CENTER FOR
DEMOCRACY
& TECHNOLOGY

Internet Universality Indicators: Online Consultation Phase II

Response from the Center for Democracy & Technology

March 15, 2018

The Center for Democracy & Technology is a public interest advocacy organization that works to promote human rights and civil liberties in technology policy around the world. CDT has been working for nearly 25 years to promote law and policy frameworks that support individuals' access to information and opportunities for speech online. In our view, these key policies include substantive and procedural protections for fundamental rights, limits on the liability that internet intermediaries face for hosting and facilitating access to user-generated content, and guarantees of network neutrality to ensure that the choice of what information to access and what services to use remains in the hands of the individual. We commend UNESCO for undertaking this effort to enable coherent assessment of the environment for free expression online in countries around the world, and welcome the opportunity to comment on the Internet Universality Indicators.

Are there any additional themes, questions or indicators which you believe should be included in the framework?

Content Monitoring, Filtering, and Removal

One area of internet freedom that would benefit from a greater emphasis in the set of proposed questions is the issue of obligations on internet intermediaries to monitor, filter, remove, or otherwise censor content. These requirements fall into two categories:

> 1) Mandates that companies block content from being made publicly available (i.e., to filter out the content before it reaches the platform); and
>
> 2) Mandates that companies remove content after being notified of its existence (e.g., traditional notice-and-takedown schemes).

One of the most prominent examples of the first scheme can be found in Article 13 of the EU's proposed Copyright in the Digital Single Market Directive, under which companies would have an obligation to ensure that copyright-infringing posts are not re-posted on their platforms. If companies are held liable for user-posted content (as the Directive proposes), they will be more inclined to "play it safe," removing content if it has *any* potential—no matter how small—to invoke the provisions of Article 13. The end result of this, or any, filtering scheme is the excessive removal of content before it has a chance to reach an audience.

Under the second scheme, it is increasingly common for states to propose that requirements for companies to remove particular content be coupled with a specified time limit. For example, in Germany, the NetzDG law mandates that companies remove illegal hate speech within 24 hours or face a €50 million fine. While there is an exception for exceedingly difficult scenarios, the quick turnaround time required raises serious concerns about the chilling effects of filtering content and responding to reports of alleged illegal speech: Companies forced to comply with a given time frame may well default to removing flagged content—legal or illegal—in the interest of complying with the law, rather than risk the liability of taking too long to decide or making a mistaken judgment. While the NetzDG law is one example, public statements from high-profile politicians in other countries, as well as recommendations from the European Commission, suggest that the question of time limits is an important dynamic to consider as countries search for new approaches to curb terrorism.

Both types of content removal obligations pose their own challenges and, therefore, the two might best be addressed in separate questions with separate (though likely similar) indicators.

Artificial Intelligence and Automation

One unfortunate consequence of the combined advances in technology and the embrace of that technology by terrorist groups is the misguided belief that internet and tech companies must have (or be close to developing) a new-fangled service or product capable of automating the search and removal of illegal content online. There have been repeated calls from a variety of politicians around the world for tech companies to remove terrorist speech quickly and *en masse*.

These calls assume, however, that there is or will be a technical solution to the daunting, complex task of accurately distinguishing “terrorist propaganda” or “illegal hate speech” from lawful expression. The reality, however, is that such technology simply does not yet exist—and that it may never exist. While automation and artificial intelligence are good at a number of tasks, developing an effective machine-learning tool requires a clear definition of the material to be identified or the issue to be resolved. Given the variability of national definitions of propaganda and hate speech; the differences in communication styles and patterns across different online and offline communities; and the challenge of interpreting context in online discussion fora that combine local, regional, and international audiences, it is likely that there will never be a magic-bullet AI tool that can accurately identify terrorist propaganda. In recognition of the role that censorship and freedom of expression will play in any measure of internet universality, UNESCO may consider devoting thoughtful attention to the actual—rather than perceived—capabilities of bots and AI as the organization fleshes out these IU indicators.

Transparency

We are glad to see that the indicators recognize the fundamental importance of transparency in fostering a rights-respecting approach to internet policy. One further addition to the document might take the form of a question or two that address transparency around government requests to private companies. One question might get at the larger issue of gag orders and what companies are or are not allowed to say, by including indicators that ask whether there are restrictions on reporting around:

1. Government requests for user information;
2. Government requests for direct access to companies' networks
3. Government requests for content removal.

The flip side—government reporting of this information—presents an opportunity for a second question. There is a lack of consistent, regular reporting by governments worldwide about the content they target for removal. This lack of transparency can easily translate to a lack of accountability. For that reason, a question that addresses the importance of "transparency reports" published by government agencies would be a strong addition to the document.

Due Process

Underscoring the significance of the above issues is the role of due process. One of the trends that we see around the world is governments increasingly putting companies in the decision-making position when it comes to application of law. These kinds of determinations are beyond private companies' capacity to make, and creating these structures breaks the link between the individual and his or her government, eliminating crucial opportunities for citizens to hold their governments accountable for the application of the law and seriously risking the loss of any due process guarantees for individuals who are censored. While a number of the suggested questions address or otherwise take into account due process, we encourage UNESCO to ensure that the significance of due process guarantees is clear throughout.

Are there any suggestions that you wish to make in respect of the proposed themes, questions and indicators which are included in the framework as it stands?

Rights

A3: We fully support the inclusion of a question that addresses the role of due process. In our work, however, we (and others) have identified a gap between the stated availability of a means of redress on a given platform and its actual existence and/or use. Any evaluation of due processes and terms of service will necessitate capturing that gap in some way.

B3: This is potentially a key question in UNESCO's consideration of the state of internet universality in any given country. For that reason, greater specificity and consideration may be necessary. For example, it is unclear whether the question and indicators are aimed at censorial actions by state actors, non-state actors, or both. Further, the single indicator here could draw out any number of distinctions or issues for evaluators: How will “significant” censorship be defined or identified? Are there multiple qualitative and quantitative sources that cover all of the potential countries to be evaluated? This question presents a great opportunity for UNESCO to dig deep into the role of state and non-state censorship in any given country. For that reason, we encourage UNESCO to ensure that this question reflects that importance.

B5: A major factor in the number of individuals who use social media and/or generate online content is access and availability of the necessary tools. For that reason, this question may fit better under the “Accessibility to All” goal. In addition, we want to flag that measuring internet users can be exceedingly difficult, particular with respect to making distinctions between active users, registered users (on a site), passive users and/or lurkers. Given this, we encourage UNESCO to further clarify what measurement(s) this question is aimed to address.

B6: As with Rights question B5 (above), this question may also fit better under the “Accessibility to All” goal. Regardless, this question could benefit from greater specificity and/or elaboration. For example, “low-cost online services” could describe a number of things: Low-cost ISPs, cyber cafes, mobile service providers, web hosts, etc. UNESCO should further elaborate on and specify the breadth and meaning of “low-cost online services” to be considered under this question.

C2: We fully support the inclusion of a question that addresses both formal and informal restrictions on internet access and use. The question might be improved by clearly describing the focus and elaborating on the ways in which government actors can engage in censorship. As written, the question could speak to direct government demands on intermediaries *and/or* the existence of a liability framework (which can act as an indirect form of censorship). We suggest rewording the question to the following: “Does the government block or filter access to the internet or to specific online services, applications or websites—either directly (e.g., through licensing requirements or other laws) or indirectly—and on what grounds is this exercised?”

C3: If possible, this question and the associated indicators should capture the reality that in some countries, journalists (as well as non-journalists, citizens, bloggers, etc.) face intimidation and detention for their work, however, the government in question may come up with a pretext for detaining the journalist rather than explicitly stating that the individual is being punished for accessing information online.

C4: There is no Crosscutting D7 indicator as referenced.

Openness

A2: We suggest that government be included alongside "business, academic, and civil society.” While government may be seen as the source of the regulatory framework(s) in question, government can also have a major role in innovating and developing opportunities online (e.g., e-governance).

A3: As written, it is unclear if this question is getting at the presence of an anti-competition regulatory framework or simply addressing whether there are minimum requirements for those looking to establish ISPs and internet-enabled services. We suggest rewording and/or adding an additional indicator that will clarify the intent of this indicator.

Accessibility to All

A1: This question could benefit from clarity regarding its intent/purpose. As written, the question could be read as speaking to the need for surveillance of internet access and use and/or as speaking to the need for high-quality survey and census data about internet access and use.

Multistakeholder Participation

C2: Consider including (explicitly naming) the Freedom Online Coalition as one example of a major international fora and organization focused on ICTs and the internet.

Cross-Cutting Indicators

E4: UNESCO might consider whether there is an objective way to measure self-reported instances of harassment or abuse, and if there is a way to do so consistently across the globe. An alternative approach could note that conversations around this issue go hand in hand with (potentially very risky) policing and monitoring obligations on platforms. A rights-respecting way of dealing with this issue would look for resources for law enforcement and voluntary efforts/best practices by companies to be responsive to issues their users face. For example, are law enforcement officials adequately trained and resourced to pursue cases and claims of harassment and abuse online?

What sources and means of verification would you recommend, from your experience, in relation to any of the questions and indicators that have been proposed?

Research and Studies

- Ranking Digital Rights Corporate Accountability Index
  rankingdigitalrights.org

- Digital Standard
  thedigitalstandard.org
- Freedom House’s (Annual) *Freedom on the Net* Report
  freedomhouse.org/report-types/freedom-net
- Freedom House’s (Annual) *Freedom of the Press* Report
  freedomhouse.org/report-types/freedom-press

Guiding Principles Documents

- Universal Declaration of Human Rights, Article 19
  un.org/en/universal-declaration-human-rights/
- International Covenant on Civil and Political Rights (ICCPR)
  ohchr.org/en/professionalinterest/pages/ccpr.aspx
- Manila Principles on Intermediary Liability
  manilaprinciples.org
- Tallinn Agenda for Freedom Online
  freedomonline.ee/foc-recommendations
- UN Guiding Principles on Business and Human Rights
  ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf
- European Convention on Human Rights
  echr.coe.int/Documents/Convention_ENG.pdf
- GNI Principles on Freedom of Expression and Privacy
  globalnetworkinitiative.org/principles/index

Additional Resources

- Access Now’s Shutdown Tracker
  accessnow.org/keepiton/#take-action
- Berkman Klein Center for Internet & Society’s Internet Monitor
  https://cyber.harvard.edu/research/internetmonitor