# Contributions for the democratic regulation of big platforms to ensure freedom of expression online

A Latin American perspective for content moderation processes that are compatible with international human rights standards

Elaboration









Support / Adhesion









This document contains measures that we have developed with input from various organizations, activists and experts over a number of months, in the hope that it will make a valuable contribution to the discussion of this issue. Although this proposal seeks to provide answers that are as comprehensive and specific as possible, it is still "under construction", for which reason we have opened it up to public consultation.

A solution for such a controversial and sensitive issue as regulation needs broader and more open processes with the participation of multiple stakeholders. Only then can it become the democratic frame of reference that we believe is necessary for the present times.

Collaborated with the writing (in a personal capacity):

## Javier Pallero

Public Policy Coordinator for Latin America, Access Now

### Joan Barata

Member of Plataforma de Defensa de la Libertad de Información

### Valeria Betancourt

Association for Progressive Communications (APC)

## Andrés Piazza

Consultant, ExLACNIC, ExLACTLD

#### Guillermo Mastrini

Research Professor at the National University of Quilmes (UnQ), University of Buenos Aires (UBA), Conicet

#### Martín Becerra

National University of Quilmes (UnQ), University of Buenos Aires (UBA), Conicet

## Juan Ortiz Freuler

Researcher affiliated with the Berkman Klein Center Berkman Klein Center (2019-2020)

## **ABSTRACT**

This document offers recommendations on specific principles, standards and measures designed to establish forms of public co-regulation and public regulation that limit the power of major Internet platforms (such as social networks and search engines).

The purpose of this effort is to protect users' freedom of expression and guarantee a free and open Internet. Such intermediaries increasingly intervene in online content, through the adoption of terms of service and the application of business moderation policies. Such forms of private regulation affect public spaces which are vital for democratic deliberation and the exercise of fundamental rights.

The proposal seeks to align with international human rights standards and takes into account existing asymmetries related to large internet platforms without limiting innovation, competition or start-up development by small businesses or community, educational or nonprofit initiatives.

## INTRODUCTION

The growing intervention of Internet platforms in the contents of their users has become an issue of concern throughout the world. In fact, "private control" is considered by none other than the Rapporteurs on Freedom of Expression<sup>1</sup> as one of the three main challenges over the next decade and a "threat to freedom of expression". For said rapporteurs, "a transformative feature of the digital communications environment is the power of private companies, and particularly social media, search platforms and other intermediaries, over communications, with enormous power concentrated in the hands of just a few companies".

This concern is nothing new, given that on many occasions both international bodies and digital rights organizations have questioned such practices and made recommendations for corporations to make a change in policies and practices in order to align with international human rights standards.<sup>2</sup>

For its part, the United Nations (UN) Office of the Special Rapporteur on Freedom of Opinion and Expression has published several reports on the issue<sup>3</sup>, while the Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights (IACHR) has observed for many years that "Intermediaries must thus keep their activities from provoking or helping to provoke negative consequences on the right to freedom of expression" in their voluntary measures for content moderation, which "can only be considered legitimate when those restrictions do not arbitrarily hinder or impede a person's opportunity for expression on the Internet."

There is also a growing interest among governments and legislative congresses, including both authoritarian regimes and consolidated democracies, to regulate activities and the distribution of content, particularly through the regulation of content disseminated via social networks. However, most of these legal initiatives configure solutions that are illegitimate or disproportionate, assigning responsibilities and obligations that transform the platforms into judges or even a private police force over the contents of third parties disseminated via the Internet.

The undersigned oppose such regulations and will continue to maintain this stance. However, we believe that the self-regulation model that has prevailed until now has run its course in the current development of the Internet, where a handful of corporations have centralized and concentrated the circulation, exchange or search for information and opinions,

\_

<sup>&</sup>lt;sup>1</sup> Joint Declaration: Challenges to freedom of expression in the next decade of the United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression; the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media; the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, 2019

<sup>&</sup>lt;sup>2</sup> Including the Santa Clara Principles

<sup>&</sup>lt;sup>3</sup> Internet Content Regulation, Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 2018

<sup>&</sup>lt;sup>4</sup> Freedom of Expression and the Internet, IACHR Special Rapporteur for Freedom of Expression 2013, Para. 111

thus distorting the idea of the decentralized, free and open Internet that we have long struggled to create.

In response to this polarized scenario of "corporate self-regulation vs. authoritarian or excessive regulation", several Latin American organizations believe that a third way is not only necessary but indeed possible. This third way involves building a proposal for democratic, adequate and intelligent regulation that can ensure adequate regulatory environments to protect human rights from the interventions of technological giants, while respecting international human rights standards.

The gatekeeper role of these companies requires that democratic societies set limits on such powers to ensure the effectiveness of historically recognized rights and freedoms, as well as the predominance of the general and public interest.

However, the proposals are not intended to cover all Internet intermediaries. Rather, they are limited to certain types of platforms and applications whose main purpose is to enable or facilitate access to information on the Internet and/or to provide support for expressions, communications and exchanges of content among its users (including social networks, search engines and video sharing platforms, but not messaging services).

A principle of "progressive regulation" is proposed based on the impact that the measures taken by intermediaries have on the exercise of fundamental rights on the Internet, in particular freedom of expression. That is, regulations should be stricter for major platforms that, for its size and reach, have become public spaces of deliberation and/or are the main entry routes for access to information, and which have an excessive level of concentration.

In view of these special characteristics, the aim is to create a regulatory environment that is appropriate for the functioning and characteristics of the Internet and that includes mechanisms of self-regulation, co-regulation and public regulation. This with the understanding that the challenges presented by the new digital scenario (including the speed and volume of information) do not allow for the application of unique and equal solutions as in other support systems for information and communication.

The document does not propose legislation that determines which content can be disseminated on the Internet. Nor does it require that platforms moderate their content. However, if they decide to do so, a series of conditions should be established so that their users' fundamental rights are not violated. Thus, proposals are included on what are the limits for content moderation that these platforms already implement, so that their terms of services, criteria and procedures are compatible with international human rights standards, taking specially into account the protection of minorities and vulnerable groups.

A democratic and balanced regulatory system should also protect platforms from the illegitimate pressures of governments and other stakeholders. As intermediaries, they are key to facilitating the exercise of these rights, and therefore the proposals include recommendations so that the regulatory frameworks allow such platforms to fulfill that role in

an appropriate manner: no legal responsibility for third-party content or prohibition of obliging them to undertake generic monitoring or supervision of contents represent some of these recommendations.

Private regulation of the Internet is produced and aggravated by a context of the robust concentration of power among a small handful of international corporations. Public regulation of the activities of these platforms should adopt anti-monopoly measures to counter such an environment of concentration and the absence of competition, but such proposals are not included in this document. Nevertheless, the simple fact that the main public spaces for the circulation of information and opinions can all be controlled by just one company, should oblige the antitrust bodies of the United States and other countries to act.

Neither have we included in this proposal important issues such as mechanisms to guarantee pluralism and diversity on the Internet or the tax issue. Rather, the document focuses on issues related to content moderation, offering principles that can be generally applied. Moreover, the specificities of certain services justify specific approaches. For example, for cultural services, positive obligations for the protection and promotion of cultural diversity should be adopted in line with the UNESCO Convention on the Protection and Promotion of the Diversity of Cultural Expressions.

Lastly, any norms and institutional designs that are adopted must be adequately developed. This should take into account the needs of market regulations subject to continuous development, the specific characteristics of the digital environment in each country, and the unique requirements of Latin America within the context of international human rights standards.

The document is organized in the following chapters:

- 1. Scope of Regulation
- 2. Service Terms and Conditions
- 3. Transparency
- 4. Application of Policies and Due Process
- 5. Right to defense and repair
- 6. Accountability
- 7. Approval and application of regulation

# 1. SCOPE OF REGULATION

- 1.1 This regulation proposal reaches online service providers when they act as intermediaries or storage platforms, search or exchange of information, opinions, expressions and other content generated by their users and that perform some type of healing or moderation of those contents (in this document: "content platforms"). Among them: search engines, social networks and other platforms for exchanging texts, images and videos.
- 1.2 Limits to the power of large content platforms should be structured on the basis of a co-regulation model, where self-regulation and public regulation structures are complemented to formulate legal, contractual and technical solutions that guarantee freedom of expression online, in balance with other fundamental rights<sup>5</sup>. Regulatory and regulatory instruments should be the result of a multisectoral governance process that takes into account local and regional contexts.
- 1.3 Platforms should directly incorporate into their conditions of service and their community standards the relevant human rights principles that ensure that the measures related to the content will be guided by the same criteria that govern the protection of expression by any means<sup>6</sup>. These principles include: transparency, accountability, due process, necessity, proportionality, non-discrimination and the right to defense and reparation. Platforms should also ensure full respect for consumer rights.
- 1.4 Content platforms that have significant market power in a given field of action<sup>7</sup> ("big content platforms") should have asymmetric regulation with respect to other providers, in view of the importance and impact that their business decisions may have on the exchange of information, opinions and cultural property, as well as the exercise of freedom of expression and public debate.
- 1.5 A smart regulation is one that considers in adequate and differentiated way content platforms that do not meet the above criteria and are not for profit purposes, the ones that have scientific or educational purposes, as well as those that bring together small and closed user groups of private and homogeneous condition.

<sup>6</sup> Regulation of online content generated by users, Special Rapporteurship on the Promotion of the Right to Freedom of Opinion and Expression of the United Nations, 2018

<sup>&</sup>lt;sup>5</sup> The institutional design and division of responsibilities is developed in chapter 7 of this proposal.

<sup>&</sup>lt;sup>7</sup> The definition of "Significant Market Power", made by an independent regulatory body, must take into account definitions of relevant markets updated for the current scenario, considering the specific function of the service, the dispute for advertising revenue and user attention time, the number absolute user and service substitutability by others.

# 2. SERVICE TERMS AND CONDITIONS

- 2.1 The terms of service (TOS) of all content platforms, as well as other complementary documents (such as guides or content application guidelines) should be written in a clear, precise, intelligible and accessible way for all users. Large platforms should also present them in the user's national language.
- 2.2 All content platforms should establish and implement TOS that are transparent, clear, accessible and in accordance with international human rights norms and principles, including the conditions under which interference with the right to freedom of expression or User privacy<sup>8</sup>. In particular, the user should be informed of the conditions that may lead to the termination of the contract (account drop, for example) as well as the elimination, de-indexing or significant reduction of the scope of their expressions and contents from unilateral modifications of algorithms of healing.
- 2.3 No content platform should be able to unilaterally change the terms of service and conditions of use, or apply new terms, without clearly informing the user of the justification and without giving, under reasonable notice, the possibility of canceling the contract, without having economic or legal consequences derivative of that contract<sup>9</sup>.
- 2.4 Platforms should not establish abusive or asymmetric conditions on the use and ownership of the content generated and published by its users, respecting their copyright in the same way that users should comply with regulations regarding content generated by third parties. In this regard, restrictions arising from copyright protection should consider the limitations and exceptions recognized in international treaties and national laws.
- 2.5 The terms of services should not grant unlimited and discretionary power to the platforms to determine the suitability of user-generated content<sup>10</sup>. In particular, the TOS that could imply limitations in the exercise of the right to freedom of expression and access to the information of its users should not be formulated in a vague or broad way, in such a way that they allow an arbitrary interpretation and application by the platform.

Freedom of Expression and Internet Platforms | CONTRIBUTIONS TO A DEMOCRATIC REGULATION

<sup>&</sup>lt;sup>8</sup> Freedom of expression and Internet, Office of the Special Rapporteur for Freedom of Expression IACHR, 2013, para. 112

<sup>&</sup>lt;sup>9</sup> Agreement between EU with Facebook, Google and Twitter (2018) - "Better social media for European consumers" <sup>10</sup> Last sentence taken from the EU Agreement with Facebook, Google and Twitter in 2018 "Better social media for European consumers"

- 2.6 Regarding the curation / prioritization of the visualization of the content generated by its users (in news feeds, search results, news access services and other similar), large platforms should:
  - A. Transparent the criteria used by the algorithms for ordering or addressing, if possible, explaining the effects to the user.
  - B. Do not use arbitrary or discriminatory criteria that could illegitimately affect the freedom of expression and the right to information of its users.
  - C. Provide customized filtering mechanisms in a clear, transparent, explicit, revocable / editable manner and under user control, so that they decide what content they want to prioritize and how they want to do it (eg, chronological order).
  - D. Respect the user's right to know and define which of their personal data are collected and stored, and how they are used in the content addressing, respecting the principle of informational self-determination.
- 2.7 In case the large platforms decide, by themselves, to incorporate in their TOS certain restrictions and even prohibitions to the publication of contents or expressions generated by their users they could only do so with the following limitations, so that they are compatible with the international standards of human rights:
  - A) They could prohibit, even by automatic filtering, those contents that are clearly and manifestly illegal and that, at the same time, are recognized as legitimate limitations to freedom of expression in international human rights declarations or treaties such as: sexual abuse or exploitation of minors, or propaganda in favor of war and any apology of national, racial or religious hatred that constitutes incitement to violence or any other similar illegal action against any person or group of people, for any reason, including those of race, color, religion, language or national origin<sup>11</sup>.
  - B) They could restrict, as a non-definitive precautionary measure, contents that, even if they are not recognized as illegal, cause serious, imminent and irreparable damage or difficult reparation, to other persons such as: unauthorized dissemination of sexual content or acts of violence or explicit and excessive or aberrant cruelty. In these cases, the list and definitions of restricted content should be included in the TOS in a restrictive, clear, precise manner and considered, in the analysis of the measure to be taken, the context of the expression published and not used in the framework of legitimate expressions (educational, informative or other content).

<sup>&</sup>lt;sup>11</sup>American Convention on Human Rights, art. 13, inc. 5

- C) Contents such as cyberbullying or explicit and abusive drug use could be restricted to specific audiences, such as children and teenagers.
- D) Any other measure of prioritization or restriction to expressions and other content generated by its users that the platform could consider —for commercial or other reasons— "offensive", "inappropriate", "indecent" and similar vague or broad definitions that could illegitimately affect the Freedom of expression, large platforms should provide mechanisms and notices for other users—voluntarily and based on their moral, religious, cultural, political or other preferences—who decide what content they want to have access to and what not. These contents should not be prohibited, deleted or reduced in scope by default disproportionately affecting the right to freedom of expression of its users.

## 3. TRANSPARENCY

- 3.1 Platforms should publish their content restriction policies online, in clear language and in accessible formats, keep them updated as they evolve, and notify users of changes as appropriate<sup>12</sup>.
- 3.2 When content is restricted in a product or service of the intermediary that allows to display a notice when trying to access, the intermediary should display a clear notice that explains what content was removed and why<sup>13</sup>.
- 3.3 In the actions of prioritization of online content accessible to the user (feeds, search results and others) the commercial nature of the communications, the sponsored content as well as the electoral or political propaganda should be clearly identified, identifying the contracting party and without raising doubts about its meaning <sup>14</sup> and being transparent about the content metadata (prices, etc.).
- 3.4 Platforms should inform their users in a clear, explicit and accessible way<sup>15</sup>, at least on what respects to:
  - A. What types of content and activities are prohibited in your services?
  - B. What are the criteria and mechanisms for curation and moderation of content? Which are directly controlled by the user and which are not?
  - C. In what cases, when and how does content analysis automation apply?<sup>16</sup>
  - D. In what cases, when and how does the human review of content apply? This question makes particular reference to the criteria for decision making to avoid affecting human rights, taking into account the context, the wide variation of idiomatic nuances and the meaning and the linguistic and cultural peculiarities of the contents subject to a possible restriction<sup>17</sup>
  - E. How many moderators do you have, describing in detail your professional profile (experience, specialization or knowledge), your spatial location and your distribution of tasks (in terms of themes, geographical areas, etc.)?
  - F. What are the rights of users regarding the content generated and published by themselves and the policies applied by the company in this regard?

<sup>13</sup> Manila Principles

<sup>17</sup> Ídem

<sup>&</sup>lt;sup>12</sup> Manila Principles

<sup>&</sup>lt;sup>14</sup> Based on Agreement between EU with Facebook, Google and Twitter, 2018 "Better social media for European consumers"

<sup>&</sup>lt;sup>15</sup> To "allow users to predict with reasonable certainty what content places them on the dangerous side of the line" (Regulation of content on the Internet, Special Rapporteurship on the Promotion and Protection of the Right to Freedom of Opinion and Expression, 2018)

<sup>&</sup>lt;sup>16</sup> Regulation of content on the Internet, Special Rapporteurship on the Promotion and Protection of the Right to Freedom of Opinion and Expression, 2018

# **APPLICATION OF POLICIES AND DUE PROCESS**

- 4.1 In the design and application of their community content management policies, platforms should seek that any restriction arising from the application of the terms of service does not unlawfully or disproportionately restrict the right to freedom of expression<sup>18</sup> for which they must respect the requirements of searching for an imperative purpose, as well as the need, suitability and proportionality of the measure to achieve the intended purpose<sup>19</sup>.
- 4.2 The criteria for making decisions, so as not to affect human rights, should take into account the context, the wide variation of idiomatic nuances and the meaning and the linguistic and cultural peculiarities of the contents subject to a possible restriction<sup>20</sup>.
- 4.3 In addition, in the analysis of the content restriction measures applicable in each case, the principles of proportionality and progressivity should be respected, weighing the severity of the damage, the recurrence of the violations and the impact that such restrictions could have on the Internet capacity to guarantee and promote freedom of expression with respect to the benefits that the restriction would bring to the protection of other rights<sup>21</sup>.
- 4.4 Users should always have the right to have the content restriction decisions of the platforms themselves respected in due process, particularly when it comes to measures that could affect their right to freedom of expression. As a general principle, and except for duly justified exceptional cases, people affected by a restriction or interference measured by the platforms and, where appropriate, the general public, must be notified in advance about the restriction measures that affect them<sup>22</sup>.
- 4.5 In view of the aforementioned principles of necessity and proportionality, in case of possible breaches of the TOS the platforms should adopt less burdensome measures than the removal or others of similar effects, opting for warning or notification mechanisms, flagging, linking with opposing information or other.
- 4.6 The most drastic unilateral measures that were adopted without notice or due to prior process, such as lowering of accounts or profiles, removal of content, or other

<sup>20</sup> Tomado del Informe 2018 David Kaye (en el original hace referencia a recomendaciones de transparencia a las plataformas

<sup>&</sup>lt;sup>18</sup> Freedom of expression and Internet, Office of the Special Rapporteur for Freedom of Expression IACHR, 2013, para. 112

<sup>&</sup>lt;sup>19</sup> Ídem, para. 55

<sup>&</sup>lt;sup>21</sup>Freedom of expression and Internet, Office of the Special Rapporteur for Freedom of Expression IACHR, 2013, para. 54 <sup>22</sup> Ídem, para. 115

measures that have a similar impact of exclusion from the possibilities of participating in the platform should be taken by large platforms only under the following conditions:

- A. When dealing with non-arbitrary or discriminatory technical management interventions (such as spam, fake accounts, malicious bots or similar),
- B. When dealing with duplicates or reiterations in "crude" (not commented on or edited for journalistic or informative purposes or other legitimate purposes) of other contents and expressions of obvious illegality that were already restricted after human evaluation following the aforementioned standards,
- C. When the following situations are identified:
  - a. The grounds set forth in section 2.7 A;
  - the fulfillment of orders of competent authorities of immediate withdrawal and the consummation of common crimes already typified in national legislation;
  - c. serious, imminent and irreparable damage or of difficult reparation to the rights of other persons as in the cases listed in point 2.7 literals B and C.

In all these cases, except in the case of orders from competent authorities, the platform should proceed to the immediate subsequent notification, with the possibility of appeal for a possible revision of the measure under the terms of chapter 5 of this document.

- 4.7 Upload filtering and filtering is only legitimate and compatible with international human rights standards when it comes to child pornography<sup>23</sup> or in the first two situations described in the previous point. Otherwise, it should be considered as an act of censorship, under the terms established by the American Convention on Human Rights.
- 4.8 Any other measure of restriction of contents or expressions that the platform intends to adopt in the event of a possible breach of the TOS or the complaint of third parties (for example with regard to an impact on copyright), the content questioned should be kept published on the platform until a final decision arising from a due process where, after notification a) the voluntary withdrawal of the questioned content was promoted b) the exercise of the right to defense of the user will be guaranteed allowing a counter-notification with filing of discharges, before taking a decision.
- 4.9 No content platform should have legal responsibility for content generated by third parties, as long as they do not intervene by modifying or editing those contents, nor do they refuse to execute judicial orders or from competent and independent official authorities that comply with appropriate due process guarantees.

\_

<sup>&</sup>lt;sup>23</sup>American Convention on Human Rights, art. 13, inc. 4

4.10 Content platforms can only be held responsible depending on their actions or negligence in the prioritization or active promotion of expressions that could affect the rights of third parties if they depart from the principles established in point 2.6. In these cases the legal responsibility that may correspond to the platforms for third party expressions or healing activities should not be of the objective type.

# 5. RIGHT TO DEFENSE AND REPAIR

- 5.1 All content platforms should clearly explain to users why their content has been restricted, limited or removed; or your account or profile suspended, blocked or deleted:
  - A. The notifications should include, at least, the specific clause of the community rules that the user is supposed to have violated.
  - B. The notification should be detailed enough to allow the user to specifically identify the restricted content and should include information on how the content or account was detected, evaluated and deleted or restricted.
  - C. People should have clear information on how to appeal the decision<sup>24 25</sup>.
  - 5.2 Content platforms should not delete publications or other user-generated content without being notified, without providing clear justification and without giving users the possibility to appeal<sup>26</sup>, so that they can exercise their right to defense and prevent abuse. In this regard, the platforms must provide users with the opportunity to appeal content moderation decisions, under the following conditions:
    - A. Appeal mechanisms should be very accessible and easy to use.
    - B. Appeals should be subject to review by a person or panel of people who were not involved in the initial decision.
    - C. Users should have the right to propose new evidence or material to be considered in their opinion.
    - D. Appeals should result in prompt determination and response to the user.
    - E. Any exception to the principle of universal appeals should be clearly disclosed and compatible with international human rights principles<sup>27</sup>.
- 5.3 Users affected by any measure of restriction of their freedom of expression as a result of the decisions of the platforms, depending on the specific regulations of domestic law, must have the right to access legal resources to dispute said decision and reparation mechanisms in relation to the possible violation of your rights<sup>28</sup>. In this regard, content platforms may not prevent their users from taking legal action against

<sup>25</sup> Manila Principles – "The notification on content restriction decisions adopted by a platform must, at a minimum, have the following information:

The reasons why the content in question violates the intermediaries restriction policies.

The Internet identifier and a description of the alleged violation of the content restriction policies.

The contact details of the issuing party or its representative, unless this is prohibited by law.

A statement in good faith that the information provided is accurate "

<sup>26</sup> EU agreement with Facebook, Google and Twitter in 2018 "Better social media for European consumers"

<sup>&</sup>lt;sup>24</sup> Santa Clara Principles

<sup>&</sup>lt;sup>27</sup> Santa Clara Principles

<sup>&</sup>lt;sup>28</sup> Freedom of expression and Internet, Office of the Special Rapporteur for Freedom of Expression IACHR, 2013, para. 115

them in their country of residence, which would imply a denial of their right to access Justice <sup>29</sup> in a subsidiary or parallel way to claims through the internal appeal mechanisms. For this purpose, the contract concluded between the user with a content platform must expressly include that the disputes will be governed by the law and the Justice of the country where the user has his/her habitual residence and not by the place where the offices of the platform are located<sup>30</sup>.

-

<sup>&</sup>lt;sup>29</sup> EU agreement with Facebook, Google and Twitter in 2018 "Better social media for European consumers"

<sup>&</sup>lt;sup>30</sup> Ídem

# **ACCOUNTABILITY**

- 6.1 Content platforms should publish transparency reports that provide specific information about all content restrictions adopted by the intermediary, including actions taken before government requests, court orders, private requirements, and on the implementation of their restriction policies of content<sup>31</sup>.
- 6.2 Content platforms should issue periodic transparency reports on the application of their community rules that include at least:
  - A. Full data describing the categories of user content that are restricted (text, photo or video; violence, nudes, copyright violations, etc.), as well as the number of pieces of content that were restricted or deleted in each category, detailed by country<sup>32</sup>.
  - В. Data on how many content moderation actions were initiated by a user's report (flag), a trusted flagger program or by the proactive application of community standards (for example, through the use of a machine learning algorithm)<sup>33</sup>.
  - C. Data on the number of decisions that were effectively appealed or determined to have been made in error<sup>34</sup>.
  - Data reflecting whether the company performs a proactive audit of its D. non-appealed moderation decisions, as well as the error rates that the company found<sup>35</sup>.
  - E. Aggregated data that illustrate trends in the field of monitoring compliance with standards and examples of real cases or detailed hypothetical cases that clarify the nuances of the interpretation and application of specific standards<sup>36</sup>.

<sup>31</sup> Manila Principles

<sup>32</sup> Santa Clara Principles

<sup>33</sup> Santa Clara Principles

<sup>34</sup> Santa Clara Principles

<sup>35</sup> Santa Clara Principles

<sup>&</sup>lt;sup>36</sup> Regulation of content on the Internet, Special Rapporteurship on the Promotion and Protection of the Right to Freedom of Opinion and Expression, 2018

## 7. APPROVAL AND APPLICATION OF REGULATION

- 7.1 While these are measures that could affect fundamental rights, the substantive aspects of the regulation proposed in this document should be adopted beforehand and by formal law, that is, a law approved by the legislative body (Congress, Parliament, National Assembly or similar), after public and open consultation. When necessary, regulatory delegations in enforcement agencies should be carefully established by law.
- 7.2 Content platforms should not depend on licenses for their operation in a given country, but there must be an obligation to identify legal officers and effective forms of communication and response for users and the respective authorities such as an email account, an electronic form, or equivalent means<sup>37</sup>.
- 7.3 Content platforms should not be obliged to monitor or supervise content generated by third parties, in a generic way, in order to detect alleged current violations of the law or to prevent future infractions<sup>38</sup>.
- 7.4 The operation of the content platforms should be framed in an environment of coregulation appropriate to the characteristics of the digital environment:
  - A. The principles and standards included in this proposal should be included by the content platforms in their terms of service and other complementary documents (such as guidelines);
  - B. Platforms should apply these principles and standards without prior intervention by state agencies;
  - C. Platforms should have internal and effective mechanisms of appeal, as well as external bodies independent of the companies for the review of cases and policies adopted, with the understanding that state regulation should act only when the instances of self-regulation do not work;
  - D. There should be a specialized regulatory body that operates with sufficient guarantees of independence, autonomy and impartiality and that has the capacity to evaluate the rights at stake and offer the necessary guarantees to the user<sup>39</sup> on the policies and practices of implementation of the terms of service of the platforms, and indicate the adequacy of them when appropriate. In case of non-compliance with the obligations of transparency, due process, right to defense and others, the agency must have sufficient enforcement capacity, being able to apply

٦.

<sup>&</sup>lt;sup>37</sup> Extracted from the draft Law of Intermediaries of Argentina, art. 7

<sup>&</sup>lt;sup>38</sup> Extracted from the draft Law of Intermediaries of Argentina, art. 5

<sup>&</sup>lt;sup>39</sup>In terms expressed in Freedom of expression and the Internet, Office of the Special Rapporteur for Freedom of Expression IACHR, 2013, para. 56

- sanctions, if necessary. However, it should not evaluate or have a binding decision in individual cases;
- E. Individual cases where there is a violation of user rights and that are not satisfactorily resolved within the internal scopes and mechanisms for dispute resolution should be resolved by judicial bodies, Public Defenders or similar independent and specialized public bodies. -of the country where the user has his habitual residence- by means of an abbreviated procedure, of digital procedure and electronic notification (fast track) with guarantees of revision by an impartial authority. Other authorities or state agencies should not be able to force platforms to remove or process specific content.

#### Elaboration









Support / Adhesion







