NetzDG Transparency Report

July 2023 (reporting period 1 January to 30 June 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 Instagram 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 Instagram platform.

First, we maintain a set of globally applicable Community Guidelines that define what is and isn't allowed on Instagram, and that in many instances tracks, to some degree, what is unlawful under German law. These guidelines apply to content worldwide and are integral to protecting both expression and personal safety on Instagram. Our Community Guidelines 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 Guidelines 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 Guidelines report does not trigger a legal review by us – in accordance with the user's Community Guidelines report, we review for violation of our Community Guidelines. We maintain separate reporting mechanisms for users to report content they believe violates the local law (described in further detail below).

Logged-in Instagram users can report Community Guidelines 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 Guidelines and remove it if it does. In addition, we use technology to help detect potentially Community Guidelines 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 Instagram is reported to us as violating German law, but doesn't go against our Community Guidelines, 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 Guidelines, 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 Guidelines 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 Guidelines. 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 Guidelines 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 Guidelines 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 Guidelines. 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 Guidelines, 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 Instagram platform for virtually all Community Guidelines violations. Second, as the user expectations change based on product area and audience size (e.g. groups, friends) we limit by severity the Community Guidelines policies enforced using automated means.

2B. TRAINING DATA USED

Rate limits: We set our rate limiting thresholds by observing how people use the Instagram 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.

<u>Machine learning</u>: 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 Guidelines. 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 Guidelines.

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 Guidelines 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 Guidelines violating content has abnormal signals, we will re-review the entry to confirm it violates our Community Guidelines. 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 Guidelines violating content used by our matching technology.

Our automated removals avoid discrimination primarily through quality controls to assure adherence to our Community Guidelines. Our Community Guidelines 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 Guidelines, 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:

- XLM-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.

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 last 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 Instagram is unlawful under one or more of the German Criminal Code provisions covered by NetzDG, they can report it by using Instagram'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 or when accessing a piece of content on the Instagram 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. This field is filled in automatically. Also, within the form it is clearly highlighted which information is requested by us optionally.

Every user in Germany (registered users as well as non-registered users) can also access the NetzDG complaint form in another way: The complaint form can be reached via a special page in the Instagram Help Center - the [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 the button "Submit Report". The NetzDG Help Center page itself can be reached in various ways. The NetzDG Help Center page can be found in the [Instagram Help Center](#) under the "Terms and Policies" tab, subsection "Network Enforcement Act ("NetzDG")". Moreover, it is searchable via the Help Center search bar. Another way to reach the NetzDG Help Center page is on the Instagram homepage, to click the link titled "NetzDG/UrhDaG/ranking of content" and, then, in the section "Network Enforcement Act ("NetzDG")" click on "NetzDG Help Center".

The NetzDG complaint form accessible via the NetzDG Help Center page 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 Instagram.

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 Guidelines. If it violates our Community Guidelines, we ensure that it is removed from the Instagram platform globally. Second, if the reported content does not violate our Community Guidelines, 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 Instagram's NetzDG complaint form between January 1, 2023 and June 30, 2023. 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 January 1, 2023 and June 30, 2023, there were 99,825 NetzDG complaints identifying a total of 75,760 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	21,958
Complaints from Other Individuals	77,867
Total	99,825

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)	6,937	11,450	18,387
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	6,157	7,695	13,852
Preparation of a serious violent offense endangering the state (sec. 89a)	3,741	19,268	23,009
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	3,317	18,690	22,007
Treasonous forgery (sec. 100a)	3,424	19,603	23,027
Public incitement to crime (sec. 111)	3,655	23,745	27,400
Breach of the public peace by threatening to commit offenses (sec. 126)	3,594	21,194	24,788
Forming criminal organizations (sec. 129)	3,858	21,937	25,795

German Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Forming terrorist organizations (sec. 129a)	3,295	18,276	21,571
Forming criminal and/or terrorist organizations abroad (sec. 129b)	3,216	18,226	21,442
Incitement to hatred (sec. 130)	6,865	16,282	23,147
Dissemination of depictions of violence (sec. 131)	4,367	9,029	13,396
Rewarding and approving of offenses (sec. 140)	3,403	23,108	26,511
Defamation of religions, religious and ideological associations (sec. 166)	5,942	13,121	19,063
Distribution, acquisition, and possession of child pornography content (sec. 184b)	3,793	4,642	8,435
Insult (sec. 185)	8,177	18,546	26,723
Defamation (sec. 186)	6,863	14,635	21,498
Intentional defamation (sec. 187)	5,316	11,085	16,401
Disparagement of the memory of the deceased (sec. 189)	3,676	4,272	7,948
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	4,795	7,767	12,562
Threatening the commission of a felony (sec. 241)	3,939	22,732	26,671
Forgery of data intended to provide proof (sec. 269)	4,111	21,525	25,636

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 Instagram and Facebook platforms.

First, content reported via the Instagram 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 Instagram's Community Guidelines (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 Guidelines, then the content is removed globally.

Second, all NetzDG complaints containing content that was not removed for violating Community Guidelines 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 Guidelines 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 Guidelines 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 June 30, 2023, 208 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 June 30, 2023, there were 160 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 June 30, 2023, 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 June 30, 2023, 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 Instagram only review these complaints for violations of Community Guidelines, their training is focused on developing operational skills and expertise in the implementation of Community Guidelines (rather than training to assess legality of content). They undergo several weeks of training in content review under Instagram's Community Guidelines. 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 mental health professional
- 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.

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 January 1, 2023 and June 30, 2023, we referred 3 NetzDG complaints to Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) for a decision on the legality of content identified in the complaints.

In none of the 99,825 complaints received between January 1, 2023 and June 30, 2023, 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.

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 Instagram's NetzDG complaint form between January 1, 2023 and June 30, 2023. 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 January 1, 2023 and June 30, 2023, 8,872 NetzDG complaints resulted in the removal or blocking of content. This amounted to a total of 10,474 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)	454	1,078	1,532

Criminal Code Provision	Reporter Type		
	Complaints from Complaints Bodies	Complaints from Other Individuals	Total
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	395	1,165	1,560
Preparation of a serious violent offense endangering the state (sec. 89a)	273	890	1,163
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	184	790	974
Treasonous forgery (sec. 100a)	226	1,159	1,385
Public incitement to crime (sec. 111)	231	1,753	1,984
Breach of the public peace by threatening to commit offenses (sec. 126)	216	1,173	1,389
Forming criminal organizations (sec. 129)	273	1,260	1,533
Forming terrorist organizations (sec. 129a)	215	714	929
Forming criminal and/or terrorist organizations abroad (sec. 129b)	222	757	979
Incitement to hatred (sec. 130)	427	1,693	2,120
Dissemination of depictions of violence (sec. 131)	303	970	1,273
Rewarding and approving of offenses (sec. 140)	192	1,573	1,765
Defamation of religions, religious and ideological associations (sec. 166)	304	1,012	1,316
Distribution, acquisition, and possession of child pornography content (sec. 184b)	286	447	733
Insult (sec. 185)	514	1,327	1,841
Defamation (sec. 186)	396	996	1,392
Intentional defamation (sec. 187)	346	819	1,165
Disparagement of the memory of the deceased (sec. 189)	175	238	413
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	269	650	919
Threatening the commission of a felony (sec. 241)	247	1,207	1,454
Forgery of data intended to provide proof (sec. 269)	235	1,236	1,471

Of the 10,474 removed or blocked pieces of content, 9,940 were removed globally for a violation of our Community Guidelines, and 534 did not violate our Community Guidelines, 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: 0.

We referred 3 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 Guidelines 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 Instagram's NetzDG complaint form between January 1, 2023 and June 30, 2023.

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 8,872 complaints that led to a block or removal, our last block/removal action occurred within 24 hours 8,495 times, within 48 hours 170 times, within 7 days 196 times, and after 7 days 11 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

	24 Hours	48 Hours	7 Days	> 7 Days
Complaints from Complaints Bodies	1,361	16	23	3
Complaints from Other Individuals	7,134	154	173	8

	24 Hours	48 Hours	7 Days	> 7 Days
Total	8,495	170	196	11

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)	452	1	0	1	1,040	15	21	2
Using symbols of unconstitutional and terrorist organizations (sec. 86a)	389	2	4	0	1,113	17	34	1
Preparation of a serious violent offense endangering the state (sec. 89a)	271	1	1	0	849	32	9	0
Encouraging the commission of a serious violent offense endangering the state (sec. 91)	183	1	0	0	751	31	8	0
Treasonous forgery (sec. 100a)	223	1	2	0	1,123	33	3	0
Public incitement to crime (sec. 111)	230	1	0	0	1,682	46	25	0
Breach of the public peace by threatening to commit offenses (sec. 126)	214	1	1	0	1,103	44	26	0
Forming criminal organizations (sec. 129)	272	1	0	0	1,215	37	8	0
Forming terrorist organizations (sec. 129a)	214	1	0	0	681	32	1	0
Forming criminal and/or terrorist organizations abroad (sec. 129b)	221	1	0	0	724	31	2	0
Incitement to hatred (sec. 130)	411	8	7	1	1,596	37	59	1
Dissemination of depictions of violence (sec. 131)	294	5	3	1	909	23	36	2
Rewarding and approving of offenses (sec. 140)	190	2	0	0	1,488	50	35	0
Defamation of religions, religious and ideological associations (sec. 166)	294	3	6	1	965	17	29	1

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
Distribution, acquisition, and possession of child pornography content (sec. 184b)	279	1	5	1	404	21	21	1
Insult (sec. 185)	496	9	8	1	1,249	27	49	2
Defamation (sec. 186)	378	7	9	2	926	26	42	2
Intentional defamation (sec. 187)	335	4	6	1	758	21	39	1
Disparagement of the memory of the deceased (sec. 189)	168	5	2	0	232	2	4	0
Violation of intimate privacy or personality rights by taking photographs or other images (sec. 201a)	263	4	2	0	608	20	21	1
Threatening the commission of a felony (sec. 241)	245	2	0	0	1,131	45	31	0
Forgery of data intended to provide proof (sec. 269)	234	1	0	0	1,202	32	2	0

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. 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 Guidelines.** In this circumstance, we inform the reporting party that the content was removed because it violated our Community Guidelines. We also provide links to our NetzDG Help Center page and to our Community Guidelines 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 Guidelines 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 Instagram's Community Guidelines, and (7) where individuals can find Instagram'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 Guidelines, we inform the posting user of our action via a personal notification on Instagram and let them know their content violated our Community Guidelines. 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. 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 researchers 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 Guidelines, 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 June 30, 2023, 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 June 30, 2023, there were around 161 registered active German users, including 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.

- **Reports:** We publish on the Meta Transparency Center the Community Standards Enforcement Reports (which also contain information on Instagram) 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.

In addition to the data we make available, as detailed in the Transparency Report for the 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).

12. 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 Guidelines 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.). Companies of the Meta group maintain a close working relationship with Jugendschutz.net. We also partner with the program Troubled Desire. 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 extensive information and tools to provide practical support for those affected, including:

- Anti-Bullying Hub, with information around features and tools we've developed to protect against bullying, such as how to limit unwanted interactions, block people, report content, or limit the visibility of comments under one's own posts.

- Family Centre, which provides parents with tools and resources to help support their teenagers' online experience. This includes an Education Hub and a range of supervision tools for parents. Teenagers between the ages of 13 and 18 can ask parents to supervise their account so parents have peace of mind. At the same time, their teenagers' autonomy is respected. For example, parents can inform themselves there when their child reports an account, for example for bullying or hate speech.

- A Parent and Carer's Guide to Instagram, developed in cooperation with the EU initiatives Klicksafe.de and SaferInternet.at as well as the Stop Hate Speech project, which includes, for example, explanations of safety features on Instagram, such as the possibility to deactivate the

comment function, filter unpleasant comments and messages and set whether you can be tagged or mentioned in a comment, caption or story.

We also offer educational events in the form of workshops or discussions both in-person or virtually, like our SafeTeens event in Berlin in November 2022, where we invited policy makers, media outlets and parent creators to discuss media literacy and online youth safety and to educate about our safety tools.

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 (Iamhere 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. In the last year, Instagram and Facebook have been joined by Bumble, OnlyFans, Reddit and TikTok, providing and empowering victims with the ability to remove content across multiple apps.

We have also been thinking through how we can help support these efforts while reducing the barriers minors face to reaching out and seeking support – similar to work we have done to prevent the non-consensual sharing of intimate images for adults with StopNCII. To that end, in December 2022, NCMEC launched Take It Down. Powered by Meta funding, minors will now be able privately generate a hash of their images or videos directly on their own devices, without having to upload their content to the platform. NCMEC will securely house the hashes, at which point participating tech companies like ours can take those hashes and use them to remove and thwart the sharing of that content – helping to return power and control back to the victim. Initially available in English and Spanish and already used globally, it will be available in other languages in coming months.

13. SUMMARY: COMPLAINTS

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, 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 and Respective Outcome

		January to June 2023 ("H1/2023")	July to December 2022 ("H2/2022")	January to June 2022 ("H1/2022")
Complaints	Number	99,825	57,541	63,696
	Complaints with Removals/Blockings	8.9%	8.9%	9.5%

As showcased by the table, the most significant difference is that there is a significant increase in the number of complaints from H2/2022 to H1/2023. A factor for the increase in H1/2023 was a few reporters sending an unusually significant amount of complaints. The difference to the slightly higher volume in H1/2022 compared to H2/2022 can be explained due to various factors such as current events triggering higher complaint volumes. We 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.

14. 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 Instagram. For consumers, the following policies are of central importance:

• Terms of Use
• Community Guidelines

Furthermore, depending on how Instagram is used, the Music Guidelines may also be relevant; they emphasize the responsibility of the users who post content with music, stating, for example, that content with music can be removed if the use of the music is not properly authorized.

• 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. Branded Content Policies, Commerce Policies, Advertising Policies).

In the following, we will provide an overview of the Terms of Use and the Community Guidelines (as valid in the relevant reporting period).

14A. TERMS OF USE

The Terms of Use of Instagram prohibit certain types of conduct under the heading "Your Commitments", including conduct related to the sharing of content. Examples of prohibited conduct include the following:

- Users are not allowed to impersonate others or provide inaccurate information.
- Users are not allowed to do anything unlawful, misleading, or fraudulent or for an illegal or unauthorized purpose on Instagram.
- Users are not allowed to violate the Terms of Use or policies of Instagram, including in particular the Community Guidelines, Meta Platform Terms, or Music Guidelines. Users are not allowed to help or encourage others to engage in such violations. Users who post so-called branded content must comply with our Branded Content Policies.
- Users are not allowed to post someone else's private or confidential information without permission or do anything that violates someone else's rights, including intellectual property rights (e.g., copyright infringement, trademark infringement).

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

According to the Terms of Use, we can remove any content shared on our Service in the event of violations of our Terms of Use or policies, or if we are required to do so by law. We may completely or partially refuse to provide or stop providing the Service (including terminating or disabling access) if users violate our Terms of Use or policies, if users repeatedly infringe other people's intellectual property rights, or where we are required to do so by law.

14B. COMMUNITY GUIDELINES

The Community Guidelines are the key set of rules for users outlining what content is allowed to be distributed on Instagram. The goal of the Community Guidelines is to create and maintain a safe and open environment for everyone. We want Instagram to continue to be an authentic and safe place for inspiration and expression. In some cases, we allow content which would otherwise go against our Community Guidelines. This concerns content that is newsworthy and 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.

Specifically, the Community Guidelines include the following rules:

1) Share only photos and videos that you've taken or have the right to share.

The Community Guidelines prohibit the posting of content that users are not authorized to post.

2) Post photos and videos that are appropriate for a diverse audience.

Nudity is not allowed on Instagram. This includes photos, videos, and some digitally-created content that show sexual intercourse, genitals, and close-ups of fully-nude buttocks. For safety reasons, we also can remove images that show nude or partially-nude children.

There is some leeway for the representation of nudity, for example, in artistic contexts or - especially in the case of images of uncovered female nipples - in the context of breastfeeding, birth giving, health-related situations or acts of protest.

3) Foster meaningful and genuine interactions.

To prevent spam, the Community Guidelines prohibit, for example, attempts to artificially collect likes, followers, or shares, posting repetitive comments or content, or repeatedly contacting people for commercial purposes without their consent. The Community Guidelines also prohibit posting content that engages in, promotes, encourages, facilitates, or admits to the offering, solicitation or trade of fake and misleading user reviews or ratings. It is also not allowed to impersonate others and to create accounts for the purpose of violating our guidelines or misleading others.

4) Follow the law.

The Community Guidelines explicitly require users to follow the law. This includes, for example, that it is not allowed to use Instagram to support or praise terrorism, organized crime, or hate groups. In addition, we prohibit certain content related to certain transactions or goods (e.g., sexual services, non-medical or pharmaceutical drugs, firearms, endangered species). Moreover, accounts promoting online gambling, online real money games of skill or online lotteries need our prior written permission before using any of our products. We also prohibit sharing sexual content involving minors or threatening to post intimate images of others.

5) Respect other members of the Instagram community.

We remove content that contains credible threats (e.g., of physical harm or vandalism) or hate speech, The same applies to content that targets private individuals to degrade or shame them, personal information meant to blackmail or harass someone, and repeated unwanted messages. Open and controversial discussions about people in the public eye remain allowed, of course. Encouraging violence or attacking anyone based on certain protected characteristics (such as ethnicity, national origin, sex, sexual orientation, religious affiliation, disabilities, or diseases) is not allowed.

6) Maintain our supportive environment by not glorifying self-injury.

We remove content that encourages or urges people to embrace self-injury, or disable related accounts if it is reported to us. We may also remove content identifying victims or survivors of self-injury if the content targets them for attack or humor.

7) Be thoughtful when posting newsworthy events.

We may remove videos of graphic violence. For example, sharing images with clearly visible violence for sadistic pleasure or to glorify violence is never allowed.

15. LEGAL COMPLIANCE OF TERMS & CONDITIONS

> Description of the extent to which the agreement of the provisions under item 14 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 Instagram under the Terms of Use and the policies described above, such prohibition or restriction is in compliance with sections 307 to 309 BGB and other applicable law.

The Terms of Use and the policies referenced in the Terms of Use such as the Community Guidelines are effectively incorporated into the contractual relationship between us and our users. Users expressly agree to the Terms of Use when registering on the platform. The Terms of Use and the other rules thereby become an effective part of the contract and regulate the permissible user behavior on Instagram.

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, sections 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 Use 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.

Nor are any other violations of the law discernible. On the contrary, the Community Guidelines and other rules relating to the permissible content serve in particular to create an authentic and safe environment for the users. 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 Instagram. However, even where our Terms of Use 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 (Bundesgerichtshof) 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 Use provide for the removal of content in the event of violations of the Terms of Use or other policies that apply to the use of our products (including the Community Guidelines) or in cases where we are required to do so by law.

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

The removal of content that violates the Terms of Use is supported by objective reasons. For example, there is already no interest worthy of protection in maintaining unlawful, misleading 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 that the rights of another person are violated; in particular, they must be able to comply with potential removal obligations. In these cases, 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 the Community Guidelines or other policies also constitutes a legitimate concern. The policies serve legitimate purposes. They protect the interests of Instagram users, third parties, and ourselves, and may outweigh the interest in distributing certain content. The Community Guidelines 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 Guidelines 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 rules. One important concern of the Community Guidelines, for example, is to prevent the glorification or promotion of self-injury such as eating disorders. Nevertheless, we want Instagram 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 Instagram. 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 Guidelines generally impose restrictions in this regard.

Insofar as the Federal Court of Justice has also ordered in connection with the Facebook platform 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 update of our Terms of Use of July 26, 2022.

The provisions in each case are also clear and understandable. In particular, our Terms of Use and the Community Guidelines are written in simple and generally understandable language.

The Terms of Use list the key rules and contain, in particular, links to the Community Guidelines. The transparency of the central Community Guidelines is ensured by a visually clear structure and easily understandable wording. The specific rules are each preceded by a simply and aptly worded basic rule which is then explained in more detail in an easily understandable manner. This enables every user to form a reliable picture of which content is allowed on Instagram and which is not.