

COMPLAINT

against

Google Ireland Limited

regarding

infringements of Article 17 and 24 (5) of
Regulation (EU) 2022/2065 of the European Parliament
and of the Council of 19 October 2022 on a Single Market for
Digital Services and amending Directive 2000/31/EC
(Digital Services Act).

Filed by the Gesellschaft für Freiheitsrechte e.V. / Society for Civil Rights

Table of contents

A. General Remarks	3
1. Introduction.....	3
2. Applicable Legal Provisions.....	3
3. Design of the DSA Transparency Database	6
B. Evaluation of Statements of Reasons	6
1. Implementation	6
2. Assessment	9

A. General Remarks

1. Introduction

The Gesellschaft für Freiheitsrechte e.V. (Society for Civil Rights, GFF) is a non-profit organization based in Berlin that has been properly constituted in accordance with the German law. According to its statutory objectives (**Exhibit 01**)¹ GFF aims to defend fundamental and civil liberties through legal means. One of its key focus areas is digital rights in the modern age. To enhance the enforcement of online rights, GFF established the Center for User Rights, which aims to uphold user rights under the Digital Services Act (Regulation (EU) 2022/2065, DSA), among other initiatives.

Liz Carolan, a Dublin resident and user of the online platform YouTube (**Exhibit 02**) has authorized GFF under Article 86 (2) DSA to exercise her right to lodge a complaint against Google Ireland Limited alleging an infringement of Article 17 and Article 24 (5) DSA with the Coimisiún na Meán on her behalf (**Exhibit 03**). She is a Board member of the Irish Council for Civil Liberties and has been active in the field of platform regulation and protection of user rights as a tech and democracy strategist and campaigner.

This complaint concerns the respondent's systematic failure to provide adequate statements of reasons for its content moderation decisions, as required by Article 17 DSA. Under this provision, providers of hosting services must offer sufficient justification for their moderation decisions. However, the respondents' current practice is to provide only generic explanations, typically referencing a violation of national law or the platform's terms of service in an abstract manner, without addressing the specific circumstances of each case.

As a result, affected users are often unable to effectively challenge moderation decisions. In many cases, it remains unclear why an account was suspended or why a post was deemed to violate the platform's rules. Even when users seek clarification, the respondents frequently fail to provide a more detailed explanation or responds with significant delays. This lack of transparency makes it difficult for affected individuals to effectively lodge complaints through the internal complaint-handling system of the platforms. Without sufficient information, users have little ability to argue against an erroneous moderation decision.

2. Applicable Legal Provisions

a) Article 17 DSA

According to Article 17 (1) DSA, providers of hosting services must provide all affected users a clear and specific statement of reasons for any restrictions imposed on the grounds that the content they provided is either illegal or violates the platform's terms of service.

¹ § 3 (2) (a) (b), (3) (a), (4) (a) Articles of Association of Gesellschaft für Freiheitsrechte e.V.

The detailed obligations set out in Article 17 DSA primarily serve to ensure transparency and constitute an essential aspect of procedural fundamental rights protection.² These requirements safeguard users from negative consequences of moderation decisions, particularly regarding their freedom of expression (Recital 54 (1) DSA). Only when hosting services provide clear and specific statements of reasons for their moderation decisions users would be able to understand the exact justifications behind a restriction and to make an informed decision about whether to adapt their behavior or challenge the restriction.³

Beyond protecting individual rights, Article 17 DSA also aims to foster a predictable and trustworthy online environment (Recitals 2, 3, and 4 DSA) and harmonize justification obligations across EU member states, providing legal certainty for providers of hosting services in their communication with users.⁴

A clear and specific statement of reasons requires detailed explanations, making generic or vague justifications inadequate, such as standardized responses with only a general reference to the violation of the terms of service.⁵ Providers of hosting services must provide information that is tailored to the specific case⁶ and explicitly state which aspect of the content is being disputed.⁷ The explanation must be as detailed and specific as reasonably possible under the given circumstances. The statements of reasons must also enable affected users to assess and to effectively exercise their available legal remedies (Article 17 (4) DSA). To this end, the provided explanation must be clear and easy understandable.⁸

Among the minimum requirements for adequate information set out in Article 17 (3) DSA, Article 17 (3) (b) DSA is particularly noteworthy, as it requires a statement of the facts and circumstances underlying the decision. This means that the statements of reasons must include sufficient details about the factual background that led to the moderation decision, clearly identifying the specific content subject to restriction, and ensuring that users can understand the key factual basis for the decision.⁹

In cases where content is classified as illegal, the specific legal basis must be provided. According to Article 17 (3) (d) DSA, a reference to the legal ground relied on has to be stated and an explanation has to be given as to why the information is considered to be illegal content on that ground. This requires an explicit connection between the content in question and the content moderation decision, allowing users to determine whether they can challenge the restriction.¹⁰ A clear reference to the content and

² *Raue*, in: Hofmann/Raue, DSA, 2023, Art. 17 para. 1; Cf. *Barudi*, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 17 para. 2.

³ *Ibid.*; cf. *Barudi*, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 17 para. 2; cf. Gerdemann/Spindler GRUR 2023, 3 (9).

⁴ Cf. *Barudi*, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 17 para. 2.

⁵ *Raue*, in: Hofmann/Raue, DSA, 2023, Art. 17 para. 48.

⁶ *Ibid.* para. 56.

⁷ *Ibid.* para. 48.

⁸ *Barudi*, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 17 para. 29.

⁹ Cf. *ibid.* para. 32.

¹⁰ Cf. *ibid.* para. 35.

its classification under the relevant legal provision is necessary; generic, standardized responses are insufficient.¹¹

When a moderation decision is based on a violation of the platform's terms of service, the provider of a hosting service must, according to Article 17 (3) (e) DSA, specify the relevant contractual provision and explain why the content is deemed non-compliant. It is important to note that the term "terms of service" is broadly defined (Article 3 (u) DSA). If the provider of a hosting service extends its terms of service across multiple documents, such as general terms of use, corresponding community standards, community guidelines, etc., the relevant provisions may derive from different policies.¹² Therefore, the exact contractual provision must be specified¹³; standardized text modules or general references to "violations of the terms of service" are not sufficient.¹⁴

b) Article 24 (5) DSA

According to Art. 24 (5) DSA the providers of online platforms shall, without undue delay, submit to the Commission the decisions and the statements of reasons referred to in Article 17 (1) DSA for the inclusion in a publicly accessible machine-readable database managed by the Commission.

The purpose of the disclosure obligation is to ensure transparency and enable public oversight of content moderation decisions (Recital 66 (1)).¹⁵ The comprehensive and collective publication in the database enables the analysis of trends and patterns in the platforms' content removal practices. According to Recital 66 (3) the structured database should allow access to, and queries for the relevant information, in particular as regards to the exact type of alleged illegal content at stake.

The moderated content itself does not need to be disclosed (Article 24 (5) (2) DSA). However, both the decision and its statement of reasons must be fully communicated. In cases where the decision is based on alleged illegality, the legal basis and an explanation of how the content violates the law must be provided. Similarly, when the decision is based on a terms of service violation, the specific contractual provision and its application to the case must be disclosed and subsequently published. Furthermore, the transparency database must allow users to search for the specific type of violation, whether it pertains to national law or the platform's policies.¹⁶ The only limitation set forth by Article 24 (5) DSA is that the disclosed information must not contain personal data.

c) Article 20 DSA

Art. 20 (1) DSA requires providers of online platforms to provide their users for a period of at least six months following the decision referred to in this paragraph, access to an effective internal complaint-

¹¹ Grünwald/Hackl: Inthaltmoderation bei Online-Plattformen, MMR 2024, 532 (535); Cf. Raue, in: Hofmann/Raue, DSA, 2023, Art. 17 para. 57.

¹² Cf. Barudi, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 17 para. 36.

¹³ Raue, in: Hofmann/Raue, DSA, 2023, Art. 17 para. 52.

¹⁴ Grünwald/Hackl: Inthaltmoderation bei Online-Plattformen, MMR 2024, 532 (535); Cf. Raue, in: Hofmann/Raue, DSA, 2023, Art. 17 para. 57.

¹⁵ Cf. Raue, in: Hofmann/Raue, DSA, 2023, Art. 24 para. 1.

¹⁶ Cf. Barudi, in: Müller-Terpitz/Köhler, DSA, 1st edition 2024, Art. 24 para. 18.

handling system that enables them to lodge complaints, electronically and free of charge, against certain content moderation decisions taken by the provider of the online platform.

The content moderation decisions subject to complaint under Article 20 (1) DSA largely correspond to the restrictions that the hosting service provider must inform users about and justify in detail under Article 17 (1) DSA. Providing a clear and specific statement of reasons required under Article 17 DSA is primarily intended to enable the affected users to effectively exercise their right to appeal through the internal complaint management system (Article 17 (4) (2) DSA).¹⁷ Users must be able to explain why the content referenced in the complaint is neither illegal nor in violation of the terms of service.¹⁸

The internal complaint management system is intended to enable affected users to challenge particularly restrictive decisions in an "easy and effective" manner (Recital 58 (1)). It allows for "quick and straightforward clarification of any misunderstandings regarding content" and for errors to be corrected.¹⁹ Thus, it also serves to clarify the facts of the case. The complaint procedure plays a crucial role in protecting users' rights and preventing chilling effects on their exercise of civil liberties in the digital realm.²⁰

3. Design of the DSA Transparency Database

The DSA transparency database is structured as follows: It features entry fields for submitting moderation decisions and their underlying statements of reasons, corresponding to the minimum information requirements outlined in Article 17 (3) DSA.

For each submitted moderation decision, one of 15 categories has to be selected to classify the type of violation (e.g. scope of platform service, illegal or harmful speech, violence etc.). For an overview of the categories, please refer to the DSA Transparency Database website: <https://transparency.dsa.ec.europa.eu/page/documentation>.

Further, among other requirements, providers of online platforms must use a text field to specify the facts and circumstances on which their decision is based. Additionally, they must provide a reference to the contractual basis and, in a separate text field, explain why the moderated content is deemed illegal or in violation of the platform's terms of service.

B. Evaluation of Statements of Reasons

1. Implementation

The deficiencies in Google's statements of reasons for moderation decisions on YouTube can be demonstrated through its content related decisions that are published in the DSA Transparency Database.

¹⁷ *Raue*, in: Hofmann/Raue, DSA, 2023, Art. 20 para. 39.

¹⁸ *Ibid.* para. 24.

¹⁹ Cf. BGH ruling from 29.7.2021 – III ZR 179/20, NJW 2021, 3179, para. 84 – hate speech.

²⁰ Cf. *Raue*, in: Hofmann/Raue, DSA, Art. 20 para. 1, 3.

Google consistently keeps its statements of reasons for their content moderation decisions vague, citing a general violation of the platform's terms and conditions and categorizing the content under abstract violation types (e.g. violent and graphic content, age restricted content, hates speech or dangerous products and services) without providing any further explanation for why the content is deemed a violation in the individual case.

The data in the DSA Transparency Database reveals that Google categorizes its statements of reasons for content moderation decisions on YouTube over the past six months into 11 different types of rule violations according to the set categories that the database provides:

- Scope of platform services
- Illegal or harmful speech
- Scams and/or fraud
- Violence
- Data protection and privacy violations
- Protection of minors
- Unsafe and/or illegal products
- Pornography or sexualized content
- Negative effects on civic discourse or elections
- Unsafe and/or illegal products
- Intellectual property infringements

For a comprehensive overview of all statements of reasons and their categorizations please consult the [DSA Transparency Database](#).²¹

An analysis of all moderation decisions submitted and published on the DSA transparency database from October 1st to December 31st, 2024, revealed that Google cited a general violation of platform rules as the secondary reason for deleting or restricting posts on YouTube. Specifically, 43,5 % of cases were based on a general violation of platform rules, 46,7 % on fraud and manipulation, 3,9 % on intellectual property and privacy issues, 5,6 % on violence and harm and 0,3% on pornography and sexual content.²²

Within each of the aforementioned categories Google narrows down the specific type of violation for each moderation decision, providing them as the contractual ground that is supposedly breached. Examples for classifications are “dangerous products and services”, “healthcare and medicines” or “abusing the ad network”. However, Google consistently omits a further explanation on why the content is considered as incompatible on that ground and assigned to a certain type of prohibited behavior in the relevant text field. For instance, for violations of its Advertising Policies it only points to

²¹ DSA Transparency Database, statements of reasons LinkedIn, available at: https://transparency.dsa.ec.europa.eu/statement?s=&platform_id%5B%5D=27&platform_id-27=on&created_at_start=&created_at_end=

²² Helen Bielawa, Nina Krug, Rina Wilkin, Was wir nicht sehen sollen from January 25, 2025, available at: <https://www.spiegel.de/netzwelt/netzpolitik/geloeschte-inhalte-auf-tiktok-x-instagram-youtube-was-wir-nicht-sehen-sollen-datenanalyse-a-ca0629a1-16c4-47a3-a9fa-10b7c6b44ac2>.

the corresponding section of its policies there. Further, Google fails to provide more detailed information related to the specific facts and circumstances of the individual case. Instead, regardless of the classification, for violation of its Advertising Policies Google generically states in the relevant text field that

“When reviewing content or accounts to determine whether they violate our policies, we take various information into consideration when making a decision, including the content of the creative (e.g. ad text, keywords, and any images and video) as well as the associated ad destination. We also consider account information (e.g., past history of policy violations) and other information provided through reporting mechanisms (where applicable) and our investigation.”

Evidence:

Screenshot [“DSA Transparency Database Google: dangerous products and services”](#)

- **Exhibit 04**

Screenshot [“DSA Transparency Database Google: healthcare and medicines”](#)

- **Exhibit 05**

Screenshot [“DSA Transparency Database Google: abusing the ad network”](#)

- **Exhibit 06**

The corresponding sections of Google’s terms of use for the relevant type of violation describe them only in abstract terms, covering a wide range of violations. For instance, there is a section on “dangerous products or services” in its Advertising Policies which contains a broad, non-exhaustive list of prohibited items such as explosives, guns, gun parts, & related products as well as any other weapons or recreational drugs.

Evidence:

Screenshot [“Google List of ad policies: dangerous products or services”](#)

- **Exhibit 07**

Similarly, as the following evidence shows, for violations of other policies Google just provides a generic and broad explanation of what generally would constitute such a violation in the text field provided for an explanation on why the content is considered as incompatible on the referenced contractual ground. For instance, when a moderation decision is taken based on a violation of its hate speech policy, only the following information is provided:

“Content that incites hatred against individuals or groups based on their protected group status isn’t allowed on YouTube. This may include dehumanization, using slurs and stereotypes, inferiority claims, and/or conspiracy theories. We review educational, documentary, artistic, and scientific content on a case-by-case basis. Limited exceptions are made for content with sufficient and appropriate context.”

Likewise, the facts and circumstances do not include any case-specific information but simply repeat this abstract explanation verbatim.

Evidence:

Screenshot "[DSA Transparency Database Google: hate speech policy](#)"

- **Exhibit 08**

Screenshot "[DSA Transparency Database Google: harassment policy](#)"

- **Exhibit 09**

Screenshot "[DSA Transparency Database Google: harmful or dangerous content policy](#)"

- **Exhibit 10**

Screenshot "[DSA Transparency Database Google: violent or graphic content policy](#)"

- **Exhibit 11**

Screenshot "[DSA Transparency Database Google: violent criminal organizations policy](#)"

- **Exhibit 12**

All these policies cover a broad spectrum of prohibited behaviors and are further divided into several subcategories that each contain its own list of prohibited acts. Additionally, some policies overlap by incorporating behaviors outlined in other policies and referencing them. For instance, the Hate Speech Policy classifies threats as prohibited behavior and directs users to the "Harassment & Cyberbullying Policies" for further details. The following evidence illustrates the diverse range of behaviors encompassed by three of YouTube's policies. For additional information on its other policies please consult the [YouTube policies page](#).²³

Evidence:

Screenshot "[YouTube hate speech policy](#)"

- **Exhibit 13**

Screenshot "[YouTube harassment & cyberbullying policies](#)"

- **Exhibit 14**

Screenshot "[YouTube harmful or dangerous content policy](#)"

- **Exhibit 15**

2. Assessment

Google systematically fails to meet the requirements for its statements of reasons of moderation decisions on YouTube as set out in Article 17 DSA. Contrary to Article 17 (3) (b) DSA, Google does not provide any specific facts or circumstances related to the individual case on a regular basis. Instead, it offers only a generic classification of the alleged violation by referring to the relevant section of its policies and provides a general description of what kind of behavior would constitute such a violation. Based on the wording and purpose of Article 17 (1) (3) DSA such abstract categorizations and explanations without further case-specific details do not meet the requirements of a clear and specific statement of reasons.

The categories Google employs to classify violations are overly broad and encompass a wide range of different situations, including illegal acts as well as other breaches of its terms of use. The lack of clarity

²³ YouTube Policies, available at:

https://support.google.com/youtube/topic/2803176?hl=en&ref_topic=6151248&sjid=13595234492736021864-EU

regarding the exact behavior being addressed is further exacerbated by the partial incorporation of other policies. Also, many of the prohibited actions are open to interpretation and require clarification and an evaluation from the platform to determine their applicability to a specific case, particularly in relation to context and language and with regard to the permissible exercise of fundamental rights in each case, in accordance with Art. 14 (4) DSA. For instance, what may initially appear to be an insult could, in fact, constitute permissible criticism within the scope of satire. These ambiguities create even greater uncertainty for affected users.

Further, Art. 17 (3) (e) DSA makes it clear that simply citing the contractual provision that has been violated is not sufficient. Instead, additional information must be provided to explain why the specific post falls under the prohibited behavior outlined in the provision. This is also demonstrated by the structure of the DSA Transparency Database which entails dedicated text fields for providing such explanations. Google consistently fails to comply with this requirement, as it merely repeats – verbatim – the description it provided in the text field for facts and circumstances for what kind of behavior constitutes a violation of a given section of its policies in general.

For instance, if Google removes a post on the grounds that it allegedly constitutes a dangerous product or service, it must identify and name the particular good or service in question (e.g. a firearm or tobacco products). If Google justifies a removal by claiming it constitutes misinformation, it must specify which part of the content is factually inaccurate and provide further details to support that claim. None of these information are provided by Google's statements of reasons.

As a result, users of YouTube systematically lack critical information about the precise reasons behind content moderation decisions. This prevents them from adjusting their behavior accordingly or effectively using the internal complaint mechanism, as guaranteed under Article 20 (1) DSA.

Finally, if the statements of reasons provided to users should differ from those submitted to the transparency database, this would in any case constitute a violation of Article 24 (5) DSA. The provision explicitly mandates a direct transmission of all content moderation decisions and their statements of reasons, including all the information required by Article 17 (3) DSA, ensuring full transparency. Only personal data is exempt from this obligation. However, this exemption does not warrant the systematic omission of any case-specific details. For instance, it is possible to specify which part of a post is being reprimanded, identify the specific violation (e.g. an insult), and provide an explanation of why a particular post qualifies as such a violation, all without disclosing any personal information.