brazil, European countries and the USA suggest that satisfactory minimum criteria for data access and quality need to be established in a fair and equal manner in all the countries where these companies operate.

With CrowdTangle, Facebook and Instagram offered regular access to data of public interest. With the tool discontinued, scores tend to fall.

CrowdTangle was discontinued by Meta in August 2024, the month in which the Brazilian election began. Even with problems of completeness and consistency, without the tool, the scenario for research has worsened. Meta's restrictions on data scraping exacerbate the problem of lack of transparency.

X/Twitter, Telegram, Kwai and WhatsApp do not provide transparency reports on moderation actions and specific governance on the Brazilian context.

Although Kwai presents biannual transparency reports on moderation activity, it does not provide specific data on Brazil. Since 2021, X/Twitter has not released any transparency reports for Brazil, joining Telegram and WhatsApp, which have also not released any such reports.

TikTok does not offer collection tools and access to data in Brazil, but stands out for their transparency moderation and governance activity.

TikTok is the only platform analyzed that provides information on the volume and types of violations infringed and moderated by the platform, as well as requests for moderation and data requests made by the Brazilian state.

However, it does not provide information on the metadata of the content and users moderated, making it impossible to audit these actions.

The granularity of the data in transparency is unsatisfactory, making it difficult to analyze by country and identify the types of content violations removed.

Since the platforms do not provide detailed data on their moderation activities, such as metadata on specific cases of publications removed, it is impossible to evaluate the application of their governance policies.

Recommendations *Priorities*

Allow for collecting the universe of public data

- It is recommended that an official and free **API be made available for full, programmatic and free access** to the universe of public data for researchers.
- In addition, we recommend offering an **interface** for data collection that is easy to use and navigate and does not require technical programming knowledge.
- Another factor to be improved on all platforms is the **permission to scrape data** for **research** purposes, offering more legal certainty for researchers and facilitating the process of systematic data collection.

Quality of the data made available

- Researchers must be given access to temporary data and historical data, as well as providing unexpired links so that permanent photo and video content can be analyzed.
- It is also important that platforms have transparency about when posts are removed and users are suspended, providing access to their metadata, even if the content of the posts is restricted.
- Offering customized keyword searches is essential for collecting relevant data for research purposes.

Clear API documentation and terms of use

- It is important that the API documentation is **made publicly** available, without the need for individual requests, in **Portuguese**, and with a clear description of its **Kso terms**.
- The official API documentation should list the **possible errors** of each available endpoint and offer **clear and understandable examples** how to make the requests to get the data.

Provide detailed transparency reports for Brazil

- Transparency reports on content moderation should be published periodically and with a guarantee of data quality. In order to analyse the moderation and governance policies of all platforms in the country, the data must be made available in greater detail and with a specific focus on Brazil.

In order to improve the moderation and governance of all platforms in the country, data should be made available in greater detail and with a specific focus on Brazil.

The Importance of Searching with Digital Data

There are many different reasons why digital data has the potential not only to identify, describe and understand relevant social phenomena and problems, but also can help deal with them or even solve them. Aspects such as data volKle, sample granularity, timeliness and non-intrusiveness in when collecting this data allow Computational Social Sciences - as well as areas such as Biomedicina (Correia *et al.*, 2020), Economics (Tumasjan, 2023), Public Health (Terry *et al.*, 2023), Biodiversity (Chowdhury, 2024) and the Environment (Ghermandi *et al.*, 2023), among others - to carry out research with sufficient relevance and connection to social reality. In this way, digital data can **inform public and governance policies** that will have an impact on the development of the technologies themselves and the type of social spaces they offer.

A key aspect of understanding the relevance of social media data for research is its growing influence on public life, as it is now one of the main arenas for public debate (Staab; Thiel, 2022; Yasseri, 2023). This involves not only political and social topics that circulate organically, but also planned campaigns to influence tastes, consumption habits, lifestyles, opinions and behaviors.

However, ordinary citizens are subject to the use of their data for a multitude of commercial actions that are currently difficult to audit, even though they have such an impact on individual and collective life. When the user accepts a contract signing up to use the platform and agreeing to its terms of use, a commercial link and a consumer relationship is created which, like all other relationships of this kind, must offer means of protecting the consumer.

The role of the academic community is to reduce the **asymmetry of information** that tends to be detrimental to the consumer: the platform knows a lot about the user, but the user knows little about what is done with the data extracted and mined by the platform.

This is, therefore, the most far-reaching gain of research with digital data: the possibility of **measuring the impact of the technologies themselves** on social life, since much of the recent development of computational methods "is aimed at analyzing the structure and dynamics of human communication" (Van Atteveldt; Peng, 2018, p. 81). Natively digital data are not produced specifically for research, but rather as a result of interactions and social processes that take place online (Edwards et al., 2013; Lee et al., 2008; Marres, 2017). In other words, the data is produced and collected without the researchers interfering with the object study, using a **non-obstructive approach** (Rogers, 2009).

In order to analyze this data, researchers must have in-depth knowledge of the workings of the digital environment from which it was extracted, since the "tends to be strongly marked by the effects of the platform, such as the search for suggested terms in autocomplete functions or the use of hashtags that are trending on the platform in question" (Shaw, 2015, p. 2). This is one of the reasons why digital data has brought new types of information to be analyzed, as well as new research challenges. New methods have emerged to explore them and to understand the different characteristics of digital platforms and, in particular, social media platforms, which shape the social dynamics in these environments (Rogers, 2009). These new methods have also

brought challenges for researchers, who must apply them with a focus on the public interest and contemporary social problems.

As digital methods are geared toward processing and finding relevant information among massive amounts of data, they have also been widely used in various areas of knowledge. This is the aspect that places the **Computational Social Sciences** as a scientific discipline that is also **instrumental** for other areas (Cioffi-Revilla, 2018). Epidemiological and pharmacovigilance studies, for example, have been using digital and computational methods for years (El-Sayed *et al.*, 2012; Pappa; Stergioulas, 2019), approaches that were useful during the Covid-19 pandemic in the early 2020s in aspects such as crowd control (Cecilia *et al.*, 2020), prediction of disease progression, data visualization and prognosis definition (Senthilraja, 2021).

Why is digital data research important?

- It allows for measuring the imimpact of the technologies themselves on social life;
- It enables analysis of behavior in a non obtrusive way, in addition to traditional methods that depend on participant self declaration;
- It facilitates collaborative work, as it makes it easier to share and reuse data;
- It provides a foundation for updating public policies based on a new type of evidence;
- It increases the scale from a relatively small sample universe to possibilities of large data samples, facilitating the study of subtle and niche relationships or effects;
- It provides a foundation for based on the variety and granularity of the samples;
- It makes it easier to observe people's behavior in real environments, rather than simulating it in a laboratory.

SOURCES:: Van Atteveldt and Peng (2018) and Marres (2017)

One of the most pressing objectives in this context is research that helps to produce diagnoses and overviews, collect evidence and indicate solutions to relevant current phenomena, such as disinformation, the use of technologies by children and adolescents, cybercrime, the impact of social networks on mental health and new consumer habits, among others. These and other aspects affect Brazilian society in different ways and require in-depth understanding, which can be provided by quality research.

DTI aims to contribute to research with digitaldata, especially in the Applied Social Sciences. Its mission is to improve free and universal access to any public data on social media platforms that is useful for **academic research of public interest**.

Although the concept of "public interest" is a broad one, the understanding of the National Data Protection Authority (ANPD) to apply special conditions to the processing of personal data is that of the public interest.

According to the LGPD (*Lei Geral de Proteção de Dados Pessoais* – Data Protection Act), this is part of the nature of the body carrying out the research – an understanding that is also adopted here to define **research in the public interest**. Thus, we understand that access to public data from social media platforms should be granted to "direct or indirect public administration bodies or entities or non-profit private legal entities legally constituted under Brazilian law, with headquarters and jurisdiction in the country, which **include basic or applied research of a historical, scientific, technological or statistic nature in their institutional mission or corporate purpose**" (ANPD, [*S.d.*], p. 26). This concept includes public and private non-profit universities, research foundations and institutes and public bodies such as Ipea, IBGE and Fiocruz, for example.

Content data that is public, open and accessible by any user of social media platforms must be available for research, **analysis and evaluation in the public interest**. The **Brazilian Data Protection Act** stipulates that data made manifestly public by its owner **does not require consent for Kso**, a definition that makes it possible, within the scope of the purpose of this index, to conduct research with data originating in public opinion spaces with the potential to impact countless citizens. Thus, public data from social media platforms can be accessed by non-profit public administration bodies or entities, or by legally constituted entities based in Brazil, whose **research mission is "of a historical, scientific, technological or statistical nature"**(Brasil, 2018).

This definition also includes public groups and channels on WhatsApp and Telegram, when they are shared on the internet to attract new members and which any user can join (Evangelista; Bruno, 2019; Garimella; Tyson, 2018; Resende *et al.*, 2019). Since anyone with the link to access the group can join, researchers have established ethical criteria about what can be considered public on these apps. Online listings of public groups and search tools have been useful in allowing reaserchers to find and identify them (Melo, 2022; Garimella, Tyson, 2018; Resende *et al.*, 2019).

In the first decade of the 2000s, research into messaging apps was limited, due to small samples and qualitative methodologies, such as interviews and surveys with group participants, for example (Garimella; Tyson, 2018; Rosenfeld *et al.*, 2018). However, the widespread use of messaging apps in the organization and flow of political mobilizations has increased interest in studying large volumes of data from public WhatsApp and Telegram groups (Moura; Michelson, 2014; Treré, 2020). As they are important channels for the propagation of political information, these apps are becoming increasingly relevant for social research (Evangelista; Bruno, 2019; Ozawa *et al.*, 2023; Melo *et al.*, 2019; Wendratama; Yusuf, 2023; Calvo-Gutiérrez; Marín-Lladó, 2023; Smith *et al.*, 2023).

Public digital data

- Visible content made public by the their authors;
- Temporary content defined as publicly visible for as long as are on the air;
- Messages sent in WhatsApp and Telegram groups and channels posted in open internet spaces.

Non-public digital data

- Publications and videos defined as private by their authors;
- Content by locked accounts;
- Unlisted or private videos on YouTube
- Messenger, Instagram and X/Twitter chat messages;
- Private conversations between two or more users, in one-to-one messages or in private groups, on WhatsApp and Telegram.

Digital Data Blackout

Although social media platforms have been **gaining relevance** as spaces for shaping **public opinion**, their transparency and public awareness about how they work are not proportional to their potential impact on social life. In addition to the **opacity** of the algorithms (Lu, 2021) and the difficulty in accessing the data (Bruns, 2019), social media platforms **are also constantly changing** the operating rules **embedded in the algorithms**. This is not always done in a transparent way, generating important ethical and political concerns (Selinger; Hartzog, 2016; Guess *et al.*, 2023).

Digital data is vital to understanding how online platforms impact political and social dynamics, but access to complete, **up-to-date**, **consistent, timely, legally compliant** and research-relevant data is increasingly limited, hampered or even extinguished by ***big tech*** (Greene; Martens; Schmueli, 2022). When analyzing the Facebook, Instagram and X/Twitter scenario, Bruns (2019) considers that big tech seeks to limit the development of critical data-based research on the platforms. At the same time, they launch tools access incomplete and inconsistent data which serve only to generate positive publicity based on a false image of active transparency and self-regulatory capacity.

Among the most emblematic examples is the **closure of CrowdTangle** by Meta (Soares, 2024), **in August 2024**; the tool guaranteed access to Facebook and Instagram data. This move goes hand in hand with X/Twitter which, since March 2023, has been charging a high fee for access to its API, which was previously free (Mozelli, 2023). The trend, however, is not restricted to recent years; shortly after the Cambridge Analytica scandal. In 2018, **Facebook** modified and restricted features of its API without prior notice, hindering the development of academic research (Bruns, 2019).

Faced with this scenario, researchers from various institutes and universities around the world frequently criticize the new restrictions imposed on access to APIs from social media platforms, claiming that the decisions could make it **impossible to carry out research** (Coalition for Independent Technology research, 2023; Mozilla Foundation, 2024). However, the lack of binding criteria on the subject enables *big tech* to ignore this type of request on a recurring basis.

Digital platforms' dominace in content distribution has allowed them to establish strong policies and discourses in defense of their prominent position in the market and actively influence the way they are seen and understood in the public eye. While social media platforms seek to be perceived as facilitating the production and distribution of content user-generated content, they try to distance themselves from having any responsibility for what these users publish and from the perception that they are algorithmically curating content (Gillespie, 2010). Reducing access to data is often part of these efforts. Bossetta (2020) points out that, although they sometimes adopt measures that can help increase transparency in the online sphere, the platforms' operations are motivated by economic interests. They also have the material conditions to quickly change their organizational bodies and software architecture in the face of scandals, remedying reputational crises in which they become involved.

"The information that platforms decide to include – and, more importantly, exclude – [...] can be seen as a positioning imposed through design"

(Bossetta, 2020, p. 2).

Despite the great influence of social media platforms on social and political life since the 2010s (Zuboff, 2021; van Dijck; Nieborg; Poell, 2019; Gerbaudo, 2021), they still have **no legal obligation as to how their data should be made available** to interested parties, such as researchers, for example. The difficulty in accessing and/or using data from social media platforms is currently a problem.

This is a **major obstacle for researchers**, especially in the Global South, who receive lower quality and even more incomplete data than their counterparts in the Global North (Orembo; Berger; Simon, 2023; Lurie, 2023).

The discriminatory practice of offering databases with different levels of quality and completeness according to geographic location suggests creating regulation mechanisms to force big tech to open up their data for public scrutiny by researchers is not sufficient — satisfactory access and quality criteria must also be added to the data access process. For example, studies on the application of the **Law on Access to Information (LAI)** in Brazil(2011) show that the legal obligation to provide information of public or individual interest is not always sufficient (Transparência Brasil, 2018; Braga; Cunha, 2022; Santos, 2023), if it is not accompanied by **objective data quality criteria**. Standardizing acceptable criteria for quality, transparency and portability is important to ensure predictability for researchers about the **longevity of data quality**.

Against this backdrop, the index proposed here is based on the premise that there is an urgent need to reduce the asymmetry in how much access different countries have to data produced on social media platforms. National differences between access to and the quality of the data made available in Brazil and in countries such as the USA and the European bloc nations were considered in the analysis of the results, as well as contributing to the development of the evaluation form. We found that a considerable number of the parameters **not met** by the platforms in the Brazilian operation **are made available in the European bloc** by these same platforms, showing that the low quality of the data bases made available here is politically, and not technically motivated.

Thus, in parallel with the effort to guarantee and expand access to data on social media platforms for academic research, the definition of quality indicators is fundamental for **monitoring and evaluating the transparency** offered by these platforms. In this way, it is possible to provide evidence of incomplete, inconsistent and lowquality data that impedes access or makes it difficult to **systematically analyze**, **diagnose risks** and **take responsibility for possible damage**. The **standardization of quality** and transparency **criteria** also establishes common parameters that enable **systematic comparison and monitoring** of different platforms.

Glossary

API: official access to data programmatically

API stands for **Application Programming Interface.** Its basic functionality is to set up **communication between two software components** (Goodwin, 2024) – such as a database and a researcher's device - by means of data requests that must comply with their own definitions and protocols. APIs thus make it possible to **share data** (PostMan, [*S.d.*]) between applications, systems, devices and platforms of different kinds, facilitating interoperability between different systems.

In ordre to gain access to an API, tokens must firstly be generated. **These are small codes given to each** registered **programmer** or researcher and which act as passwords to authenticate and validate requirements. Other specific components of the API include the endpoints, with which the programmer **makes a request to the server and receives the response** (CloudFlare, [*S.d.*]), according to the instructions indicated in its documentation.

Data collection interface: official access to data

This refers to the application that allows **visualization, consulting, exploring** and **downloading** publications, making data available in an accessible way and with specific solutions.

The data collection interface should allow you to search for publications using specific search terms or published by pages of interest, generate summaries and visualizations, and export them in a format which is suitable for more in-depth analysis.

Unlike the API, the data collection interface facilitates access to the publishing ecosystem as it does **not require programming knowledge**, but on the other hand, **does not allow the automation of data collection processes and systematic monitoring of posts.**

Platform user interface
X Data collection interface

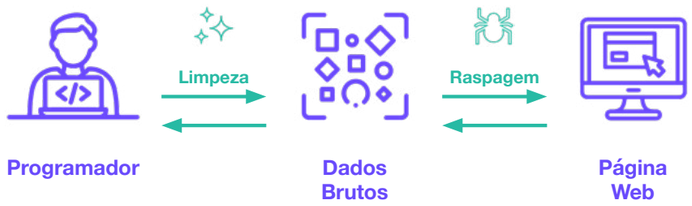
An interface is an online environment that allows interaction between a user and a database through graphic elements such as icons, windows, menus and other visual indicators. In DTI, we refer to two types of interface

- **Platform user interface:** the platform environment accessed by users to view content, interact with other users and maintain their respective profiles, whether through a website or through an application.
- **Data collection interface:** an environment-designed for researchers and other parties interested in consistently exploring and monitoring discussions around relevant topics on social media platforms with the aim of understanding social phenomena and market trends. It refers to the application that allows data from user-generated posts to be viewed, consulted, explored and downloaded and made available in a user-friendly manner. It should be accessible and have specific solutions. Ideally, the data collection interface should make it possible to search for publications using specific search terms.

Scraping techniques: unofficial access to data programatically

Web scraping is a **process of extracting and combining online content** (Bar-Ilan, 2001; Mooney; Westreich; El-Sayed, 2015). Because of the limitations found in official means of data collection, such as APIs, it is not uncommon for researchers to have to resort to web scraping-based solutions to develop studies. Often, scraping techniques are not allowed in the **platforms' usage policies**, and there are **technical obstacles to prevent them**, which requires increasing and continuous effort from researchers to collect and analyze data.

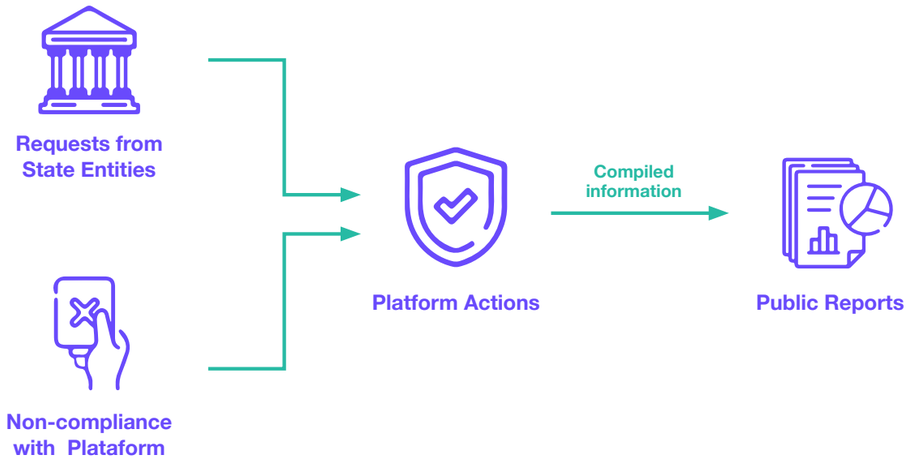
In addition, various cases of legal disputes imposed by platforms on people who have scraped data (Roth, 2023; Conger, 2016) point to the problem of **legal uncertainty** that this type of methodology imposes. Platforms claim that they make scraping difficult in order to prevent wrongdoing and guarantee the integrity of users. The reasons listed include protecting individuals' data (Meta, [S.d.]), ensuring the authenticity of their user base, preventing bots (X/Twitter, [S.d.]) and possibly preventing other companies Artificial Intelligence (AI) models from being fed public user data (Mehta, 2023).



What are transparency reports?

Traditionally, transparency reports are **voluntary** documents produced by platforms. They may come as a result of requests from authorities or noncompliance with the platform's terms of use. The decisions behind these actions can include, for example, the **provision (or not)** of non-publicly available information to **state entities** and the **moderation** or **removal** content.

Some of the frequent criticisms of the reports are related to the availability of aggregated data with unsatisfactory granularity and the absence of concrete examples that allow moderation policies to be detailed. The excessive focus on government actions, the under-reporting of platform actions and the lack of standardization between different reports are also negative points.



Why Measure Transparency from Quality Data?

"The ability to measure [data] quality is dependent on the ability to measure the degree to which the data meets the stipulated requirements"
(Benson, 2019)

Ideals of transparency have long been part of models of understanding and accountability for different systems and involve, on the one hand, knowing **how they work** and, on the other, **how it is possible to govern them** (Crawford, 2021). Broadly speaking, data governance involves both the **governance of information technology resources** of companies and organizations and **corporate governance** and "focuses on **principles of organization** and control over these inputs [data]" (Barbieri, 2019, p. 36) in order to enable the best use of the information and knowledge they can generate.

Data governance is not a new area, since organizations of all sizes have always had to deal with data. However, the digital revolution has come to involve the intersection of new disciplines, always "with a **central focus on data quality** in the broadest sense", which involves improvements in data production, monitoring its use and different critical aspects of security, privacy, ethics and compliance (Barbieri, 2019, p. 36).

Social media platforms are already **actively developing data governance processes and policies**, because they are essential for extracting useful information and knowledge for their businesses. However, it is necessary that, amid corporate and technological governance, social media platforms centrally consider the public interest in **data governance**. To this end, it is essential to define **mechanisms** that allow researchers to analyze the processes **that provide or produce data in these spaces**.

One of the difficulties of managing data is that it tends to seem "abstract" and "immaterial" and, as a result, data-based systems can "easily fall outside (...) responsibilities of care, consent and risk" (Crawford, 2021, p. 113). Recent legislative initiatives seek to circumvent these difficulties and mitigate the **negative social effects of the high influence of social media platforms on social life** by increasing their responsibility as intermediaries, regulating their services and imposing binding legal obligations with regard to data transparency for research. Examples include **PL 2.630/2020**, a bill of law which has been under discussion in Brazil for four years and was withdrawn from the agenda in April 2024, the **Digital Services Act (DSA)** in force in the European Union (European Commission, [S.d.]) and the **Platform Accountability and Transparency Act**, currently before the American Senate, after being reintroduced in June 2023 (Perrino, 2023).

Transparency, a central aspect of the current debate on the governance of digital platforms, is understood in this context as the practice of platforms to provide information that are in the public interest (Urman; Makhortykh, 2023). Accessing information on the functionalities and application of platform policies is fundamental to enabling society, governments and other stakeholders assess the performance of these companies (Urman; Makhortykh, 2023).

Data-based systems, such as digital platforms and particularly social media platforms, are highly dependent not only on the quality of the data that feeds them, but also on **how that data is used throughout its life cycle**. IIn general terms, data quality measures establish objective criteria that help identify how suitable an organization's data is for its intended purpose (Mahanti, 2018).

These measures help to assess whether the data complies with defined, internationally recognized and employed dimensions, such as those of **ISO 8000** ([S.d.]).

Previous Experiences with Data Transparency Indexes

One source of important lessons in handling data in the public interest in Brazil is the **Access to Information Law (LAI)** (Brasil, 2011), which has been in force since 2012. The LAI resulted in advances in the transparency of nonconfidential public data and led to the accumulation of **good references in the standardization and evaluation of quality criteria in public data, which can inform similar initiatives for data from social media platforms.**

In the **public sector**, some problems concerning access to information have been identified using various transparency indexes that evaluate the portals of the Executive, Legislative and Judiciary branches at municipal, state and federal levels, as well as Public Prosecutors and Audit Courts. Examples of transparency and quality indexes are:

- **Transparency Index for Legislative Portals (ITpL)**: created by Brazilian Congress (2022), is applicable to all legislative houses in Brazil and has five evaluation criteria, with different weights: availability of information, timeliness information, use of Simple Language techniques in texts; existence of historical series; and availability of information in open data.
- **Public Transparency Quality Seal:** created by Atricon (Association of Members of the Brazilian Court of Auditors) (2023), the seal is based an index of "active transparency" on the portals of the Executive, Legislative and Judiciary branches, as well as the Courts of Auditors, Public Prosecutors' Offices and Public Defenders' Offices, at the federal, state, district and municipal levels.
- **Active(ITA) Transparency Index)**: created by the Government of Brasilia (2023), This index measures the degree to which bodies and entities of the Federal District's executive branch are in compliance with the LAI. In 2023, the ITA also assessed passive transparency.

***

In addition, organized civil society has also developed indexes to assess the transparency of data of public interest, mobilizing action by state entities and fostering public debate on sensitive topics. One successful example was the Covid-19 Transparency Index:

- **Covid-19 Transparency Index:** created by Open Knowledge Brasil ([S.d.]), evaluates the transparency of the data disclosed on contagion and health infrastructure for dealing with the Covid-19 pandemic. The index assigns a score and compares the transparency initiatives of states and capitals.

ISO 8000

ISO 8000 is a set of international standards that specifies the requirements for data quality and integrity in portability between different software or applications. Developed by the **International Organization for Standardization (ISO)**, these standards cover aspects of data quality management, such as measurement methods, validation processes and record-keeping practices.

ISO 8000 is important because it establishes a common standard across different countries for the data and information portability with quality guaranteed. From a technical point of view, the standard defines portable data as: "data that can be moved from one software application to another without losing meaning" (Benson, 2019). The standard addresses aspects such as the ability of data to be read in any application without losing information or having to pay fees or royalties. To this end, it defines standards for data encoding, the presentation of metadata and the types of data that should be included in cases of portability.

Guaranteeing the quality of the data collected and analyzed is a way of ensuring the **reliability** and **reproducibility** of studies, as well as enabling important generalizations for social research (Srivastava; Mishra, 2021). Ultimately, it increases the **level of transparency for the functioning of the system as a whole**. In addition, the quality of the data allows academic research to verify that the transparency measures announced by social media platforms are in fact being implemented, if not throughout the entire data life cycle.

Although the adoption of standardizations such as ISO 8000 is optional, they provide an international benchmark for quality and transparency that brings greater confidence to action taken in the public interest and the data-driven economy, as well as making it possible to maintain the usefulness and quality of data in the long term.

According to the Electronic Code Management Association (ECCMA)[1], "poor quality data is (ECCMA)[1], "the main cause of transparency problems increasing the costs of regulatory compliance" ([*S.d.*]).

The high costs of regulatory compliance to which the ECCMA refers are likely to get even higher if the issue of transparency is not solved in parallel with the growing legal demands for protection of personal data, following the arrival of the so-called fourth-generation data protection laws, such as the **General Data Protection Regulation (GDPR)** in the European Union (2016), and the **General Data Protection Law (LGPD)** in Brazil (2018), aas well as the growing demands related to generative AI.

One of the compliance measures most widely used by platforms today are **transparency reports**, which are regularly made available by companies. These contain information on

[1] *A ECCMA is a non-profit organization and the administrator of the International Organization for Standardization's ISO) technical advisory group in the USA.*

content removal, including, those removed at the request of governments (Urman; Makhortykh, 2023). Traditionally, transparency reports on moderation are voluntary documents with no mandatory format. In many cases, they only promote "moderate visibility" (Wagner *et al.*, 2020), since, without the means for an **independent audit**, it is not possible to know whether only certain information and that promote a **favorable perspective** are released **to the platforms**. Recently, the DSA made it compulsory for large online platforms and search engines to publish this type of report on their activities in European Union countries. The first reports after the new rule were made available in November 2023.

The history of transparency reports produced by social media platforms is still recent. The first was published by Google in 2010 (Google, [*S.d.*]), followed by Twitter (X/Twitter, [*S.d.*]), in 2012. The platforms found themselves under pressure from different companies and by local governments, to take down copyrighted content and to take down illegal content. But the requests also involved restrictions on citizens' freedom of expression on sensitive political issues, especially (but not only) in countries with authoritarian governments. This prompted Google, and later other platforms, to publicly report cases in which the removal of content was related to court orders or government requests.

One of the common criticisms of the platforms' transparency reports is their focus on **aggregate data, with unsatisfactory granularity**, and without the inclusion of concrete cases that allow for a more detailed examination of the policies applied in content moderation (Kosta; Brewczyńska, 2019; Suzor *et al.*, 2019). In other words, the concepts and justifications for content moderation are usually not transparent. Another criticism from researchers is that platforms tend to make public and give a lot of visibility to government requests to remove content and profiles, however there is a deliberate underreporting of moderation actions decided on and carried out by themselves on a daily basis (Hovyadinov, 2019; Kosta; Brewczyńska, 2019). These actions raise suspicion over "*transparency-washing*" (Zalnieriute, 2021), a term used to describe the corporate practice of distracting the public with certain issues in order to minimize the accountability of digital platforms, including social media platforms, for the content moderation decisions they make arbitrarily and proactively, moving them away from their commitment to transparency (Urman; Makhortykh, 2023).

The most recent transparency reports on content moderation in Europe, after the legal binding DSA came into force, have also received criticism. One of them is that there is a lack of standardization between different reports, making it impossible to make any analytical comparisons between platforms. Among the most notable aspects are the **differences in the quality and level of granularity** of the data made available (Miller, 2023).

Therefore, in order to assess different perspectives on transparency policies and practices on platforms, it is important to carry out a comprehensive analysis of the resources available, such as:

- the **availability and quality of the data, through official collection methods** (such as the API and the collection interface). data);
- the possibility of using **unofficial methods** for this purpose (such as data scraping);
- the details of the **official documentation** to guide the use of these resources; and transparency reports.

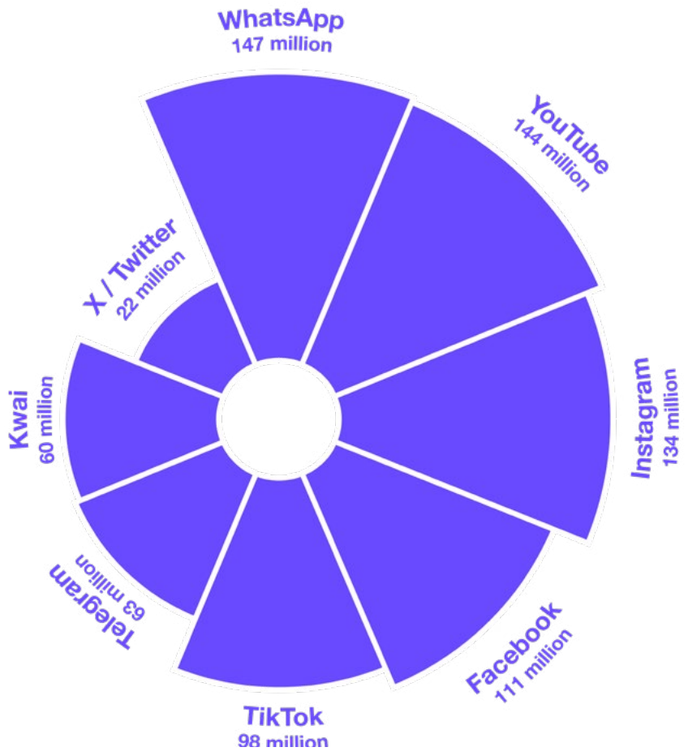
All these aspects are assessed in this **Data Transparency Index (DTI)**, which seeks to reveal the level of transparency of social media platforms in Brazil, using objective criteria that apply to all platforms evaluated. **DTI** is based on the premise that t**he standardization of criteria to systematically evaluate** the transparency and quality of **data on social media platforms**

and those of interest for research tends to have a positive impact on the transparency of these spaces, which affect public life so greatly.

To this end, the DTI is based on quality dimensions recommended by scientific literature, regardless of the particularities of each platform. The **endogenous dimensions** of the data include **completeness, timeliness, consistency and accessibility** (Batini; Scannapieca, 2006; McGilvray, 2008; Loshin, 2008; Barbieri, 2019). Other dimensions, such as compliance and relevance (Barbieri, 2019), depend on exogenous factors and can therefore vary according to the legal regulations in force in each country or the specific objectives of the research. Seen together, endogenous and exogenous dimensions of data quality can indicate the level of transparency of the platforms.

Scope of *DTI*

The **Data Transparency Index for Social Media Platforms** assesses the accessibility and quality of data from **organic and public publications** on the main social media and messaging platforms: **YouTube, Facebook, Instagram, X/Twitter, Telegram, TikTok, Kwai and WhatsApp.**



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

Number of users in Brazil on each platform analyzed

The DTI aims to identify

What are the **transparency** and data **access** measures of the main social media platforms in Brazil?

What is the **quality of** the data made available by these platforms for research activities?

Objectives

Standardize

Define **evaluation parameters** for access and quality of public interest data from social media platforms.

Evaluate

Systematically and consistently identify the strengths and weaknesses of data access and quality.

Compare

Benchmark the performance of each platform using common criteria and a standardized methodology.

Improve

Publicly and objectively indicate, what needs to be improved in the provision of data for research.

Justification: Which platforms are we looking at?

DTI includes the **social media platforms** that are most popular in Brazil for the dissemination of **user-generated content** (dos Santos, 2021). We consider social media platforms to be spaces where users produce and consume content, interacting and connecting with other users (Ellison; Boyd, 2013).

This concept includes so-called messaging apps such as Telegram and WhatsApp. These apps also break down the barrier between interpersonal and mass communication, playing an important role in spreading information and sharing content produced by users in public groups and channels. They also allow for the creation of interaction networks based on common affinities (Júnior et al., 2021; Rogers, 2020), in which users become both empowered content producers and targets for exploitation by companies.

We based our study on the premise of the ***Digital Services Act*** **(DSA)**, which establishes **accountability and transparency** measures for digital platforms that reach **more than 10% of the European Union.**

This criterion was used to select the platforms evaluated in the index, considering the Brazilian context, including the platforms with the **greatest social impact.** As they are the largest and most relevant, they must guarantee investment in a **robust** transparency **infrastructure** and follow the best market practices, as they have the necessary resources to do so.

This version of the Index does not include evaluations of **Pinterest and LinkedIn.** Although they have a significant user base, these platforms have little presence in **academic and scientific research** in Brazil and around the world (Kapoor et al., 2017; Zuckerman, 2021). These and other platforms may be included in future DTI updates.

Methodological Approach

In order to measure the transparency of the main platforms available in Brazil, an evaluation roadmap was drawn up in an iterative and deliberative process that established the parameters and criteria for evaluation and their conceptual definitions.

The parameters were evaluated and justified by eight researchers from NetLab UFRJ, divided into pairs that included a specialist in data collection, infrastructure and processing and another with experience in data analysis and research design in Computational Social Sciences. The teams were also responsible for reviewing peer's responses, as shown in Table 1. Researchers were assigned platforms to analyse based on their previous knowledge and experiience researching data on that platform.

Throughout the process of drawing up the index, the suitability of the parameters and the relevance of their justifications were continually deliberated jointly by the evaluators and other researchers involved in the study. The evaluation was carried out and revised during the first half of 2024.

Platform	Researchers Responsible for Response	Researchers Responsible for Reviewes
YouTube	E1 and E2	E7 and E5
Facebook	E3 and E4	E6 and E2
Instagram	E3 and E4	E6 and E2
X/Twitter	E3 and E5	E6 and E8
Telegram	E6 and E2	E3 and E4
Kwai	E3 and E2	The answers were deliberated together
TikTok	E7 and E5	E1
WhatsApp	E6 and E8	E3 and E4

Table 1: Breakdown of parameter responses by expert pairs (E_n)

The roadmap consists of **40 parameters** that analyze six dimensions of data quality: **accessibility, conformity, completeness, consistency, relevance and quality**. The assessments were made and justified on the basis of five different sources of information: the platform's official documentation, academic literature, access and data collection tests, the accumulated experience of NetLab UFRJ and transparency reports.

Out of all these, we mainly consider **API documentation**. A software's documentation reports, details and explains how it works, showing users how to use it. Platforms that make APIs available often include documentation so that developers can understand them during requests. In addition to the documentation, we consulted the platforms' **policies and terms of use** as well as APIs in order to respond to the parameters assessed. In exceptional cases, we contacted platform support directly for clarification

For the evaluations, we also took into account national and international academic production published in impact journals, with methodologies developed, tested and approved by peers. The academic literature was mainly used to identify and confer the use of unofficial collection methods.

We also conducted controlled experiments between December 2023 and January 2024, simulating real situations of use and data collection, to test and verify consistency of API responses, the persistence of removed content, the viability of browser scraping and the occurrence of data scraping blocks.

We also took into account the obstacles faced and the solutions developed by NetLab UFRJ in building its own customized infrastructure to ensure the monitoring of different platforms over the years.

Since 2020, the laboratory has been developing and maintains a continuous and uninterrupted collection for the **constant monitoring** of different social media platforms. Building this collection of infrastructure depends on knowing how to handle different systems, as well as fully understanding the types of data available on the platforms and how to collect them.

Finally, we also assessed the platform's **transparency reports**: how readily available they are, where they are available, how frequently they are published and how detailed their information is.

Evaluation Criteria: *The Dimensions of Data Uniqueness*

Accessibility (16 parameters):

The most important dimension analyzed, since the evaluation parameters for the others dimensions are dependent on access to data. Accessibility refers to the availability of data and the ease with which it can be located, accessed, obtained and exploited for a particular purpose (Mahanti, 2018). Therefore, it is not enough to make them accessible. They also need to be **easily understood and analyzed** by researchers with varying degrees of technical knowledge, especially with regard to programming.

This dimension analyzed factors such as whether the API is **free or paid** for and whether it allows for the **full or partial extraction** of data of public interest.

Five parameters of this dimension are broken down into two special assessment criteria, which make **up 50% of the grade:**

Special Criterion 1: Is it possible to access the universe of public data via the API free of charge for research purposes?

The universe **to be monitored** consists of **public posts that are immediately locatable, accessible and retrievable by any user**. Therefore, we believe that, if the interested party so wishes, the platforms should make **the entire set of data available from the public posts corresponding to the requests made**, rather than just making sets available which have been cut out, in order to allow for the reproducibility of collections, analyses and the advancement of research, thus avoiding possible bias.

The highest score is given to a platform that guarantees **full access to the universe to be Monitored through a free API**. If there is no data collection API available free of charge to the general public, we assess whether at least researchers have specific means of accessing the data programmatically.

Based on access to data, the parameters that assess universal access free of charge by researchers and/or other interested parties make up 25% of each platform's final score.

Evaluation parameters that make up Special Criterion 1

Q1: Does the platform provide an official API for accessing public data published by users?

Q2: Is the universe to be monitored retrievable by the platform's API?

Q3: Is access to the platform's API free?

Q4: The platform offers researchers free and specific access to the API?

Special criterion 2: Does the platform offer an interface for collecting data by customized search?

It is expected that the platform will provide a graphic interface so that **interested parties with little or no technical programming knowledge** can collect the same data returned by the API.

This type of tool contributes to the **democratization of research** and the **transparency of data for citizens**. However, having an interface of this type does not do away with the need for APIs, since APIs guarantee greater customization of data collection processes and allow these processes to be automated and scaled up. This parameter makes up 25% of the platforms' final score.

Evaluation parameter that makes up Special Criterion 2

Q5: Does the platform offer an interface for collecting data through customizable searches?

Other parameters that make up the Accessibility dimension

Q6: Is it possible to extract the requested data directly from the platform's API response?

Q7: Does the platform's API provide a form of authentication that allows automatic renewal, without blocking data acquisition?

Q8: Can tokens to access the platform's API be created free of charge?

Q9: Is it possible to create new tokens to access the platform's API without quantity limitations?

Q10: Is the process for researchers to access the platform's API clear, uncomplicated and with a well-defined deadline?

Q11: Does the platform's API provide an endpoint to retrieve data from a specific publication?

Q12: Does the platform's API provide an endpoint to retrieve data from a specific author?

Q13: Does the platform's API provide an endpoint for retrieving data using search terms?

Q14: Is it possible to acquire data by scraping, without the need for authentication, via the platform's user interface?

Q15: Is it possible to acquire data by scraping without the need for other devices?

Q16: It is possible to recover data by scraping, without having to circumvent tools and techniques that prevent programmatic access to the data?

Compliance (11 parameters):

This dimension assesses whether the official documentation and the data retrieved are appropriate in terms of the forms adopted and the legal standards in force in the country (Mahanti, 2018). This is an exogenous dimension, i.e. related more to the **"'environment' of the data** than to the data themselves" and therefore more linked to "their **governance and management** than to their content itself" (Barbieri, 2019). This dimension assesses, for example, whether specific data, such as **dates** and **URLs**, are delivered in accordance with international standards. We also analyze the documentation provided by each platform regarding the means of data collection made available, in order to understand if they are **easily accessible and understandable**, if they present **clear descriptions and examples** of use and what their conditions and **terms of use are**. In addition, we also assessed the disclosure and detailing of **transparency reports** on moderation actions by the platforms analyzed.

Parameters that make up the Compliance dimension

Q17: Is the structure of the data provided by the platform's API stable?

Q18: Is the data returned by the platform's API in a standardized format?

Q19: Is the platform's API documentation published in open access?

Q20: Is the platform's API documentation clearly written and exemplified?

Q21: Does the platform's API documentation describe its terms of use?

Parameters that make up the Compliance dimension

(continued)

Q22: Does the documentation describe the format used in the platform's endpoint response?

Q23: Is the platform API documentation available natively in Portuguese?

Q24: Does the platform allow scraping and other types of automatic access in its terms of use?

Q25: Does the platform produce periodic transparency reports on content moderation in Brazil and make them publicly available, without the need for a request?

Q26: In its transparency reports, does the platform indicate the volume of each type of violation identified in Brazil according to the moderation policies in place?

Q27: Do the transparency reports specify information on the number and type of requests made to the platform by entities of the Brazilian state, as well as the number and type of requests that are complied with?

Completeness (6 parameters):

This dimension indicates whether the data retrieved has the **essential attributes for comprehension** and whether it is possible to carry out complete monitoring on each of the platforms analyzed when collecting the data, mainly considering the criteria required for academic and scientific research (Mahanti, 2018). In this dimension, data is considered complete when it can be used and applied in different research situations, even if some optional data fields have not been filled in.

Thus, the frequency allowed for data collection through official channels is assesed, and whether this enables consistent monitoring and the retrieval of data on comments and temporary content, for example.

Parameters that make up the Completeness dimension

Q28: Is it possible to retrieve data from a publication's comments via the platform's API?

Q29: Is it possible to recover data from temporary content via the platform's API?

Q30: Is it possible to retrieve historical data via the platform's API?

Q31: Is the number of requests allowed by the platform's API enough to monitor more than 1 million publications in 24 hours?

Q32: Is the number of requests allowed by the platform's API enough to monitor more than 100,000 publications in 24 hours?

Q33: Is the number of requests allowed by the platform's API enough to monitor more than 10,000 publications in 24 hours?

Consistency (4 parameters):

This dimension assesses whether the **format and presentation of the data are consistent** and identical in all the databases extracted and, in particular, in requests that are identical to each other (Mahanti, 2018). It also checks whether the search terms and filters used have produced **coherent** data **without contradictions, duplications or discrepancies**. This dimension analyzes, for example, whether the same data is retrieved when requested at **different** times, avoiding **inconsistencies** in systematic monitoring. Consistency is essential for producing **accurate and agile** reports, as it avoids the need to constantly check and/or correct data and allows for greater **auditability**.

Parameters that make up the Consistency dimension

Q34: Is the data returned by the platform's API persistent?

Q35: Does the data retrieved by the platform's API reflect what is displayed on its user interface?

Q36: Is the response returned by the platform's API always what you expected?

Q37: Is the response returned by the platform's API consistent with the parameters and filters used in the request?

Relevance (2 parameters):

This dimension assesses whether the data is **relevant to the purpose** for which it is intended (Mahanti, 2018), i.e. whether it is in line with the **objectives of the research and the requisiton**. The data retrieved must also be sufficient to support a robust analysis. This dimension analyzes, for example, whether **filtering data by search terms** in the process collection is possible.

Parameters that make up the Relevance dimension

Q38: Are the entities returned by the platform's API sufficient to understand the data in all its levels of detail?

Q39: Does the platform's API allow using filters to refine the data request?

Timeliness (1 parameter):

This parameter qualifies the **impact of the passage of time** on the availability of data, in order to assess the **speed** with which the data **update** process takes place (Mahanti, 2018).

Parameters that make up the Actuality dimension

P40: Is it possible to retrieve newly published data, in near real-time upon publication, via the platform API?

Rating Composition

To calculate the index ratings, the five parameters for assessing accesibility were grouped **two special criteria, which account for 50% of the score. The other parameters correspond to the remaining 50% of the score.** Each of the remaining 35 parameters have the same weight in the composition of the final score, so that dimensions with more parameters have higher weights.

In this way, we perform a **weighted calculation in which:**

1. **25% of the score** corresponds to **Special Criterion 1: "*Is it possible to access the universe of public data via the platform's API, free of charge, for research purposes?*".**

   Only platforms that provide access to all public content and allow **full, systematized and free recovery**, avoiding bias and ensuring the reproducibility of collections and analyses score in this field. In other words, the platform needs to be positively evaluated in parameters P1 (*Does the platform provide an official API for accessing public data published by users?*) and P2 (*Is the universe to be monitored retrievable through the platform's API?*), as well as responding positively to at least one of the parameters P3 (*Is access to the platform's API free of charge?*) or P4 (*Does the platform offer researchers free and specific access to the API?*).

2. **25% of the score** corresponds to **Criterion 2: "*Does the platform offer an interface for collecting data through customizable searches?*".**

   In order to score in this criterion, the platform must respond positively to parameter P5 (*Does the platform offer an interface for collecting data?*).

   The determining factor for scoring in this parameter is to offer means of collection that are **easy to understand and navigate** for people with little technical knowledge.

3. **50% of the score** is made up of **35 parameters** with the same weight, so that **dimensions with more parameters have greater weight in the final score.** The score for each platform corresponds to the percentage of YES answers ✓ obtained in each dimension in relation to the total number of parameters in that dimension. The **graphic** on the next page details the distribution of weights and parameters that make up this portion of the grade.

The calculation of the final grade for each platform is represented by:

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

In which:

xi is the **positive or negative binary response** (0 or 1) for **Special Criterion 1;**

xii is **the positive or negative binary answer** (0 or 1) for **Special Criterion 2;**

Positives is the number of **positive responses** of each platform;

Parameters is the number of **applicable evaluation parameters in the questionnaire[2].**

[2] *In the **WhatsApp** and **YouTube** evaluations, we disregarded one of the evaluation parameters and we adjusted the rest of the calculations around this decision.*

Graphic Representation of the Rating

FREE ACCESS TO THE UNIVERSE (SPECIAL CRITERION 1)
WEIGHT: 25 PTS
Nº OF PARAMETERS: 4

ACCESS TO THE DATA COLLECTION INTERFACE (SPECIAL CRITERION 2)
WEIGHT: 25 PTS
Nº OF PARAMETERS: 1

ACCESSIBILITY
WEIGHT: 15.71 PTS
Nº OF PARAMETERS: 11

COMPLIANCE
WEIGHT:15.71 PTS
Nº OF PARAMETERS: 11

COMPLETENESS
WEIGHT: 8,57 PTS
Nº OF PARAMETERS: 6

CONSISTENCY
WEIGHT: 5.71 PTS
Nº OF PARAMETERS: 4

RELEVANCE
WEIGHT: 2.86 PTS
Nº OF PARAMETERS: 2

CURRENT
WEIGHT:1.43 PT
Nº OF PARAMETERS: 1

Data *Transparency* Levels

To make it easier to interpret the scores obtained, the index divides the platforms analyzed into five categories:

Transparency
ideal
(81 to 100 points)

Platforms with **efficient official solutions for data collection**, including **APIs and a data collection interface**, with well-documented examples and no obstacles to scraping. They usually publish **regular transparency reports** detailing violations and removals at the request of the state in Brazil.

Transparency
satisfactory
(61 to 80 points)

Platforms that make data available **without financial restrictions**, but with limitations on **the volume of data that can be requested** and/or with **quality problems**, especially **consistency**. They publish transparency reports on their moderation actions in Brazil on a regular basis.

Transparency
regular
(41 to 60 pontos)

Platforms that present some measures of transparency and access to data, but with various **limitations related to the type of content that can be accessed and the sample of the universe of public data that can be collected.** In general, they publish transparency reports with moderation actions in Brazil, but **without the expected detail.**

Transparency
precarious
(21 to 40 points)

Platforms that **impose** significant **technical**, **operational** and/or **financial barriers** to their data access measures, making monitoring unfeasible for most researchers and interested parties. They are also not in the habit of publishing periodic transparency reports on their content moderation actions in Brazil.

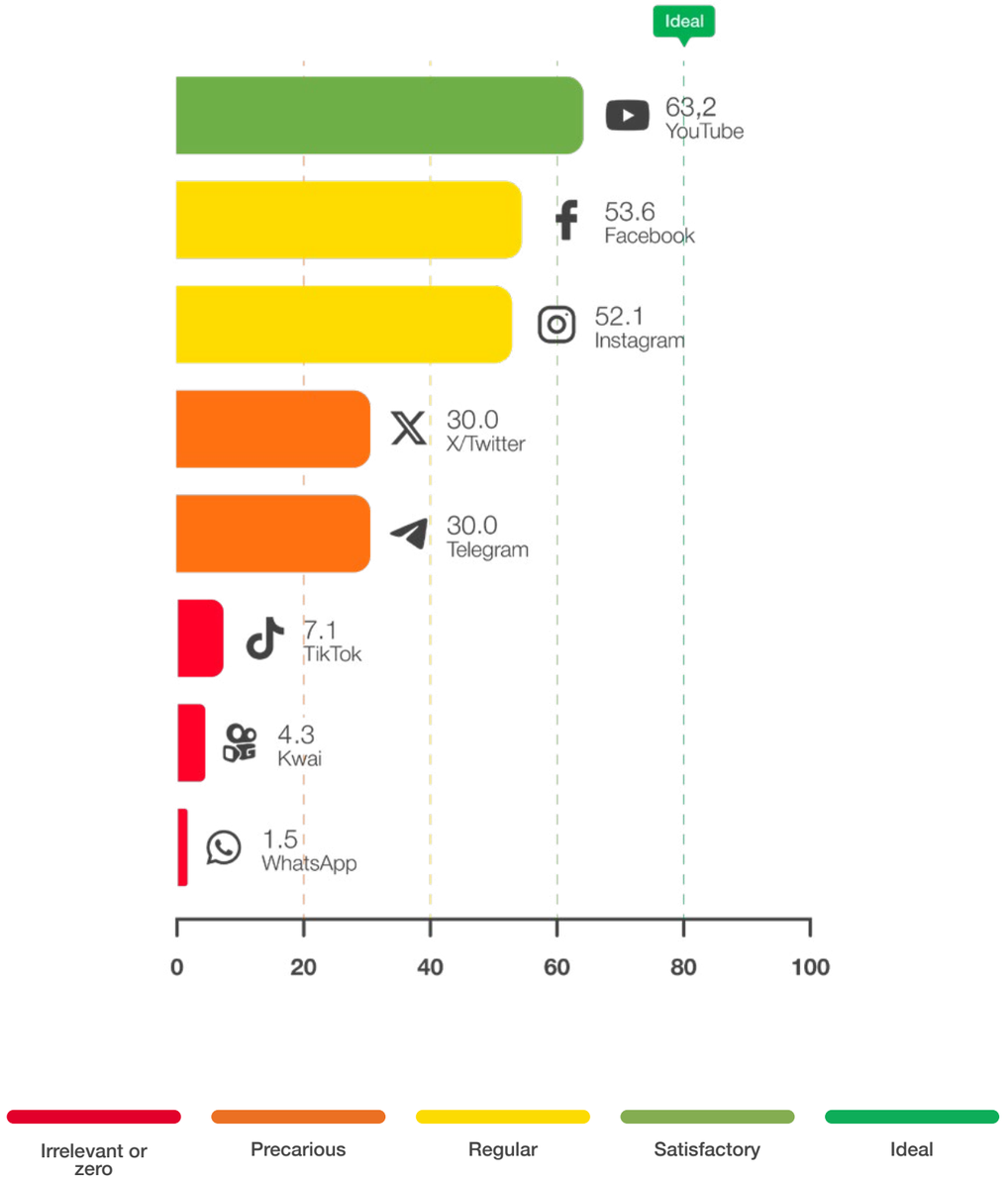
Transparency
Irrelevant or zero
(0 to 20 points)

Platforms that **don't invest in any transparency and data access measures.** They receive few points thanks to the possibilities of data scraping, which **are** generally **not officially allowed.** They don't usually publish periodic transparency reports on their content moderation actions in Brazil.

Results

The DTI results show that none of the evaluated platforms achieves an ideal score **terms of transparency measures and data access and the quality of the data returned.**

Below is an overview of what was observed on each platform analyzed. The overview of each platform, divided by dimension, as well as the specific answers and justifications for each evaluation parameter are available in the **Appendix**

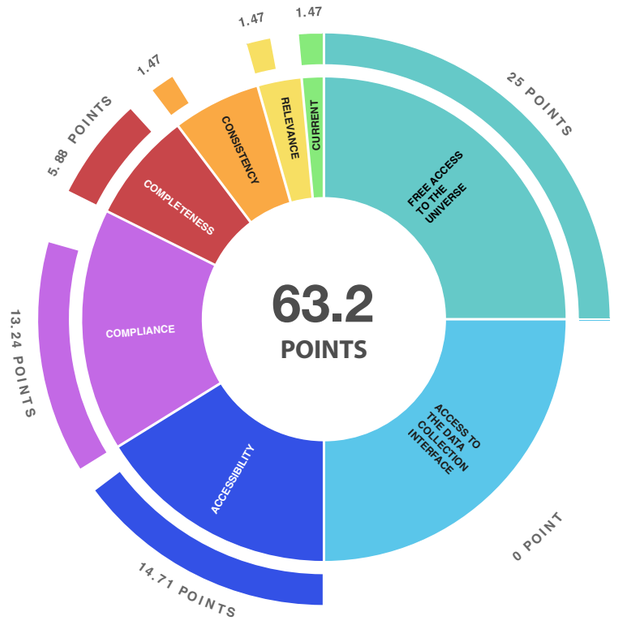


YouTube

Data transparency:
Satisfactory

Key:

- **FREE ACCESS TO THE UNIVERSE** — 25 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 14.71 OUT OF 16.18 POINTS
- **COMPLIANCE** — 13.24 OUT OF 16.18 POINTS
- **COMPLETENESS** — 5.88 OUT OF 7.35 POINTS
- **CONSISTENCY** — 1.47 OUT OF 5.88 POINTS
- **RELEVANCE** — 1.47 OUT OF 2.94 POINTS
- **CURRENT** — 1.47 OUT OF 1.47 POINTS



Among the platforms surveyed, YouTube obtained the best score (**63.2 points**), and its data transparency was considered **satisfactory**[3]. The platform has an official API (YouTube, [*S.d.*]) which is free for any user to access (Q1 and Q3), allowing the entire universe of videos classified as public to be searched. The platform also provides a specific API for researchers (Q4), access to which can be requested easily (Q10). Another positive point is that YouTube also allows quick retrieval of newly published data (Q40) and historical data (Q30).

On the other hand, the platform does not score in Special Criterion 2 due to the absence of a data collection interface (Q5). Parameters for consistency and relevance also undermine YouTube's performance, since there are several problems with inconsistent answers (Q36) and incoherent data in relation to the parameters and filters applied at the time of collection (Q37).

In addition, the data retrieved by the platform's API is not permanent (Q34): it does not provide metadata from removed publications, unlike the platform's interface, which displays a message about the removal and the reason for this, if the user tries to watch a video that is no longer on the air.

n the Report on Compliance with Community Guidelines (Google, [*S.d.*]), YouTube makes data available on a quarterly basis on the volume of removed videos published by users in Brazil which violated the platform's guidelines (Q25). However, information such as the number of channels removed, the number of removals based on complaints and aggregations by type of violation in the country are not made available (Q26), despite this type of information being provided to other countries in specific reports published by YouTube. The platform also appears in a biannual report together with other services associated with Google, which shows the volume of government requests to remove content, by type of violation, in Brazil (Q27).

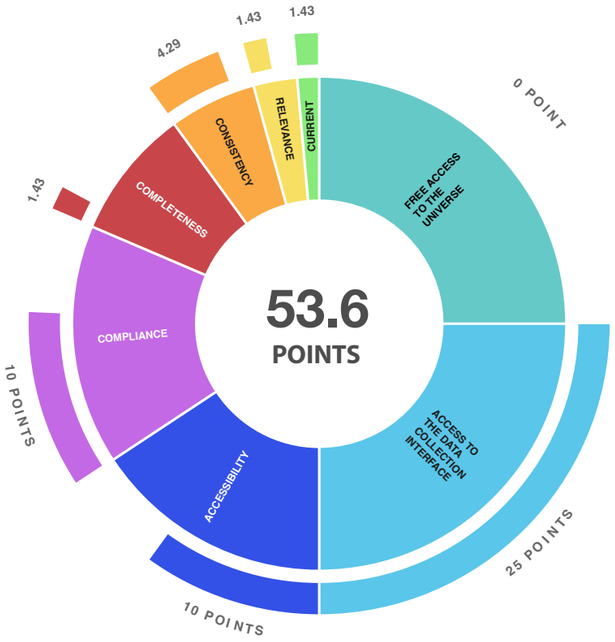
[3]*For the final calculation of YouTube's score, we disregarded the parameter "Is it possible to retrieve data from temporary content via the platform's API?", referring to the completeness dimension, since the platform does not allow users to publish temporary content. Therefore, in addition to the two special criteria, we considered 34 of the 35 parameters applicable and adjusted all the other calculations.*

Facebook

Data transparency:
Regular

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 25 OUT OF 25 POINTS
- **ACCESSIBILITY** — 10 OUT OF 15.71 POINTS
- **COMPLIANCE** — 10 OUT OF 15.71 POINTS
- **COMPLETENESS** — 1.43 OUT OF 8.57 POINTS
- **CONSISTENCY** — 4.29 OUT OF 5.71 POINTS
- **RELEVANCE** — 1.43 OUT OF 2.86 POINTS
- **CURRENT** — 1.43 OUT OF 1.43 POINTS



Facebook's data transparency score of **53.6** points is considered **regular**. Almost half of the platform's score depends on the **25 points** received by CrowdTangle, the interface and API for data collection (Q5), which Meta made available for researchers and journalists to collect public data samples from Instagram and Facebook (CrowdTangle, [*S.d.*]).

The CrowdTangle API had several limitations and did not allow access to comment data linked to original posts (Q28) or temporary content (Q29), such as stories, which hindered the completeness of the data retrieved. In addition, the tool also did not indicate when content was removed from the platform (Q34): deleted publications were treated by CrowdTangle as if they had never existed.

The CrowdTangle API also did not allow the recovery of the entire universe of public data from the platform (Q2), since only publications made: (i) by pages with more than 25 thousand followers or likes, (ii) by verified profiles, and (iii) in public groups with more than 95 thousand members could be recovered.

Facebook also loses points in parameters related to alternative collection methods for placing limitations to prevent data scraping (Q14 and Q16).

In Meta's Transparency Center (Meta, [*S.d.*]), the company provides information every six months on publications, profiles and comments moderated on its platforms in accordance with the local laws of each country in which it operates, including Brazil (Q25). Meta also publishes half-yearly data on requests made by Brazilian government entities to moderate content on its platforms (Q27). However, the total of data moderation activity, including removals for violating Meta's terms of use, is only available at a global level or for some countries with specific reports – which is not the case in Brazil (Q26).

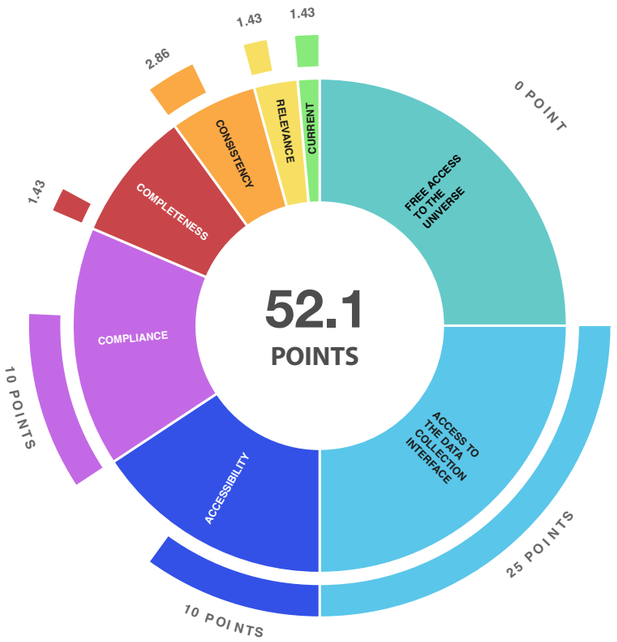
During the preparation of this index, Meta announced that the service would be officially discontinued as of August 2024 **(CrowdTangle, 2024).** If equivalent data access tools are not implemented, Facebook's score is likely to be even lower in future evaluations.

Instagram

Data transparency:
Regular

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 25 OUT OF 25 POINTS
- **ACCESSIBILITY** — 10 OUT OF 15.71 POINTS
- **COMPLIANCE** — 10 OUT OF 15.7 POINTS
- **COMPLETENESS** — 1.43 OUT OF 8.57 POINTS
- **CONSISTENCY** — 2.86 OUT OF 5.71 POINTS
- **RELEVANCE** — 1.43 OUT OF 2.86 POINTS
- **CURRENT** — 1.43 OUT OF 1.43 POINTS



Instagram's data transparency score of **52.1** points is considered regular. As with Facebook, we used the CrowdTangle tool to answer the platform's evaluation parameters. They offered an official API **(CrowdTangle, [*S.d.*])** for data extraction (Q1), as well as an interface **(CrowdTangle, [*S.d.*])** dedicated to data collection (Q5).

However, CrowdTangle did not allow the entire universe of public data from the platform to be retrieved (Q2), since only data from posts from profiles with more than 50,000 followers and/or verified followers could be retrieved and because data from reels could not be retrieved. Instagram was also penalized for the lack of completeness of the data returned by CrowdTangle, which did not allow the collection of comments and stories **(Q28 and Q29)**, for example.

As with Facebook, CrowdTangle did not indicate when a post had been removed from Instagram (Q34). As it does for Facebook, Meta makes information available on a weekly basis **(Meta, [*S.d.*])** broken down to country level about posts, profiles and moderated comments on Instagram (Q25) and a brief report on requests for moderation and/or access to data from the platform made by Brazilian state entities (Q27).

The main difference between evaluations of Facebook and Instagram in DTI lies in the consistency dimension. In the case of Instagram, CrowdTangle returned redirected links to access the images present in the posts identified and retrieved, but these expired almost instantly and could not be analyzed systematically, meaning that the data collected did not reflect what was displayed on the platform's user interface (Q35). This is especially problematic in the case of Instagram, since the content published is mostly visual, in photo and/or video formats.

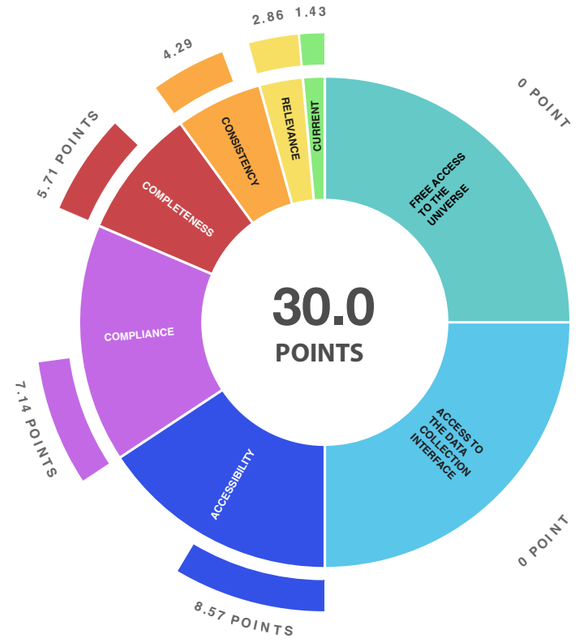
During the preparation of this index, Meta announced that CrowdTangle would be officially discontinued as of August 2024 **(CrowdTangle, 2024)**. As was observed with Facebook, if equivalent data access tools are not implemented, Instagram's score tends to be even lower in future evaluations.

X / Twitter

Data transparency:
Precarious

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 8.57 OUT OF 15.71 POINTS
- **COMPLIANCE** — 7.14 OUT OF 15.71 POINTS
- **COMPLETENESS** — 5.71 OUT OF 8.57 POINTS
- **CONSISTENCY** — 4.29 OUT OF 5.71 POINTS
- **RELEVANCE** — 2.86 OUT OF 2.86 POINTS
- **CURRENT** — 1.43 OUT OF 1.43 POINTS



Scoring only **30 points**, the platform's data transparency is considered **precarious**. In the past, X/Twitter had established itself as one of the most accessible platforms for collecting public data **(Zuckerman, 2021)**. However, since 2023, the X/Twitter API has made access to data conditional on adherence to paid plans **(X/Twitter, [*S.d.*])**, with prohibitive fees, including the *Basic plan*, which costs US$100 (around R$543.00, according to the September 2024 quote) per month, the *Pro plan*, which costs US$5.000 (around R$27,500, as quoted in September 2024) per month, and the *Enterprise plan*, which starts at US$42,000 (around R$231,400, as quoted in September 2024).

To evaluate the platform, we considered the functionalities of the *Basic* and *Pro plans*, since X/Twitter only sells the *Enterprise plan* to approved companies upon request **(X/Twitter, [*S.d.*])**. With the implementation of these paid plans, X/Twitter no longer provides a free form of data access for any user (Q3), and no longer guarantees any form of access for researchers (Q4). Although it claims to make data available to researchers investigating systemic risks in the European Union, the platform is accused of not complying with this DSA obligation **(European Commission, 2023)**.

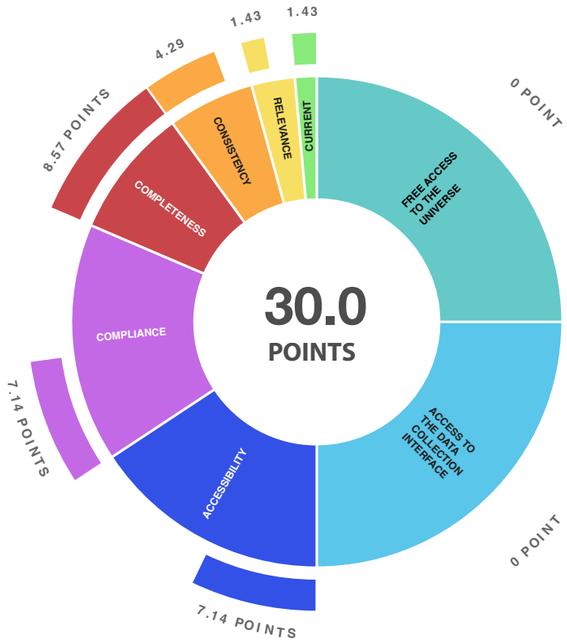
Due to the high quality of the data returned via the API, even though it is paid for, the platform gains points in the dimensions of completeness and consistency. Of all the platforms, X/Twitter is the only one that flags removed content properly (Q34). However, the platform does not provide any report with information on content removed in Brazil or subjected to other forms of moderation, as it does for other countries (Q25) **(X/Twitter, [*S.d.*])**. In the X/Twitter Transparency Center, the latest data for Brazil was published in 2021 and is currently unavailable **(X/Twitter, [*S.d.*])**.

Telegram

Data transparency:
Precarious

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 7.14 OUT OF 15.71 POINTS
- **COMPLIANCE** — 7.14 OUT OF 15.71 POINTS
- **COMPLETENESS** — 8.57 OUT OF 8.57 POINTS
- **CONSISTENCY** — 4.29 OUT OF 5.71 POINTS
- **RELEVANCE** — 1.43 OUT OF 2.86 POINTS
- **CURRENT** — 1.43 OUT OF 1.43 POINTS



Telegram scored **30 points**, with its data transfer is considered **precarious**. The platform does provide a free official API **(Telegram, [*S.d.*])** (Q1 and Q3), which enables programmatic access to data in a reliable, consistent manner and, to a large extent, in accordance with good practices.

However, the platform does not allow for retrieval of message data on a given topic without prior knowledge of the groups or channels in which they circulate. As there is no way to search the API for public groups and channels, access to the universe of data of interest is limited (Q2), which contributes to its low score. The platform also does not offer a dedicated interface for data collection (Q5).

On the one hand, Telegram is the only platform analyzed that meets all the completeness criteria, as it allows for the collection of comment data **(Q28)**, temporary content data **(Q29)** and historical data **(Q30)**, in addition to guaranteeing the retrieval of a large volume of data without difficulty **(Q31, Q32 and Q33)**.

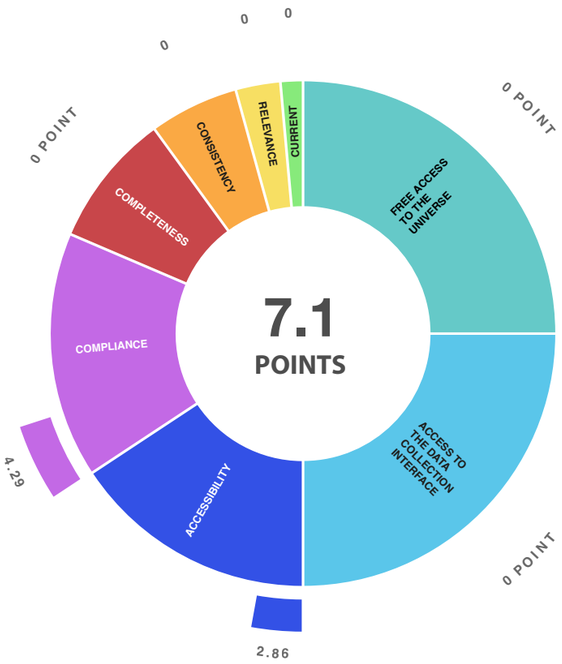
On the other hand, the platform is heavily penalized in the compliance dimension. Telegram does not publish periodic public transparency reports, which makes it impossible to identify moderated content in Brazil (Q25, Q26 and Q27). Its API documentation is also deficient because it does not directly define how to call endpoints (Q22), does not describe the formats of the data returned (Q18) and is not available in Portuguese (Q23).

TikTok

Data transparency:
Irrelevant

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 2.86 OUT OF 15.71 POINTS
- **COMPLIANCE** — 4.29 OUT OF 15.71 POINTS
- **COMPLETENESS** — 0 OUT OF 8.57 POINTS
- **CONSISTENCY** — 0 OUT OF 5.71 POINTS
- **RELEVANCE** — 0 OUT OF 2.86 POINTS
- **CURRENT** — 0 OUT OF 1.43 POINTS



TikTok does not offer access to a free official data collection API for interested parties in Brazil, only for researchers registered in the United States and Europe **(TikTok, [*S.d.*])**. Largely because of this, the platform's data transparency score is considered **irrelevant**, at just **7.1 points**.

In addition to the fact that there is no API available in the country (Q1), which makes it impossible to programmatically access the whole universe of public data (Q2), the platform also does not offer a dedicated interface for data collection (Q5).

The few points for accessibility are gained from the possibility of scraping data from the platform's interface **(Q14 and Q15)**. However, there is a *Captcha* verification processes used by the platform to prevent automated use, which substantially limits the process of collecting data by unofficial means **(Q16)**.

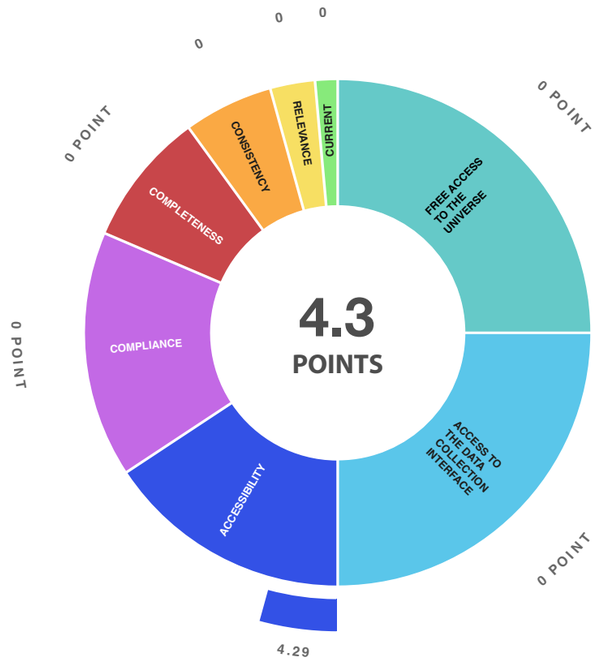
On the other hand, TikTok is the only platform to score on all three parameters relating to public transparency reports (Q25, Q26 and Q27) for providing information on the volume and types of violations moderated by the platform **(TikTok, 2024)** and for pointing out, through bi-annual reports, the requests for moderation and data requests made by the Brazilian state since 2019 **(TikTok, 2023)**.

Kwai

Data transparency:
Irrelevant

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 4.29 OUT OF 15.71 POINTS
- **COMPLIANCE** — 0 OUT OF 15.71 POINTS
- **COMPLETENESS** — 0 OUT OF 8.57 POINTS
- **CONSISTENCY** — 0 OUT OF 5.71 POINTS
- **RELEVANCE** — 0 OUT OF 2.86 POINTS
- **CURRENT** — 0 OUT OF 1.43 POINTS



Kwai does not provide any API for data collection in Brazil or anywhere else in the world (Q1), making the auditability of the universe of public posts on the platform unfeasible. (Q2). Nor does Kwai provide a dedicated interface for data collection (Q5). These factors largely contribute to the **4.3 points** achieved by the platform in the assessment of its data transparency, which is considered **irrelevant.**

The few points obtained by Kwai in the DTI evaluation stem from three parameters in the accessibility dimension, as scraping data is a possibility here **(Q14, Q15 and Q16)** with no need for authentication on the platform's user interface. However, the data extracted is limited and biased by the low volume of publications returned for each search on the platform's web interface.

Another negative point is the absence of transparency reports dedicated to removed publications and suspended users in Brazil **(Q25, Q26 and Q27)**, since the Kwai provides informations of transparency only at the continental level **(Kwai, [*S.d.*])**.

WhatsApp

Data transparency:
Irrelevant

Key:

- **FREE ACCESS TO THE UNIVERSE** — 0 OUT OF 25 POINTS
- **ACCESS TO THE DATA COLLECTION INTERFACE** — 0 OUT OF 25 POINTS
- **ACCESSIBILITY** — 1.47 OUT OF 16.8 POINTS
- **COMPLIANCE** — 0 OUT OF 16.18 POINTS
- **COMPLETENESS** — 0 OUT OF 7.35 POINTS
- **CONSISTENCY** — 0 OUT OF 5.88 POINTS
- **RELEVANCE** — 0 OUT OF 2.94 POINTS
- **CURRENT** — 0 OUT OF 1.47 POINTS

WhatsApp comes last in the DTI assessment, with data transparency considered **irrelevant** and only **1.5 point**[4] earned in the accessibility dimension. The platform does not have an official API. (Q1), unlike its competitor Telegram, nor does it have a dedicated interface for data collection (Q5).

WhatsApp had a positive response in only one of the evaluation parameters, relating to the data scraping process **(Q16)**. In any case, even if it is possible to obtain data through scraping, it is not possible to have full access to the universe of messages of interest on WhatsApp, since the collection is limited to a sample of public groups previously selected by the researchers.

The lack of transparency, also due to there being no public reports on moderation activity by the platform (Q25), raises alarm bells, since WhatsApp is seen as one of the main platforms for the dissemination of information of disinformation in the Global South, especially in Brazil **(Kalogeropoulos; Rossini, 2023)**.

[4] *For the final calculation of WhatsApp's score, we disregarded the parameter "Is it possible to retrieve data from a publication's comments the platform's API?", which refers to the completeness dimension, the platform does not support comments in its messages. Therefore, in addition to the two special criteria, we considered 34 of the remaining 35 applicable parameters and adjusted all the other calculations.*

Good and Bad Practices in Data Availability

Based on the evidence observed in the analysis of each platform, we present an overview of measures, **divided by data quality dimensions**, that should be widely adopted or avoided by platforms in order to guarantee an **ideal level of transparency and availability of data for research in the public interest.**

Good practices That Can Be Replicated

Accessibility:

YouTube is the only platform analyzed that guarantees **full access to the universe of public data through a free API**. Any API user can retrieve all the main video data, including identifier, title, description, author and comments, making the platform the most accessible of those evaluated.

A good practice adopted by **YouTube, X/Twitter, Facebook and Instagram** is allowing for the programmatic retrieval of data using specific search terms.

It is possible to collect data from **YouTube, WhatsApp and Telegram** using scraping techniques via the platform's user interface, in a relatively uncomplicated way.

Compliance:

YouTube provides the API documentation translated into Portuguese and describes its terms of use clearly and accessibly, without redirecting to other pages.

Although **TikTok** does not provide official and programmatic means of collecting data, it does provide highly detailed transparency reports on moderation activity undertaken in Brazil.

Completeness:

YouTube and **Telegram** offer official APIs for collecting different types of data, such as **comments, temporary publications and historic data.**

These platforms also enable massive and rapid data collection, with **Telegram** standing out as the only one that meets all the parameters of this dimension.

Consistency:

The **Telegram**, **X/Twitter and Facebook** APIs follow the filters applied in researcher's requests. Requests made to the APIs at different times or by different users retrieved practically identical data, as should be expected.

Relevância:

As well as offering efficient collection filters in its API, such as language and time range, **X/Twitter** returns data that allows you to map the **relationships between different pieces of content**, including shares and replies. It is the only platform that meets both parameters of this dimension.

Current:

YouTube, Telegram, X/Twitter, Facebook and Instagram, the platforms that, at the time of the analysis, offered official APIs in Brazil. for data collection, made data available in (near) real time for collection and analysis.

Bad Practices Which Should Be Avoided

Accessibility:

TikTok, Kwai and WhatsApp do not provide an official API for collecting public data from their publications in Brazil, although TikTok does so in the United States and Europe.

X/Twitter, Facebook, Instagram and Telegram have limited the number of tokens that can be created, making it difficult to carry out parallel collection processes. With the exception of Telegram, these platforms are also active in restricting data scraping.

Compliance:

X/Twitter, Facebook, Instagram and Telegram have not made their API documentation available in open access and fully translated into Portuguese.

Telegram stands out for its lack of examples clear rules for its use and poorly written documentation.

X/Twitter, Telegram, WhatsApp and Kwai do not publish transparency reports specifically for Brazil.

Completeness:

Facebook and Instagram's data collection APIs did not allow for exporting of more than 10,000 posts at a time, hindering the collection of large volumes of data.

In addition, **comment data and temporary content** were inaccessible on both platforms.

Consistency:

The **YouTube** API shows **inconsistent results** for collections with identical parameters. Data sets retrieved in identical searches, carried out over a short period of time during the tests, were considerably different from each other and did not comply with the criteria defined in the requests.

None of the differences seen in the tests were due to the inclusion of more recent data.

Relevance:

Despite having official collector APIs, Telegram and YouTube have limitations when it comes to applying filters. YouTube allows for filtering videos by language, but this does not work in practice. In fact, the API documentation recognizes that it **can return videos in any language if** it deems them relevant to a search, according to criteria that are not clear to users. **Telegram**, on the other hand, does not allow results to be filtered by language.

Current:

TikTok, Kwai and WhatsApp do not provide an official API for collecting public data from their publications in Brazil.

Recommendations

The main points for improving platform transparency fall into **four areas**:

01

Collecting the universe of public data

When considering the **possibilities for data collection**, the **most problematic cases are TikTok, Kwai and WhatsApp**. It is recommended that an official, free **API be made available** for researchers to access the platforms' public data universe. In addition, we recommend offering a **dedicated interface for data collection**, which is easy to use and navigate and does not require technical knowledge of programming. Most of the parameters are not even applicable to these three platforms because it is not possible to collect data, which prevents the assessment of other quality and transparency parameters.

Although **Facebook** and **Instagram** offered a free API and interface for data collection at the time of our analysis, CrowdTangle required a reduction in restrictions to provide **full and free access** to their universe of data. In the case of **You Tube and Telegram**, even if there is broad access to data via API, it would be improved with a user friendly **data collection interface**.

As for **X/Twitter**, it is crucial that it restore a **free API**, at least for Brazilian researchers, in order to improve its transparency practices. Another factor to be improved on all platforms is **the official permission to scrape data for research** purposes, offering more legal certainty and facilitating the process of systematic analysis for those who use this method.

02

Data Quality Available

As for the quality of the data provided, all the free APIs have problems in the different dimensions evaluated. In the case of **YouTube**, various **inconsistencies** in the results delivered need to be corrected, as well as **incoerences** between results and search parameters. In addition, almost all platforms need to improve access to **temporary content** data, such as Instagram and Facebook stories, for exemple. In the case of **Instagram,** the

situation is even more **critical**, as stories are often prioritized by users when publishing. Access to this data needs to be provided, as well as **non-expiring links** to make it possible to analyze permanent photo and video content.

It is also important that platforms signal when **publications are removed** and **users suspended**, offering access to the **metadata**, even if the content of the publication is restricted. This type of data is essential for research into the circulation of **misinformation and illegal or abusive content** on social media platforms, as well as for understanding the platforms' moderation practices.

It is also worth noting that **customized keyword searches** are essential collecting relevant data for the purposes of the research, a resource that is not offered by Telegram to access the entire universe of public data, which is only accessible via previously known groups.

03

Clarity of API Documentation and Terms of Use

As for the API documentation, it is important that it is made **publicly** available, without the need for an individual request, in **Portuguese** and with a clear description of its **use terms.**

Some of these aspects, which would help **democratize** access to existing resources, have not been observed on Facebook, Instagram, X/Twitter and Telegram. In the case of Telegram, it is also necessary to improve the instructions on using the official API, which are not **clear and detailed**, nor do they describe the format of the responses expected when using each endpoint.

04

Detailed Transparency Reports on Brazil available

As far as transparency reports are concerned, the most critical point lies in making the document available on a **regular basis**. Even when the reports are made available with some regularity, their data is **of low quality** and delivered with unsatisfactory granularity, usually without a specific focus. This makes it impossible to carry out more relevant and in-depth analyses of the platforms' moderation and governance policies. Some platforms, such as **X/Twitter, Telegram, Kwai and WhatsApp** don't even publish this type of document, making the problem even worse.

Appendix I:
Overview of Evaluation

✓ Indicates a positive assessment of the parameter

✗ Indicates a negative evaluation of the parameter

— indicates parameter is non-applicable to the evaluated platform

Accessibility — Special Criteria 1	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q1	Does the platform offer an official API for published public data by users?	✓	✓	✓	✓	✓	✗	✗	✗
Q2	Can the universe to be monitored be recovered by the platform's API?	✓	✗	✗	✓	✗	✗	✗	✗
Q3	Is access to the platform's API free?	✓	✓	✓	✗	✓	✗	✗	✗
Q4	Does the platform offer researchers free and specific access to the API?	✓	✓	✓	✗	✗	✗	✗	✗

Accessibility — Special Criteria 2	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q5	Does the platform offer an interface for collecting data using customizable search?	✗	✓	✓	✗	✗	✗	✗	✗

Other Accessibility Parameters

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	OK	WhatsApp
Q6	Can the requested data be extracted directly from the platform's API response?	✓	✓	✓	✓	✓	✗	✗	✗
Q7	Does the platform's API provide a form of authentication that allows for automatic renewal without blocking data acquisition?	✓	✓	✓	✓	✗	✗	✗	✗
Q8	Can access tokens for accessing the platform's API be generated free of charge?	✓	✓	✓	✗	✓	✗	✗	✗
Q9	Can access tokens be generated to access the platform's API with no limit on the quantity?	✗	✗	✗	✗	✗	✗	✗	✗
Q10	Is the process researchers use to gain access to the platform's API clear, uncomplicated and with a well-defined deadline?	✓	✗	✗	✗	✗	✗	✗	✗
Q11	Does the platform's API provide an endpoint for retrieving data from a specific publication?	✓	✓	✓	✓	✓	✗	✗	✗
Q12	Does the platfrom's API provide an endpoint to retrieve data from a specific author?	✓	✓	✓	✓	✗	✗	✗	✗
Q13	Does the platform's API provide an endpoint for retrieving data using search terms?	✓	✓	✓	✓	✗	✗	✗	✗
Q14	Can data be aquired by scraping, without need for authentication through the platform's user interface?	✓	✗	✗	✗	✗	✓	✓	✗

Other Accessibility Parameters (continued)

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	OK	WhatsApp
Q15	Can data be aquired by scraping without the need for other devices?	✓	✓	✓	✓	✓	✓	✓	✗
Q16	Can data be recovered using scraping, without need to circumvent tools and techniques aimed at preventing programmatic access to data?	✓	✗	✗	✗	✓	✗	✓	✓

Compliance

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	OK	WhatsApp
Q17	Is the structure of the data made available on the platform stable?	✓	✓	✓	✓	✓	✗	✗	✗
Q18	Is the data returned by the platform's API in a standardized format?	✓	✓	✓	✓	✓	✗	✗	✗
Q19	Is the documentation for the platform published in open access?	✓	✓	✓	✓	✓	✗	✗	✗
Q20	Is the platform's API documentation clearly written and exemplified?	✓	✓	✓	✓	✗	✗	✗	✗
Q21	Does the platform's API documentation describe its terms of use?	✓	✗	✗	✗	✓	✗	✗	✗
Q22	Does the documentation describes the response format of the platform's API endpoints?	✓	✓	✓	✓	✗	✗	✗	✗

Compliance (continued)

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q23	Is the documentation for the platform available in Portuguese?	✓	✗	✗	✗	✗	✗	✗	✗
Q24	Does the platform allow scraping and other types of automatic access in its terms of use?	✗	✗	✗	✗	✓	✗	✗	✗
Q25	Does the platfrom produce transparency reports on content moderation in Brazil and is it made publicly available, without the need for a request?	✓	✓	✓	✗	✗	✓	✗	✗
Q26	In its transparency reports, does the platform indicate the volume of each type of violation identified in Brazil, according to the current moderation policies?	✗	✗	✗	✗	✗	✓	✗	✗
Q27	Do transparency reports specify information on the amount and type of requests made to the platform by entities of the Brazilian State, as well as the number and type of requests complied with?	✓	✓	✓	✗	✗	✓	✗	✗

Completeness

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q28	Is it possible to retrieve data from the comments of a publications using the platform's API?	✓	✗	✗	✓	✓	✗	✗	—

Completeness (continued)

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q29	It is possible to recover temporary content data through the platform's API?	—	✗	✗	✓	✓	✗	✗	✗
Q30	Can historical data be retrieved via the platform's API?	✓	✓	✓	✗	✓	✗	✗	✗
Q31	Is the number of requests allowed by the platform sufficient to monitor more than 1 million publications in 24 hours?	✗	✗	✗	✗	✓	✗	✗	✗
Q32	Is the number of requests allowed by the platform sufficient to monitor more than 100,000 publications in 24 hours?	✓	✗	✗	✓	✓	✗	✗	✗
Q33	Is the number of requests allowed by the platform sufficient to monitor more than 10,000 publications in 24 hours?	✓	✗	✗	✓	✓	✗	✗	✗

Consistency

	Evaluation parameters	YouTube	Facebook	Instagram	X	Telegram	TikTok	Kwai	WhatsApp
Q34	Is the data returned by the platform API persistent?	✗	✗	✗	✓	✗	✗	✗	✗
Q35	Does the data retrieved by The platform's API reflect what is displayed in its user interface?	✓	✓	✗	✗	✓	✗	✗	✗
Q36	Is the answer returned by the platform API always as expected?	✗	✓	✓	✓	✓	✗	✗	✗
Q37	Is the response returned by the platform's API consistent with the parameters and filters used in the request?	✗	✓	✓	✓	✓	✗	✗	✗

Relevance

Evaluation parameters								
Q38	Are entities returned by the platform's API sufficient to understand the data at all levels of detail?	✓	✗	✗	✓	✓	✗	✗	✗
Q39	Does the platform's API allow the use of filters to refine the data request?	✗	✓	✓	✓	✗	✗	✗	✗

Current

Evaluation parameters								
Q40	Can newly-published data be retrieved from the platfprm's API in real time, or just after publication?	✓	✓	✓	✓	✓	✗	✗	✗

Appendix II:
Breakdown by Evaluation Parameter

Accessibility

Special Criteria 1

Q1

Does the platform offer an official API for published public data by users? (Special criterion 1)

Here, we look at whether the platform offers an API with at least one endpoint for accessing user-generated content data.

YouTube

Since 2014, the platform has been operating with the YouTube Data v3 API. All the documentation and instructions for using the official YouTube API can be accessed **here.**

Facebook

For accessing and extracting public data from Facebook and Instagram, at the time of the analysis Meta offered the CrowdTangle tool, which, in addition to presenting a user interface for this purpose, also **provided an API.** However, Meta **closed the tool in August 2024** and replaced it with the *Meta Content Library*, which does not offer equivalent resources.

Instagram

For accessing and extracting public data from Facebook and Instagram, at the time of the analysis Meta offered the CrowdTangle tool, which, in addition to presenting a user interface for this purpose, also **provided an API.** However, Meta **closed the tool in August 2024** and replaced it with the *Meta Content Library*, which does not offer equivalent resources.

X / Twitter

V2 of the X/Twitter API features paid access plans with endpoints available for retrieving data from public publications, albeit at a fee monthly prohibitions for the collection of publication data.

Telegram

The platform **offers an API** that allows the collection of public data in Telegram conversations.

TikTok

TikTok has an **API for academic research,** but access to users' public data is only available to researchers in the United States and Europe.

Kwai

Kwai does not provide any official API for data collection anywhere in the world.

WhatsApp

WhatsApp does not provide an API to access data for research anywhere in the world.

Accessibility

Special Criteria 1

Q2

Can the universe to be monitored be recovered by the platform's API? (Special criterion 1)

In this field, the aim is to check whether the platform allows for the programmatic discovery and collection of data from the entire set of public publications of interest.

YouTube

YouTube allows you to search the entire universe of public videos available via API. In this way, we understand that it is possible, through the official API, to structure the programmatic collection of videos, search through any public channels or videos.

Facebook

Although it made a considerable amount of data available, CrowdTangle limited Facebook's monitorable universe to public posts from verified profiles, pages with more than 25,000 followers or likes and public groups with more than 95,000 members. According to the **data,** this encompasses around 7 million pages, profiles and groups worldwide.

Instagram

Although it made a considerable amount of data available, CrowdTangle limited Instagram's monitorable universe to the public publications of profiles with more than 50,000 followers and all profiles verified by the platform. According to the **latest data,** this includes around 2 million pages and profiles worldwide. CrowdTangle also **did not allow the collection of data generated from reels,** short videos introduced by the platform to compete with TikTok.

X / Twitter

Although paid for, the X/Twitter API allows you to follow the entire universe of public publications around a topic of interest, even in (almost) real time. The platform only excludes private profile activities that are not followed by the user who is doing the search, as established in its **Privacy Policy.**

Telegram

Collection is only possible within a certain group or channel, requiring prior knowledge of those profiles in order for data to be collected. It is therefore impossible to locate and retrieve all the content considered publicly available on the platform.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API to access public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Special Criteria 1

Q3

Is access to the platform's API free of charge? (Special criterion 1)

We check whether the platform offers general access to its API free of charge.

YouTube

Use of the YouTube API is limited by **daily data quotas** for each project created in Google Cloud, but obtaining **credentials for access authorization** is completely free.

Instagram

Until its closure in August 2024, CrowdTangle provided free access to its API to accredited partners, which included journalists, researchers and fact-checkers.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

Facebook

Until its closure in August 2024, CrowdTangle provided free access to its API to accredited partners, which included journalists, researchers and fact-checkers.

X / Twitter

Currently, X/Twitter only offers paid access to its data collection API, divided into different price ranges.

WhatsApp

WhatsApp does not provide an API to access public data.

Telegram

Payment is not required to use the API, as as made clear in the **homepage documentation**.

Accessibility

Special Criteria 1

Q4

Does the platform offer researchers free and specific access to the API? (Special criterion 1)

This field checks whether the platform's API provides any kind of access to researchers, either in the form of a specific token or exclusive endpoints.

YouTube

YouTube Researcher Program is a registration option for access to the YouTube Data v3 API aimed specifically at researchers..

Instagram

CrowdTangle's terms allowed free access for **researchers and students** linked to academic institutions.

TikTok

The TikTok API is only available to researchers in the United States and Europe.

Facebook

CrowdTangle's terms allowed free access for **researchers and students** linked to academic institutions.

X / Twitter

There is only privileged and free access for researchers linked to academic institutions in the European Union, due to legal impositions from the DSA – although the rule **is not being followed** properly there. So far, the platform has not announced any plans to expand access to the rest of the world.

Kwai

Kwai does not offer any official API for data collection, nor for use by researchers.

WhatsApp

WhatsApp does not offer any official API for data collection, nor for use by researchers.

Telegram

Telegram does not grant researchers specific access to its public data. However, access to the API is guaranteed to any interested party, which is taken into account when answering the special criterion.

Accessibility

Special Criteria 2

Q5

Does the platform offer an interface for collecting data through customizable search? (Special criterion 2)

This field checks whether the platform offers an interface for observation and data collection designed for users with no technical knowledge of programming.

YouTube

The platform does not offer any official method of collecting data via a graphical interface for non specialized users.

Facebook

In addition to the API, CrowdTangle offered an interface with dashboards, through which it was possible to test and develop specific searches, monitor the evolution of discussions on the platform, analyze the engagement of pages of interest and request the retrieval of data from publications.

Instagram

In addition to the API, CrowdTangle offered an interface with dashboards, through which it was possible to test and develop specific searches, monitor the evolution of discussions on the platform, analyze the engagement of pages of interest and request the recovery of publication data.

X / Twitter

The platform does not offer any official method of collecting data via an interface for non expert users.

Telegram

The platform does not offer any official method of collecting data via an interface for non expert users.

TikTok

The platform does not offer any official method of collecting data via an interface for non-expert users.

Kwai

The platform does not offer any official method of collecting data via an interface for non-expert users.

WhatsApp

The platform does not offer any official method of collecting data via an interface for non-expert users.

Accessibility

Q6

Can the requested data be extracted directly from the platform's API response?

In this field, we check whether the API returns structured data as a response to the request, rather than delivering a link that redirects to the data. Audiovisual media data, such as image, video and audio files, were not considered when evaluating this parameter.

YouTube

All the main data can be collected, including title, description, ID, author, comments and link to the audiovisual content, available in the API.

Facebook

The responses to requests to the CrowdTangle API provided the expected data in an appropriate format, without the need for redirection.

Instagram

The responses to requests to the CrowdTangle API provided the expected data in an appropriate format, without the need for redirection.

X / Twitter

In the publication retrieval *endpoints*, the data is returned in the response to the request.

Telegram

The responses to requests to the Telegram API provide the expected data in the correct format without the need for redirection.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q7

Does the platform's API provide a form of authentication that allows for automatic renewal without blocking data acquisition?

This field assesses whether the tokens made available for API use expire and whether they can be renewed automatically.

YouTube

The collection does not require the renewal of tokens, **but is limited to the daily quotas** for each project accredited by the user.

Facebook

The tokens generated by CrowdTangle did not expire and renewal was the responsibility of the users, who could generate new tokens at any time.

Instagram

The tokens generated by CrowdTangle did not expire and renewal was the responsibility of the users, who could generate new tokens at any time.

X / Twitter

Except in situations of nonpayment, X/Twitter does not block existing tokens and the user can renew them whenever they wish.

Telegram

Although Telegram **provides for automatic updating of tokens,** the experience in the tests carried out was different. There were numerous blocks to the account and therefore to the token, without Telegram clarifying the reasons for this. What's more, as of the closing date of this report, the platform had not responded to requests for clarification via email sent throughout 2022 and 2023.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q8

Can access tokens for accessing the platform's API be generated free of charge?

In this field, it was assessed whether the platform allows the creation of new tokens free of charge, without the need to register more than one account to use the API.

YouTube

Registration for access to the API can be done by **any user for free.**

Facebook

CrowdTangle did not have any paid features, including the generation of API access tokens

Instagram

CrowdTangle did not have any paid features, including the generation of API access tokens

X / Twitter

With the new paid plans for the X/Twitter API, it is no longer possible to create new tokens to access the platform's data for free. It is only possible to create free tokens for publishing, but not for collecting publications programmatically.

Telegram

Telegram provides **automatic token updates** free of charge.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q9

Can access tokens be generated to access the platform's API with no limit on the quantity?

Here we check whether the platform limits the number of API access tokens that can be created by the same user.

YouTube

Although YouTube does not use a system based on renewing tokens, access to data is via a single route and is limited to the standard quota of 10,000 units per day for each project on Google Cloud. Different developers must share the quota for the same project by registering their individual Google accounts.

Facebook

CrowdTangle allowed only one token to access the API for each team. Thus, it was possible to renew the token but each team could only use one, and its members had to share it, meaning simultaneous use by more than one user was impossible.

Instagram

CrowdTangle allowed only one API access token to be used for each work team. It was therefore possible to renew the token, but each team could only use one, and its members had to share it, making it impossible for more than one user to use it simultaneously.

X / Twitter

With the new X/Twitter API rules, only one token is allowed per application. While the Free plan allows the creation of one application, which cannot be used for data collection, the Pro (paid) plan allows you create three applications.

Telegram

Telegram only allows one data collection token per account, so it is necessary to have more than one cell phone number if the person needs to expand the collection volume. In our tests, we experienced unjustified bans that were not provided for in the API's terms of use. This justifies the need for researchers and other interested parties to diversify their collection in different accounts.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q10

Is the process researcher use to access the platform's API clear, uncomplicated and with a well-defined deadline?

Here, we check whether the process for requesting access to the API exclusive for researchers is well described by the platform. We assess whether the platform makes clear what documentation is required to prove institutional ties and how long it will take for access to be granted, if the researcher's request is approved.

YouTube

The researcher access proceedure is described **here.** In order to make a request, a form must be filled in detailing the researcher's duties at the university and information about the department to which he or she is attached. A lot of information is requested, and with a lot of detail, but there are links to clarifications in the terms of service where necessary.

Facebook

CrowdTangle hadn't released API access to **new requests** since 2022 and Meta shut down the tool in August 2024.

Instagram

CrowdTangle hadn't released API access to **new requests** since 2022 and Meta shut down the tool in August 2024.

X / Twitter

X/Twitter does not offer data access specifically for Brazilian researchers.

Telegram

Telegram does not offer data access specifically for researchers.

TikTok

TikTok does not offer data access specifically for Brazilian researchers.

Kwai

Kwai does not offer any official API for data collection or for use by researchers.

WhatsApp

WhatsApp does not offer any official API for data collection, nor for use by researchers.

Accessibility

Q11

Does the platform's API provide an endpoint for retrieving data from a specific publication?

In this field, we check if it is possible to retrieve data from a specific public publication on the platform in question, by means of a unique identifier and not necessarily by search terms or other parameters and filters.

YouTube

It is possible to collect data on a single video from its **identification record.**

Facebook

The CrowdTangle API had the **GET /post/:id,** endpoint which made it possible to retrieve specific publications, when available in the API, based on their unique identifiers or the identifiers of their authors.

Instagram

The CrowdTangle API had the **GET /post/:id,** endpoint which allowed the retrieval of specific publications, when available in the API, based on their unique identifiers or the identifiers of their authors. Unlike Facebook, which made these identifiers available to any user of the platform, Instagram required registration in its **Graph API,** separate from the CrowdTangle API, in order provide the information, which made the task of retrieving the data unnecessarily complex.

X / Twitter

You can provide a list of of specific publication IDs to be collected via the **endpoint Tweets Lookup** via the platform's API.

Telegram

As long as you know the identifier of a group or public channel, you can retrieve the data of a specific message using its unique identifier.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q12

Does the platform's API provide and endpoint for retrieving data from a specific author?

In this field, we check whether the API allows the retrieval of data from public publications made on the platform in question by a specific author, via their username or unique identifier.

YouTube

The platform's API makes it possible to collect data from the videos posted by **a single channel** without too much difficulty.

Facebook

A specific user identifier with more that 25,000 followers was considered one of the perameters parameters for retrieving publications via ***GET /posts*** endpoints and ***GET /posts/search*** from the CrowdTangle API.

Instagram

A specific user identifier with more that 50,000 followers was considered one of the perameters parameters for retrieving posts via the *endpoints* **GET /posts** e **GET /posts/search** of the CrowdTangle API.

X / Twitter

The search expression on a given topic can be constructed by indicating the author(s) of the publications to be retrieved. It is also possible to retrieve publications and information from an account through the ***endpoint Users Lookup.*** In addition, the ***endpoint Timeline*** allows you to retrieve the entire timeline of posts by a specific user.

Telegram

Telegram does not offer a dedicated endpoint for collecting messages sent by a specific author. In addition, although it is possible to identify an author's messages in a particular group or room, it is impossible to track all the public messages sent by them on the platform.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q13

Does the platform's API provide an endpoint for retrieving data using search terms?

This field checks whether it is possible to retrieve data from public publications from the from the platform in question using search terms, i.e. to assemble a database with publications mentioning these terms.

YouTube

The YouTube Data v3 API provides a search option for retrieving video data that matches specific **search** terms or expressions.

Facebook

Searching for terms was one of the possible parameters for retrieving data from public **GET /posts** available on the API, using the **GET /posts/search** and endpoints of the CrowdTangle API.

Instagram

Searching for terms was one of the possible parameters for retrieving data from public **GET /posts** available on the API, using the **GET /posts/search** and endpoints of the CrowdTangle API.

X / Twitter

The X/Twitter API makes it possible to retrieve data available on the API using search terms, with the ***endpoint Search Tweets.***

Telegram

The search for messages that mention terms of interest on Telegram is only allowed in public groups and channels **that are already known to** the researcher. As the search is limited, the platform does not achieve the minimum expected in the evaluation parameter.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Accessibility

Q14

Can data be aquired by scraping, without the need for authentication, via the platform's user interface?

This field assesses whether it is possible to collect data via scraping techniques without the need to create accounts and logins.

YouTube

Several scientific publications indicate simple methods for acquiring data from YouTube by scraping, including the automatic transcription of videos, which is not provided by the API of the platform.

Facebook

It is impossible to navigate the platform's content properly and consistently without login authentication.

Instagram

It is impossible to navigate the platform's content properly and consistently without login authentication.

X / Twitter

It is impossible to navigate the platform's content properly and consistently without login authentication.

Telegram

Without logging in, it is possible to browse a limited preview of the content of the platform's public channels, but not of public groups, which only allows partial collection of the data of interest.

TikTok

Data scraping is possible without the need for a login. However, alternative acquisition methods are limited to identifying a maximum of 1,000 videos per search, the results of which are determinated by the platform's recommendation algorithm.

Kwai

It is possible to scrape data without authentication. However, the number of publications returned by Kwai is considerably reduced. Data acquisition is limited to a maximum of a few hundred videos that can be discovered by searching the platform's user interface, which makes consistent monitoring difficult.

WhatsApp

It is not possible to scrape data without having a WhatsApp account and being added to the group whose data is to be collected.

Accessibility

Q15

Can data be acquired by scraping without the need for other devices?

This field checks whether there is a need for other electronic devices (such as cell phones) to carry out the collection by alternative means, if it is not possible to carry it out via the platform's user interface on a computer.

YouTube

Data can scraped using the platform's user interface on a computer, without the need for other devices.

Instagram

Data can scraped using the platform's user interface on a computer, without the need for other devices.

TikTok

Data can scraped using the platform's user interface on a computer, without the need for other devices, even though the volume of data returned is limited.

Facebook

Data can scraped using the platform's user interface on a computer, without the need for other devices.

X / Twitter

Data can scraped using the platform's user interface on a computer, without the need for other devices.

Kwai

Data can scraped using the platform's user interface on a computer, without the need for other devices, even though the volume of data returned is limited.

Telegram

Data can scraped using the platform's user interface on a computer, without the need for other devices.

WhatsApp

It is possible to collect data using scraping techniques, but it is necessary to use devices other than a computer

Accessibility

Q16

Can data can be recovered by scraping, without the need to circumvent tools and techniques designed to prevent programmatic access to data?

Here, we checked whether the platform uses tools such as Cloudflare to prevent access to data by means of a *paywall*, limited traffic rate and/or blocking by detecting automated behavior

YouTube

We have not identified any blockages in data collection by unofficial methods.

Instagram

Instagram blocks user activity based on the detection of automated behavior for data scraping.

TikTok

The platform periodically asks the user to solve a *Captcha* to check for automated use, preventing satisfactory scraping.

Facebook

Facebook blocks user activity based on the detection of automated behavior for data scraping.

X / Twitter

O X/Twitter blocks user activity based on the detection of automated behavior for data scraping.

Kwai

We have not identified any blockages in data collection by unofficial methods.

Telegram

Telegram provides no obstacles to users scraping the content available in public groups and channels.

WhatsApp

It is possible to scrape data from public groups on the platform, without blocks, as long as the application is running on a cell phone.

Compliance

Q17

Is the structure of the data made available on the platform stable?

In this field, the stability of the structure of the answers returned is checked by the API. It is considered stable if it does not change constantly and without at least 30 days' notice, with ample publicity given. It was also assessed whether changes to the API impact the functioning of applications integrated with them

YouTube

The API has been in version 3 since 2014 and has not undergone any major changes since then. In addition, the platform provides a well-documented revision history..

Instagram

Since CrowdTangle had not undergone any major updates since 2021, the structure of the data returned in requests to the tool's API had been stable until the tool was discontinued in August 2024.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

Facebook

Since CrowdTangle had not undergone any major updates since 2021, the structure of the data returned in requests to the tool's API had been stable until the tool was discontinued in August 2024.

X / Twitter

X/Twitter has done a lot of planning to ensure the migration from v1 to v2 of its API and there is documentation dedicated to migration processes. Although the structure of the data has changed considerably, all the processes have been well documented and exemplified.

WhatsApp

WhatsApp does not provide API access to public data.

Telegram

There are no reports from users about Telegram changing the structure of its collection API regularly without prior notice, which was confirmed in our experience of use.

Compliance

Q18

Is the data returned by the platform's API in a standardized format?

Here, we assessed whether the data types provided by the API match to consensus and/or standards in the field, such as the ISO 8601 date format, since standardized structuring makes data easier store and use.

YouTube

The dates of the publications rerendered by the YouTube API follow the ISO 8601 format and the URLs are complete.

Instagram

The dates of the publications returned by CrowdTangle were in ISO 8601 format and the URLs were complete.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Facebook

The dates of the publications returned by CrowdTangle were in ISO 8601 format and the URLs were complete.

X / Twitter

Currently, the results returned by X/Twitter show dates according to the ISO 8601 standard and the platform provides URLs in full form or using its own shortener (t.co).

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Telegram

The dates of the publications rerendered by the Telegram API follow the ISO 8601 format and the URLs are complete.

Compliance

Q19

Is the platform's API documentation published in open access?

Here, we checked whether the platform openly publishes documentation on the internet for the use of its API, without the need to register and log in.

YouTube

The platform has extensive documentation on the **most current version of the API**, as well as on previous versions.

Facebook

The **CrowdTangle API documentation** can be accessed by any user on GitHub, without the need for authentication, even though it has been discontinued.

Instagram

The **CrowdTangle API documentation** can be accessed by any user on GitHub, without the need for authentication, even though it has been discontinued.

X / Twitter

All of API X/Twitter's documentation can be accessed without the need for authentication, with the exception of details about the permissions of its paid plans, which are only available to paying users.

Telegram

Telegram's API documentation can be accessed by any user, without the need for authentication.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Compliance

Q20

Is the platform's API documentation clearly written and exemplified?

Here, we assessed whether the documentation for using the platform's API is clearly written, complete and has examples of implementation

YouTube

The YouTube API documentation provides clear examples for its use, referring to the different types of objects returned and the expected responses.

Facebook

In its API documentation, CrowdTangle provided clear examples of how requests should be made and the expected responses for each type of request.

Instagram

In its API documentation, CrowdTangle provided clear examples of how requests should be made and the expected responses for each type of request.

X / Twitter

The X/Twitter API documentation provides clear examples for its use, regarding the different types of objects returned and the expected responses.

Telegram

Telegram's API documentation does not make it clear how to call the endpoints and the standards adopted are not commonly used in other APIs, making their use dependent on solutions developed by third parties, such as **Telethon.**

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Compliance

Q21

Does the platform's API documentation describe its terms of use?

This field checks whether the API documentation clearly and unambiguously states the terms for its use and its legal aspects.

YouTube

The YouTube API documentation details its terms of use, including those **specific to regions of the world**.

Facebook

The terms of use of the CrowdTangle API were not clearly described in the official documentation, although they could be found on other pages referenced in it.

Instagram

The terms of use of the CrowdTangle API were not clearly described in the official documentation, although they could be found on other pages referenced in it.

X / Twitter

The terms of use of the X/ Twitter API are not clearly described in the official documentation, although they can be found on other pages referenced in the documentation and on the platform's developer portal.

Telegram

The Telegram API **terms of use page** is available on the platform's API homepage.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Compliance

Q22

Does the documentation describe the response format of the platform's API endpoints?

This field checks whether the API documentation describes the format each response, including examples and possible errors.

YouTube

There are descriptions and examples of API responses in various sections of the documentation, such as the ***Video List.***

Facebook

The CrowdTangle API documentation provided clear examples for each of the collected endpoints, as well as **a page dedicated to the errors** for each type of request.

Instagram

The CrowdTangle API documentation provided clear examples for each of the collection *endpoints*, as well as **a page dedicated to the possible errors** each type of request.

X / Twitter

The API documentation provides a data dictionary with examples of the expected responses for endpoints that deal with different types of objects.

Telegram

The official documentation does not describe the data types in a conventional and easily understandable way, and uses its own object terms.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Compliance

Q23

Is the documentation for the platform available in Portuguese?

Here, we checked whether the platform provides its API documentation in Portuguese, to be accessible for Brazilian users.

YouTube

Yes, the entire website on which the platform's API documentation is found is available in several languages, including Portuguese.

Instagram

All of CrowdTangle's API documentation, as well most of the user help pages, was only available in English.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Facebook

All of CrowdTangle's API documentation, as well most of the user help pages, was only available in English.

X / Twitter

X/Twitter only makes its API documentation available in English.

Kwai

Kwai does not provide an API for accessing public data.

Telegram

Telegram only makes its API documentation available in English.

WhatsApp

WhatsApp does not provide API access to public data.

Compliance

Q24

Does the platform allow scraping and other types of automatic access in its terms of use?

Here, we check whether the platform expressly prohibits data scraping techniques in its terms of use.

YouTube

In the **terms of use,** YouTube states that it is not permitted to "1) access, reproduce, download, distribute, transmit, display, sell, license, alter, modify or otherwise use any part of of the Service or any Content, except: (a) if expressly authorized by the Service; or (b) with the prior written permission of YouTube and, if applicable, the respective rights ; [...] 3) Access the Service using any automated means (such as robots, botnets or scrapers), except (a) in the case of public search engines compatible with YouTube's robots.txt file; or (b) with YouTube's prior written permission".

Instagram

According to **Instagram's Help Center,** the platform has teams dedicated to detecting and blocking accounts that show automated behavior patterns associated with improper data extraction.

TikTok

According to **plataform's terms of use,** not permitted to use automated scripts to collect information from the services available or to interact with them.

Facebook

According to **Facebook's Help Center,** the platform actively attempts to make it difficult for actors to external companies collect data from their platforms via scraping, under the justification of preventing them from unduly profiting from their users' data.

X / Twitter

In the **rules on automation account ,,** Twitter claims that the use of external scripts to automate navigation in its user interface can lead to the user's account being permanently deleted.

Kwai

The use of *web crawlers* and data mining tools violates Kwai's **terms of use**.

Telegram

In Telegram's **list of prohibitions and impermissible uses** in its terms of service (both for users and for the API), the platform does not mention scraping or other practices that made data scraping impossible.

WhatsApp

In its **terms of service,** o WhatsApp includes prohibitions that result in restrictions on the use of data collection applications, such as reverse-engineering parts of data. from the application and have a parallel system with data extracted from the application without prior authorization from the platform.

Compliance

Q25

Does the platfrom produce transparency reports on content moderation in Brazil and is it made publicly available, without the need for a request?

In this field, we checked whether the platform produces transparency reports and makes them publicly available at least every six months, without interested parties having to make a request. These reports should detail information on the application of its governance policies and moderation activity in Brazil, such as the number of publications removed or restricted and/or users suspended in the country.

YouTube

The **YouTube Community Guidelines Compliance Report** provides the number of videos removed and made by IPs in Brazil on a quarterly basis. The same report provides other data at a global level, such as the number of channels removed, the number of removals resulting from complaints, detection or other sources. It should be noted, however, that this document refers only to removals resulting from violation of the platform's guidelines: **Google's Transparency Center** does not provide information on all removals carried out by YouTube in Brazil, although this type of information is made available in specific reports for certain countries, such as **South Korea** and **Türkiye.**

Facebook

On the **Content Removals Based on Local Legislation Report** page, it is possible to download a database with Facebook's overall figures for posts, profiles, pages and comments with access restricted by Meta in Brazil (and in other countries where it operates) based on the country's legislation.
It is worth noting, however, that this data does not include removals applied as a result of violations of the Meta's guidelines, and its **Transparency Center** does not provide information on all removals and/or other moderation measures carried out by Facebook in Brazil.
This type of information is available in specific reports for certain countries, such as **South Korea, Austria, Türkiye and India.**

Instagram

On the **Content Restrictions on Local Legislation Report** page you can download a database with Instagram's general numbers for accounts and publications with access restricted by Meta in Brazil (and in other countries where it operates) based on the country's legislation.
It is worth noting, however, that this data does not include removals applied as a result of violations of Meta's guidelines, and its **Transparency Center** does not provide information on all removals and/or other moderation measures carried out by Instagram in Brazil.
This type of information is available in specific reports for certain countries, such as **South Korea, Austria, Türkiye and India.**

Compliance

Q25

The platform produces regular transparency reports about content moderation in Brazil and makes it publicly available, without the need for a request?

In this field, it is checked whether the platform produces and makes available, publicly and without the need for a request from interested parties, transparency reports at least every six months. These reports should detail information on the application of its governance policies and moderation actions in Brazil, such as the number of publications removed or restricted and/or users suspended in the country.

X / Twitter

X/Twitter does not provide transparency information on content moderation activity in Brazil. In its **Transparency Center** the section on Brazil was last updated in 2022, and only returns a message warning that no data is available.

TikTok

TikTok provides a ***Community Guidelines Enforcement Report*** which is updated every three months, with information on the volume and type of moderation actions carried out in all countries. It is also possible to export a database and filter it according to the country of origin of the content with identified violations, including Brazil.

Kwai

The **Kwai Safety Center** submits semi-annual transparency reports in which it groups together moderation activities for Latin America and other regions of the world, and it is not possible to identify specific data on Brazil.

Telegram

Telegram does not provide any transparency reports for Brazil. In its **Privacy policies,** the platform states that if it shares user data with authorities confirmation of suspected terrorism, this will be reported in a biannual transparency report, to be made available on its own @transparency channel. In any case, Telegram claims that this has never happened.

WhatsApp

WhatsApp does not provide any transparency reports on content and user moderation initiatives in Brazil or in other countries.

Compliance

Q26

In its transparency reports, does the platform indicate the volume of each type of violation identified in Brazil, according to the current moderation policies?

Here, we checked whether the platform's transparency reports present information on the volume of violations identified, divided up by type of violation, during the enforcement of its governance policies and moderation activities in Brazil. The types of violation can include, for example, the dissemination of illegal content, hate speech and false information.

YouTube

Grouping by type of violation identified is only available at a global level in the **YouTube Community Guidelines Compliance Report.** At the national level, only **governmental requests for content removal** are grouped by type of violation, not the number of violations identified of removals actually carried out – although this type of information is available in specific reports for certain countries, such as **South Korea and Türkiye.**

Facebook

None of the documents from the **Transparency Center** presents the number of moderation activities in Brazil grouped by type of violation – as is available at a global level and in specific reports for countries, such as **South Korea, Áustria** e **India.**

Instagram

No document from the **Transparency Center** presents the number of moderation actions in Brazil grouped by type of violation – as is available at a global level and in specific reports for certain countries, such as **South Korea, Austria** e **India.**

X / Twitter

X/Twitter does not provide transparency information on content moderation actions in Brazil. In **its Transparency Center,** the section on Brazil was last updated in 2022 and only returns a message warning that no data is available.

Telegram

Telegram does not provide any transparency reports for Brazil. In its **Privacy Policies,** the platform states that if it shares user data with authorities upon confirmation of suspected terrorism, this will be reported in a biannual transparency report, to be made available on the @transparency channel. In any case, Telegram claims that this has never happened.

Compliance (continued)

Q26

In its transparency reports, does the platform indicate the volume of each type of violation identified in Brazil, according to the current moderation policies?

Here, we checked whether the platform's transparency reports present information on the volume of violations identified, divided up by type of violation, during the enforcement of its governance policies and moderation activities in Brazil. The types of violation can include, for example, the dissemination of illegal content, hate speech and false information.

TikTok

In the global databases available on TikTok's ***Community Guidelines Enforcement Report*** page, it is possible to export a database showing the number of violations identified for each country where the platform operates, including Brazil. In this database, the types of violations committed are indicated, with categories such as "Misinformation", "Civic & Election Integrity", "Hate Speech & Hateful Behavior", "Youth Exploitation & Abuse", among others.

Kwai

The **Kwai Safety Center** submits half-yearly transparency reports in which it groups moderation actions for Latin America and other regions of the world, and it is not possible to identify specific data on Brazil.

WhatsApp

WhatsApp does not provide any transparency reports on content and user moderation initiatives in Brazil or in other countries.

Compliance

Q27

Do Transparency reports specify information on the number and type of requests made by Brazilian government entities to the platform, as well as number and type of requests accepted?

This field checks whether the transparency reports produced by the platform list the requests for data moderation and delivery made by Brazilian state entities, detailing the nature of the request, the total number of requests and the volume of requests granted and denied.

YouTube

YouTube is part of Google's report on **government requests to remove content** along with other services from the same business group. A general overview and a set of databases are provided which include the volume of requests by type, every six months and for each platform and country, including Brazil.

Instagram

Meta presents a report on *Government Requests for User Data* which includes requests made by state entities in Brazil and other countries for the moderation of and/or access to Instagram data.

Telegram

Telegram does not provide any transparency reports for Brazil. In its **Privacy Policies,** the platform states that if it shares user data with authorities upon confirmation of suspected terrorism, this will be reported in a biannual transparency report to be made available on the @transparency channel. n any case, Telegram claims that this has never happened.

X / Twitter

X/Twitter does not provide transparency information on content moderation actions in Brazil. In **its Transparency Center,** the section on Brazil was last updated in 2022 and only returns a message warning that no data is available.

Facebook

Meta presents a report on *Government Requests for User Data* which includes requests made by state entities in Brazil and other countries for the moderation of and/or access to Facebook data.

Compliance (continued)

Q27

Do Transparency reports specify information on the number and type of requests made by Brazilian government entities to the platform, as well as number and type of requests accepted?

This field checks whether the transparency reports produced by the platform list the requests for data moderation and delivery made by Brazilian state entities, detailing the nature of the request, the total number of requests and the volume of requests granted and denied.

TikTok

TikTok's *Government Removal Requests Report* presents information on the volume and types of data requests made by state entities in Brazil and other countries since 2019, segmented by semester.

Kwai

The Kwai *Safety Center* asubmits half-yearly reports in which it segments moderation actions for Latin America and other regions of the world, and it is not possible to identify specific data for Brazil.

WhatsApp

WhatsApp does not provide any transparency reports on content and user moderation initiatives in Brazil or in other countries.

Completeness

Q28

Is it possible to retrieve data from the comments of a publications using the platform's API?

Here, we see if comment data, including their content, can be retrieved when available on the platform, either alongside the publication data or through a specific endpoint.

YouTube

Comments can be **collected by specific endpoints of** the platform's official API.

Facebook

CrowdTangle only returned information on the volume of comments made on a publication, but it was not possible to access the content of the comments

Instagram

CrowdTangle only returned information on the volume of comments made on a publication, but it was not possible to access the content of the comments

X / Twitter

There are specific operators for constructing search expressions **to retrieve responses to a publication.** In addition, the API provides endpoints for capturing quote tweets made about a publication.

Telegram

All messages sent on public Telegram channels are returned by the platform's API, without specifying whether the message is a comment.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

All posts in WhatsApp groups are treated as "messages", with no differentiation between the original message and the reply to another post. Thus, there are no "comments" in WhatsApp group conversations.

Completeness

Q29

It is possible to recover data from temporary content through the platform's API?

In this field, we check whether the platform's API provides at least one endpoint for retrieving data from temporary publications, such as stories and temporary messages.

YouTube

It is not possible to publish temporary content on YouTube.

Facebook

As stated in the **oficial documentation** CrowdTangle did not archive any kind of temporary content, even though it is possible to publish content of this nature on Facebook, such as stories.

Instagram

As stated in the **official documentation,** CrowdTangle did not archive any kind of temporary content, even though it is possible to publish content of this nature on Instagram, such as stories.

X / Twitter

V2 of the X/Twitter API provides a specific endpoint for **capturing data for spaces**, although it is not possible to retrieve audio content.

Telegram

Telegram's API makes it possible to collect **data on stories** (temporary photos) while the content is still live.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Completeness

Q30

Can historical data be retrieved using from the platform's API?

Here, we assessed whether the API offers endpoints that allow you to indicate a specific period of time covering at least the last 365 days for data collection, counted from the time of the request.

YouTube

Yes, the API makes it possible to **collect data from any videos that are available** on the platform, based on periods indicated by the user.

Facebook

CrowdTangle allowed for the search and extraction of historical data from publications made available, using both the collection interface and in the API, for the periods indicated by the user.

Instagram

CrowdTangle allowed for the search and extraction of historical data from publications made available, using both the collection interface and in the API, for the periods indicated by the user.

X / Twitter

The X/Twitter API paid plans we reviewed only allow access to data published **within the seven days prior to the date** it is requested.

Telegram

Telegram does not stipulate a time limit for searching and collecting messages, as long as they are public in the channel or interest group to be collected.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Completeness

Q31

Is the number of requests allowed by the platform sufficient to monitor more than 1 million publications in 24 hours?

Here, we asses whether data can be recovered without interruption and losses through the platform's API for requests that accumulate more than 1 million publications in 24 hours.

YouTube

According to the **quota calculators** available in the platform's API documentation, it is not possible to collect data at this level, whether it is video data or of comments.

Facebook

As CrowdTangle documentation stated the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

Instagram

As CrowdTangle documentation stated the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

X / Twitter

As the ceiling of publications in the Pro plan is 1 million per month, it would not be possible to monitor this scale, considering the access plans analyzed.

Telegram

Based on the tests carried out and the platform's documentation, no limitations were identified regarding the message request rate for the specified quantity

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Completeness

Q32

Is the number of requests allowed by the platform sufficient to monitor more than 100,000 publications in 24 hours?

Here, we asses whether data can be recovered without interruption and losses through the platform's API for requests that accumulate more than 100,000 publications in 24 hours.

YouTube

According to the **quota calculation** provided in the platform's API documentation, it is possible to collect as much data as you want, be it videos or comments, as long as the IDs of the videos of interest are known in advance.

Instagram

As the **CrowdTangle documentation** stated, the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

Facebook

As the **CrowdTangle documentation** stated, the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

X / Twitter

As the ceiling for publications in the Pro plan is 1 million per month, it would be possible to monitor this scale, considering the access ranges analyzed.

WhatsApp

WhatsApp does not provide API access to public data.

Telegram

Based on the tests carried out and the platform's documentation, no limitations were identified regarding the message request rate for the specified quantity.

Completeness

Q33

Is the number of requests allowed by the platform sufficient to monitor more than 10,000 publications in 24 hours?

Here, we asses whether data can be recovered without interruption and losses through the platform's API for requests that accumulate more than 10,000 publications in 24 hours.

YouTube

According to the **quota calculation** provided in the platform's API documentation, it is possible to collect this volume of publications, whether videos or comments, as long as the IDs of the videos of interest are known in advance.

Instagram

As the **CrowdTangle documentation** stated, the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

Facebook

As the **CrowdTangle documentation** stated, the API was not recommended for extracting more than 10,000 posts at a time. For cases between 10,000 and 100,000 posts, the tool's collection interface was more appropriate.

X / Twitter

As the ceiling for publications in the Pro plan is 1 million per month, it would be possible to monitor this scale, considering the access ranges analyzed.

WhatsApp

WhatsApp does not provide API access to public data.

Telegram

Based on the tests carried out and the platform's documentation, no limitations were identified regarding the message request rate for the specified quantity.

Consistency

Q34

Is the data returned by the platform's API persistent?

This field checks that the data retrieved by the platform's API is immune to expiration, even if it is a link. It is expected that some metadata from publications removed from the platform will not be excluded from API response, but that there will be a sign that the content has been removed.

YouTube

The API does not return any metadata from videos that have been removed or made private, unlike the platform's user interface, which displays a message about the removal and the reasons for it if the user tries to watch a video that is no longer on the air.

Facebook

CrowdTangle treated posts deleted from Facebook as if they had never existed, even when a specific search was made for them. In addition, our tests showed that the API returned random posts when a nonexistent or misspelled post ID was entered. On a positive note, links to other content do not expire.

Instagram

Image link fields returned by CrowdTangle expired so quickly (URL Signature Expired error) that it was impossible to manipulate or analyze them, even in the case of active publications.

X / Twitter

The links returned by API requests do not expire and **compliance endpoints** guarantee a higher return of metadata, even when the publications are no longer available on the platform; status and availability are displayed for both profiles and publications.

Telegram

The Telegram API treats deleted messages as if they had never existed.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Consistency

Q35

Does the data retrieved by the platform's API reflect what is displayed on its user interface?

This field checks whether the data returned by the API corresponds to the information displayed on the platforms' user interface. It should be possible to identify in the API response information such as authorship, complete content and main interactions.

YouTube

Yes, the API allows all the main data available on the platform's user interface to be retrieved, with the exception of the video content itself. There is, however, an identifier to access the video on the platform.

Facebook

As CrowdTangle returned the main information from the publications collected, which included content, authorship, date and time of publication, as well as the volume of the main interactions, it was possible to say that the data reflected that which was displayed on the platform. It was not possible, however, to identify the users responsible for the likes, comments or shares.

Instagram

It wasn't possible to use image links systematically because they expired quickly. Due to the importance of visual content on the platform, we understood that the CrowdTangle API should provide functional links to published media. For this reason, we considered that the data did not reflect what was shown on the platform, even though the API provided data such as authorship, date and time of publication and the volume of the main interactions.

X / Twitter

It is not possible to collect all information because X/Twitter does not provide the number of impressions received by publications through its API, although it does show this information in the user interface. This data is only made available via API with the authorisation of its authors.

Telegram

There are no inconsistencies in the data returned by the platform in the collection and what is displayed on the user interface, including media (images, videos, audio), which can be downloaded and analyzed in an uncomplicated way.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Consistency

Q36

Is the answer returned by the platform's API always expected?

This field assesses whether the data retrieved through the API is always the same, according to the parameters and filters used, or at least consistent with other collections made in the same way, with the exception of publications deleted or exchanged between them.

YouTube

There are frequent inconsistencies, even in collections by fixed parameters, i.e. without the involvement of algorithmic content recommendation. Collection tests carried out a few minutes apart retrieve data that is considerably different, and this difference is not necessarily due to the inclusion of more recent data.

Facebook

Based on the tests carried out, requests to the CrowdTangle API made at different times and by different people returned consistent data, without substantial variations, with the exception of publications that have been removed or made private.

Instagram

Based on the tests carried out, requests to the CrowdTangle API made at different times and by different people returned consistent data, without substantial variations, with the exception of publications that had been removed or made private.

X / Twitter

Based on the tests carried out, requests to the X/Twitter API made at different times and by different people return consistent data, without substantial variations, with the exception of publications that have been removed or made private.

Telegram

Based on the tests carried out, requests to the Telegram API, at different times and by different people, return consistent data, without substantial variations, with the exception of publications that have been removed or made private.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Consistency

Q37

Is the response returned by the platform's API consistent with the parameters and filters used in the request?

This field checks whether the data retrieved by the platform's API reflects the choices of parameters and filters determined at the time of the request.

YouTube

There are frequent inconsistencies, even in collections using fixed parameters, without the involvement of algorithmic content recommendation. In our tests, we found that the YouTube API does not necessarily obey filters determined by users (most notably date and language filters).

Facebook

The publications returned by CrowdTangle were consistent with the parameters and filters defined. Furthermore, the data collection interface showed which publications were the result of a specific search and why each one had been returned.

Instagram

The publications returned by CrowdTangle were consistent with the parameters and filters defined. In the data collection interface, it was even possible to check which publications resulted from a specific search and confirm why each one had been returned.

X / Twitter

The API responses indicates the appropriate application of the filters and parameters indicated by the user in relation to the time interval, the language of the publications and the search terms of interest, for example.

Telegram

Based on the tests carried out, the data returned always matches the filters selected in the request to the API.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Relevance

Q38

Are the entities returned by the platform's API enough to understand the data in all its levels of detail?

Here, we assessed whether the API response delivers all the data necessary to understand the complete lifecycle of a specific publication, including comments, shares, replies and other possible relationships, as well as its authors and referenced content, as in the case of shares and mentions.

YouTube

Yes, the API returns **all the main entities seen** in a video on the platform's user interface. Although the content of the video is not returned by the API, it provides an identifier so that it can be accessed on the platform.

Facebook

As CrowdTangle **didn't return responses and comments from publications**, it wasn't possible to evaluate the data with all the expected levels of detail.

Instagram

As CrowdTangle **didn't return responses and comments from publications**, it wasn't possible to evaluate the data with all the expected levels of detail.

X / Twitter

In addition to the ID of the publication in question, v2 of the X/Twitter API **provides the identifiers of related** publications and, if they are identified in the request, the data of related publications.

Telegram

The data is technically understandable. There is one exception: the Telegram API does not, by default, allow the programmatic collection of data relating to the authors of messages, since they are not represented by their own entities. Even so, we believe that Telegram meets the criterion because it is possible to identify authors and crossreference authorship information between messages sent in different groups.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Relevance

Q39

Does platform's API allow the use of filters to refine the data request?

This field checks whether the platform's API allows the use of search filters, such as the publisher's location, language or specific period, among others.

YouTube

The API allows you to apply various search filters. However, the language filter does not guarantee that the results will be limited to the chosen language because, according to the platform's, documentation, if content in other languages is considered relevant, it will be included in the answer. Therefore, the options presented do not meet the minimum requirement for this parameter.

Facebook

The language of the publications, the country declared by their authors and the time interval in which they were made were among the **possible filters to recuperate publications** through the CrowdTangle API.

Instagram

The language of the publications, the country declared by their authors and the time interval in which they were made were among the **possible filters to recuperate publications** through the CrowdTangle API.

X / Twitter

Certain operators only work in certain **plans API access** , but in general it is possible to apply filters according to the user's objectives.

Telegram

In the platform's API documentation we did not find any mention of localization or language filters to search for publications.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Current

Q40

Can newly-published data be retrieved from the platfprm's API in real time, or just after publication?

Here, we assessed whether the platform's API allows data to be retrieved from a set of specific publications within one hour of their publication.

YouTube

Yes, tests with re-published videos show that once the video ID has been applied, collection takes place normally. The collection of videos from live broadcasts also occurs in a similar way,delivering the same types of data.

Facebook

Tests with CrowdTangle indicated that very recent publications quickly became available for collection by means of of the tool.

Instagram

Tests with CrowdTangle indicated that very recent publications quickly became available for collection by means of of the tool.

X / Twitter

Tests indicate that data is made available in the API as soon as a publication is made, and it is possible to retrieve inserted publications in a window of a few minutes.

Telegram

Tests show that the API makes the message available almost instantly.

TikTok

TikTok does not provide an API for accessing public data in Brazil.

Kwai

Kwai does not provide an API for accessing public data.

WhatsApp

WhatsApp does not provide API access to public data.

Bibliographical references

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://atricon.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/VGhlIEd1aWRlIHRvIElTTyA4MDAw/V2hpdGVwYXBlcg==>. 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>. 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: <https://independenttechresearch.org/letter-twitter-api-access-threatens-public-interest-research/>. Acesso em: 4 abr. 2024.

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-biodatasci-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 CrowdTangle | 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.; WILLIANS, M.; SLOAN, L.; WILLIANS, 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/europe-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. KoPl-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-conteudo/uploads/2018/02/Crit%C3%A9rios-Avalia%C3%A7%C3%A3o-i-TA-2016.pdf. 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_ActicismAndMisinformationOnWhatsApp.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_VeEqU2usAb7cb6fv&_nc_ht=z-p3-scontent.fsdu2-2.fna&oh=00_AfAtew8y7eJiaiUAzbhY6GJ80FpnuyFFR2_iRGVHSkU0OQ&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-do-or-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.; STRAUBHAAR, 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/fulltext>. 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-crowtangle-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 Kozhikode Society & Management Review**, Calicute, 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-twitters-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