

Best Practices on Platforms' Implementation of the Right to an Effective Remedy

I. Coordinators

Luca Belli and Nicolo Zingales

II. Background

This document represents the collective output of the *ad hoc* working group of the Dynamic Coalition on Platform Responsibility¹ (DCPR) on the implementation in the context of online platforms of the right to an effective remedy, enshrined inter alia in article 2.3 of the International Covenant on Civil and Political Rights, and several regional Human Rights instruments. The interest in elaborating this document emerged as a clear outcome of the 4th annual meeting of the DCPR, held during the 12th Internet Governance Forum, in December 2017. Many session participants expressed interest in advancing the discussion on platform responsibility, pivoted by the 2017 DCPR official outcome book² and building on the solid ground laid by the 2015 **DCPR Recommendations on Terms of Service and Human Rights** (hereinafter the “**Recommendations**”)³ which are the 2015 DCPR official outcome.

Based on the expression of interests expressed at the IGF 2017 meeting, DCPR Coordinators shared a call for participation to an ad hoc DCPR Working Group (WG) tasked with the analysis of reviewing the existing mechanisms for alternative dispute resolution offered by a selection of platforms, scrutinising due process requirements, and to identify best practices. WG members provided inputs to form a proposed Template⁴ to be used for review of existing dispute resolution mechanisms. At the RightsCon 2018 meeting of the DCPR the composition of the WG was further expanded⁵ and it was agreed to open an additional request for comments on the draft Template, to allow all DCPR members, besides the existing WG members, to provide comments for two additional weeks.⁶

¹ DCPR is a multistakeholder group of the United Nations Internet Governance Forum, dedicated to the analysis of online platforms. DCPR is commonly referred to as the IGF Coalition on Platform Responsibility.

² Specifically, the edited volume “**Platform regulations: how platforms are regulated and how they regulate us**”. The book is freely available at <http://bibliotecadigital.fgv.br/dspace/handle/10438/19402>

³ The **Recommendations** can be accessed at <https://www.intgovforum.org/cms/documents/igf-meeting/igf-2016/830-dcpr-2015-output-document-1/file>

⁴ To encourage and facilitate the inclusion of inputs and comments from WG and DCPR members, the DCPR Coordinators utilised a shared online document available at https://docs.google.com/document/d/1T-bMKnFBtsDQ_AychJl-dlzwPAletMBWJilWRyD-4IM/edit#

⁵ The list of contributing WG member is the following (in alphabetical order): Christina Angelopoulos; Luca Belli (DCPR Coordinator); Maria Bjarnadottir; Marta Cantero Gamito; Giovanni De Gregorio; Luã Fergus; Rosalie Gillett; Agnieszka Janczuck; Cynthia Khoo; Chiara Poletti; Roxana Radu; Nicolas Suzor; Ilana Ullman; Rolf Weber; Chris Wiersma; Richard Wingfield; Nicolo Zingales (DCPR Coordinator).

⁶ The Report of the DCPR meeting at RightsCon 2018 is available at http://www.intgovforum.org/multilingual/index.php?q=filedepot_download/4905/1255

IGF COALITION ON PLATFORM RESPONSIBILITY

The WG members agreed to work towards the identification of best practices, with a view to promoting due process in the context of alternative dispute resolution mechanisms offered by online platforms. The first draft was grounded on the analyses⁷ developed by the WG members and was shared on the public DCPR mailing list to collect feedback. A consolidated version was developed and shared with the wider IGF community to collect a broader range of comments, between October and 31 December 2018.⁸ This final version of the Best Practices was shared on the DCPR mailing-list to receive final comments, over the month of February 2019, and verify that the text represented a consensus document, before publishing it. No objection was raised. However, we acknowledge that the Best Practices should be considered as a living document that could be updated in the future.

III. Introduction

In accordance with the approach adopted by the **Recommendations**, this document utilises the term “**shall**” when practices correspond to minimum standards for the respect of due process by platform operators (standards that “shall” be met), while it utilises “**should**” to suggest practices which are recommended, or “should” be followed to facilitate the most “responsible” adherence to due process principles in the definition and implementation of alternative dispute resolution mechanisms.

The document is structured in four sections exploring the safeguards (a) prior to the adoption dispute resolution measures; (b) in connection with the adoption dispute resolution measures; (c) relating to dispute resolution mechanism; (d) and relating to the implementation of the remedy. Best practices have been identified by merging together solutions that appear most suitable to protect platform users’ rights, at the same time attending to considerations of viability of online platforms’ business models. Quotations of the contractual clauses that inspired the practices are included. When best practices were not identifiable, this document has suggested formulations that maximise the protection of user rights while striking a fair balance between stakeholder interests.

This document was based primarily on the analysis of the contractual agreements that Internet users are required to adhere to in order to become platform users. Platform operators typically detail in these agreements, broadly defined as “Terms of Service” (Tos),⁹ the rules and mechanisms applicable to alternative dispute resolution mechanisms. Moreover, analysts where

⁷ WG members analysed the mechanisms described in the Terms of Service (ToS) of the selected platforms. WG members considered the ToS publicly available in July 2018. All analyses performed by the WG members are available at <https://docs.google.com/spreadsheets/d/11NJr2dQvTSoHs6ZubtQvbwf4Z-h8o7FaNTzR8Qk3UFI/edit#gid=1224846873>

⁸ See <https://www.intgovforum.org/multilingual/content/dcpr-best-practices-on-due-process-safeguards-regarding-online-platforms%E2%80%99-implementation-of>

⁹ These Best Practices utilise the same definition of ToS provided by the Recommendations, thus covering not only contractual agreements available under the traditional heading of “Terms of Service” or “Terms of Use”, but also any other platform’s policy document (e.g. Privacy Policy, Community Guidelines, etc.) that is linked or referred to therein.

IGF COALITION ON PLATFORM RESPONSIBILITY

asked to verify, to the extent possible, the concrete implementation of those mechanisms by simulating a dispute in the platforms of choice.

A. Safeguards prior to the adoption of dispute resolution measures

1. Platforms should require registration in order for users to actively interact with others and to create content, within the platform. However, they should not impose the use of real name as public user login. While requiring complete and accurate information about users at the moment of registration, platforms shall not oblige users to make that information public.¹⁰

Furthermore, platforms should not permit registrations with the effect to:

- a) Creating public reliance on someone else's name, image, or other personal information, if that is liable to deceive third parties as to a user's identity. No deception arises, however, in case of clearly parodic impersonification of public figures.
- b) Misleading third parties as to a user's authority to represent a particular natural or legal person.

User information shall be shared with third parties, including state actors, only when this is justified by a court order.

Twitter

If you do choose to create an account, you must provide us with some personal data so that we can provide our services to you. On Twitter this includes a display name (for example, "Twitter Moments"), a username (for example, @TwitterMoments), a password, and an email address or phone number. Your display name and username are always public, but you can use either your real name or a pseudonym.

LinkedIn

Members cannot: a) impersonate others on the Services or mislead, confuse, or deceive others. Pretending to be someone else or to be representing a business in a way that is not truthful is not allowed. b) use someone else's name, image, or other personal information to deceive others into thinking you are someone other than the member or associated with a business or organization when the members are not. c) use or attempt to use another individual's LinkedIn account or create a member profile for anyone other than the member (a real person). d) misrepresent their identity or information or mislead, confuse, or deceive others. When choosing a profile picture, members may not use an image that is not their likeness or a head-shot photo for their profile. Also, members may not manipulate identifiers in order to disguise the origin of any message or post transmitted through the Services.

2. In case platforms aim at restricting the type of content deemed as acceptable, their terms of service shall set out detailed rules, clearly explaining what type of content can be considered acceptable.¹¹ In doing so, platforms shall bear in mind their responsibility to respect human rights, including freedom of expression. Categories of content that could be deemed as

¹⁰ See Recommendations, Section I.5

¹¹ See Recommendations, Section III.1

IGF COALITION ON PLATFORM RESPONSIBILITY

unacceptable and shall be clearly defined in the terms of service include spam, shocking and pornographic content, content instigating violence or discriminating against individuals based on race, ethnicity, national origin, sex, gender, gender identity, sexual orientation, religious affiliation, disabilities, or diseases, or content deemed as illegal in specific jurisdictions.

LinkedIn (applicable to disputes concerning: Intellectual property infringement; Revenge porn; Fake news; Terrorism-inciting content; Hate speech; Right to erasure/ right to object to processing/ right to rectify or restrict processing; Defamation; Child pornography)

Honesty and Authenticity [...] You may not use the Services to share false content or information, including news stories, that presents untrue or unverified facts or events as though they are true or acts or events as though they are true or likely true. [...] Adult Content It's not acceptable to post content containing nudity, sexually explicit material, or pornography. Some adult content may be allowed in an educational, medical, scientific, or professional artistic context so long as it is not gratuitously graphic. The Services are never to be used for sexual exploitation of children. You also may not post content that threatens sexual violence or sexual assault. You may not use the Services to engage in or promote escort services, prostitution, or human trafficking. Bullying and Harassment Bullying or harassment that targets individuals or groups to degrade or shame them is not allowed. This includes, but is not limited to, abusive or humiliating language, sexual advances and innuendo, revealing others' personal or sensitive information (aka "doxing") or posting content about them without consent, or inciting or engaging others to do any of the same. Hate, Violence, and Terrorism We do not allow organizations or groups that engage in or promote violence or property damage, organized criminal activity, prejudice, or hate. Also, you may not use our Services to express support for such groups or to post content or otherwise use the Services to incite violence or hatred against particular individuals or groups. Content that depicts terrorist activity, that is intended to recruit for terrorist organizations, or promotes or supports terrorism in any manner, is not tolerated on the Services. Harmful Content and Shocking Material You may not post violent or graphic content or otherwise use the Services with the intent to shock or humiliate others. We do not allow activities that promote, organize, depict or facilitate criminal activity. We also do not allow content depicting or promoting instructional weapon making, drug abuse, and threats of theft. Content or activities that promote or encourage suicide or any type of self-injury, including self-mutilation and eating disorders, is also not allowed. Spam Untargeted, irrelevant, unwanted, unsolicited, unauthorized, inappropriately commercial or promotional, or gratuitously repetitive messages and other similar content are considered spam and are not allowed on the Services. You may not use our invitation features to send messages to people who don't know you or who are unlikely to recognize you as a known contact. Please make the effort to create original, professional, relevant, and interesting content in order to gain popularity, instead of trying ways to artificially increase the number of views, re-shares, likes, or comments.

IGF COALITION ON PLATFORM RESPONSIBILITY

3. As a general rule, platforms should only store personal data for as long as necessary for the purpose(s) for which they were originally collected.¹² This should include retention for a period that is reasonably necessary to comply with legal obligations (e.g. law enforcement requests), meet regulatory requirements and resolve disputes, or to protect the safety or integrity of the platform. Examples of the latter are where storage helps to prevent spam and detect fraud or malicious behaviour aimed at service disruption, or to explain why platform operators removed specific content or accounts from the platform.

Airbnb

Airbnb generally retain personal information “for as long as is necessary for the performance of the contract between you and us and to comply with our legal obligations”. Users/members can request the erasure of personal information.

4. Platforms should provide meaningful notice of any changes in their ToS at least 30 days before the changes go into effect.¹³ Platforms shall provide users with the opportunity to review the changes before they become effective and changes cannot be retroactive. Notification of changes shall be communicated both via email, where practicable, and through the platform.

WordPress

WordPress uses posts/email/other communication in advance of changes - see "13. Changes." in ToS, including statement that "any dispute that arose before the changes shall be governed by the Terms (including the binding individual arbitration clause) that were in place when the dispute arose." - AND it keeps change logs –

Wikipedia

Wikipedia provides Terms of Use, as well as any substantial future revisions of these Terms of Use, to the community for comment at least thirty (30) days before the end of the comment period. If a future proposed revision is substantial, we will provide an additional 30 days for comments after posting a translation of the proposed revision in multiple languages

5. Platforms shall offer mechanisms to report behaviours categorised as abusive by the ToS, by flagging contents and or by filing predefined forms. For instance, when prohibited by the platform’s ToS, users should be able to flag:
 - Spam,
 - Content categorised as inappropriate by the terms of service
 - Profiles or groups engaging in activities forbidden by the terms of service
 - Phishing and or fraud attempt
 - Safety concerns

Specific notice-and-counter-notice mechanisms should be established for

- Intellectual property infringements

¹² See Recommendations, Section I.

¹³ See Recommendations, Section II.1

IGF COALITION ON PLATFORM RESPONSIBILITY

- Law enforcement requests for account information (routine and emergency)
- Content removal requests, based on ToS infringement
- Reporting of hacked account.

Where relevant, the abovementioned form shall include at least the following elements

- The email address of the claimant
- The description of the violation type
- The username of violating account
- The URL of post
- Any supporting material in attachments

Linkedin

Linkedin provides mechanism to report abusive behaviours by flagging contents or filing forms according to its Community Guidelines and User Agreement. In general, the following contents could be flagged by users: - Spam, inappropriate, and offensive content - Inappropriate profile photos - Inaccurate profiles - Fake profiles - Inappropriate groups - Phishing or suspicious messages - Safety concerns A specific mechanism based on notice and counter notice is established for copyrights contents (<https://www.linkedin.com/legal/copyright-policy>). Moreover, a member can report also by flagging or filling a form: - trademark infringements (see the "Trademark Infringement Form"). - fake profiles - hacked accounts (see the form "Reporting Your Hacked Account") - scams

Medium

Medium's rules state: How to report a violation If you find a post or account on Medium that violates these rules, please flag it. You can use this form to provide more detail or to report other conduct you believe violates our rules. Additionally, you can send us an email to yourfriends@medium.com. The report form asks for the following details: How can we help you? (drop down menu features: "report a rules violation.") Your email address Description Violation type Medium username of violating account URL of post Attachments Medium also provides information on filing a DMCA notice: How To File a DMCA Notice To submit a notice of claimed copyright infringement, you will need to provide us with the following information: 1. A physical or electronic signature (typing your full name will suffice) of the copyright owner or a person authorized to act on their behalf; 2. Identification of the copyrighted work claimed to have been infringed (e.g., a copy of or link to your original work or clear description of the materials allegedly being infringed upon); 3. Identification of the infringing material and information reasonably sufficient to permit Medium to locate the material on our website or services (e.g., a link to the infringing post); 4. Your contact information, including your address, telephone number, and an email address; 5. A statement that you have a good-faith belief that the use of the material in the manner asserted is not authorized by the copyright owner, its agent, or the law; and 6. A statement that the information in the notification is accurate, and, under penalty of perjury, that you are authorized to act on behalf of the copyright owner. You can report alleged copyright infringement by emailing the above information to copyright@medium.com. You can also mail a copyright notice to: Designated Copyright Agent A Medium Corporation 760 Market Street, Suite 900 San Francisco, CA 94102

IGF COALITION ON PLATFORM RESPONSIBILITY

6. Platform users shall have the right to initiate litigation and take part in class actions in their own jurisdiction.¹⁴ Such rights shall always be available in jurisdictions that are targeted by the platform services (e.g. by using local language, currency or country code domain name).

B. Safeguards in connection with the adoption of dispute resolution measures

7. As a general rule, platforms shall notify affected individuals prior to the adoption of any adverse measures, explaining the specific grounds on which such measure is taken.¹⁵ Exceptions to user notification should be narrowly circumscribed and explained in the terms of service.

Twitter

By default, Twitter will attempt to notify the reported account holder(s) of the existence of a legal request pertaining to the account(s) if we are not otherwise prohibited from doing so. Exceptions to user notice may include exigent circumstances, such as emergencies regarding imminent threats to life, child sexual exploitation, or terrorism. Twitter attempts to notify the user(s) about the legal request through a notification in the Twitter app and by sending a message to the email address associated with the account(s), if available. If we are not permitted to notify the user(s) at this step in the process (e.g., because the legal request is accompanied by a non-disclosure order), we may notify the user(s) about the existence of a legal request after Twitter has withheld the reported content or disclosed information associated with the Twitter account(s).

8. Platform should always allow affected individuals to contest a notified measure *before* adoption.¹⁶ Measures should be implemented immediately, on a temporary basis, when this is justified by particular urgency e.g. when content shall be removed before it incites others to do harm, or in case of child abuse imagery.

Medium

If you break the rules If it looks like you've violated our rules, we may send you an email and ask you to explain what you're up to and why. Context is important, and we want to understand the big picture. If you don't adequately explain yourself or fix the problem, we may suspend your account or remove your content. We strive to be fair, but we reserve the right to suspend accounts or remove content, without notice, for any reason, particularly to protect our services, infrastructure, users, or community. If you attempt to evade suspension by creating new accounts, we will suspend your new accounts.

¹⁴ See Recommendations, Section II.2

¹⁵ See Recommendations, Sections III.1 and III.2

¹⁶ See Recommendations, Section III.2

IGF COALITION ON PLATFORM RESPONSIBILITY

9. Platforms shall always notify affected individuals after the adoption of the measure, explaining the specific grounds based on which the measure was taken.¹⁷

Youtube

If a strike is issued, you'll get an email and see an alert in your account's Channel Settings with information about why your content was removed (e.g. for sexual content or violence).

10. Furthermore, platforms shall always allow affected individuals to contest a measure *after* adoption.¹⁸

Twitter

Violators can appeal permanent suspensions if they believe we made an error. They can do this through the platform interface or by filing a report. Upon appeal, if we find that a suspension is valid, we respond to the appeal with information on the policy that the account has violated."

"File an appeal and we may be able to unsuspend your account. If you are unable to unsuspend your own account using the instructions above and you think that we made a mistake suspending or locking your account, you can appeal. First, log in to the account that is suspended. Then, open a new browser tab and file an appeal.

Instagram

Instagram complies with the notice and takedown procedures defined in section 512(c) of the Digital Millennium Copyright Act ("DMCA"), which applies to content reported and removed for infringing United States copyrights. If your content was removed under the notice and counter-notice procedures of the DMCA, you will receive instructions about the counter-notification process, including how to file a counter-notification, in the warning we send you. When we receive an effective DMCA counter-notification, we promptly forward it to the reporting party. If the reporting party does not notify us that they have filed an action seeking a court order to restrain you from engaging in infringing activity on Instagram related to the material in question within 10-14 business days, we may restore or cease disabling eligible content under the DMCA". "Similarly, if the content was removed based on U.S. trademark rights, and if you believe the content should not have been removed, you will be provided an opportunity to submit an appeal. In these cases, you'll receive further instructions about this process in the notification you receive from Instagram.

11. To ensure the effectiveness of contestation, time limits to contest any measure shall be clearly specified.

Twitter

¹⁷ See Recommendations, Section III.1

¹⁸ Id.

IGF COALITION ON PLATFORM RESPONSIBILITY

A time limit is mentioned only in the copyright procedure but not for the contestant, only for the original claimant. "What Happens After I Submit a Counter-notice? Upon receipt of a valid counter-notice, we will promptly forward a copy to the person who filed the original notice. If we do not receive notice within 10 business days that the original reporter is seeking a court order to prevent further infringement of the material at issue, we may replace or cease disabling access to the material that was removed.

C. Safeguards relating to the dispute resolution mechanism

12. Platforms should have in place a specific mechanism in their websites allowing users to resolve disputes arising between them in relation to their platform activity,¹⁹ besides the mechanisms allowing users to solve disputes between the platform and its users, as specified in paragraph 16.

Airbnb (only concerning claims on security deposits)

The procedure for claims on security deposits proceeds as follows: - Airbnb will ask for documentation from the host, and as soon as it is received, Airbnb will ask the host to contact the guest through Airbnb's Resolution Center to discuss the claim. - When the host sends a request, the guest will be notified by email and through an alert on Airbnb Dashboard. - The guest will have to reply to the host's request in the Resolution Center within 72 hours. The guest's response will depend on whether or not the guest agrees to the amount requested by the host: o Agree to the amount:

- *Click Accept in the Resolution Center. In such case, Airbnb will process the payment and send it to the host (usually within 5 to 7 business days).*
- *Don't agree to the amount: Click Involve Airbnb in the Resolution Center. The guest must provide reasons the invalidity of the host's claim. In such event, Airbnb will contact the guest and provide 72 hours to respond so that Airbnb can mediate.*

The Help Center signals that, in any case, they will make sure both guest and host are represented fairly and gather any details and documentation needed to reach a resolution. It states that most security deposit claims will be resolved within one week.

13. Platforms should provide detailed and clear explanations to users on the significance of any requests for initiation of disputes that is notified to them, and actions that may be taken in response to those.²⁰ Platforms should also offer additional assistance, for example by providing a channel for interaction with customer service, or listing contact information of the relevant non-governmental organisations.

Twitter (general guidance)

¹⁹ See Recommendations, Section II.2

²⁰ Id.

IGF COALITION ON PLATFORM RESPONSIBILITY

In case of suspension of account, they describe the procedure to unblock/unsuspend the account and explain the possible reasons (e.g. Your account has been locked for security purposes, Your account is limited because it may have violated the Twitter Rules)

- *"You may be able to unsuspend your own account. If you log in and see prompts that ask you to provide your phone number or confirm your email address, follow the instructions to get your account unsuspended." <https://help.twitter.com/forms/general?subtopic=suspended> - "Are you seeing a message that your account is locked? Your account may also be temporarily disabled in response to reports of spammy or abusive behavior. For example, you may be prevented from Tweeting from your account for a specific period of time or you may be asked to verify certain information about yourself before proceeding. Get help unlocking your account. File an appeal and we may be able to unsuspend your account. If you are unable to unsuspend your own account using the instructions above and you think that we made a mistake suspending or locking your account, you can appeal. First, log in to the account that is suspended. Then, open a new browser tab and file an appeal. Source (<https://help.twitter.com/en/managing-your-account/locked-and-limited-accounts>) - Help with locked or limited account We may lock an account or place temporary limitations on certain account features if an account appears to be compromised or in violation of the Twitter Rules or Terms of Service. If you log in or open your app and see a message that your account is locked or that some of your account features have been limited, follow the instructions to restore it or continue reading for more information. In case of legal requests in the US they offer the contact of two NGOs specialised in freedom of expression (ACLU and EFF). "Unfortunately, we cannot provide you with any legal advice and cannot provide any further information beyond what we provided in our notice. If you wish to seek legal counsel, here are some resources that may help. For U.S. legal requests, you might consider contacting the American Civil Liberties Union (<http://www.aclu.org/affiliates>, +1 212-549-2500) or the Electronic Frontier Foundation (<https://www.eff.org/pages/legal-assistance>, info@eff.org, +1 415-436-9333). In other countries For non-U.S. legal requests, you might consider contacting a local attorneys' association or law school, which may be able to provide you with contact information for specialised legal assistance on free expression issues or reduced-cost legal aid services available in your location*
- Twitter also has a social media account (@Twittersupport) which is the official source for 24/7 Twitter support.

14. Platforms should inform complainants of counternotices and other defenses raised in response to their requests, so as to enable a meaningful contestation.²¹

Linkedin (only for copyright)

Yes, LinkedIn has included a note in the counter-notice form which explains the time for the complainant to commence a formal judicial action upon receipt of a copy of the counter-notice (<https://www.linkedin.com/help/linkedin/ask/TS-CNRCCI?lang=en>). "Note: Under the Digital Millennium Copyright Act, upon receipt of a copy of this Counter-Notice, the Complainant has 10 business days to commence a formal judicial action against the User in relation to the User's infringing activity. If such action is filed, the allegedly infringing content will be removed or will remain removed from the LinkedIn and/or SlideShare site until the matter is resolved. If no action is filed, we will

²¹ See Recommendations, Section II.2

IGF COALITION ON PLATFORM RESPONSIBILITY

re-post, or allow you to re-post, the content 10-14 business days after receipt of this Counter-Notice.

15. Platforms that receive requests for content removal shall only implement permanent deletion after an internal (human) review. Users shall always have the possibility to challenge automated deletion and the right to have the deletion decision reviewed by an independent expert or a panel of experts.

Youtube

Reported content is reviewed along the following guidelines: Content that violates our Community Guidelines is removed from YouTube. Content that may not be appropriate for all younger audiences may be age-restricted." However, in its most recent transparency report, YouTube stated that 74.2% of videos are removed before any views thanks to automated flagging.

16. Platforms shall provide an alternative dispute resolution mechanism, designed in a flexible way based on generally accepted procedural rules, for disputes between a user and the platform.²²

The rules for such a mechanism should encompass at least the following elements:

- Procedure for the appointment of the adjudicator
- Necessary independence and qualifications of the adjudicators
- Choice between 1 adjudicator and panel of 3 adjudicators
- Procedural principles of such mechanisms should enshrine the right to be heard, equal treatment of parties, access to information, acting in good faith.

Wikimedia

We hope that no serious disagreements arise involving you, but, in the event there is a dispute, we encourage you to seek resolution through the dispute resolution procedures or mechanisms provided by the Projects or Project editions and the Wikimedia Foundation.

Tumblr

The Terms of Service state that: "You and Tumblr agree that we will resolve any claim or controversy at law or equity that arises out of this Agreement or the Services in accordance with this Section or as you and Tumblr otherwise agree in writing. Before resorting to formal dispute resolution, we strongly encourage you to contact us to seek a resolution.

Snapchat (only for businesses)

Not if the user is an individual, Yes if the user is a business. Then the dispute will be settled under LCIA Arbitration Rules. "One arbitrator (to be appointed by the LCIA), the arbitration will take place in London, and the arbitration will be conducted in English. If you do not wish to agree to this clause, you must not use the Services.

²² Id.

IGF COALITION ON PLATFORM RESPONSIBILITY

17. Platforms should offer alternative dispute resolution mechanisms as an option, but not as an inderogable pre-requisite or substitute for litigation.²³ Platform users shall always have a meaningful opportunity to opt out from the use of such mechanisms.

Amazon (only for small claims)

Any dispute or claim relating in any way to your use of any Amazon Service, or to any products or services sold or distributed by Amazon or through Amazon.com will be resolved by binding arbitration, rather than in court, except that you may assert claims in small claims court if your claims qualify. The Federal Arbitration Act and federal arbitration law apply to this agreement.

Reddit (informal process, not specified)

Yes. In their User Agreement, par. 13. Governing Law and Venue they specify that "if you have an issue or dispute, you agree to raise it and try to resolve it with us informally. You can contact us with feedback and concerns here or by emailing us at contact@reddit.com.

eBay (opt out available)

Opt-Out Procedure IF YOU ARE A NEW USER OF OUR SERVICES, YOU CAN CHOOSE TO REJECT THIS AGREEMENT TO ARBITRATE ("OPT-OUT") BY MAILING US A WRITTEN OPT-OUT NOTICE ("OPT-OUT NOTICE"). THE OPT-OUT NOTICE MUST BE POSTMARKED NO LATER THAN 30 DAYS AFTER THE DATE YOU ACCEPT THE USER AGREEMENT FOR THE FIRST TIME. YOU MUST MAIL THE OPT-OUT NOTICE TO EBAY INC., ATTN: LITIGATION DEPARTMENT, RE: OPT-OUT NOTICE, 583 WEST EBAY WAY, DRAPER, UT 84020.

Uber (small claims, & equitable relief against possible IP infringement)

However, you and Uber each retain the right to bring an individual action in small claims court and the right to seek injunctive or other equitable relief in a court of competent jurisdiction to prevent the actual or threatened infringement, misappropriation or violation of a party's copyrights, trademarks, trade secrets, patents or other intellectual property rights.

18. Platforms should set a reasonable time limit (e.g. 30 days) for the resolution of any controversy, with the possibility to extend such period upon mutual agreement between the disputing parties. Furthermore, platforms should only set a time limit (e.g., 1 year) to the initiation of claims that have arisen in the past.

Lyft

Before initiating any arbitration or proceeding, you and Lyft may agree to first attempt to negotiate any dispute, claim or controversy between the parties informally for 30 days, unless this time period is mutually extended by you and Lyft.

Tumblr

Time Limitation on Claims and Releases From Liability | You agree that any claim you may have arising out of or related to this Agreement or your relationship with Tumblr must be filed within one year after such claim arose; otherwise, your claim is permanently barred.

Tumblr (for copyright)

²³ Id.

IGF COALITION ON PLATFORM RESPONSIBILITY

The original Notifying Party (or the copyright holder he or she represents) will then have ten (10) days to notify us that he or she has filed legal action relating to the allegedly infringing material. If Tumblr does not receive any such notification within ten (10) days, we may restore the material to the Services.

19. Platforms shall ensure that adjudication of disputes conforms to established standards of independence and impartiality,²⁴ for example by reference to rules and procedures adopted by recognised mediation or arbitration associations.

Ebay

The arbitration will be conducted by the American Arbitration Association ("AAA") under its rules and procedures, including the AAA's Consumer Arbitration Rules (as applicable), as modified by this Agreement to Arbitrate. The AAA's rules are available at www.adr.org or by calling the AAA at 1-800-778-7879. The use of the word "arbitrator" in this provision shall not be construed to prohibit more than one arbitrator from presiding over an arbitration; rather, the AAA's rules will govern the number of arbitrators that may preside over an arbitration conducted under this Agreement to Arbitrate.

User Privacy Notice

If you have an unresolved privacy or data use concern that we have not addressed satisfactorily, please contact our U.S.-based third party dispute resolution provider (free of charge) at <https://feedback-form.truste.com/watchdog/request>. eBay is committed to your privacy. This privacy notice explains our collection, use, disclosure, retention, and protection of your personal information.

Amazon

The arbitration will be conducted by the American Arbitration Association ("AAA") under its rules, including the AAA's Supplementary Procedures for Consumer-Related Disputes. The AAA's rules are available at www.adr.org or by calling 1-800-778-7879. Payment of all filing, administration and arbitrator fees will be governed by the AAA's rules. Amazon will reimburse those fees for claims totaling less than \$10,000 unless the arbitrator determines the claims are frivolous. Likewise, Amazon will not seek attorneys' fees and costs in arbitration unless the arbitrator determines the claims are frivolous. You may choose to have the arbitration conducted by telephone, based on written submissions, or in person in the county where you live or at another mutually agreed location.

20. Platforms shall provide sufficient reasons to appreciate the rationale of the decision taken by the appointed adjudicator, and should provide an updated list of factors elucidating the application of their terms of service (i.e., their implementation criteria).²⁵

Twitter

Our enforcement philosophy

We empower people to understand different sides of an issue and encourage dissenting opinions and viewpoints to be discussed openly. This approach allows many forms of speech to exist on our platform and, in particular, promotes counterspeech: speech that

²⁴ See Recommendations, Section II

²⁵ Id.

IGF COALITION ON PLATFORM RESPONSIBILITY

presents facts to correct misstatements or misperceptions, points out hypocrisy or contradictions, warns of offline or online consequences, denounces hateful or dangerous speech, or helps change minds and disarm.

Thus, **context matters**. When determining whether to take enforcement action, we may consider a number of factors, including (but not limited to) whether:

- *The behavior is directed at an individual, group, or protected category of people;*
- *The report has been filed by the target of the abuse or a bystander;*
- *The user has a history of violating our policies;*
- *The severity of the violation;*
- *The content may be a topic of legitimate public interest.*

Is the behavior directed at an individual or group of people?

To strike a balance between allowing different opinions to be expressed on the platform, and protecting our users, we enforce policies when someone reports abusive behavior that targets a specific person or group of people. This targeting can happen in a number of ways (for example, @mentions, tagging a photo, mentioning them by name, and more).

Has the report been filed by the target of the potential abuse or a bystander?

Some Tweets may seem to be abusive when viewed in isolation, but may not be when viewed in the context of a larger conversation or historical relationship between people on the platform. For example, friendly banter between friends could appear offensive to bystanders, and certain remarks that are acceptable in one culture or country may not be acceptable in another. To help prevent our teams from making a mistake and removing consensual interactions, in certain scenarios we require a report from the actual target (or their authorized representative) prior to taking any enforcement action.

Does the user have a history of violating our policies?

We start from a position of assuming that people do not intend to violate our Rules. Unless a violation is so egregious that we must immediately suspend an account, we first try to educate people about our Rules and give them a chance to correct their behavior. We show the violator the offending Tweet(s), explain which Rule was broken, and require them to delete the content before they can Tweet again. If someone repeatedly violates our Rules then our enforcement actions become stronger. This includes requiring violators to delete the Tweet(s) and taking additional actions like verifying account ownership and/or temporarily limiting their ability to Tweet for a set period of time. If someone continues to violate Rules beyond that point then their account may be permanently suspended.

What is the severity of the violation?

Certain types of behavior may pose serious safety and security risks and/or result in physical, emotional, and financial hardship for the people involved. These egregious violations of the Twitter Rules — such as posting violent threats, non-consensual intimate media, or content that sexually exploits children — result in the immediate and permanent suspension of an account. Other violations could lead to a range of different

IGF COALITION ON PLATFORM RESPONSIBILITY

steps, like requiring someone to delete the offending Tweet(s) and/or temporarily limiting their ability to post new Tweet(s).

Is the behavior newsworthy and in the legitimate public interest?

Twitter moves at the speed of public consciousness and people come to the service to stay informed about what matters. Exposure to different viewpoints can help people learn from one another, become more tolerant, and make decisions about the type of society we want to live in.

To help ensure people have an opportunity to see every side of an issue, there may be the rare occasion when we allow controversial content or behavior which may otherwise violate our Rules to remain on our service because we believe there is a legitimate public interest in its availability. Each situation is evaluated on a case by case basis and ultimately decided upon by a cross-functional team.

Some of the factors that help inform our decision-making about content are the impact it may have on the public, the source of the content, and the availability of alternative coverage of an event.

*Public impact of the content: A topic of legitimate public interest is different from a topic in which the public may be curious. We will consider what the impact is to citizens if they do **not** know about this content. If the Tweet does have the potential to impact the lives of large numbers of people, the running of a country, and/or it speaks to an important societal issue then we may allow the the content to remain on the service. Likewise, if the impact on the public is minimal we will most likely remove content in violation of our policies.*

Source of the content: Some people, groups, organizations and the content they post on Twitter may be considered a topic of legitimate public interest by virtue of their being in the public consciousness. This does not mean that their Tweets will always remain on the service. Rather, we will consider if there is a legitimate public interest for a particular Tweet to remain up so it can be openly discussed.

Availability of coverage: Everyday people play a crucial role in providing firsthand accounts of what's happening in the world, counterpoints to establishment views, and, in some cases, exposing the abuse of power by someone in a position of authority. As a situation unfolds, removing access to certain information could inadvertently hide context and/or prevent people from seeing every side of the issue. Thus, before actioning a potentially violating Tweet, we will take into account the role it plays in showing the larger story and whether that content can be found elsewhere.

D. Safeguards relating to the implementation of the remedy

- 21.** Platforms should clarify both in their ToS and in the implementation of their practices the territorial scope of any remedy that can be sought or imposed.

IGF COALITION ON PLATFORM RESPONSIBILITY

Twitter (global remedy unless it is a request by a government or third party in which case it is local)

If content violates their ToS they remove the content from the platform (globally) otherwise if content are removed on the basis of legal requests they remove it only on the country. For content removal requests, this may mean the reported content violates Twitter's Terms of Service or Rules, and the content will be removed from the Twitter platform. Or, perhaps the content is determined to be illegal in a particular jurisdiction and Twitter will withhold access to the identified content in the location in which it is alleged to be in violation of local law. For information requests, Twitter may file or serve objections for requests that are legally defective, overly broad, and/or appear to impermissibly burden free expression. Twitter also checks whether the user(s) filed any objections with the appropriate court. For valid and properly scoped information requests where there has not been a successful objection by Twitter or the user(s), a Twitter agent will assemble the required account records and produce them electronically through our secure LRS site to the requester. Once the records have been produced, the case is considered completed and closed unless we're able to provide delayed notice to affected users after the expiration of an associated non-disclosure order. Source: <https://help.twitter.com/en/rules-and-policies/twitter-legal-faqs>

22. Platforms should offer the possibility to request the adoption of temporary measures prior to resolution of a dispute. The provisions of set out in paragraph 16 apply for such procedures mutatis mutandis.

Wordpress (only useful answer, based on the analyst's personal experience)
Occasionally, WordPress responds to reports by suspending a blog(-post);

23. Platforms should give users the opportunity to request a review of any implemented measures.²⁶ This includes the right to appeal against the assessment of the factual context in which a decision was taken and its consistency with the factors laid out in the platform's ToS (i.e. the enforcement philosophy referred to in C9). Platforms should also provide the possibility to request a review to account for supervened circumstances, as well as representative examples of the types of circumstances (e.g. court decisions) that qualify for the granting of such requests.

Twitter

If content that was withheld in response to a legal request becomes allowed in the future, where we can, we will restore access to it so anyone in the world can view it. Some circumstances in which we have un-withheld content in the past include: An objection filed by Twitter against a court order deeming certain content was illegal was accepted by a higher court. An objection filed by a user against a court order deeming certain content was illegal was accepted by a higher court. The validity period of a court order prohibiting publication of certain material expired. An official judicial body expressed an opinion that a request made by an administrative authority was invalid.*

²⁶ See Recommendations, Section II.2

IGF COALITION ON PLATFORM RESPONSIBILITY

Airbnb (yes, about the facts- but do not allow to challenge their interpretation of standards and expectations)

Following Airbnb's Standards and expectations (<https://www.airbnb.com/help/article/1199/what-are-airbnb-s-standards-and-expectations>): enforcement teams are made up of dedicated professionals, "but they're still human". Therefore, they acknowledge potential incorrect decisions. ("So, in rare cases, enforcement decisions may be incorrect"). In the event of disagreement with a decision, users are invited to contact Airbnb directly, and then the platform commits to "re-review the decision carefully". However, as it is specified, the definitions of the standards and expectations themselves aren't subject to review.

24. Platforms should have flexible rules allowing for different types of arrangements regarding the allocation of costs in relation to the implementation of a remedy. These rules may include an indication of the amount of claim below which a platform will reimburse users for filing, administration, and arbitrator fee; and should include penalties in case a claim is established to be frivolous.

eBay

Costs of Arbitration Payment of all filing, administration and arbitrator fees will be governed by the AAA's rules, unless otherwise stated in this Agreement to Arbitrate. If the value of the relief sought is \$10,000 or less, at your request, eBay will pay all filing, administration, and arbitrator fees associated with the arbitration. Any request for payment of fees by eBay should be submitted by mail to the AAA along with your Demand for Arbitration and eBay will make arrangements to pay all necessary fees directly to the AAA. If (a) you willfully fail to comply with the Notice of Dispute requirement discussed above, or (b) in the event the arbitrator determines the claim(s) you assert in the arbitration to be frivolous, you agree to reimburse eBay for all fees associated with the arbitration paid by eBay on your behalf that you otherwise would be obligated to pay under the AAA's rules.

Amazon

Payment of all filing, administration and arbitrator fees will be governed by the AAA's rules. We will reimburse those fees for claims totaling less than \$10,000 unless the arbitrator determines the claims are frivolous.

Lyft

Lyft attributes costs in the event of passenger cancellations (by charging a fee, which the driver receives). Drivers are not charged a fee for cancelling on passengers, but are penalized on performance or ratings: Passengers: <https://help.lyft.com/hc/en-ca/articles/115012922687-Cancellation-policy-for-passengers> "Cancel fees | You may be charged a fee for cancelling a ride when both of the following occur: - 2 minutes or more pass since a driver accepts your ride request - Your driver is on time to arrive within 5 minutes of the original estimated arrival time In most cities, you'll be charged \$10 for cancelling a scheduled ride." "No-show fee | No-show fees are charged under these circumstances: 1. Your driver arrived to pick you up 2. Your driver waited 5 minutes or more 3. Your driver tried to contact you" Drivers: <https://help.lyft.com/hc/en-ca/articles/115012922847> "Cancellation and no-show fee policy for drivers | As consideration for your time and effort, drivers receive cancellation and no-show fees. Fees are based on your region and ride type, so use our cities page to see specific

IGF COALITION ON PLATFORM RESPONSIBILITY

amounts." *Damage Fee is attributed to passengers: "Damage Fee. If a Driver reports that you have materially damaged the Driver's vehicle, you agree to pay a "Damage Fee" of up to \$250 depending on the extent of the damage (as determined by Lyft in its sole discretion), towards vehicle repair or cleaning."* <https://www.lyft.com/terms> In the event of a dispute going to arbitration, Lyft will compensate users for all but \$50 of filing fee, unless claim is for \$5000 or more (section 17(e)), if user initiates, or compensate entirety of filing and arbitration fees if Lyft initiates: <https://www.lyft.com/terms>. Lyft also agrees not to seek attorneys' fees and non-filing expenses if it wins in arbitration (section 17e(6)), but will also not pay user's legal fees in any event.

25. Platforms should set out rules mentioning the possible consequences of repeated infringement of ToS, specifying any significant variations in those consequences depending on the type of violation. They should also make clear that such consequences may only arise in case of established, rather than merely asserted, violations.

YouTube (for copyright)

If you receive more than one strike in the same three-month period, here's what happens: Second strike: If your account receives two Community Guidelines strikes within a three-month period, you won't be able to post new content to YouTube for two weeks. If there are no further issues, full privileges will be restored automatically after the two-week period. Each strike will remain on your account and expire three months after it was issued. Each strike expires separately. Third strike: If your account receives three Community Guidelines strikes within a three-month period, your account will be terminated.

Wikimedia

There are detailed policies relating to blocking users from editing content, and banning users from the platform.

In an unusual case, the need may arise, or the community may ask us, to address an especially problematic user because of significant Project disturbance or dangerous behavior. In such cases, we reserve the right, but do not have the obligation to:

- *Investigate your use of the service (a) to determine whether a violation of these Terms of Use, Project edition policy, or other applicable law or policy has occurred, or (b) to comply with any applicable law, legal process, or appropriate governmental request;*
- *Detect, prevent, or otherwise address fraud, security, or technical issues or respond to user support requests;*
- *Refuse, disable, or restrict access to the contribution of any user who violates these Terms of Use;*
- *Ban a user from editing or contributing or block a user's account or access for actions violating these Terms of Use, including repeat copyright infringement;*
- *Take legal action against users who violate these Terms of Use (including reports to law enforcement authorities); and*
- *Manage otherwise the Project websites in a manner designed to facilitate their proper functioning and protect the rights, property, and safety of ourselves and our users, licensors, partners, and the public.*

In the interests of our users and the Projects, in the extreme circumstance that any individual has had his or her account or access blocked under this provision, he or she is

IGF COALITION ON PLATFORM RESPONSIBILITY

prohibited from creating or using another account on or seeking access to the same Project, unless we provide explicit permission. Without limiting the authority of the community, the Wikimedia Foundation itself will not ban a user from editing or contributing or block a user's account or access solely because of good faith criticism that does not result in actions otherwise violating these Terms of Use or community policies.

The Wikimedia community and its members may also take action when so allowed by the community or Foundation policies applicable to the specific Project edition, including but not limited to warning, investigating, blocking, or banning users who violate those policies. You agree to comply with the final decisions of dispute resolution bodies that are established by the community for the specific Project editions (such as arbitration committees); these decisions may include sanctions as set out by the policy of the specific Project edition.

Especially problematic users who have had accounts or access blocked on multiple Project editions may be subject to a ban from all of the Project editions, in accordance with the Global Ban Policy. In contrast to Board resolutions or these Terms of Use, policies established by the community, which may cover a single Project edition or multiple Projects editions (like the Global Ban Policy), may be modified by the relevant community according to its own procedures.

The blocking of an account or access or the banning of a user under this provision shall be in accordance with Section 12 of these Terms of Use.

Section 12: Though we hope you will stay and continue to contribute to the Projects, you can stop using our services any time. In certain (hopefully unlikely) circumstances it may be necessary for either ourselves or the Wikimedia community or its members (as described in Section 10) to terminate part or all of our services, terminate these Terms of Use, block your account or access, or ban you as a user. If your account or access is blocked or otherwise terminated for any reason, your public contributions will remain publicly available (subject to applicable policies), and, unless we notify you otherwise, you may still access our public pages for the sole purpose of reading publicly available content on the Projects. In such circumstances, however, you may not be able to access your account or settings. We reserve the right to suspend or end the services at any time, with or without cause, and with or without notice. Even after your use and participation are banned, blocked or otherwise suspended, these Terms of Use will remain in effect with respect to relevant provisions, including Sections 1, 3, 4, 6, 7, 9-15, and 17.

Twitter

Note: If your account appears to have engaged in repeated violations of the Twitter Rules, or has aggressively engaged with other accounts, you may not be presented with the option to verify by phone. In this case, you will only be able to use Twitter in a limited state for the specified time listed." <https://help.twitter.com/en/managing-your-account/locked-and-limited-accounts> - If someone repeatedly violates our Rules then our enforcement actions become stronger. This includes requiring violators to delete the Tweet(s) and taking additional actions like verifying account ownership and/or temporarily limiting their ability to Tweet for a set period of time. If someone continues to violate Rules beyond that point then their account may be permanently suspended.

26. Platforms should have in place mechanisms allowing to complement the above-mentioned measures with e.g. public apologies, commitments to review internal policies and processes,

IGF COALITION ON PLATFORM RESPONSIBILITY

which may be more effective and suitable to redress, in fulfillment of their corporate social responsibility to:

- make a policy commitment to the respect of human rights
- adopt a human rights due-diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights
- have in place processes to enable the remediation of any adverse human rights impacts they cause or to which they contribute