



# NetzDG Transparency Report

January 2023

## 1. GENERAL OBSERVATIONS

General observations outlining the efforts undertaken by the provider of the social network to eliminate criminally punishable activity on the platforms.

Meta Platforms Ireland Limited, which provides the social network Facebook for users in Germany, aims to create a safe and trusted platform, where people can feel free to express themselves. But we are clear that we do not allow people to post content that is against the law or encourages criminal behavior. We also do not allow for bullying or harassment in any form. In order to achieve this balance, we take a multi-faceted approach to addressing potentially criminal or harmful activity on the Facebook platform.

First, we maintain a set of globally applicable Community Standards that define what is and isn't allowed on Facebook, and that in many instances tracks, to some degree, what is unlawful under German law. These standards apply to content worldwide and are integral to protecting both expression and personal safety on Facebook. Our Community Standards prohibit a wide range of objectionable or harmful content, including content that:

- Promotes violent and criminal behavior
- Threatens the safety of others
- Is considered hate speech
- Is considered graphic violence
- Is considered spam
- Is considered bullying or harassment

The [Community Standards](#) are created by a global team with a wide array of backgrounds, including those who have dedicated their careers to issues like child safety, hate speech, and terrorism. This team regularly seeks input from outside experts and organizations to help balance the different perspectives that exist on free expression and safety, and to better understand the impacts of our policies on different communities globally. Our reviewers enforce these standards using comprehensive guidelines, which ensure that decisions are as consistent as possible. A Community Standards report does not trigger a legal review by us – in accordance with



the user's Community Standards report, we review for violation of our Community Standards. We maintain separate reporting mechanisms for users to report content they believe violates the local law (described in further detail below).

Logged-in Facebook users can report Community Standards violations in a variety of ways, including through an option appearing with each piece of content. When content is reported through these tools, we review the reported content (automated or manually) to determine whether it violates our Community Standards and remove it if it does. In addition, we use technology to help detect potentially Community Standards violating content, before people see it and report it to us (see Section 2 for details on this).

Second, we provide an array of options for people to report content they believe violates local law. When something on Facebook is reported to us as violating German law but doesn't go against our Community Standards, we may block the content from being available in Germany. People in Germany may use various channels to report alleged violations of German law such as:

- [NetzDG complaint form](#) — This form allows people in Germany to report content they believe violates one or more of the German Criminal Code provisions set forth in NetzDG (more information on this form can be found in Section 3 of this transparency report).
- [Intellectual property reporting forms](#) — These forms allow rights owners and their authorized representatives to report content they believe violates copyright or trademark rights.
- [Defamation reporting form](#) — This form allows injured parties and certain authorized representatives to report content they believe to be defamatory. This may include content that is a false assertion of fact, leading to injury to reputation under the law. When people select Germany as the country for which they want to claim rights, they are provided an option to report through our NetzDG complaint form should they believe the content constitutes criminal insult, defamation or intentional defamation.
- [Legal removal request form](#) — This form allows individuals in European Union Member States to report content they believe violates their personal legal rights or applicable local laws. When people select Germany as the country for which they want to claim rights, they are, first, provided an option to report through our NetzDG complaint form. If they opt not to do so, this leads to access to the other forms mentioned above, as well as to a separate form for Right to Privacy/Erasure. The form also provides a means for people to report content they believe violates other laws not covered by the categories above.

## 2. PROCEDURES USED FOR THE AUTOMATED DETECTION OF CONTENT TO BE REMOVED OR BLOCKED

Type, principles of functioning and scope of any possibly used procedures for the automated detection of content to be removed or blocked, including general



information on training data used and on the provider's examination of the results of these procedures, as well as information on the extent to which scientific and research communities are being supported in evaluating these procedures, and the extent to which they have been granted access to information of the provider for this purpose.

## **2A. TYPE, FUNCTIONING AND SCOPE FOR AUTOMATED DETECTION OF CONTENT TO BE REMOVED OR BLOCKED**

To enforce our Community Standards, we employ a combination of human review and technology (see already Section 1). Every day, we remove millions of violating pieces of content and accounts on Facebook and Instagram. In most of the cases, this happens automatically, with technology such as artificial intelligence working behind the scenes to detect and remove Community Standards violating content. In other cases, our technology selects content for human review. Our review teams review a blend of user reports described in Section 1 and posts surfaced by our artificial intelligence tools. Our technology also supports the review teams by prioritizing the most critical content to be reviewed, based on severity, virality and likelihood of a violation. Our review systems use technology to prioritize high-severity content with the potential for offline harm (e.g. posts related to terrorism and suicide) and viral content which is spreading quickly and has the potential to reach a large audience, in order to prevent as much harm as possible.

Our technology is set to identify violations of our Community Standards. It is not geared to directly find unlawful content within the meaning of section 1 subsection (3) NetzDG, even though there are, of course, overlaps between what is prohibited under our Community Standards and what constitutes unlawful content within the meaning of section 1 subsection (3) NetzDG (e.g. in the field of hate speech). There are three primary forms of technology used to detect Community Standard violations.

- First, we employ rate limits (speed limits) on how fast accounts can do actions on our platforms, including making posts, to prevent the usage of bots.
- Next, we have matching technology that identifies identical or near identical copies of URLs, text, images, audio and videos which we have previously identified as violating our Community Standards. This matching technology can work even if there are some minor modifications to the original content. When we match the content exactly or we determine it is near identical we will typically remove the content. Sometimes this technology is referred to as content hashing or content digital fingerprinting.
- Finally, we also use artificial intelligence in a narrow sense (i.e. machine learning and rules based systems - in the following: artificial intelligence, AI), in two ways: like with the matching technology, when confident enough that a post violates one of our Community Standards, the artificial intelligence will typically remove the content. We also use artificial intelligence to select the content for human review on the basis of severity, virality and likelihood of a violation. As with matching technology, artificial intelligence operates on URLs, text, images, audio and videos. Unlike technologies that can only match violations they've seen before, artificial intelligence has the



potential to identify certain violations it has never seen before.

First, we use our technology on content available to everyone in all areas of the Facebook platform for virtually all Community Standards violations. Second, as the user expectations change based on product area and audience size (e.g. groups, friends) we limit by severity the Community Standards policies enforced using automated means.

## **2B. TRAINING DATA USED**

Rate limits: We set our rate limiting thresholds by observing how people use the Facebook platform and then setting conservative thresholds that allow us to address the worst bot behavior while only infrequently affecting legitimate behavior.

Matching technology: Our lists of known violating content powering the matching technology are typically created after the same content has been labeled as violating by our human reviews multiple times.

Machine learning: The development of techniques used to train machine learning models is a fast moving area of study by industry and academia. Primarily, Meta uses two techniques to train its machine learning models.

The first technique is largely referred to as supervised learning. Meta's models for content moderation use variations on the same general technique for training these supervised models. Meta selects a statistically random sample of all content that users have viewed, which is the same method we use when we calculate our publicly reported prevalence of violations measurements, or a statistically random sample of all reports by our community. Human reviewers label the selected content as either benign or violating one or more of our Community Standards. As part of this process, the same content may be reviewed multiple times for quality control. We then combine these benign and violating examples as inputs into machine learning training algorithms. The output of these machine learning training algorithms is called a "model" often referred to as a "classifier". We can then use this classifier to determine if a post is likely to violate our Community Standards.

The second technique is referred to as self supervised learning. In this training technique, the machine learning model removes a word from a sentence and then attempts to see if it can predict the missing word. This is a recently developed technique that Meta uses in more limited contexts, primarily to train language machine learning models.

## **2C. EXAMINATION OF RESULTS**

To ensure and improve the quality, i.e. how accurate the technologies are in detecting Community Standards violations, there are ongoing quality evaluation processes in place. Meta uses overlapping techniques and systems for maintaining a high overall accuracy for our automated content moderation.



Prior to fully launching any new rate limit, matching technology, or artificial intelligence, we use the technology to only log what it would have deleted instead of deleting it. We then use human reviewers to determine the accuracy rate against real time content rather than just historical content as we did during technology's training. Technology is often able to achieve higher levels of accuracy than human reviewers.

After launching rate limits, matching technologies, or artificial intelligence, we monitor the volumes of removals and objections by the user who posted the content as well as the rate at which objections are granted. If any of the metrics we monitor are abnormal, our engineering teams investigate. For rate limits engineers reevaluate if the limit is preventing bot behavior. If an entry in our list of known Community Standards violating content has abnormal signals, we will re-review the entry to confirm it violates our Community Standards. If one of our artificial intelligence tools has abnormal signals, we will either send a sample of the artificial intelligence tool's recent results to human labeling to confirm the accuracy rate or deprecate the artificial intelligence tool.

In addition, many of our machine learning classifiers are automatically reassessed for accuracy after each human review inside of our severity weighted viewership prevalence reduction system (using a [multi armed bandit algorithm](#)). This classifier reassessment is an example of the general feedback loop between human review and technology. The content labeling decisions taken by human reviewers are used to train and refine our technology. As a part of this process the review teams manually label the policy guiding their decision, i.e. they mark the policy that the content, account or behavior violates. This helps to improve the quality of our artificial intelligence algorithms and our lists of known Community Standards violating content used by our matching technology.

Our automated removals avoid discrimination primarily through quality controls to assure adherence to our Community Standards. Our Community Standards avoid discrimination through processes such as stakeholder consultations also described in Section 1 and extensive analysis of our internal signals. Our internal signals include user research, large community surveys, and detailed analysis of what our community is reporting using the mechanisms described in Section 1. Our stakeholder consultations include active engagement with NGOs, governments, individual activists, and academia.

Beyond adherence to our Community Standards, we have a Responsible AI team. The team is a multidisciplinary team composed of ethicists, social and political scientists, policy experts, artificial intelligence researchers and engineers. The team's overall goal is to develop guidelines, tools, and processes to tackle issues of artificial intelligence responsibility and help ensure these systemic resources are widely available across Meta. This way, we want to address fairness and inclusion concerns associated with the deployment of artificial intelligence in Meta technologies. Finally, the independent [Oversight Board](#) can issue recommendations concerning the enforcement of our policies, which may also concern the enforcement through automated means.



## 2D. SUPPORT AND ACCESS FOR SCIENTIFIC AND RESEARCH COMMUNITIES

Meta believes openness and collaboration with the academic community will spur research and development, create new ways of detecting and preventing harmful content, and help keep people safe. We collaborate with these communities in various ways within the constraints of respecting the privacy of our individual users. We are open and transparent about our content moderation in three important ways: (1) the algorithms, (2) the individual results of those algorithms, and (3) aggregated results of all of our content moderation efforts.

Since our Facebook AI research (FAIR) lab (today: Meta AI) was founded in 2013, we have committed to an open science-based approach. Our research model revolves around publishing source code and methodologies, collaborating with other researchers across industry and academia, and creating open benchmarks and challenges. In addition, our affiliated researchers frequently publish the results of our source code and methodologies applied on our platforms. Some examples of our methodologies and technologies published in recent years, include, but are not limited to, the following:

- [XML-R](#), [Linformer](#), and [RoBERTa](#): We have open-sourced our models and code so the research community can evaluate our natural language, especially multilingual, understanding machine learning models.
- [Faiss](#), [PDQ](#) and [TMK+PDQF](#): We have published our research and released the code for three of our algorithms used for finding identical and near identical copies of known photo and video content.
- [CLARA](#): Confidence of Labels and Raters. We published our methodology and evaluation of boosting the accuracy of human reviewer labels.

Information access for the purpose of evaluating the procedures for the automated detection of content specifically on the basis of the new section 5a subsection (2) NetzDG was not provided during the reporting period.

Irrespective of that, Meta maintains [Facebook Open Research and Transparency \(FORT\)](#), a research and transparency initiative allowing controlled independent researcher access to datasets covering ads targeting, URL share activity, civic engagement and others. The URL share might potentially, to a certain degree, be used by researchers to identify potential gaps in our Community Standards and our enforcement, including automated enforcement.

We publish our aggregated enforcement numbers across various areas of the Facebook Community Standards and the Instagram Community Guidelines in our quarterly [Community Standards Enforcement Report \(CSER\)](#). In addition, for many violations, we publish measured viewership prevalence. Viewership prevalence is how many views of violating content we didn't prevent – either because we haven't caught the violations early enough or we missed them altogether on Facebook and Instagram. We publish more about prevalence measurement [here](#). But transparency is only helpful if the information we share is useful and accurate. In the context of the



Community Standards Enforcement Report, that means the metrics we report are based on sound methodology and accurately reflect what's happening on our platform. To this end, we worked with international experts in measurement, statistics, law, economics and governance to provide an independent, public assessment of whether the metrics we share in the Community Standards Enforcement Report provide accurate and useful measures of Meta's content moderation challenges and our work to address them. They broadly agreed that we are looking at the right metrics and provided some recommendations for improvement. You can read the full report [here](#).

In August of 2020, we also committed to undertaking and releasing an independent, third-party assessment of our Community Standards Enforcement Report and this year we delivered on that commitment by [publishing EY's independent findings](#).

### 3. COMPLAINT MECHANISMS/CRITERIA

Description of the mechanisms for submitting complaints about unlawful content, description of the criteria applied in deciding whether to remove or block unlawful content and description of the review procedure including the sequence of the review as to whether there is unlawful content or whether there is a violation of the contractual provisions between provider and user.

If someone believes content on Facebook is unlawful under one or more of the German Criminal Code provisions covered by NetzDG, they can report it by using Facebook's dedicated [NetzDG complaint form](#). This form has been in place for people in Germany since January 1, 2018, and is intended for complaints claiming violations of the German Criminal Code provisions listed in NetzDG.

The NetzDG complaint form can be accessed via a link available next to a piece of content on the Facebook platform.

So that we can properly evaluate the claim and comply with section 2 ("Reporting obligation") of NetzDG, people are prompted to provide the following information in our NetzDG complaint form:

- Complete contact information, including whether the reporter is a complaints body (*Beschwerdestelle*) under NetzDG and/or is reporting the content on behalf of a client
- Section(s) of the German Criminal Code alleged to have been violated by the reported content
- Why the reported statements or images are alleged to be unlawful under NetzDG
- A court order, if available

Since people reach the complaint form directly from a piece of content, they do not need to provide links to the reported content or contact information (if they are reporting while signed into their Facebook account). These fields are filled in





automatically. Also, within the form it is clearly highlighted which information is requested by us optionally.

Every user in Germany (registered users and non-registered users) can also access the NetzDG complaint form in other ways. One way to do so is on the Facebook homepage, to click the link titled "Impressum/Terms/NetzDG/UrhDaG", choose the section "Network Enforcement Act ("NetzDG") & Impressum" and click the link "submit a report". Another way is via Facebook's [NetzDG Help Center page](#), which is dedicated to helping individuals understand how to submit NetzDG complaints. On that page, people can reach our complaint form by clicking "Submit Report". This NetzDG Help Center page can be found via a link in the aforementioned section "Network Enforcement Act ("NetzDG") & Impressum" as well as in the [Facebook Help Center](#) under the "Policies and Reporting" tab, subsection "Network Enforcement Act ("NetzDG")" and is otherwise searchable via the Help Center search bar.

The NetzDG complaint form accessible in the ways mentioned in the preceding paragraph is aligned with the one available directly from the content. The only difference is that users who do not submit their NetzDG complaint via the complaint form available directly from the content are asked (in addition to the aforementioned points) to provide the link to the content in question on Facebook.

Once an individual has completed the NetzDG complaint form (via one of the ways described above) and clicks "send", their complaint automatically reaches us through an internal review tool – no additional action is required by the reporting party. We then take a two-step approach to reviewing content that is reported through the NetzDG complaint form. First, we review the reported content under our Community Standards. If it violates our Community Standards, we ensure that it is removed from the Facebook platform globally. Second, if the reported content does not violate our Community Standards, we review it for legality based on the information provided in the complaint. Specifically, we assess whether the reported content violates the relevant provisions of the German Criminal Code listed in NetzDG. If the reported content is deemed to be unlawful under NetzDG, we will disable access to that content in Germany (see further information on how we handle NetzDG complaints in Section 5). We also communicate with the reporting party to provide updates on their complaint, request additional information if necessary, and inform them about our decision once we have finished reviewing the complaint (see further information on this in Section 10).

## 4. COMPLAINT VOLUMES

Number of incoming complaints about unlawful content in the reporting period, broken down according to whether the complaints were submitted by complaints bodies (Beschwerdestelle) or by users; according to the reason for the complaint.

The following tables show the number of complaints submitted through Facebook's NetzDG complaint form between July 1, 2022 and December 31, 2022. In principle, only one piece of content can be reported per NetzDG complaint. In some cases, however,





users cite multiple pieces of content in a single NetzDG complaint (e.g. by mentioning multiple URLs in an attachment uploaded to the complaint). The numbers reflected in the tables below pertain to complaints submitted rather than unique pieces of content identified in the complaints. It is worth noting that in the period between July 1, 2022 and December 31, 2022, there were 125,195 NetzDG complaints identifying a total of 126,208 pieces of content. When the same piece of content was reported to us multiple times, we counted it as one piece of content.

The tables cover two different categories of numbers:

#### 4A. NETZDG COMPLAINTS BY REPORTER TYPE

- This Section breaks down the number of complaints according to whether they were submitted by complaints bodies or other individuals.

Table 1. NetzDG Complaints by Reporter Type

Complaints from Complaints Bodies	21,903
Complaints from Other Individuals	103,292
Total	125,195

#### 4B. NETZDG COMPLAINTS BY CRIMINAL CODE PROVISION(S) CITED

- This Section breaks down the number of complaints according to the provision(s) of the German Criminal Code cited by the reporting party.
- Please note that a NetzDG complaint may cite multiple reasons for illegality. Therefore, the sum of complaints listed in the rows per reporter type in the table below exceeds the total number of complaints submitted per reporter type as outlined in Section 4A.

Table 2. NetzDG Complaints by German Criminal Code Provision(s) cited

German Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Dissemination of propaganda material of unconstitutional and terrorist organizations (sec. 86)	6,243	15,228	21,471
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	4,886	8,582	13,468
Preparation of a serious violent offense endangering the state (sec. 89a)	4,925	6,614	11,539
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	4,425	6,252	10,677
Treasonous forgery (sec. 100a)	4,693	6,237	10,930



German Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Public incitement to crime (sec. 111)	5,129	11,866	16,995
Breach of the public peace by threatening to commit offenses (sec. 126)	4,883	8,965	13,848
Forming criminal organizations (sec. 129)	4,702	7,442	12,144
Forming terrorist organizations (sec. 129a)	4,327	5,076	9,403
Forming criminal and/or terrorist organizations abroad (sec. 129b)	4,368	5,048	9,416
Incitement to hatred (sec. 130)	8,607	25,985	34,592
Dissemination of depictions of violence (sec. 131)	6,115	12,247	18,362
Rewarding and approving of offenses (sec. 140)	4,515	11,268	15,783
Defamation of religions, religious and ideological associations (sec. 166)	6,474	16,840	23,314
Distribution, acquisition, and possession of child pornography content (sec. 184b)	4,614	5,256	9,870
Insult (sec. 185)	10,217	40,750	50,967
Defamation (sec. 186)	7,961	29,178	37,139
Intentional defamation (sec. 187)	7,645	22,668	30,313
Disparagement of the memory of the deceased (sec. 189)	4,670	7,237	11,907
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	7,154	11,605	18,759
Threatening the commission of a felony (sec. 241)	5,006	9,668	14,674
Forgery of data intended to provide proof (sec. 269)	5,780	11,016	16,796

## 5. ORGANIZATION, PERSONNEL RESOURCES, PERSONNEL EXPERTISE, TRAINING, AND SUPPORT

Organization, personnel resources, specialist and linguistic expertise in the units responsible for processing complaints, as well as training and support of the persons responsible for processing complaints.



## 5A. ORGANIZATION

NetzDG complaints are reviewed in two steps by teams of trained professionals and lawyers, who cover both the Facebook and Instagram platforms.

First, content reported via the Facebook NetzDG complaint form is reviewed by members of our Global Operations team. Our Global Operations team is a mix of full-time employees and personnel of companies we partner with. Each NetzDG complaint is reviewed by an individual member of this team through our contractual partnerships with Majorel in Berlin and Telus/CCC in Essen (Germany) to determine whether the reported content violates Facebook's Community Standards (as opposed to reviewing the content for potential unlawfulness, which as discussed below is handled by separate teams). If the content is found to violate the Community Standards, then the content is removed globally.

Second, all NetzDG complaints containing content that was not removed for violating Community Standards undergo a legal review process that can consist of multiple stages (see below) handled by our Scaled Regulatory Operations team.

This team is made up of two groups – a group of employees based out of Ireland and Sunnyvale, California (and one employee working remotely from Austin) and a group of contractors based out of Dublin, Ireland and Austin, Texas. NetzDG complaints containing content that was not removed for violating Community Standards are first reviewed by one of the contractors. The purpose of this review stage is to ensure that manifestly unlawful content is blocked within 24 hours. Each complaint is reviewed by an individual member of that team, who is tasked with identifying and blocking manifestly unlawful content, and corresponding with the reporting party, including when the complaint lacks critical context. All of this is done in accordance with guidance developed by our in-house lawyers and external legal counsel. Should the complaint require more granular investigation, it is enqueued for review by one of the Scaled Regulatory Operations team employees. That individual will then carefully review the complaint and take appropriate action in instances where illegality or legality can be determined on the basis of guidance prepared for the team by our in-house lawyers and external legal counsel.

Where the legality of reported content is still unclear, the complaint is then escalated to our in-house lawyers for review. In particularly complex cases, our in-house lawyers may obtain a legal assessment from outside counsel in Germany.

Once a decision on how to handle the content is made, the Scaled Regulatory Operations team handles any advised content actions and corresponds directly with the reporting party and – if content is blocked – the reported user.

There are open lines of communication between content reviewers at each stage of the review process.

In the Global Operations team the Majorel and Telus/CCC teams work closely with the Global Operations team employees in Dublin. The Majorel and Telus/CCC teams receive training, additional guidance, and Community Standards expertise on NetzDG



cases as needed from specially trained contractors, who in turn are trained by different Meta teams.

The group of Scaled Regulatory Operations vendor teams work closely with our Scaled Regulatory Operations team employees, who provide training, guidance, and assistance on challenging or unique complaints. A similar line of communication is also open between the Scaled Regulatory Operations employees and a team of our in-house lawyers. These two teams meet multiple times a week and maintain open lines of communication to discuss legally complex NetzDG complaints. Every month, employees from the Scaled Regulatory Operations team send out an internal update to a cross-functional team on complaint volumes, trends, questions that arose while handling the complaints, and training of our teams.

## 5B. PERSONNEL RESOURCES AND EXPERTISE

As of December 31, 2022, 178 individuals spread across three teams are eligible and have been trained accordingly (as further set out in Section 5C) to process NetzDG complaints. These individuals also engage in work outside of NetzDG complaints, which allows for flexible staffing. When complaints volumes are low, only a subset of these individuals process NetzDG complaints. When volumes increase, additional trained members of the teams can be utilized to process complaints as well. The personnel resources and expertise of each team are as follows:

- **Global Operations team:** As of December 31, 2022, there were 130 reviewers on this team eligible to handle NetzDG complaints. Prior to handling NetzDG complaints, all of these individuals were required to display proven operational efficiency over the course of at least 3-6 months on other types of content takedown requests. In addition, all of these individuals are fluent in German, and have received NetzDG training (as further discussed in Section 5C below). As noted above, these individuals also engage in other queues focused on content moderation outside of NetzDG to help balance their workload depending on NetzDG complaints volumes.
- **Scaled Regulatory Operations team:** As of December 31, 2022, there were a total of 46 individuals eligible to handle NetzDG complaints on the Scaled Regulatory Operations team. 31 of these individuals were contractors, and 15 were employees. The 31 contractors are all fluent in German and have received NetzDG training (as further discussed in Section 5C below). The Scaled Regulatory Operations team employees eligible to handle NetzDG complaints are fluent in a wide variety of languages, including German, English, French, and Turkish, and can rely on other members of the team for expertise in other languages, such as Bosnian, Polish, Spanish, Russian, and Dutch. Complaints are generally reviewed by members of the team that are fluent in both German and English, with occasional exceptions for complaints submitted in other languages or where the reported content does not require German language expertise. The 15 employees have varied backgrounds to account for both the legal and operational complexities of NetzDG complaints. Specifically, 5 have law degrees, and 14 had operational experience in other roles before joining the team. All of these individuals have received NetzDG training (as further discussed in Section 5C below).



- **Legal:** As of December 31, 2022, there were 2 in-house lawyers involved in handling NetzDG-related complaints (among other work). These 2 lawyers are specialists for the assessment of potentially problematic content and have extensive experience regarding handling legal questions concerning takedown requests. These lawyers work closely with the Scaled Regulatory Operations team employees and correspond regularly with German external legal counsel who provide advice on specific NetzDG complaints.

## 5C. TRAINING AND SUPPORT

The teams which handle NetzDG complaints receive distinct types of training based on the nature of their respective work.

As Global Operations team members who review NetzDG complaints on Facebook only review these complaints for violations of Community Standards, their training is focused on developing operational skills and expertise in the implementation of Community Standards (rather than training to assess legality of content). They undergo several weeks of training in content review under Facebook's Community Standards. Importantly, as noted above in Section 5B, NetzDG complaints are only handled by a select group of Global Operations team members who reach a certain level of tenure on the team and demonstrate consistently high operational proficiency prior to being considered for the work. Once selected, those individuals then receive training that provides background information on NetzDG to help contextualize their work.

Members of the Scaled Regulatory Operations team review content for illegality and therefore receive greater levels of training on NetzDG and the German Criminal Code provisions that make up the definition of "unlawful content" under the law. Every member of this team who is eligible to handle NetzDG complaints receives several weeks of training focused heavily on operational proficiency and NetzDG ahead of them processing NetzDG complaints. The NetzDG-specific trainings include background material on the law, detailed breakdowns of every Criminal Code provision referenced in the law, and instruction on how to correspond with reporting parties and users, whose content was reported. The team members also receive refresher trainings at least once every half year. These refresher trainings provide reminders and updates on operational best practices, as well as breakdowns of common types of complaints and commonly cited Criminal Code provisions. The refresher training for the group of contractors is conducted by a Scaled Regulatory Operations team employee, and the refresher training for the group of Scaled Regulatory Operations team employees is conducted by our in-house lawyers.

There is a robust and diverse program to support our Global Operations and Scaled Regulatory Operations team members who review NetzDG complaints. This program currently offers e.g. the following services for team members:

- Psychological support
- One-on-one sessions with a full-time in-house psychologist



- Group therapy
- On-site counseling sessions to support emotional well-being

We are dedicated to providing our content reviewers with a high-quality, diverse support program and will continue to add new services and improve on existing services to accomplish this goal.

## 6. INDUSTRY ASSOCIATIONS

Membership of industry associations with an indication as to whether these industry associations have a complaints body.

Companies of the Meta group are a member of the following industry associations in Germany:

- Eco - Verband der Internetwirtschaft e.V. (which includes an internal complaints body)
- Bitkom - Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (which does not have an internal complaints body)
- BVDW - Bundesverband Digitale Wirtschaft e.V. (which does not have an internal complaints body)

## 7. EXTERNAL CONSULTATION

Number of complaints for which an external body was consulted in preparation for making the decision.

Between July 1, 2022 and December 31, 2022, we referred 29 NetzDG complaints to Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) for a decision on the legality of content identified in the complaints.

In 3 of the 125,195 complaints received between July 1, 2022 and December 31, 2022, we consulted external legal counsel (who we do not consider to be an “external body” but rather an extension of our legal team) to assist in making a decision on individual complaints. All of these consultations were with German outside counsel, who we work with for clarification on the law for complaints when needed.

## 8. REMOVAL/BLOCKING VOLUMES

Number of complaints in the reporting period that resulted in the deletion or blocking of the content at issue, by total number as well as broken down according to whether the complaints were submitted by complaints bodies or by users, according to the reason for the complaint, according to whether the case fell under section 3 subsection (2) number (3) letter (a), and if so, whether the complaint was forwarded to the user, which step of the review sequence according to number (3) led to the



removal or blocking, and whether the matter was referred to a recognized self-regulation institution pursuant to section 3 subsection (2) number (3) letter (b).

The following table depicts the number of times content was removed or blocked following complaints submitted through Facebook's NetzDG complaint form between July 1, 2022 and December 31, 2022. Please note the following about this table:

- This table breaks down the number of times a complaint led to the removal or blocking of content according to the provision(s) of the German Criminal Code cited by the reporting party.
- In principle, only one piece of content can be reported per NetzDG complaint. In some cases, however, users cite multiple pieces of content in a single NetzDG complaint (e.g. by mentioning multiple URLs in an attachment uploaded to the complaint). The numbers reflected in the tables below pertain to complaints submitted rather than unique pieces of content identified in the complaints. It is worth noting that in the period between July 1, 2022 and December 31, 2022, 17,242 NetzDG complaints resulted in the removal or blocking of content. This amounted to a total of 34,806 removed or blocked pieces of content.
- Individuals may cite multiple reasons for illegality in a single NetzDG complaint. If we took action on content pursuant to a complaint, it is listed in the table under every provision cited in the complaint. Therefore, the sum of removals/blockings listed in the table below exceeds the total number of complaints that led to the removal or blocking of content.

Table 3. Number of Complaints Resulting In Removal/Blocking

Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Dissemination of propaganda material of unconstitutional and terrorist organizations (sec. 86)	788	1,611	2,399
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	500	1,331	1,831
Preparation of a serious violent offense endangering the state (sec. 89a)	536	418	954
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	448	390	838
Treasonous forgery (sec. 100a)	400	330	730
Public incitement to crime (sec. 111)	644	1,270	1,914
Breach of the public peace by threatening to commit offenses (sec. 126)	456	735	1,191
Forming criminal organizations (sec. 129)	482	528	1,010
Forming terrorist organizations (sec. 129a)	558	261	819





Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Forming criminal and/or terrorist organizations abroad (sec. 129b)	594	236	830
Incitement to hatred (sec. 130)	902	3,167	4,069
Dissemination of depictions of violence (sec. 131)	821	1,313	2,134
Rewarding and approving of offenses (sec. 140)	350	1,267	1,617
Defamation of religions, religious and ideological associations (sec. 166)	530	1,670	2,200
Distribution, acquisition, and possession of child pornography content (sec. 184b)	434	656	1,090
Insult (sec. 185)	1,403	6,772	8,175
Defamation (sec. 186)	810	3,130	3,940
Intentional defamation (sec. 187)	816	2,072	2,888
Disparagement of the memory of the deceased (sec. 189)	326	483	809
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	735	1,004	1,739
Threatening the commission of a felony (sec. 241)	475	942	1,417
Forgery of data intended to provide proof (sec. 269)	437	817	1,254

Of the 34,806 removed or blocked pieces of content, 33,700 were removed globally for a violation of our Community Standards, and 1,106 did not violate our Community Standards, but were blocked in Germany due to a violation of a provision of the German Criminal Code listed in the NetzDG.

Number of complaints which fell under section 3 subsection (2) number (3) letter (a) and in which we reached out to the user who posted the reported content for additional facts: 8.

We referred 29 complaints to a recognized self-regulation institution.

## 9. REMOVAL/BLOCKING TURNAROUND TIMES

The number of complaints about unlawful content which, once received, have resulted in the removal or blocking of the unlawful content within 24 hours, within 48 hours, within a week or at a later point in time, respectively, additionally broken down according to complaints from complaints bodies and users and each broken down according to the reason for the complaint.



The following Section contains an overview of the time taken to remove or block content deemed unlawful or in violation of our Community Standards after receiving a NetzDG complaint, in terms of total numbers and broken down according to the reporter type (complaints body or user) and the reason for the complaint. This Section concerns complaints submitted through Facebook's NetzDG complaint form between July 1, 2022 and December 31, 2022.

Please note the following about this Section:

- As set forth under section 2 subsection (2) number (9), our removal time is divided into: (a) within 24 hours, (b) within 48 hours, (c) within a week, and (d) at a later date.
- The numbers reflected in this Section below pertain to complaints submitted rather than unique pieces of content identified in the complaints.
- The time periods refer to the time between when the complaint was submitted and the last action we took in response to the complaint. As an example, if a complaint identified two pieces of content and we actioned one piece within 24 hours and the other within 7 days, then that complaint would be listed in the table as taking action within 7 days.
- Individuals may cite multiple reasons for illegality in a single NetzDG complaint. If we took action on content pursuant to a complaint, it is listed in table 5 under every provision cited in the complaint. Therefore, the sum of removals/blockings listed in table 5 below exceeds the total number of complaints that led to the removal or blocking of content.

Turnaround Time for total number of NetzDG complaints that led to a block or a removal: Of the 17,242 complaints that led to a block or removal, our last block/removal action occurred within 24 hours 16,017 times, within 48 hours 430 times, within 7 days 741 times, and after 7 days 54 times.

In addition, table 4 shows the following details for each complaint where a piece of content was removed or blocked: (1) the type of reporter who submitted the complaint, and (2) the time it took to remove or block the reported content.

Table 4. Turnaround Time for Complaints with Removal/Blocking broken down per Reporter Type

	<b>24 Hours</b>	<b>48 Hours</b>	<b>7 Days</b>	<b>&gt; 7 Days</b>
Complaints from Complaints Bodies	2,138	39	76	6
Complaints from Other Individuals	13,879	391	665	48
<b>Total</b>	<b>16,017</b>	<b>430</b>	<b>741</b>	<b>54</b>

In addition, table 5 below shows the following details for each complaint where a piece of content was removed or blocked: (1) the type of reporter who submitted the complaint, (2) the time it took to remove or block the reported content, and (3) the provision(s) of the German Criminal Code cited by the reporter in the complaint.



Table 5. Turnaround Time for Complaints with Removal/Blocking broken down per Reporter Type and Reason

Criminal Code Provision	Reporter Type							
	Complaints from Complaints Bodies				Complaints from Other Individuals			
	24 Hours	48 Hours	7 Days	> 7 Days	24 Hours	48 Hours	7 Days	> 7 Days
Dissemination of propaganda material of unconstitutional and terrorist organizations (sec. 86)	752	14	21	1	1,422	68	114	7
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	470	15	15	0	1,100	92	128	11
Preparation of a serious violent offense endangering the state (sec. 89a)	509	14	13	0	394	12	12	0
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	425	12	11	0	366	14	9	1
Treasonous forgery (sec. 100a)	376	11	13	0	315	8	7	0
Public incitement to crime (sec. 111)	620	13	11	0	1,208	26	34	2
Breach of the public peace by threatening to commit offenses (sec. 126)	428	14	14	0	697	17	20	1
Forming criminal organizations (sec. 129)	459	12	11	0	495	12	19	2
Forming terrorist organizations (sec. 129a)	538	12	8	0	248	8	5	0
Forming criminal and/or terrorist organizations abroad (sec. 129b)	571	12	10	1	221	9	5	1
Incitement to hatred (sec. 130)	860	16	25	1	2,869	101	185	12
Dissemination of depictions of violence (sec. 131)	789	18	14	0	1,251	22	35	5
Rewarding and approving of offenses (sec. 140)	326	12	12	0	1,145	43	74	5
Defamation of religions, religious and ideological associations (sec. 166)	490	17	22	1	1,579	34	49	8
Distribution, acquisition, and possession of child pornography content (sec. 184b)	399	18	16	1	623	11	21	1
Insult (sec. 185)	1,333	25	42	3	6,335	154	266	17
Defamation (sec. 186)	757	15	34	4	2,835	98	177	20



Criminal Code Provision	Reporter Type							
	Complaints from Complaints Bodies				Complaints from Other Individuals			
	24 Hours	48 Hours	7 Days	> 7 Days	24 Hours	48 Hours	7 Days	> 7 Days
Intentional defamation (sec. 187)	766	15	32	3	1,858	78	122	14
Disparagement of the memory of the deceased (sec. 189)	292	13	18	3	437	16	26	4
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	684	22	27	2	954	18	29	3
Threatening the commission of a felony (sec. 241)	447	13	15	0	900	21	19	2
Forgery of data intended to provide proof (sec. 269)	410	12	14	1	787	15	14	1

## 10. CORRESPONDENCE

Measures to inform the person who submitted the complaint and the user for whom the content at issue was saved about the decision on the complaint.

We take a variety of measures to correspond with people who report content through the NetzDG complaint form, and with members of our community whose content is removed or blocked under NetzDG.

### 10A. CORRESPONDENCE WITH REPORTING PARTY

When someone submits a complaint, we correspond with them over email and through their Facebook Support Inbox (if the reporter is logged into their Facebook account and provides an email address associated with that account when submitting their complaint). Immediately after receiving a complaint, we send the reporting party an automatic response that informs them their complaint is being reviewed, and this correspondence provides them with a reference number and link to our NetzDG Help Center. If we are still looking into a complaint 24 hours after submission, we inform the reporting party that we are continuing to review the complaint. If at any point we need additional information from the reporting party to review the complaint, we will ask the reporting party for that information.

Once we decide what (if any) action to take on the reported content, we provide the reporting party with tailored correspondence that informs them of why we did or didn't take action on the reported content. This can take a variety of forms depending on the result of our review, but can broadly be categorized into the following:

- **Reported content is removed for violating Community Standards.** In this



circumstance, we inform the reporting party that the content was removed because it violated our Community Standards. We also provide links to our NetzDG Help Center page and to our Community Standards in case the reporting party would like additional information.

- **Reported content is blocked for violating a German Criminal Code provision covered by NetzDG.** In this circumstance we inform the reporting party that the content is no longer accessible in Germany, and state the specific Criminal Code provision(s) under which we determined the content was unlawful. We also provide a link to our NetzDG Help Center page in case the reporting party would like additional information.
- **Reported content does not violate Community Standards or a German Criminal Code provision covered by NetzDG.** In this circumstance we inform the reporting party that we have reviewed the complaint, but are not in a position to remove the content because we determined it was not unlawful.
- **Variety of actions taken on reported content.** Individuals may identify in some cases multiple pieces of content in a single NetzDG complaint. We review each piece of content individually and act accordingly. In order to limit the amount of correspondence we send the reporting party, we generally will review and act on each piece of reported content before informing them of our decisions. If we take different actions on content identified in a complaint (e.g., we block one piece of content and do not take any action on another piece of content), we will provide the reporting party with specific information on what we did, using a hybrid of the responses explained above.

We received a few complaints that used email addresses that almost certainly did not belong to the reporting party as contact information. To protect the true owners of those email accounts from spam, we reviewed the complaints and actioned the reported content according to the result of our review, but did not send notifications about the results of the review to the email addresses.

In addition to the information we provide in response to specific complaints, individuals can also find comprehensive educational information about NetzDG in our [NetzDG Help Center](#). The NetzDG Help Center contains information about (1) the type of content individuals can report under NetzDG, (2) what happens after an individual submits a NetzDG complaint, (3) how to report content that an individual believes is subject to NetzDG, (4) what information to include in a NetzDG complaint, (5) in which cases an individual can turn to the Zustellungsbevollmächtigter under section 5 subsection (1) NetzDG, (6) the difference between NetzDG and Facebook's Community Standards, and (7) where individuals can find Facebook's NetzDG Transparency Reports.

## 10B. CORRESPONDENCE WITH POSTING USER

We also correspond with users whose content is removed or blocked following a NetzDG complaint. If we remove content for violating our Community Standards, we inform the posting user of our action through the Facebook Support Inbox and let them



know their content violated our Community Standards. If we block content for violating a German Criminal Code provision covered by NetzDG, we inform the posting party via email of our action and of the specific German Criminal Code provision the content violated. If a complaint results in no action being taken on a user's content, we don't notify the user.

## 11. APPEALS PURSUANT TO SECTION 3B SUBSECTION (1) SENTENCE (2)

Number of incoming appeals pursuant to section 3b subsection (1) sentence (2) in the reporting period, by total number as well as broken down according to appeals by the person who submitted the complaint and by users for whom the reported content was saved, in each case with information on the number of cases in which the appeal was remedied.

We offer the possibility to appeal the decisions we take on the removal or blocking of access to content following a complaint about unlawful content. A user who reported content through the NetzDG complaint form (reporting party) and which has not been removed or blocked can appeal the decision taken on the initial complaint. If content is removed for an infringement of our Community Standards following a NetzDG complaint, the user who posted this content (posting user) can generally ask for a review of that decision, except in certain cases, for example in connection with terrorism or child sexual abuse and exploitation. If content is blocked following a decision that the content violates the relevant provisions of the German Criminal Code listed in NetzDG, the posting user can always appeal that decision.

The following table shows the number of appeals received following a complaint submitted through Facebook's NetzDG complaint form between July 1, 2022 and December 31, 2022 as well as in how many cases the appeal was remedied. In principle, only one piece of content can be reported per NetzDG complaint, resulting in one piece of content being subject to each appeal. In some cases, however, reporting parties cite multiple pieces of content in a single NetzDG complaint (e.g. by mentioning multiple URLs in an attachment uploaded to the complaint). In these cases, the reporting party can only appeal the decision we took on all cited pieces of content and we will count this appeal as one incoming appeal. If we remedy our initial decision on all or parts of the mentioned URLs, this will be counted as one remedied appeal.

Table 6. NetzDG Appeals by Applicant

	Number of incoming appeals	Number of appeals remedied
Appeals by reporting party	14,614	1,584
Appeals by posting user	6,989	1,589
Total	21,603	3,173



## 12. APPEALS PURSUANT TO SECTION 3B SUBSECTION (3) SENTENCE (1)

Number of appeals received in the reporting period pursuant to section 3b subsection (3) sentence (1), in each case with information on the number of cases in which the provider refrained from a reassessment pursuant to section 3b subsection (3) sentence (3) and the number of cases in which the appeal was remedied.

Usually, users who report content other than through a NetzDG complaint (reporting party) and which has not been removed or blocked in response, and users whose content is removed or blocked as a consequence of such report or without a report (posting user) can also inform us that they do not agree with the decision we took to remove or block the content.

Between July 1, 2022 and December 31, 2022 for 1,117,128 individual pieces of content, we received objections in relation to decisions we took to remove or block content with a nexus to Germany that were not based on a NetzDG complaint. Regarding 155,674 individual pieces of content we revised our decision following the objection.

## 13. ACCESS TO INFORMATION FOR SCIENTIFIC AND RESEARCH COMMUNITIES

Information on whether and to what extent scientific and research communities were granted access to the provider's information in the reporting period in order to allow them an anonymized evaluation to what extent a) removed or blocked unlawful content is linked to characteristics within the meaning of section 1 of the General Equal Treatment Act [Allgemeines Gleichbehandlungsgesetz] of August 14, 2006 (Federal Law Gazette I p. 1897), last amended by Article 8 of the Act of April 3, 2013 (Federal Law Gazette I p. 610), in its uptodate version, b) the dissemination of unlawful content results in specifically affecting certain user groups, and c) the dissemination is based on organized structures or concerted practices.

Through our Data Transparency efforts, Meta supports qualified academics by sharing Facebook and Instagram data (in protected form, i.e. by means of identification and privacy risk mitigation through a combination of programmatic (access criteria, contractual terms) and technical controls) to support the study of key social issues. Our data sharing efforts have facilitated the study of e.g. content containing mis/disinformation, which infringes our Community Standards, as well as other key social issues. Examples of our Data Transparency efforts are noted below. More information can be found via the [Meta Transparency Center](#).

- **IO Research Archive:** We created the IO (Influence Operations) Research Archive to facilitate information sharing with independent researchers and academics studying influence operations about networks Meta has removed for engaging in Coordinated Inauthentic Behavior (CIB), which we define as "coordinated efforts to manipulate public debate for a strategic goal, in which fake accounts are central to





the operation.” The IO Research Archive allows researchers to study public data about CIB networks on Facebook and Instagram that Meta teams have identified and disrupted. In late 2020, Meta launched a beta archive with a small group of researchers who study and counter-influence CIB operations. This beta group included the Stanford Internet Observatory, Atlantic Council’s Digital Forensic Lab, Graphika, Cardiff University, and the Australian Strategic Policy Institute. As of December 31, 2022, Meta has shared data from over 100 CIB networks that were removed since June 2020, to facilitate independent analysis and study of these operations, their tactics, techniques, and procedures.

- **CrowdTangle:** CrowdTangle is a content discovery and social monitoring platform that provides access to a small subset of public data on Facebook and Instagram. On December 31, 2022, there were around 144 registered active German accounts, including accounts by universities, with access to CrowdTangle. Researchers use CrowdTangle to study a variety of key topics of social interest, including misinformation, elections, Covid-19, and racial justice.
- **Researcher API (beta):** On December 31, 2022, 13 researchers worldwide had beta access to the Researcher API. The Researcher API is a beta tool that provides a closed group of researchers access to certain public data on Facebook in near-real time as well as historic public data. Researchers use the tool to study social issues such as misinformation, elections, and Covid-19. We intend to broaden access to this platform over time.
- **Reports:** We also publish [Community Standards Enforcement Reports](#) and the [Widely Viewed Content Report](#). We publish on the Meta Transparency Center the Community Standards Enforcement Report on a quarterly basis to more effectively track our progress and demonstrate our continued commitment to transparency and supporting research using Facebook and Instagram data. We publish the Widely Viewed Content Report which aims to provide more transparency and context about what people are seeing on Facebook by sharing the most-viewed domains, links, pages and posts for a given quarter on Newsfeed in the United States.

In addition to the data we make available, as detailed in the Transparency Report for the previous reporting period 1 January to 30 June 2022 (hereinafter "H1/2022 Transparency Report"), we have also been active participants in the European Digital Media Observatory’s Working Group on Researcher Access to Platform Data. We have committed significant resources to this work because we consider it important to find the correct balance between individuals’ privacy and the social benefit of research. This working group issued in 2022 a [report](#) on how to achieve this balance, within the legal obligations of the General Data Protection Directive (GDPR).

## 14. PROTECTION MEASURES

Other measures by the provider to protect and support those affected by unlawful content.



In addition to the measures already described in this report (such as the measures we use to detect content that infringes our Community Standards and easily accessible reporting channels), Meta takes a number of measures to protect and support those affected by harmful content - including unlawful content - distributed online.

Companies of the Meta group are working with a number of organizations and institutions which operate in the field of combating illegal content and promoting safety online. These include e.g. FSM (Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e. V.) as well as DsiN (Deutschland sicher im Netz e. V.). Companies of the Meta group maintain a close working relationship with [Jugendschutz.net](https://www.jugendschutz.net). We also partner with the youth safety organization [Troubled Desire](https://www.troubled-desire.org). A key goal of the program is to prevent child sexual abuse and the use of child abuse images.

Furthermore, we report all apparent instances of child sexual exploitation appearing on our site from anywhere in the world to the National Center for Missing and Exploited Children (NCMEC), including content drawn to our attention by government requests. NCMEC coordinates with law enforcement authorities from around the world.

Furthermore, companies of the Meta Group closely cooperate with law enforcement, who may use the dedicated Law Enforcement Online Request System (LEORS) for the submission, tracking and processing of requests, including emergency request response.

Meta also offers comprehensive information and practical assistance to support affected people, which include:

- [Bullying Prevention Hub](#), a resource for teens, parents and educators seeking support and help for issues related to bullying and other conflicts.
- [Stop Sextortion Safety Center](#), a resource for anyone seeking support and information related to threats to reveal intimate images to get you to do something you don't want to do.
- [Educational Resources](#) including Safety Tools with tips to stay safe in interactions with sharing, friending and reporting, securing your account, and protecting your information with specific sections for the safety of [children](#), [women](#), and the [LGBTQ+-Community](#). Through [Get Digital!](#), our digital and well-being resource, we provide lesson plans, conversation starters and other resources to help young people become empowered in a digital world.

We strongly advocate for women safety online and to protect those women who become the target of attacks on our platforms. In this regard, we have previously launched the program "Starke Frauen, starke Politik" (strong women, strong politics), which we designed together with the counterspeech Initiative #Iamhere (#ichbinhier) as well as a specialized psychologist to provide support for women active in politics as well as their teams, activists, and other politically active women. Together with our partners, it e.g. offers that members of the program can book 1:1 coaching sessions or order information material on e.g. content moderation.

Meta also formally launched the [Courage Against Hate report](#) (CAH Report), in partnership with the European Commission's Directorate-General for Justice &



Consumers. Courage Against Hate (CAH) is an initiative brought together by Meta (at the time still “Facebook”) for the purpose of sparking cross-sector, pan-European dialogue and action to combat hate speech and extremism. The [CAH Report](#), published on 13 July 2021, brings together four research organisations (the [Centre for the Analysis of the Radical Right, CARR](#); [HOPE Not Hate](#); the [Jena Institute for Democracy and Civil Society, Institut für Demokratie und Zivilgesellschaft, IDZ-Jena](#); and the [Swedish Defence Research Agency/Uppsala University](#)) and eight practitioner NGOs/companies ([lamhere International](#); [Galop UK](#); [Moonshot](#); the [Media Diversity Institute/Textgain](#); the [Institute for Strategic Dialogue, ISD](#); [Zivilcourage & Anti-Rassismus-Arbeit, ZARA](#)) with the aim of helping to develop a mapping of both trends in hate speech and extremism as well as effective programs and initiatives countering hate in Europe. The report prompts a multi-disciplinary conversation around what policies, further analysis and programs are needed for the fight against hate, extremism and terrorism to be truly effective. This collection of articles unites European academic analysis with practitioners who are actively working on countering extremism within civil society, and demonstrates our continued commitment to tackling these issues.

We have also facilitated the launch of a website created by the UK NGO “The Revenge Porn Helpline (RPH)”, which is available in a range of countries, including Germany (see already the H1/2022 Transparency Report). RPH is an organization that supports adult victims (i.e. individuals over the age of 18) of intimate image abuse and is a leading NGO in this space. The platform ([StopNCII.org](#)) is the first of its kind and RPH has designed it with the specific goal of empowering victims, by giving them a private and secure tool to proactively stop the proliferation of their non-consensual intimate image (NCII) online. It uses technology that hashes images and videos directly on the platform user’s device, so that victims are not required to share the original content with either the NGO or the StopNCII platform. Once the hashes are shared with Facebook and Instagram, we use technology to detect identical or similar content as it is being uploaded to the platform and action that content accordingly.

## 15. SUMMARY: COMPLAINTS & APPEALS

A summary with a summary table contrasting the total number of complaints received about unlawful content, the percentage of content removed or blocked as a result of these complaints, the number of appeals pursuant to section 3b subsection (1) sentence (2) and pursuant to section 3b subsection (3) sentence (1) and in each case the percentage of decisions modified as a result of these appeals with the corresponding figures for the two previous reporting periods, together with an explanation of significant differences and their possible reasons.



Table 7. Number of Complaints, Appeals/Objections and Respective Outcome

		July to December 2022 ("H2/2022")	January to June 2022 ("H1/2022")	July to December 2021 ("H2/2021")
Complaints	Number	125,195	170,233	115,085
	Complaints with Removals/Blockings	13.8%	14.2%	15.5%
Appeals pursuant to section 3b subsection (1) sentence (2) NetzDG (Section 11)	Number	21,603	11,093	N/A
	Appeals remedied as a result	14.7%	13.2%	N/A
Pieces of content for which objections were filed following a content decision taken without a NetzDG complaint (Section 12)	Number	1,117,128	1,493,668	N/A
	Objections remedied as a result	13.9%	12.0%	N/A

As showcased by the table, the most significant difference is that there is an increase in the number of complaints from H2/2021 to H2/2022, whereas H1/2022 saw even higher volumes. The increase from H2/2021 to H2/2022 is in line with the consistent increase of complaints visible also from transparency reports previous to the reporting periods showcased in this table. This consistent increase can be explained by various factors, such as the users getting more used to the NetzDG reporting options and the addition of section 189 GCC as provision to be reportable under the NetzDG as of February 1, 2022. The difference to the volume in H1/2022 can also be explained due to various factors such as current events triggering higher complaint volumes. The increase in the numbers of appeals pursuant to section 3b subsection (1) sentence (2) NetzDG (Section 11) from H1/2022 to H2/2022 can be explained by various factors, such as the users getting more used to the appeal option. Please note that no comparison can be made to the reporting period July to December 2021 in view of numbers of appeals/objections, as these are being reported on for the first time since the reporting period January to June 2022 as per section 6 subsection (3) NetzDG. We further note that we count (initial) complaints, for which the initial decision was to remove or block the content (see Section 8), separate from a potential appeal which may lead to an alteration of the initial decision. This means that the number of complaints that we report here and in Section 8 will not change if following an appeal the initial decision was altered.

## 16. TERMS & CONDITIONS

Explanation of the provisions in the provider's general terms and conditions on the permissibility of distributing content on the social network which are used by the provider for contracts with consumers.



We use a strategy called “remove, reduce, inform”. This includes that we remove content that violates our policies when it is reported to us or identified by our systems as problematic and that we reduce the distribution of certain content, or inform people by providing additional context so they can choose what to click, read, or share. To help with this strategy, we have introduced policies which, depending on the user group and type of use, may govern the distribution of content on Facebook. For consumers, the following policies are of central importance:

- [Terms of Service](#)
- [Community Standards](#)

Furthermore, depending on which features are used on Facebook, the following policies may also be relevant:

- [Facebook Pages, Groups and Events Policies](#)
- [Music Guidelines](#)
- [Live Policies](#)

In addition, there are a number of other rules which may also result in restrictions on what content is allowed but which are primarily aimed at commercial users and companies, or relate to the sale of goods (e.g. [Commerce Policies](#), [Advertising Policies](#)).

In the following, we will provide an overview of the Terms of Service and the Community Standards as well as the Facebook Pages, Groups and Events Policies, Music Guidelines and Live Policies.

Our Terms of Service, Community Standards, and other policies are updated from time to time. For example, since July 26, 2022, we have been using a revised version of the Terms of Service. This report is based on this version, as it was last valid in the relevant reporting period (July to December 2022) (see also H1/2022 Transparency Report for the previous version).

## 16A. TERMS OF SERVICE

Section 3.2 of the [Terms of Service](#) of Facebook prohibits actions, including the sharing of content, in the following cases:

- The conduct or content breaches the Terms of Service, the Community Standards, or any other terms and policies that apply to the use of our products.
- The conduct or content is unlawful, misleading, discriminatory or fraudulent.
- The user does not own the content or does not have the necessary right to share it.
- The conduct or content infringes someone else’s rights, including their intellectual property rights.

These basic rules are specified in more detail, in particular, in the Community Standards but, depending on the type of use, also in other policies.



According to the Terms of Service, we can remove or block content or disable user accounts or restrict the use of certain features for a limited period of time if users breach the applicable provisions.

Details on disabling or restricting the use of user accounts are set out in the Community Standards.

## 16B. COMMUNITY STANDARDS

The [Community Standards](#) are the key set of rules for users outlining what content is allowed to be distributed on Facebook. The stated goal of the Community Standards is to create a place for expression and give people a voice. We want users to be able to talk openly about the issues that matter to them, even if some may disagree or find content objectionable. In some cases, we allow content – which would otherwise go against our standards – if it is newsworthy and if its publication is in the public interest. We do this only after weighing the public interest value against the risk of harm, and we look to international human rights standards to make these judgments.

Our commitment to freedom of expression is paramount, but we recognize the internet creates new and increased opportunities for abuse. For these reasons, when we limit expression, we do it in service of one or more of the following values:

- **Safety:** We remove content that could contribute to a risk of harm to the physical security of persons. Content that threatens users has the potential to intimidate, exclude or silence others. Such content is not allowed on Facebook.
- **Dignity:** We expect that users will respect the dignity of others and not harass or degrade others.
- **Privacy:** We are committed to protecting privacy and personal information.
- **Authenticity:** In order to make sure the content users see on Facebook is authentic, we want to prevent people from using Facebook to misrepresent who they are or what they are doing.

The Community Standards apply to every user, all around the world, and to all types of content. They are divided into individual sections or policies.

Each section of our Community Standards, i.e., each policy, starts with a description of the “Policy Rationale” underlying the relevant policy. This description is followed in most policies by detailed explanations which typically concretize and, also by use of examples, illustrate the Policy Rationale of the relevant policy. For this purpose,

- most policies contain a section marked with a red octagon with a white crossbar containing a description of content that is not allowed (“Do not post”),
- and in many cases a section marked with a yellow triangle with a white exclamation mark containing a description of content
  - that requires additional information or context to enforce,
  - that is allowed with a warning screen, or
  - that is allowed but can only be viewed by people over a certain age.



In order to further enhance transparency for users, the relevant sections/policies on our website include a change log that enables users to access previous versions of our policies in addition to the current version.

In addition, there is an interactive “User Experience” section on the website with each policy that allows users to get an overview of the user experience in specific situations (“Reporting”, “Post-report communication”, “Takedown experience” and “Warning screens”).

The individual policies are grouped under parent categories:

## **1) Violence and Criminal Behavior**

The “Violence and Criminal Behavior” category comprises the following policies:

- **Violence and Incitement**

The [“Violence and Incitement”](#) policy provides, in particular, in which cases threats of or calls for violence (including kidnapping), statements admitting to violence, statements of intent to commit violence, and statements advocating for violence are not allowed on Facebook. For example, content that asks or offers services for hire to kill others is prohibited.

- **Dangerous Individuals and Organizations**

The [“Dangerous Individuals and Organizations”](#) policy contains a detailed set of rules aimed at not providing a platform to certain dangerous organizations and related individuals. The policy contains, among other things, guidelines for taking action against the presence of organizations that organize or advocate for violence or engage in systematic criminal operations or repeatedly dehumanize people, as well as for the removal of content such as praise, substantive support, and representation of dangerous organizations, their activities or their leaders, founders or prominent members, or for the removal of content that praises, substantively supports, or represents certain events such as terrorist attacks or hate events, their perpetrator(s) or hate ideologies.

- **Coordinating Harm and Promoting Crime**

The [“Coordinating Harm and Promoting Crime”](#) policy prohibits certain forms of facilitating, organizing, promoting or admitting to certain criminal or harmful activities targeted at people, businesses, property or animals. At the same time, the policy also protects the integrity of elections from fraud or intimidation attempts, for example.

- **Restricted Goods and Services**

The [“Restricted Goods and Services”](#) policy restricts or prohibits the distribution of content in connection with certain transactions or goods (e.g., non-medical drugs, pharmaceutical drugs, firearms, endangered species) for reasons such as safety.

- **Fraud and Deception**

The [“Fraud and Deception”](#) policy is primarily aimed at preventing fraudulent activity by removing content that, for example, purposefully deceives, willfully





misrepresents or otherwise defrauds or exploits others for money or property or coordinates or promotes these activities using Facebook.

## 2) Safety

The “Safety” category comprises the following policies:

- **Suicide and Self-Injury**

The [“Suicide and Self-Injury”](#) policy contains differentiated rules aimed at preventing the promotion of suicide or self-injury (including eating disorders). For example, content that identifies or negatively targets victims or survivors of self-injury, suicide or suicide attempts, as well as real time depictions of suicide or self-injury, is prohibited.

- **Child Sexual Exploitation, Abuse and Nudity**

The [“Child Sexual Exploitation, Abuse and Nudity”](#) policy contains, in particular, detailed rules prohibiting content that sexually abuses, exploits or endangers children or depicts child nudity. In addition, it also prohibits, for example, imagery that depicts (non-sexual) abuse of children, as well as content that praises, supports, promotes, advocates for, provides instructions for or encourages participation in (non-sexual) child abuse.

- **Adult Sexual Exploitation**

The [“Adult Sexual Exploitation”](#) policy provides, among other things, that content which depicts, threatens or promotes sexual violence, sexual assault, or sexual exploitation will be removed. The same applies to content which displays, advocates for or coordinates sexual acts with non-consenting parties, or certain intimate images shared without the consent of the person(s) pictured or corresponding threats.

- **Bullying and Harassment**

The [“Bullying and Harassment”](#) policy contains differentiated rules aimed at preventing behavior and content that constitutes bullying or harassment as defined by the Community Standards.

- **Human Exploitation**

The [“Human Exploitation”](#) policy prohibits specifically defined content that facilitates or coordinates the exploitation and the abuse of humans, including human trafficking.

- **Privacy Violations**

The [“Privacy Violations”](#) policy contains rules prohibiting certain content containing personally identifiable information or other private information, including financial, residential, and medical information, as well as private information obtained from illegal sources.

## 3) Objectionable Content

The “Objectionable Content” category comprises the following policies:



- **Hate Speech**

The [“Hate Speech”](#) policy contains detailed rules on content that is, in particular, prohibited as hate speech, i.e. as a direct attack against people on the basis of what we call protected characteristics, such as ethnicity, national origin, disability, religious affiliation, sexual orientation, sex and serious disease.

- **Violent and Graphic Content**

The [“Violent and Graphic Content”](#) policy contains rules on certain graphic or violent content. According to these rules, for example, content that glorifies violence or celebrates the suffering or humiliation of others will be removed.

- **Adult Nudity and Sexual Activity**

The [“Adult Nudity and Sexual Activity”](#) policy restricts the display of nudity and sexual acts on Facebook. According to this policy, for example, imagery with certain depictions of nudity or more closely defined sexual content will generally be removed.

- **Sexual Solicitation**

The [“Sexual Solicitation”](#) policy contains restrictions on content that facilitates, encourages or coordinates sexual encounters or commercial sexual services between adults. The aim is to avoid facilitating transactions that may involve trafficking, coercion and non-consensual sexual acts. In addition, it also restricts the use of sexually-explicit language or the offer and ask of pornographic material.

#### 4) Integrity and Authenticity

The “Integrity and Authenticity” category comprises the following policies:

- **Account Integrity and Authentic Identity**

The [“Account Integrity and Authentic Identity”](#) policy provides that we may prohibit the use of the services or restrict or disable accounts and other entities (such as pages, groups, and events) under certain circumstances. These include, for example, severe or persistent violations of the Community Standards or the impersonation of others. More information can be found on the [“Disabling accounts”](#) page linked from the “Account Integrity and Authentic Identity” policy.

- **Spam**

The [“Spam”](#) policy provides, among other things, that it is prohibited to post, share, engage with content or create accounts, Groups, Pages, Events or other assets at very high frequencies. In addition, the policy also prohibits certain other abusive conduct to artificially increase viewership or distribute content en masse (for example, attempts to sell site privileges or product features, such as accounts, admin roles, permission to post, Pages, Groups, likes, etc., or certain deceptive or misleading conduct encouraging likes, shares, follows, clicks or the use of apps or websites).

- **Cybersecurity**

The [“Cybersecurity”](#) policy prohibits, in particular, attempts to compromise user accounts, profiles or other Facebook entities or to gather sensitive information or to engage in unauthorized access through the abuse of our platform, products, or services.



- **Inauthentic Behavior**

The [“Inauthentic Behavior”](#) policy does not allow people, in the interest of authenticity, to misrepresent themselves, use fake accounts, artificially boost the popularity of content or engage in behaviors designed to enable other violations under the Community Standards.

- **Misinformation**

The [“Misinformation”](#) policy formulates different categories of misinformation and provides guidance about how we treat such content, taking into account the aim to balance the values of expression, safety, dignity, authenticity, and privacy.

- **Memorialization**

The [“Memorialization”](#) policy contains rules for cases in which users pass away. It provides that certain content may be removed or changed upon request/notice from the legacy contact or family members (for example, content related to the death of the deceased in the case of victims of murder and suicide).

## 5) Respecting Intellectual Property

The “Respecting Intellectual Property” category comprises the following policy:

- **Intellectual Property**

The [“Intellectual Property”](#) policy provides that infringements of other people’s intellectual property rights, including copyrights and trademarks, are not tolerated on the platform. It states, in particular, that Facebook’s Terms of Service do not allow people to post content that violates someone else’s intellectual property rights, including copyright and trademark. Furthermore, it describes that, upon receipt of a report from a rights holder or an authorized representative, we will remove or restrict content that engages in copyright infringement or trademark infringement.

## 6) Content-Related Requests and Decisions

The “Content-Related Requests and Decisions” category comprises the following policies:

- **User Requests**

The [“User Requests”](#) policy provides for the removal of accounts upon the request of the users themselves, verified immediate family members (in the case of deceased users), and authorized representatives (in the case of incapacitated users).

- **Additional Protection of Minors**

The [“Additional Protection of Minors”](#) policy clarifies that we comply with certain requests for removal of accounts or other content in the interest of protecting minors (for example, requests for removal of an underage account).

## 16C. FACEBOOK PAGES, GROUPS AND EVENTS POLICIES

The [Facebook Pages, Groups and Events Policies](#) contain additional rules for users who create or administrate a Facebook page, group, or event, or who use Facebook to



communicate or administer a promotion. For example, pages, groups, and events must not falsely represent a brand, entity, or public figure or without our prior written permission promote online gambling. In addition, there are, for example, certain restrictions on the use of page names, cover photos and profile pictures for pages. The Community Standards also apply to pages, groups, and events.

## 16D. MUSIC GUIDELINES AND LIVE POLICIES

The [Music Guidelines](#) are supplemental terms which apply when users distribute content containing music on Meta products such as Facebook. They emphasize the responsibility of the users who post the content in question, stating, for example, that content with music can be removed if the use of the music is not properly authorized. The [Live Policies](#) contain additional restrictions on the distribution of content via the Facebook Live feature. The aim of these policies, as well, is less to prohibit certain content as such but rather to prevent users from being misled about the real-time nature of the broadcast or location of the broadcasting user and to preserve the real-time nature of the feature.

## 17. LEGAL COMPLIANCE OF TERMS & CONDITIONS

Description of the extent to which the agreement of the provisions under item 16 is consistent with the requirements of sections 307 to 309 of the German Civil Code and other law.

Insofar as we prohibit or restrict the distribution of certain content on Facebook under the Terms of Service and the policies described above, such prohibition or restriction is in compliance with sections 307 to 309 BGB and other applicable law.

The Community Standards and other relevant policies referenced in the Terms of Service under section 3.2 and section 5 are effectively incorporated into the contractual relationship between us and our users. Users expressly agree to the Terms of Service when registering on the platform. The Terms of Service and the other rules thereby become an effective part of the contract and regulate the permissible user behavior on Facebook.

The provisions regarding permissible content are not captured by the provisions of sections 308 and 309 BGB. Insofar as the provisions have any point of reference regarding the permissible content at all, first, section 308 number (3) BGB is not relevant here: Insofar as the contractual provisions entitle us to deactivate user accounts in cases of violations, ultimately releasing ourselves from our duty of performance, this right does not exist independently of objectively justified reasons that are indicated in the Terms of Service and the other policies. In addition, the usage agreement establishes a continuous contractual relationship to which section 308 number (3) BGB does not apply from the outset. Section 308 number (4) BGB is also not relevant: The blocking or removal of content does not in itself present a modification of or deviation from the promised performance within the meaning of the



provision, but is precisely intended to enforce the previously agreed communication standards. The same applies to the potential restriction or disabling of user accounts in the event of violations. Moreover, our duty of performance is not modified in the process, but is cancelled, as a last resort, in whole or in part. These measures are in any case reasonable for the affected users, also in view of our interests; therefore, even if the provision was applicable, these measures would be permissible. The Federal Court of Justice also examined section 308 number (5) BGB and section 138 BGB in connection with the update of our Terms of Service relevant to the legal dispute therein and did not establish any infringement of these provisions or of section 308 or section 309 BGB in general.

Nor are any other violations of the law discernible. On the contrary, the Community Standards and other rules relating to the permissible content serve in particular to protect the safety, dignity and data of users and to prevent deceptive and misleading conduct. In doing so, they also prevent the violation of laws and infringement against the rights of third parties.

It goes without saying that content that is unlawful or even criminally punishable is also prohibited on Facebook. However, even where our Terms of Service and policies declare content to be prohibited that does not or has not yet exceeded the threshold of criminal liability or other unlawfulness, such restrictions are permissible. Section 307 subsection (1) BGB does not preclude this. Pursuant to section 307 subsection (1) BGB, provisions in general terms and conditions are invalid if, contrary to the requirement of good faith, they unreasonably disadvantage the contractual partner (in this case: the users) of the provider of the general terms and conditions (in this case: us, i.e. Meta Platforms Ireland, Ltd.). An unreasonable disadvantage may also arise from the fact that the respective provision lacks clarity and comprehensibility.

The Federal Court of Justice has clarified that social network providers have the right to define the communication standards applicable to their platform themselves, and that they may also prohibit content which is not punishable or otherwise unlawful (FCJ, judgment of July 29, 2021, file no. III ZR 192/20, para. 71 - available [here](#); FCJ, judgment of July 29, 2021, file no. III ZR 179/20, para. 59 - available [here](#)). The Federal Court of Justice correctly considered the opposing position to be unconvincing and expressly rejected the view that the providers would be bound by the fundamental rights in the same way as if they were a state. The providers may, among other things, reserve the right to remove individual content or block access to the user account in the event of a violation of their communication standards. This is an expression of the providers' freedom of profession and freedom of expression and also corresponds to the interests of the other users, who are generally equally interested in a respectful culture of discourse and a safe communication space. This also applies in particular with regard to the protection of underage users. It goes without saying that due consideration must be given to the freedom of expression of the content creators and the significance of our services for the exercise of freedom of expression. The removal of content is therefore not to be carried out in an arbitrary manner. Pursuant to the rulings of the Federal Court of Justice, the removal of content requires an objective reason and must also be linked to objective, verifiable conditions.



These requirements are fulfilled. The Terms of Service provide for the removal of content in the event of violations of the Terms of Service, the Community Standards or other terms and policies that apply to the use of our products, in cases of unlawful, misleading, discriminatory or fraudulent content and conduct, in cases where the user does not have the necessary rights to share, and in cases of the violation of the rights of another persons.

These are objective, verifiable conditions, which in turn are based on objective reasons.

There is already no interest worthy of protection in maintaining unlawful, misleading, discriminatory or fraudulent content. Rather, the removal or blocking of such content is in the interest of all users and in our own interest. Similarly, providers are not required to accept the fact that the rights of another person are violated. In this case, as well, an objective reason and an overriding interest in the removal or blocking exist.

The removal of content that violates the rules specified in Facebook's Community Standards or other policies also constitutes a legitimate concern. The policies from each of the aforementioned categories of the Community Standards - Violence and Criminal Behavior, Safety, Objectionable Content, Integrity and Authenticity, Respecting Intellectual Property, Content-Related Requests and Decisions - serve legitimate purposes. These policies protect the interests of Facebook users, third parties, and ourselves, and may outweigh the interest in distributing certain content. The Community Standards are particularly concerned with preventing damage in the offline world, threats to the safety and dignity of users, and violations of the law. In addition, a communicative environment in which all users - including minors - can feel safe, is to be created and maintained.

It is important to bear in mind that the Community Standards are by no means formulated in a rigid and inflexible manner, but in many cases leave room for consideration of the diversity of the factual circumstances and nuances in the content, and also explicitly take into account legitimate concerns such as information, discussion and criticism. The necessity for a differentiated view is expressed various times in the relevant policies. One important concern of the Community Standards, for example, is to prevent the glorification or promotion of self-harm such as eating disorders. Nevertheless, we want Facebook to be a place where people can share their experiences in these areas and bring attention to these issues. This is why discussions about these important topics are, of course, allowed. Further, we remove certain depictions of violence because we do not condone the glorification of violence on Facebook. In contrast, depictions of violence are allowed, with certain restrictions, if they raise awareness about these issues. Similarly, we allow certain depictions of nudity, for example in the context of art and protest or in health-related contexts, although the Community Standards generally impose restrictions in this regard.

Insofar as the Federal Court of Justice has also ordered that the requirement of an objective reason for the removal of content and for the blocking of user accounts be secured by procedural provisions directly in the Terms of Service, namely by means of specific rights to information, consultation, and objection, we have made the necessary



clarifications and adjustments in the course of the most recent update of our Terms of Service (applicable since July 26, 2022).

The provisions in each case are also clear and understandable. Our Terms of Service, Community Standards, and other relevant policies are written in simple and generally understandable language.

The Terms of Service list the key rules in section 3.2 and contain links to the Community Standards as well as to the overview of other potentially relevant policies contained in section 5. The transparency of the central Community Standards is ensured, first, by the fact that the individual policies are arranged in thematically appropriate categories and, second, that each policy is given an appropriate heading. In addition, the individual policies of the Community Standards follow a fundamentally uniform and visually clear structure: The universally comprehensible description of the basic ideas behind each policy is followed by more detailed rules and explanations, the comprehensibility of which is often further enhanced by examples. This enables every user to form a reliable picture of which content is allowed on Facebook and which is not. The design in terms of the transparency of the Community Standards is further enhanced by the fact that any previous versions of the provisions, including the date, are made available to users in proximity to each individual policy.