BSR®

HUMAN RIGHTS ASSESSMENT

Global Internet Forum to Counter Terrorism

Contents

1. Executive Summary 3
2. Project Overview and Methodology 13
3. GIFCT Operating Context 19
4. GIFCT Mission and Goals 22
5. Human Rights Impacts 25
6. Terrorist and Violent Extremist Content 32
7. Content Removal and Preservation 37
8. Theory of Change and Programmatic Priorities 48
9. GIFCT Membership 52
10. Stakeholder Engagement 58
11. Governance, Accountability, and Transparency 62
12. Organizational Issues 67
13. Conclusion 70

ABOUT THIS REPORT

The Global Internet Forum to Counter Terrorism (GIFCT) commissioned BSR to conduct a human rights assessment of its strategy, governance, and activities. The purpose of this assessment is to identify actual and potential human rights impacts (including both risks and opportunities) arising from GIFCT's work and make recommendations for how GIFCT and its participants can address these impacts. BSR undertook this human rights review from December 2020 to May 2021.

This assessment combines human rights assessment methodology based on the UN Guiding Principles on Business and Human Rights (UNGPs) with consideration of the human rights principles, standards, and methodologies upon which the UNGPs were built. This review was funded by GIFCT, though BSR retained editorial control over its contents.

ACKNOWLEDGMENTS

This report was researched and written by Dunstan Allison-Hope, Lindsey Andersen, and Susan Morgan. BSR wishes to thank all GIFCT staff, members, and stakeholders for their participation in this assessment.

DISCLAIMER

The conclusions presented in this report represent BSR's best professional judgment, based upon the information available and conditions existing as of the date of the review. In performing its assignment, BSR must rely upon publicly available information, information provided by GIFCT, and information provided by third parties. Accordingly, the conclusions in this report are valid only to the extent that the information provided to BSR was accurate and complete. This review is not intended as legal advice, nor is it an exhaustive review of legal or regulatory compliance. BSR makes no representations or warranties, express or implied, about the business or its operations. BSR maintains a policy of not acting as a representative of its membership, nor does it endorse specific policies or standards. The views expressed in this publication are those of its authors and do not necessarily reflect those of BSR members.

SUGGESTED CITATION

BSR, 2021. "Human Rights Assessment: Global Internet Forum to Counter Terrorism."

1.Executive Summary

IN THIS SECTION

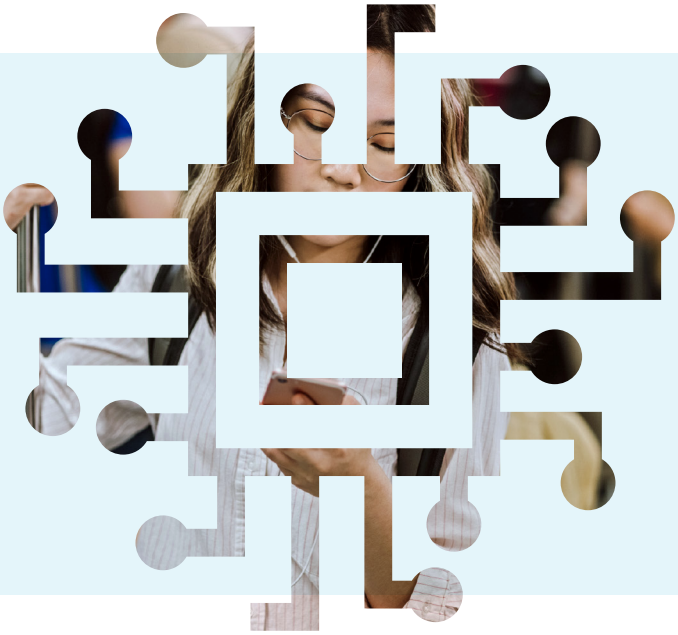
1.1 Project Overview

1.2 Project Methodology

1.3 GIFCT Operating Context

1.4 Summary of Analysis

1.5 Summary of Recommendations



1.1 Project Overview

The Global Internet Forum to Counter Terrorism (GIFCT) states that the protection and promotion of human rights—meaning the rights of the victims of terrorism and violent extremism, and those impacted by efforts to address terrorism and violent extremism—is central to its efforts. With this in mind, GIFCT commissioned this human rights assessment to:

- Identify actual and potential human rights impacts (including both risks and opportunities) arising from GIFCT's work.
- Make recommendations for how GIFCT and its participants can address these impacts.

The desired outcome is that this assessment (1) provides GIFCT with the knowledge, insights, and actions recommended to embed human rights into its strategy, governance, and actions; (2) establishes a clear understanding of the roles and responsibilities of different participants in GIFCT as they relate to human rights; and (3) increases trust and collaboration between participating organizations and with other stakeholders with an interest in GIFCT.

The scope of this assessment is GIFCT, not the actions of individual GIFCT member companies. Further, the scope of this assessment is primarily forward looking, rather than a review of prior activities.

1.2 Project Methodology

BSR deployed a methodology that combined a human rights assessment approach based on the UN Guiding Principles on Business and Human Rights (UNGPs) with a deeper consideration of the human rights principles, standards, and methodologies upon which the UNGPs were built.

Given the role of governments in GIFCT, we have considered both the "first pillar" of the UNGPs (the state duty to protect human rights) and the "second pillar" (the corporate responsibility to respect human rights). Key assumptions in BSR's methodology include:

All human rights are potentially relevant for GIFCT.

BSR used the international legal human rights framework as the baseline, including the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and the International Covenant of Economic, Social, and Cultural Rights, as well as other international human rights instruments and standards of international humanitarian law relevant to GIFCT.

Human rights are interconnected.

All human rights are indivisible, interdependent, and interrelated—the improvement of one right can facilitate advancement of others; the deprivation of one right can adversely affect others.

A stakeholder-inclusive process is essential.

Effective human rights due diligence requires meaningful engagement with rightsholders whose human rights may be impacted, or such reasonable alternatives as independent expert resources, human rights defenders, and other representatives from civil society. In addition to GIFCT member companies, BSR engaged with around 40 individuals and organizations in workshops, interviews, and written submissions.

Vulnerability must be prioritized.

A human rights-based approach implies greater attention should be paid to impacts on individuals from groups or populations that may be at heightened risk of vulnerability or marginalization. In this context, both the victims of terrorism and violent extremism and the victims of efforts to address terrorism and violent extremism are especially important.

BSR undertook this assessment from December 2020 to May 2021. This review took place concurrently with GIFCT reviewing its overall strategy and plan, and the observations and recommendations in this human rights review informed GIFCT's thinking in real time.

1.3 GIFCT Operating Context

GIFCT exists in the context of a growing number of institutions and organizations spanning national and international boundaries designed to counter terrorism. The lack of a globally agreed upon definition of terrorism and the highly politicized context within which counterterrorism takes place have resulted in government overreach and a disproportionate focus on organizations that have distorted Islam. This creates bias throughout the system in a way that has negative human rights impacts on freedom of expression, privacy, freedom of religion, freedom from discrimination, freedom of association and assembly, and the ability to participate in government, among other rights.

Further, while preventing terrorists and violent extremists from exploiting digital platforms clearly has a significant role to play in countering terrorism, there is also a tendency to overemphasize the role of the internet at the expense of attention on other systematic political, economic, and diplomatic shortcomings.

Adding to the challenging subject matter addressed by GIFCT are important questions posed by the concept of multi-stakeholderism. Three issues are particularly noteworthy: placing affected communities at the heart of multi-stakeholder efforts is both important and hard to achieve given the challenges of limited resources and often complex processes; power and information asymmetries are a reality in multi-stakeholder initiatives; and independence can be difficult to achieve in the early years of new organizations.

This assessment does not attempt to address the full range of challenges that exist in the counterterrorism field or in multi-stakeholder initiative design; however, these challenges exist as a backdrop for this assessment, and many of BSR's observations and recommendations are directly relevant to addressing them.

The lack of a globally agreed upon definition of terrorism and the highly politicized context within which counterterrorism takes place have resulted in government overreach and a disproportionate focus on organizations that have distorted Islam. This creates bias throughout the system in a way that has negative human rights impacts.

1.4 Summary of Analysis

BSR's analysis is organized around nine key themes and 35 questions, and these are addressed in detail in the main body of this report.

Mission and goals

While the specific purpose of GIFCT is to prevent terrorists and violent extremists from exploiting digital platforms, in doing so GIFCT enhances the protection, fulfillment, and realization of human rights. In other words, human rights for GIFCT is about more than simply "avoiding harm" while pursuing its mission. We conclude that GIFCT would benefit from a clearer description of the interdependent relationship between human rights and the GIFCT mission that conveys human rights as a deeply embedded, complementary, and reinforcing objective in counterterrorism and violent extremism efforts.

Human rights impacts

The main human rights impacted by GIFCT's activities include life, liberty, and security of person; nondiscrimination and equality before the law; access to effective remedy; freedom of opinion, thought, conscience, and religion; freedom of expression; freedom of assembly, and association; and privacy. We note that GIFCT is one step removed from human rights impacts—since they largely result from a member company's own activities and decisions, rather than from those of GIFCT itself—but emphasize the importance of GIFCT maintaining a system of human rights due diligence, embedding human rights across its activities, and engaging with affected stakeholders.

A human rights-based approach requires a clear understanding of which rightsholders are impacted by GIFCT activities, and these include both the actual and potential victims of terrorism and violent extremism, and victims of efforts to counter terrorism and violent extremism.

Terrorist and violent extremist content

The lack of consensus around definitions of terrorist and violent extremist content, and the prevalence of bias in the counterterrorism field—manifested in a disproportionate focus on Islamist extremist content—influence GIFCT's human rights impacts. Here we conclude that the multi-stakeholder status of GIFCT provides an opportunity to create a common understanding of terrorist and violent extremist content. We explore the benefits that could arise from this common understanding, such as pushing back against overbroad definitions deployed by governments, improving the capability of smaller companies without extensive policy teams to establish their own definitions, and establishing a bulwark against "slippery slope" definitions that may extend too far into other forms of speech.

Content removal and preservation

Several of the human rights impacts we identify concern the risk of overbroad removal of content by companies making use of the hash-sharing database or URL sharing, resulting in adverse impacts on freedom of expression and other rights. Here we consider a range of restrictions and controls on who should be able to add hashes to the database, the ability to challenge and remove hashes, maintaining a repository of content tied to hashes in the database, governance and oversight of the database, and transparency.

This section also considers questions arising from the need to preserve removed content that may have value as evidence in legal and access-to-remedy processes at a later date. Here we believe the most appropriate role for GIFCT is to advocate in favor of a "legal carve-out" or "safe harbor" for companies storing removed content that may be used as evidence later.

Theory of change and programmatic priorities

Terrorism and violent extremism are "whole of society" problems that require whole of society solutions, of which GIFCT is one part. While recognizing that there will always be resource constraints, BSR's instinct is toward a strategic, holistic, and systematic approach to preventing terrorists and violent extremists from exploiting digital platforms, and this integrated approach would encompass research, expertise development, proactive engagement with relevant policymakers, and services (such as the hash-sharing database) provided to members. It was frequently noted during this assessment that while most public attention has been focused on the hash-sharing database, the opportunities provided by GIFCT for field building can be one of its most valuable contributions.

In this section we explore the question of whether GIFCT should express an independent and expert point of view on rights-based approaches to addressing the exploitation of digital platforms by terrorists and violent extremists, such as what policies, actions, and strategies governments and companies should deploy. We conclude that the most important role GIFCT can play in public policy dialogue will be to combine (1) evidence based, nuanced, specialist, substantive, and technical insights into how to prevent terrorists and violent extremists from exploiting digital platforms, with (2) a strong point of view that counterterrorist activities should be undertaken in ways that respect human rights, and (3) the perspective that taking action online is only one element of a more holistic approach. The focus should be on increasing the quality of the dialogue and raising awareness of the technical feasibility of different policy solutions; our emphasis is not lobbying on specific laws and regulations.

GIFCT Membership

During this assessment BSR encountered considerable debate around whether GIFCT should actively increase its company membership, especially with companies headquartered outside the US. BSR notes that the UNGPs emphasize the importance of prioritizing the most severe impacts, influencing our conclusion—in the context of GIFCT's global mission and remit—that a human rights-based approach should focus on the locations where impacts are most severe, rather than where they have the highest profile or media coverage. GIFCT will be better positioned to prevent terrorists and violent extremists from exploiting digital platforms through more engagement with companies and organizations outside the US and Europe, rather than less, and we recommend a proactive effort to recruit new member companies from non-US locations.

However, expanding GIFCT membership also presents human rights risks that would need to be addressed. Here BSR's recommendations address five considerations: (1) the policies and commitments of applicant companies; (2) contextual factors, such as local laws, government pressure, and ownership that may impact the ability of a company to fulfill those commitments; (3) different membership categories that maximize benefits while minimizing harms; (4) enhanced transparency requirements; and (5) the ability to expel companies from membership in certain circumstances. Further, we highlight the complexity of any additional human rights-based criteria for GIFCT membership—for example, while it might be reasonable to require member companies to achieve certain business and human rights benchmarks, in reality these criteria can be subject to local realities outside of the companies' control, such as local legal requirements. Here it is BSR's instinct that the right approach for GIFCT is not to avoid these companies but rather to go in "eyes wide open" with appropriate measures in place to avoid, prevent, or mitigate adverse human rights impacts that may arise. Further, BSR is cognizant of the need not to fall into the trap of viewing non-US companies and locations as having inherently greater risk; many stakeholders we engaged with for this assessment emphasized that the reverse can be equally true.

Stakeholder engagement

GIFCT contains some features of a multi-stakeholder initiative (i.e., non-companies actively participate in the work of GIFCT) but lacks others (i.e., decision-making power rests solely with companies). For this reason, it is especially important to implement a robust approach to stakeholder engagement where mutual expectations are widely shared. Stakeholder engagement plays a central role in a human rights-based approach, and for this reason BSR concludes that GIFCT would benefit from a more deliberate and meaningful integration of affected stakeholders into its work, including by (1) establishing greater clarity around the channels for stakeholder input, (2) broadening the range of groups engaged, and (3) clarifying the role of governments in GIFCT.

On the latter point, engagement with governments presents enormous opportunities to develop more holistic, preventative, and rights-respecting approaches to terrorist and violent extremist content, but also presents human rights risks, both real and perceived. During this assessment, BSR encountered a range of different perspectives on government involvement in GIFCT as it relates to human rights, such as the absence of government participation from outside the US, Western Europe, Australasia, and Japan, or the concern that some governments may "use" GIFCT to pressure companies into overbroad restrictions on speech.

BSR makes recommendations to address these concerns, while also pursuing the opportunity for GIFCT to support enhanced dialogue between companies, civil society, and governments on rights-respecting approaches, especially given the plethora of other channels and venues for interaction between companies and governments (such as internet referral units) that already exist. We also emphasize the need for greater interaction with the UN Special Procedures system, Office of the High Commissioner for Human Rights, and UN Office of Counter-Terrorism, and for greater support for multilateralism generally.

Governance, Accountability, and Transparency

The concepts of governance, accountability, and transparency feature prominently in the UNGPs and BSR's analysis. On governance, we make several recommendations to clarify, strengthen, and formalize the role of the Independent Advisory Committee (IAC). We also conclude that an Operating Board of four founding member companies is not a sustainable model over the medium-to-long term, and recommend that GIFCT review (in greater depth than was possible in this assessment) the merits of transitioning to a multi-stakeholder decision-making model in two years. On transparency, it is BSR's instinct that a strategic and deliberate approach to transparency

will (1) enable enhanced accountability, (2) spread expertise, insight, and learning on how to prevent terrorists and violent extremists from exploiting digital platforms, and (3) address various myths and misunderstandings that exist about GIFCT. Given GIFCT's connection to human rights impacts exists primarily through its member companies, we place special emphasis on the transparency requirements of GIFCT member companies, in addition to GIFCT itself.

Organizational Issues

GIFCT is a small, young, and newly-independent organization, and BSR is very cognizant of the challenges associated with the combination of high expectations and finite resources. It is a challenge we have seen repeated across many multicompany and multi-stakeholder collaborations that we have been involved in over the past three decades, and throughout this assessment we've been attentive to the need to make actionable and practical recommendations. The final section of the assessment makes recommendations relating to the GIFCT organization itself, such as increasing the diversity of staff, securing funds to provide stipends that support civil society participation, and hiring a human rights and stakeholder engagement director.

GIFCT would benefit from a clearer description of the interdependent relationship between human rights and the GIFCT mission that conveys human rights as a deeply embedded, complementary, and reinforcing objective in counterterrorism and violent extremism efforts.

1.5 Summary of Recommendations

BSR makes recommendations across each of these nine themes—however, many recommendations are relevant for multiple themes, and there are strong relationships and interdependencies between them. The main body of the report contains a more detailed description of each recommendation accompanied by a human rights-based rationale.

BSR highlights three important elements about these recommendations.

First, it is our hope that in addition to addressing human rights impacts, these recommendations also increase the "connective tissue" across different segments of GIFCT's work, such as the Operating Board, IAC, working groups, and the different stakeholder constituencies. This takes time to achieve in multi-stakeholder efforts, but is essential for impact and effectiveness.

Second, we note that not all these recommendations can or should be implemented immediately; rather, we assume that GIFCT should embark on a three to five year plan of continuous improvement.

Third, we note that some recommendations will need to be implemented with the different needs, resources, and capabilities of smaller companies taken into account. This is consistent with Principle 14 of the UNGPs, which states that "the means through which a business enterprise meets its responsibility to respect human rights will be proportional to, among other factors, its size. Small and medium-sized enterprises may have less capacity as well as more informal processes and management structures than larger companies, so their respective policies and processes will take on different forms." We have also recommended that GIFCT provide technical assistance to smaller companies on how to address the human rights risks associated with their counterterrorism and violent extremism activities.

THEME	RECOMMENDATION	
Mission and goals	• Create a human rights policy for GIFCT.	• Embed a commitment to human rights into other relevant GIFCT governing documents.
Human rights impacts	• Ensure that addressing the full range of GIFCT human rights impacts is embedded into the GIFCT's work plan. • Create a framework for ongoing human rights due diligence.	• Conduct a stakeholder mapping to identify organizations and experts that would increase the diversity of rightsholders whose voices are heard in GIFCT activities.
Terrorist and violent extremist content	• Develop a common understanding of terrorist and violent extremist content. • Build the common understanding of terrorist and violent extremist content on "behavior" rather than "groups."	• Participate in efforts to pursue counterterrorism and violent extremism priorities from a holistic and strategic perspective.

THEME	RECOMMENDATION
Content removal and preservation	• Convene multi-stakeholder discussions to advance acceptance and adoption of legal carve-outs for evidentiary content. • Conduct a review after the Content Incident Protocol (CIP) has been initiated / completed to identify lessons learned, human rights impacts, content for use later, and opportunities for transparency. • Publicly disclose information about the CIP process and relevant metrics when activated. • Enhance disclosure about collaboration that occurs for incidents not triggering the CIP. • Use a GIFCT "common understanding" of terrorist and violent extremist content to determine inclusion in the hash-sharing database in the medium-to-long term. • Introduce and expand transparency and oversight mechanisms alongside the extension of content in the hash-sharing database. • Require contributing companies to conduct human review and approval prior to adding hashes to the database. • Do not allow governments to directly add hashes to the database, and ensure any content that governments recommend for inclusion undergoes a review against the standards outlined in the taxonomy. • Require companies utilizing the hash-sharing database for content moderation decisions to have an adequate appeals mechanism, and build in relevant remedy capabilities into the functioning of the database itself. • Investigate how to enable third-party reviews of the hash-sharing database to assess whether hashes are consistent with the GIFCT taxonomy. • Develop a process for enabling researcher access to the hash-sharing database and associated content. • Publish a detailed explanation of how the hash-sharing database functions and conduct webinars for interested stakeholders. • Establish a multi-stakeholder process to develop metrics on how the hash-sharing database is used. • Require companies that contribute to and utilize the hash-sharing database to commit to specific disclosure. • Enable multi-stakeholder governance of the hash-sharing database to the extent possible under the current management model (i.e., hosted by Facebook Threat Exchange), and develop a plan for long-term governance and oversight.
Theory of change and programmatic priorities	• Develop position statements on the rights-based laws, policies, regulations, and strategies needed to more effectively address the exploitation of digital platforms by terrorists and violent extremists. • Proactively express this point of view with relevant governments, policymakers, and regulators.

THEME	RECOMMENDATION	
GIFCT membership	• Refine and publish human rights-based membership criteria. • Refine the due diligence process for new company membership applications. • Establish a tiered membership structure for GIFCT. • Establish a process to (1) expel companies not living up to their membership commitments and/or (2) alter a company's membership tier.	• Actively recruit new member companies, especially from non-US locations. • Actively recruit new member companies from elsewhere in the technology "stack." • Provide technical assistance to smaller companies to address human rights risks.
Stakeholder engagement	• Continue mapping stakeholders to further identify organizations and experts that would increase the diversity of rightsholders whose voices are heard in GIFCT activities—and create plans for their involvement.	• Establish and maintain closer relationships with the United Nations system. • Train GIFCT participants in principles of good stakeholder engagement. • Consider geographic diversity when rotating government membership of IAC.
Governance, Accountability, and Transparency	• Institute a system of formal recommendations from the IAC to the Operating Board, and formal responses from the Operating Board to the IAC. • Task the IAC with publishing an annual statement about the performance of GIFCT. • Publish minutes of Operating Board and IAC meetings.	• In two years, review the merits of transitioning to a multi-stakeholder Operating Board. • Establish a multi-stakeholder process to develop metrics about how the hash-sharing database is used. • Require GIFCT member companies to publish insights into their use of the hash-sharing database as part of their transparency reports, or similar.
Organizational Issues	• Create a diversity, equity, and inclusion ambition for (1) GIFCT staff and (2) GIFCT participants. • Establish a mechanism to provide stipends for non-company / nongovernment participants in GIFCT.	• Continue enhancing GIFCT staff support for the IAC and working groups. • Hire a "Human Rights & Stakeholder Engagement" Director. • Create a GIFCT government and management chart.

2. Project Overview And Methodology

IN THIS SECTION

2.1 Project Overview

2.2 Project Methodology

2.3 Assessment Themes and Questions

2.4 Recommendations

2.1 Project Overview

The mission of the Global Internet Forum to Counter Terrorism (GIFCT) is to prevent terrorists and violent extremists from exploiting digital platforms. GIFCT was founded by Facebook, Microsoft, Twitter, and YouTube in 2017, and since then GIFCT's membership has expanded to include over a dozen content platforms. GIFCT was launched as an independent, not-for-profit organization in 2020.

Today, GIFCT brings together the technology industry, government, civil society, and academia with the aim of fostering collaboration and information-sharing to counter terrorist and violent extremist activity online. Four foundational goals shape GIFCT's work:[1]

- Empower a broad range of technology companies, independently and collectively, with processes and tools to prevent and respond to abuse of their platforms by terrorists and violent extremists.
- Enable multi-stakeholder engagement around terrorist and violent extremist misuse of the internet and encourage stakeholders to meet key commitments consistent with the GIFCT mission.
- Promote civil dialogue online and empower efforts to direct positive alternatives to the messages of terrorists and violent extremists.
- Advance broad understanding of terrorist and violent extremist operations and their evolution, including the intersection of online and offline activities.

GIFCT has stated that the protection and promotion of human rights—meaning the rights of both the victims of terrorism and violent extremism, and the rights of those impacted by efforts to address terrorism and violent extremism—is central to its efforts. With this in mind, GIFCT commissioned this human rights assessment to:

- Identify actual and potential human rights impacts (including both risks and opportunities) arising from GIFCT's work.
- Make recommendations for how GIFCT and its participants can address these impacts.

The desired outcome is that this assessment (1) provides GIFCT with the knowledge, insights, and actions recommended to embed human rights into its strategy, governance, and actions; (2) establishes a clear understanding of the roles and responsibilities of different participants in GIFCT as they relate to human rights; and (3) increases trust and collaboration between participating organizations and with other stakeholders with an interest in GIFCT.

The scope of this assessment is GIFCT, not the actions of individual GIFCT member companies. Further, the scope of this assessment is primarily forward looking, rather than a review of prior activities.

2.2 Project Methodology

BSR deployed a methodology that combined a human rights assessment approach based on the UN Guiding Principles on Business and Human Rights (UNGPs) with a deeper consideration of the human rights principles, standards, and methodologies on which the UNGPs were built. While the UNGPs were written for use by companies rather than multi-stakeholder efforts,[2] their overall spirit and approach can be applied to GIFCT. In addition, the full members of GIFCT are companies, and so it can be assumed that the UNGPs apply for that reason too.

Given the role of governments in GIFCT, we have considered the "first pillar" of the UNGPs (the state duty to protect human rights) in addition to the traditional focus of a human rights assessment on the "second pillar" (the corporate responsibility to respect human rights). The first pillar emphasizes that states must protect against human rights abuse within their territory, provide guidance to companies on how to respect human rights, and support business respect for human rights in conflict-affected areas. The second pillar emphasizes the responsibility of business to avoid infringing on the human rights of others and address the adverse human rights impacts with which they are involved, including via policy commitments, human rights due diligence, and providing access to remedy.

Key assumptions in BSR's methodology include:

All human rights are potentially relevant for GIFCT.

In this assessment BSR used the international legal human rights framework as the baseline, including the International Bill of Human Rights (the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and the International Covenant of Economic, Social, and Cultural Rights), as well as other international human rights instruments and standards of international humanitarian law relevant to GIFCT.

Human rights are interconnected.

All human rights are indivisible, interdependent, and interrelated—the improvement of one right can facilitate advancement of others; the deprivation of one right can adversely affect others. This means that human rights can come into conflict with one another for legitimate reasons, and it is important to deploy rights-based methods when two conflicting rights cannot both be achieved in their entirety. Rather than "offsetting" one right against another, it is important to pursue the fullest expression of both and identify how potential harms can be addressed. Key principles to apply when rights come into conflict include the following:

- **Legitimacy**—Restrictions to a right must pursue an objectively legitimate purpose and address a precise threat.
- **Necessity and proportionality**—Only restricting a right when the same goal cannot be achieved by other means, and using restrictions that are the least intrusive to achieve the legitimate purpose.
- **Nondiscrimination**—Restrictions to a right must be implemented in a nondiscriminatory manner.
- **Lawful**—Restrictions to a right should be contained in a relevant law that is clear and accessible to the public.
- **Reverting to principle**—Focusing on the underlying principle of the right being limited and identifying ways to uphold the core principle, even if not the exact right.

A stakeholder-inclusive process is essential.

Effective human rights due diligence requires meaningful engagement with rightsholders whose human rights may be impacted, or with reasonable alternatives such as independent expert resources, human rights defenders, and other representatives from civil society.[3] In addition to GIFCT member companies, during this assessment BSR engaged with around 40 individuals and organizations in workshops, interviews, and written submissions, including civil society organizations, academics, governments, and victims of both terrorism and violent extremism and of overbroad restrictions on content. The names of individuals and organizations consulted have been withheld to respect confidentiality and to enable more candid, direct, and constructive engagement. BSR notes that the feedback we received during these engagements varied both between and within different "constituencies," such as business, civil society, academia, and government.

Vulnerability must be prioritized.

A human rights-based approach implies greater attention should be paid to impacts on individuals from groups or populations that may be at heightened risk of vulnerability or marginalization. In this context both the victims of terrorism and violent extremism and the victims of efforts to address terrorism and violent extremism are especially important, including independent expert resources, human rights defenders, and others from civil society with insights into their interests.

BSR undertook this assessment from December 2020 to May 2021 following four main phases, described in the table below. This review took place concurrently with GIFCT reviewing its overall strategy and plan, and the main observations and recommendations in this human rights review informed GIFCT's thinking in real time.

2.3 Assessment Themes and Questions

To structure the human rights assessment BSR identified nine key themes and 35 questions that we believe provide a useful framework for reviewing GIFCT's governance, strategy, and activities from a human rights perspective. There is significant overlap and interdependency between these themes.

THEME	QUESTIONS	
GIFCT Mission and Goals	• Is "human rights at GIFCT" about addressing potential adverse human rights impacts arising from its work, or should human rights be positioned as central to GIFCT's mission and goals?	• Does commitment and reference to international human rights instruments need to be built into GIFCT founding documents, such as the governance charter?
Human Rights Impacts	• What human rights are impacted by GIFCT activity? • Which rightsholders are impacted by GIFCT activities? • What is GIFCT's role in the connection between online content and offline harm (human rights impacts)? • In practical terms, how can GIFCT address the issue of counterbalancing competing human rights? Which human rights need counterbalancing?	• What methods of ongoing human rights due diligence are appropriate for GIFCT, such that GIFCT can anticipate the challenges of the future, not just the challenges of today? • What research and dialogue methods should GIFCT deploy to stay ahead of emerging trends in the exploitation of digital platforms by terrorists and violent extremists?
Terrorist and Violent Extremist Content	• Should GIFCT seek to create shared definitions of terrorist and violent extremist content? • How should GIFCT address the discrimination and bias that exists in the counterterrorism field, such as the disproportionate focus on Islamist extremist content, rather than white supremacist content?	• In addition to terrorist and violent extremist content, should GIFCT address issues "higher up the chain," such as the radicalization and recruitment process?
Content Removal and Preservation	• How should the freedom of expression risks associated with the hash-sharing database be addressed?	• What is the right model for transparency for the hash-sharing database?

THEME	QUESTIONS	
Content Removal and Preservation (Cont)	• How should GIFCT address the need to preserve removed content that may be used as evidence later?	• How should GIFCT address the risk that the Content Incident Protocol (CIP) results in over-removal of content?
Theory of Change and Programmatic Priorities	• What is the relative role of different GIFCT workstreams as it relates to protecting and respecting human rights? • Should GIFCT have an independent and expert point of view on rights-based approaches to addressing the exploitation of digital platforms by terrorists and violent extremists, such as what policies, actions, and strategies governments and companies should deploy?	• Where, when, and how should this point of view be conveyed, and who should decide what this point of view is? • What research agenda would support the development of a point of view on rights-based approaches to addressing the exploitation of digital platforms by terrorists and violent extremists? • What other human rights-focused activities might GIFCT undertake to achieve its mission?
GIFCT Membership	• How important is increasing company membership outside the US to the achievement of GIFCT's mission? • What minimum human rights-based criteria should be established for company membership? • Under what human rights-based circumstances should companies be removed from membership?	• Are there options other than "full membership" that can help advance GIFCT's mission and grow membership while addressing human rights risk? • Are there companies elsewhere in "the stack" that can contribute toward the achievement of GIFCT's mission?
Stakeholder Engagement	• What is the right stakeholder engagement model for GIFCT—for example, how should the IAC and working groups evolve? • Which civil society organizations and academic experts should participate to ensure coverage of the most relevant human rights impacts? • What risks and opportunities exist with government involvement in GIFCT? How can they be addressed?	• What are the human rights risks and opportunities associated with GIFCT becoming "the go-to place" for governments during a crisis, and how should they be addressed? • Is there a size limit for civil society and government participation—what is the right balance between size and securing diverse voices?

THEME	QUESTIONS	
Governance, Transparency, and Accountability	• Should GIFCT transition to a multi-stakeholder board? • What qualitative and quantitative disclosures from GIFCT would be useful for those with an interest in the human rights impacts of GIFCT?	• Should there be any transparency requirements for member companies and other participants? If yes, what would they be?
Organizational Issues	• What level of resources does GIFCT need to effectively address human rights?	• Which working group (if any) should carry forward this work?

2.4 Recommendations

A core feature of this assessment is recommendations across each of the nine themes, and these are found in sections 4-12 of this report. Here BSR highlights three important elements about these recommendations.

First, it is our hope that in addition to addressing human rights impacts, these recommendations also increase the "connective tissue" across different segments of GIFCT's work, such as the Operating Board, IAC, working groups, and the different stakeholder constituencies. This takes time to achieve in multi-stakeholder efforts, but is essential for impact and effectiveness.

Second, we note that not all these recommendations can or should be implemented immediately; rather, we assume that GIFCT should embark on a three to five year plan of continuous improvement.

Third, we note that some recommendations will need to be implemented with the different needs, resources, and capabilities of smaller companies taken into account. This is consistent with Principle 14 of the UNGPs, which states that "the means through which a business enterprise meets its responsibility to respect human rights will be proportional to, among other factors, its size. Small and medium-sized enterprises may have less capacity as well as more informal processes and management structures than larger companies, so their respective policies and processes will take on different forms. But some small and medium-sized enterprises can have severe human rights impacts, which will require corresponding measures regardless of their size."

It is BSR's intention that some recommendations (e.g., transparency requirements) are implemented in ways that both address severe human rights impacts and consider the likelihood that implementation at smaller companies may take a different form. We have also recommended that GIFCT provide technical assistance to smaller companies on how to address the human rights risks associated with their counterterrorism and violent extremism activities. There is no single global definition of small and medium-sized enterprises, though "less than 250 employees" is often used—such as by the European Union, for example.

3. GIFCT Operating Context

IN THIS SECTION

3.1 The counterterrorism landscape

3.2 Challenges faced by multi-stakeholderism

3.3 Role of this assessment

GIFCT pursues its mission in an extremely complex context. Limits to transparency in counterterrorism efforts, the transnational nature of counterterrorism work, and the increasingly contested concept of multi-stakeholderism shape the backdrop against which GIFCT operates. This introduction highlights some of the human rights challenges of this context, with implications for the assessment that follows.

3.1 The counterterrorism landscape

GIFCT exists in the context of a growing number of institutions and organizations spanning national and international boundaries designed to counter terrorism. This ecosystem has grown hugely in the 20 years since 9/11, and many of these institutions are relatively opaque, operate outside traditional multilateral boundaries, and offer few opportunities for human rights oversight or the involvement of civil society.

Providing security is a fundamental obligation of states so that individuals are able to realize their human rights, including the right to life. However, many countries have challenged or ignored other human rights and the rule of law in the name of security. The lack of a globally agreed upon definition of terrorism and the highly politicized context within which counterterrorism takes place have resulted in many examples of government overreach, with political opponents, human rights defenders, and environmental groups wrongly designated as terrorist groups. The expansion of efforts to counter violent extremism have made this already complex field even more difficult to define.

The lists of designated terrorist organizations defined by individual countries, the European Union, and the UN disproportionately include those that have distorted Islam. This creates bias throughout the system in a way that has potential negative human rights impacts on freedom of religion, freedom from discrimination, and the right to assembly. During this assessment, BSR engaged with several affected stakeholders who described the adverse impact online counterterrorism and violent extremism efforts have had on their ability to organize, advocate, share information, and participate in political dialogue—for example, where political activists have been subject to content restrictions as a result of posts using certain terms or referencing specific organizations.

There is significant public debate about the threat posed by right-wing extremism and various ethnic nationalist groups, and some countries have designated a small number of these groups as terrorist organizations. However, other countries make a distinction between foreign and domestic terrorist groups and do not designate those that are domestic. Given the existing challenges created by the designated lists system, it is difficult to know whether increasing the range of groups on the lists to include a wider variety of organizations across the ideological spectrum would reduce overall bias or create more potential for adverse human rights impacts. There are no easy answers or choices on these big questions.

In her latest report to the UN Human Rights Council, the UN Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism highlights the particular impacts on the rights of women, girls, and the family created by counterterrorism and countering violent extremism policies and practices. While noting that there have been efforts to include a gender perspective into some national security efforts, she highlights continuing issues due to a fundamental lack of diversity in institutional decision-making and the disproportionate targeting of women human rights defenders. The regulation of family life through counterterrorism efforts is also an ongoing concern.

Power and information asymmetries are a reality in the makeup of multi-stakeholder efforts, with some stakeholders having more insight, control, and resources than others. This makes governance structures particularly important as they offer a pathway toward more equitable approaches through institutional decision-making authority.

3.2 Challenges faced by multi-stakeholderism

Adding to the challenging subject matter addressed by GIFCT are important questions posed by the concept of multi-stakeholderism that have become more pressing as the number of these organizations grows. Three issues are particularly noteworthy:

First, placing affected communities at the heart of these efforts is both increasingly important and hard to achieve given the challenges of limited resources and accessing often complex processes and organizations. As more initiatives emerge, small and often resource-constrained civil society groups need to make choices about where they will and won't engage. External pressure on multi-stakeholder efforts to demonstrate progress can lead those efforts to short-cut the necessary steps to identify the best groups to engage with and invest sufficient time to build trusted relationships.

Second, power and information asymmetries are a reality in the makeup of multi-stakeholder efforts, with some stakeholders having more insight, control, and resources than others. This makes governance structures particularly important as they offer a pathway toward more equitable approaches through institutional decision-making authority.

Third, independence is important from several different angles in the context of multi-stakeholder efforts, but can be difficult to achieve in the early days of new organizations. For civil society groups, ensuring they continue to have an independent voice is vital, both individually and collectively within the initiative. Further, initiatives that are heavily reliant on company funding at the beginning need to diversify their funding over the medium term as an important signal of independence, an issue we address later in the report.

3.3 Role of this assessment

This assessment focuses on identifying the actual and potential human rights impacts (including both risks and opportunities) arising from GIFCT's work, and makes recommendations for how GIFCT and its participants can address these impacts. As such this assessment does not attempt to address the full range of challenges that exist in the counterterrorism field or in multi-stakeholder initiative design; however, these challenges exist as a backdrop, and many of BSR's observations and recommendations are directly relevant to addressing them.

Further, this assessment should be read as an assessment of GIFCT, and not as an assessment of the counterterrorism and violent extremism efforts of individual GIFCT member companies. However, we have made recommendations for greater transparency from GIFCT member companies in how they utilize GIFCT resources, so that those actions may be assessed by others.

4. GIFCT Mission And Goals

IN THIS SECTION

4.1 Analysis

4.2 Recommendations

4.1 Analysis

The UNGPs state that "as the basis for embedding their responsibility to respect human rights, business enterprises should express their commitment to meet this responsibility through a statement of policy."[4] The term "statement" in the UNGPs is used generically to describe whatever means an organization employs to publicly communicate its responsibilities, commitments, and expectations. In the case of the case of GIFCT, BSR interprets this expectation as having implications for how GIFCT's mission and goals are expressed.

Is "human rights at GIFCT" about addressing potential adverse human rights impacts arising from its work, or should human rights be positioned as central to GIFCT's mission and goals?

Some interviewees described GIFCT to BSR as "pursuing counterterrorism objectives while avoiding adverse impacts on human rights when doing so," such as on the right to freedom of expression. By contrast, other interviewees highlighted that the mission of GIFCT itself has a clear human rights purpose, such as protecting the right to life, liberty, and security of person. Given this discrepancy, there is an opportunity to express more clearly the relationship between human rights and the GIFCT mission and goals.

BSR's concludes that while the "narrow" purpose of GIFCT is to prevent terrorists and violent extremists from exploiting digital platforms, in doing so GIFCT enhances the protection, fulfillment, and realization of human rights. In other words, human rights for GIFCT is about more than simply "avoiding harm" while pursuing its mission.

For this reason, we believe that while the precise mission statement of GIFCT should remain as currently stated ("prevent terrorists and violent extremists from exploiting digital platforms"), GIFCT would benefit from a clearer description of the complementary relationship between human rights and the GIFCT mission.

Today, human rights is too often conveyed as an extra factor in counterterrorism and violent extremism, rather than a deeply embedded objective. As the UN Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism emphasizes, human rights are inherent in and reinforcing of counterterrorism efforts. According to the UN Special Rapporteur, terrorism poses a serious challenge to the protection of human rights, and "effectively combatting terrorism and ensuring respect for human rights are not competing but complementary and mutually reinforcing goals."[5]

A new language taxonomy for GIFCT can underpin greater conceptual clarity throughout GIFCT's work, and several stakeholders emphasized to BSR the significant benefit that would arise from a clear and shared articulation of rights-based approaches to countering terrorism and violent extremism online.

Does commitment and reference to international human rights instruments need to be built into GIFCT founding documents, such as the governance charter? What should be contained in a GIFCT statement of human rights policy?

The UNGPs state that "the responsibility of business enterprises to respect human rights refers to internationally recognized human rights—understood, at a minimum, as those expressed in the International Bill of Human Rights and the principles concerning fundamental rights set out in the International Labour Organization's Declaration on Fundamental Principles and Rights at Work."[6] The UNGPs go on to reference the relevance of other international human rights instruments and international humanitarian law.

There is growing momentum behind the notion that a company's own content policies (sometimes called community standards, community guidelines, or rules) should also be grounded in international human rights standards and make reference to relevant international human rights instruments and principles.[7]

BSR believes this forms an important element to a human rights-based approach to content governance, and we believe that a similar approach can apply to GIFCT. Taking this approach will ensure alignment with internationally agreed norms, support consistency of approach across borders, and allow GIFCT to refer to authoritative sources when faced with proposals or demands that present risks to human rights.

Today, human rights is too often conveyed as an extra factor in counterterrorism and violent extremism, rather than a deeply embedded objective.

4.2 Recommendations

01 RECOMMENDATION

Create a human rights policy for GIFCT

This human rights policy would be written by GIFCT staff, informed by the IAC and engagement with stakeholders, signed off by the GIFCT Operating Board, and used to enhance an informed commitment to human rights across all GIFCT activities. It would be publicly available.

The human rights policy would express a commitment to the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, the International Covenant on Economic, Social, and Cultural Rights, and the principles of International Humanitarian Law. It would reference other relevant international human rights instruments, such as the Convention on the Elimination of Discrimination against Women and the Convention on the Elimination of All Forms of Racial Discrimination.

The human rights policy would describe the relationship between GIFCT's mission, goals, and human rights, such as:

- The rights protected by the GIFCT mission, such as the rights to life, liberty, security of person, nondiscrimination, freedom of religion, and participation in government.
- The rights potentially adversely impacted in pursuit of the GIFCT mission, such as freedom of expression, freedom of religion, nondiscrimination, and access to remedy.
- Key elements of the GIFCT plan to address these rights, such as controls on technology innovation, expertise development, and sharing authoritative research.
- The view that combatting terrorism and violent extremism and ensuring respect for human rights are complementary and mutually reinforcing goals.
- A commitment to transparency.

BSR recommends consulting with the UN Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism and the UN Special Rapporteur on the promotion and protection of freedom of opinion and expression when drafting this policy.

01 EXPLANATION

Principle 16 of the UNGPs states that "as the basis for embedding their responsibility to respect human rights, business enterprises should express their commitment to meet this responsibility through a statement of policy."

Many stakeholders interviewed for this assessment referenced the central importance of human rights to the successful pursuit of counterterrorism objectives and emphasized the benefits of greater conceptual clarity between GIFCT's mission and human rights.

02 RECOMMENDATION

Embed a commitment to human rights into other relevant GIFCT governing documents.

Documents such as the GIFCT governance charter and bylaws, IAC terms of reference, and working group mandates should make appropriate reference to GIFCT's human rights commitments.

02 EXPLANATION

Principle 16 of the UNGPs states that the statement of human rights policy should be "reflected in operational policies and procedures necessary to embed it throughout the business enterprise."

Further, Principle 16 of the UNGPs expands that "just as States should work toward policy coherence, so business enterprises need to strive for coherence between their responsibility to respect human rights and policies and procedures that govern their wider business activities and relationships."

5. Human Rights Impacts

IN THIS SECTION

5.1 Analysis

5.2 Recommendations

5.1 Analysis

The UNGPs imply that GIFCT should have a clear understanding of the human rights impacts with which it might be involved and which rightsholders may be impacted.[8]

BSR notes that GIFCT is one step removed from these impacts—since they largely result from company's own activities and decisions, rather than those of GIFCT itself—but that it is important to understand these impacts given GIFCT's potential connection to them and influence on company actions. Specifically, the UNGPs encompass responsibility for "adverse human rights impacts that are directly linked to their operations, products, or services by their business relationships," and GIFCT is certainly linked to adverse impacts that may arise from member company decision making informed by their participation in GIFCT.

What human rights are impacted by GIFCT activity?

While considerable attention has been paid to the impact GIFCT may have on freedom of expression, BSR concludes that many other rights may be impacted by GIFCT too. Further, while the UNGPs are oriented toward addressing adverse impacts, we believe it is important to consider opportunities to support the fulfillment, realization, and enjoyment of rights as well.

- **Life, liberty, and security of person (UDHR 3; ICCPR 6, 9):** By helping prevent terrorist and violent extremists from exploiting social media platforms, GIFCT is facilitating the increased protection of this right. This includes both reducing the likelihood of physical harm and the psychological and mental health impacts that can arise from content.

- **Nondiscrimination and equality before the law (UDHR 1, 2, 7; ICCPR 2, 3, 26; ICESCR 2, 3; CEDAW 2; CERD 2):** There is significant anti-Islamic bias in the counterterrorism field, with efforts often focused on Islamist extremist terrorism, rather than other forms of terrorism, such as white supremacist terrorism or terrorism that has exploited Hindu nationalism. As a result, there is a risk that GIFCT efforts—such as the hash-sharing database—embed similar discrimination, resulting in discrimination based on race, religion, national origin, or other status. Conversely, GIFCT has the opportunity to proactively address bias in the counterterrorism field, as it has already sought to do with its research, learning, and expertise development activities, and has initiated with its review of hash-sharing database taxonomy.

- **Access to effective remedy (UDHR 8; ICCPR 2):** If GIFCT and/or member company actions (e.g., via the hash-sharing database, URL sharing, content incident protocol) result in content being both removed and deleted, this may adversely impact access to remedy in cases where that content has evidential value for serious crimes, including war crimes. This risk is especially prominent if content is taken down so rapidly that relevant law enforcement agencies or other entities do not learn of its existence and therefore cannot request that the content is retained. Conversely, GIFCT has the opportunity to support the establishment of methods to preserve content, enhancing the ability of rightsholders to use digital evidence in pursuit of remedy—an issue we consider later in this assessment.

- **Freedom of opinion, thought, conscience, and religion (UDHR 18, 19; ICCPR 18, 19):** Everyone has the right to hold views on any issue without fear of punishment or censure, the right to believe in any religion (or none), and the right to change religion or belief. Everyone is free to hold views and religion in public or private, and to practice, teach, and observe these religions and beliefs. If GIFCT and/or member company actions (e.g., via the hash-sharing database, URL sharing, content incident protocol) lead to overbroad restrictions on content, then users may be deprived of exposure to the content necessary to freely form these beliefs. This can be especially important for communities that have been forced to disperse over multiple different geographies, and for whom online spaces represent an especially important place to form beliefs and opinions.[9]

- **Freedom of expression (UDHR 19; ICCPR 19):** While Article 20 of the ICCPR makes clear that advocacy of national, racial, or religious hatred that constitutes incitement to discrimination, hostility, or violence shall be prohibited by law, there is a risk that legitimate content may be removed in error as part of "overbroad" efforts to remove terrorist and violent extremist content. This risk is heightened in situations where companies "apply" the hash-sharing database without first scrutinizing the content being removed for noncompliance with their terms of service. Conversely, many affected stakeholders emphasized to BSR the significant opportunity to enhance freedom of expression by fostering a safer online space for marginalized populations and vulnerable groups too often targeted by terrorists and violent extremists.

- **Freedom of assembly and association (UDHR 20; ICCPR 21, 21):** If specific communities are disproportionately impacted by overbroad content removals, then their ability to freely assemble and associate will be adversely impacted. This can be especially important for communities that have been forced to disperse over multiple different geographies, and for whom online spaces represent an especially important assembly place. Conversely, the creation of safer online spaces may enhance the ability for some marginalized populations and vulnerable groups to engage, collaborate, and organize online.

- **Privacy (UDHR 12; ICCPR 17):** While focused on content removal today, GIFCT's Technology Innovation Program could deploy other methods in the future, such as metadata and behavioral signals analysis, or the sharing of Machine Learning (ML)-based detection and pattern analysis techniques. These methods may be accompanied by risks to privacy if GIFCT member companies use or share data in ways that result in arbitrary interference with family, home, or correspondence.

- **Freedom from arbitrary arrest, detention, and exile; right to a fair trial; innocence before being proven guilty (UDHR 9, 10, 11; ICCPR 14; CERD 5):** If member companies use resources from GIFCT's Technology Innovation Program to share insights with governments and law enforcement agencies in ways that are subsequently misused by that government or agency, then these rights would be placed at risk.[10]

Which rightsholders are impacted by GIFCT activities?

A human rights-based approach requires a clear understanding of which rightsholders are impacted by GIFCT activities. This in turn informs which rightsholders, stakeholders, and experts GIFCT should seek to engage with and include in its activities. BSR identifies the following categories of impacted rightsholder:

- **Actual and potential victims of terrorism and violent extremism:** These rightsholders have an interest in the activities of GIFCT in two major ways: first, they have a clear interest in the rapid removal of online content that may incite terrorist and violent extremist activity offline, or that may cause psychological harm; second, they have an interest in content being appropriately archived for use in criminal or other relevant proceedings and research.

- **Victims of efforts to counter terrorism and violent extremism:** Several groups of rightsholders may be adversely impacted by efforts to address terrorist and violent extremist content, such as individual victims of overbroad or wrongful content removal and other actions (e.g., account suspensions, down-ranking), and communities whose content is disproportionately impacted by these actions. To date the victims of overbroad restrictions to speech have tended to be from Islamic communities, whose content has been disproportionately targeted by efforts to remove terrorist and violent extremist content—including, for example, the hash-sharing database. In addition, families of rightsholders can be caught up in overbroad efforts to counter terrorism.

- **Human rights defenders:** Human rights monitors, lawyers, professional and citizen journalists, civil society organizations, nonviolent political activists, and members of vulnerable groups advocating for their rights all have an interest in both the speedy removal of terrorist and violent extremist content and the archiving of content relevant for their ability to access remedy for themselves or for others. During this assessment, affected stakeholders described both (1) how the removal and deletion of content can hinder efforts to expose human right violations and secure access to remedy, and (2) the sheer volume of removed and potentially useful content can present challenges.

- **Women, girls, families, men, and boys:** Terrorist and violent extremist content efforts can have a disproportionate impact on women, girls, and families, for example arising from the problematic use of gender stereotypes in efforts to counter terrorism and violent extremism. While women and girls carry the heaviest burden, men and boys also experience gender stereotyping in the counterterrorism field.[11] Several interviewees emphasized to BSR the male-dominated nature of the sector (e.g., law enforcement, military,

intelligence services) and the bias that likely results from this culture. Conversely, interviewees also emphasized the "toxic masculinity" that exists in terrorist and violent extremist content narratives, and the opportunity to foster safer online spaces by addressing content that incites violence against women, girls, and the LGBTI+ community.

BSR notes that impacted rightsholders will vary across geographies and contexts, and that it is beyond the scope of this paper to provide a comprehensive mapping. However, we believe it is important for GIFCT member companies to identify impacted rightsholders using the following criteria:

- **Formal discrimination**—laws or policies that favor one group over another. In the GIFCT context this can include the fact that lists of designated terrorist organizations defined by individual countries, the European Union, and the UN disproportionately impact Islamic communities.
- **Societal discrimination**—cultural or social practices that marginalize some and favor others. In the GIFCT context this can include efforts to address terrorist and violent extremist content in specific geographies where stereotypes suggest some communities are more closely associated with terrorism, resulting in an undue focus on some groups rather than others.
- **Practical discrimination**—marginalization due to life circumstances, such as poverty. In the GIFCT context this can include those unable to access or use appeals mechanisms that exist to address overbroad content removals.
- **Hidden groups**—people who might need to remain hidden and consequently may not speak up for their rights, such as undocumented migrants. In the GIFCT context, this might include those less able to advocate for their rights online for fear of retaliation.

We propose a variety of actions throughout this report to address these situations, such as greater hash-sharing database transparency, clearer and limited definitions of terrorist and violent extremist content, seeking to diversify the range of terrorist and violent extremist content targeted, and promoting practical solutions to the challenge of storing and saving content for use as evidence.

What is GIFCT's role in the connection between online content and offline harm (human rights impacts)?

During this assessment BSR encountered a variety of different views about the connection between online content and offline harm, with different interpretations on the extent to which removing content addresses broader terrorism and violent extremism challenges.

Despite this variety, there was significant interest in the potential for GIFCT to play the role of "brain trust for the online-offline nexus" and explore with membership and stakeholders how they intersect in practice. Playing the role of a brain trust for the online-offline nexus implies that GIFCT should be cognizant of its position in the broader counterterrorism field, and that GIFCT should play its appropriate part in "whole of society" approaches to counterterrorism and violent extremism.

There is interest in GIFCT clearly conveying the potential and limits of countering terrorist and violent extremist content. Taking a strategic, holistic, and preventative approach to human rights implies that GIFCT should continue to deliberately address topics adjacent to terrorist and violent extremist content, such as how the internet is used in the recruitment and radicalization process, or the role of counterspeech.

In practical terms, how can GIFCT address the issue of counterbalancing competing human rights? Which rights need counterbalancing?

Human rights can come into conflict with one another for legitimate reasons, and it is important to deploy rights-based methods when two conflicting rights cannot both be achieved in their entirety. Rather than "offsetting" one right against another, it is important to pursue the fullest expression of both and identify how potential harms can be addressed.

In the case of GIFCT, the rights most frequently referenced as being in need of counterbalancing are (1) the right to life, liberty, and security of person with (2) the rights to freedom of expression, association, and assembly, and access to remedy. However, before discussing how to counterbalance these rights, it is important to emphasize that most of the time these rights need not be in conflict—removing terrorist and violent extremist content from the internet and protecting life enhances the ability we all have to more fully realize our rights online. As several interviewees noted, removing terrorist propaganda and violent extremist content is not in itself a human rights violation, provided the removal is a lawful, legitimate, necessary, and a proportionate restriction on speech.

Nevertheless, there are times where the need for urgent action to protect life by removing content rapidly may result in errors or the overbroad removal of content (e.g., legitimate political speech) that adversely impact the rights to freedom of expression, association, and assembly, or that may make access to remedy more challenging if content of evidential value is deleted.

Here the cases seem less about rights being inherently in tension—the accurate removal of genuine terrorist and violent extremist content does not adversely impact other rights—but rather about establishing the mitigating actions to address errors that will arise when taking action at speed and in contexts with significant nuance.

Contrary to perception, these rights may not be in conflict in principle, even if they often are in practice. We propose a variety of actions throughout this report to address these situations, such as greater hash-sharing database transparency, clearer and limited definitions of terrorist and violent extremist content, seeking to diversify the range of

terrorist and violent extremist content targeted, and promoting practical solutions to the challenge of storing and saving content for use as evidence.

Finally, BSR observes that broadening the range of stakeholders participating in GIFCT will be essential for a more successful approach to counterbalancing conflicting rights over time. During this assessment we engaged with a range of affected stakeholders not currently participating in GIFCT, and their stories provided significant insight into how GIFCT and member company activities may impact their rights in practice.

What methods of ongoing human rights due diligence are appropriate for GIFCT, such that GIFCT can anticipate the challenges of the future, not just the challenges of today? What research and dialogue methods should GIFCT deploy to stay ahead of emerging trends in the exploitation of digital platforms by terrorists and violent extremists?

The UNGPs emphasize that human rights due diligence "should be ongoing, recognizing that the human rights risks may change over time as the business enterprise's operations and operating context evolve." In the context of GIFCT, there is a need to continue efforts (such as the research workstream) designed to help ensure that the organization remains relevant by addressing emerging trends in human rights and the exploitation of digital platforms by terrorists and violent extremists over time—that GIFCT doesn't get stuck only addressing "the risks of yesterday" rather than "the risks of tomorrow."[12]

A specific priority for attention is GIFCT's Technology Innovation workstream, and in particular the potential to deploy additional technology-based methods (such as metadata analysis, machine learning techniques, or geolocation insights) to pursue GIFCT's mission. Assessing the human rights risks and opportunities associated with innovations in technology—and the different human rights impacts, such as privacy, that may become more significant as a result—will be important for anticipating the human rights challenges of the future.

It was also raised with BSR that GIFCT could have a role in threat monitoring, and engaging with a range of experts and stakeholders on how terrorist and violent extremist threats may evolve over time and how to address them in a rights-respecting manner.

5.2 Recommendations

01 RECOMMENDATION

Ensure that addressing the full range of GIFCT human rights impacts are embedded into the GIFCT's work plan.

For example, as the GIFCT working groups evolve and new working groups are formed, GIFCT should be deliberate in ensuring that human rights impacts are appropriately included in their terms of reference, and that when taken in combination, the various GIFCT workstreams cover all of GIFCT's priority human rights impacts.

01 EXPLANATION

Principle 19 of the UNGPs states that, "in order to prevent and mitigate adverse human rights impacts, business enterprises should integrate the findings from their impact assessments across relevant internal functions and processes, and take appropriate action."

02 RECOMMENDATION

Create a framework for ongoing human rights due diligence.

GIFCT's ongoing human rights due diligence should begin with an action plan to implement the recommendations in this report that GIFCT will take forward.

Other important elements for ongoing human rights due diligence include (1) reviewing the contents of this report on an annual basis to assess progress, (2) conducting strategic foresight / futures processes that seek to identify evolving human rights risk over time, especially as technology-based methods evolve, and (3) commissioning research and analysis into emerging human rights trends and risks of relevance to counterterrorism.

The annual review of this report could be overseen by the IAC, and should involve active participation of affected stakeholders, civil society organizations, and other experts.

02 EXPLANATION

A key feature of terrorism is that its threats, methods, and manifestations evolve over time—and as a result, the human rights impacts of terrorist and violent extremist content efforts will evolve too.

Several interviewees emphasized the importance of GIFCT remaining relevant by identifying and addressing the threats of tomorrow, not simply those of today.

Principle 17 of the UNGPs states that human rights due diligence "should be ongoing, recognizing that the human rights risks may change over time as the business enterprise's operations and operating context evolve."

Principle 20 of the UNGPs states that "to verify whether adverse human rights impacts are being addressed, business enterprises should track the effectiveness of their response. Tracking should...draw on feedback from both internal and external sources, including affected stakeholders."

03 RECOMMENDATION

Conduct a stakeholder mapping to identify organizations and experts that would increase the diversity of rightsholders whose voices are heard in GIFCT activities.

In particular, GIFCT would benefit from increased participation from groups representing both the victims of terrorism and the victims of efforts to address terrorism, organizations from underrepresented regions (notably Africa, the Middle East, and South Asia), interfaith organizations, and those able to share a gender-perspective on terrorist and violent extremist content. These stakeholders could join GIFCT in a formal capacity (e.g., IAC) or in a more informal capacity via participation in working groups or other activities.

03 EXPLANATION

Principle 18 of the UNGPs states that when assessing human rights impacts, business enterprises should "involve meaningful consultation with potentially affected groups and other relevant stakeholders."

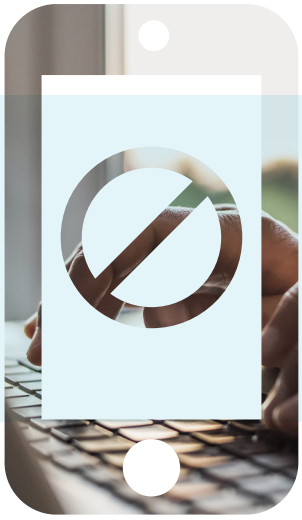
Principle 20 of the UNGPs states that tracking progress on human rights should "draw on feedback from both internal and external sources, including affected stakeholders."

6. Terrorist and Violent Extremist Content

IN THIS SECTION

6.1 Analysis

6.2 Recommendations



6.1 Analysis

The UNGPs state that companies should take appropriate action to address the potentially adverse human rights impacts with which they are involved.[13] As section 5 (above) describes, many of GIFCT's human rights impacts (both risks and opportunities) are associated with issues relating to terrorist and violent extremist content—such as how it is defined, how it is addressed, and bias that exists in the counterterrorism field.

Should GIFCT seek to create shared definitions of terrorist and violent extremist content?

BSR's assessment surfaced a range of views on this question, with some expressing interest in shared definitions as a means for improving overall expertise and consistency, and others expressing skepticism at the scale and complexity of reaching consensus, and even whether it is appropriate for nongovernmental actors to create definitions in the first place. GIFCT recently announced an effort toward developing an expanded taxonomy for the hash-sharing database, which will consider this question.

We also observed that assessment participants may be referring to very different models of shared definition. Some pointed to Article 3(1) of the EU regulation on preventing the dissemination of terrorist content online[14] or the definition proposed by the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism[15] as existing definitions to adopt—while others had in mind the much longer and detailed definitions that are actionable by companies.

We note the debate about competing and conflicting definitions of terrorist content, with concern that some existing government definitions of terrorist content may encompass legitimate expression protected under international human rights law.[16]

Overall, this assessment surfaced widely-held skepticism toward GIFCT creating a shared definition of terrorist and violent extremist content, and BSR has come to agree with this sentiment. We believe this task properly resides with governments, and acknowledge both the huge challenge of reaching consensus and the fact that companies operating in very different contexts may need different definitions.

However, while stopping short of a shared definition, we do believe that creating a common understanding of terrorist and violent extremist content—even if companies choose to adapt their own precise definitions—would have considerable value. Benefits would include (1) pushing back against overbroad definitions of terrorist and violent extremist content deployed by governments; (2) improving the capability of smaller companies without extensive policy teams to establish their own definitions; (3) establishing a bulwark against "slippery slope" definitions of terrorist and violent extremist content that may extend too far into other forms of speech, thereby presenting risks for freedom of expression; and (4) improving shared awareness of the relationship between human rights and terrorist and violent extremist content.

In addition, we note that the multi-stakeholder setting of GIFCT is an excellent opportunity to improve understanding, generate increased consensus, and enhance the shared mission of GIFCT participants.

BSR recommends starting with a common understanding of terrorist content, then moving on to violent extremist content, which presents a more complex challenge given its adjacency with broader notions of hate speech and extremist (but not violent) content.

How should GIFCT address the discrimination and bias that exists in the counterterrorism field, such as the disproportionate focus on Islamist extremist content, rather than white supremacist content?

The principle of nondiscrimination underpins international human rights law, and Article 2 of the UDHR states that "Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, color, sex, language, religion, political or other opinion, national or social origin, property, birth or other status."

In addition, the UNGPs state that they "should be implemented in a nondiscriminatory manner, with particular attention to the rights and needs of, as well as the challenges faced by, individuals from groups or populations that may be at heightened risk of becoming vulnerable or marginalized."

However, GIFCT operates in a context of historical discrimination and bias in the counterterrorism field, which has tended to focus predominantly on terrorism that has distorted Islam, rather than other forms of terrorism, such as white supremacist terrorism or terrorism that has exploited Hindu nationalism. To address potential adverse human rights impacts, GIFCT has a responsibility to help address (rather than entrench) this bias, such as through research, advocacy, or by engaging with affected stakeholders throughout its work. BSR identified significant consensus on this point throughout the assessment.

BSR acknowledges that this discrimination has not been caused by GIFCT; however, GIFCT may contribute to or be directly linked to its continuation if GIFCT omits to take action to address it, and undertakes activities that continue to perpetuate this discrimination—such as a hash-sharing database focused predominantly on organizations that distort Islam. In these circumstances the UNGPs imply that business entities should use

leverage (defined as "the ability to affect change in the wrongful practices of an entity that causes a harm") to address the adverse impacts, including collaborating with others to increase this leverage.

GIFCT does not currently have a framework in place for responding to terrorist and violent extremist content from all parts of the ideological spectrum in the hash-sharing database, though it has launched a new effort to address this. BSR notes that addressing bias in the hash-sharing database likely involves expanding the content of the hash-sharing database beyond content related to organizations on the United Nations Security Council's consolidated sanctions list, and this will present its own risks that need to be addressed—for example, to prevent the overbroad inclusion of content.

In addition to terrorist and violent extremist content, should GIFCT address issues "higher up the chain," such as the radicalization and recruitment process?

Principle 19 of the UNGPs emphasizes the responsibility to exercise leverage to address adverse human rights impacts, where leverage is defined as "the ability to affect change in the wrongful practices of an entity that causes a harm." In BSR's experience, it is important to take a systems-based approach to leverage, and in the case of GIFCT, this may involve not simply issues relating to content removal, but interventions further up the chain, such as the radicalization and recruitment process. Indeed, GIFCT's strategy already includes a "prevention" pillar, designed to "equip digital platforms and civil society groups with awareness, knowledge, and tools to develop sustainable programs to disrupt terrorist and violent extremist activity online."

The majority of external stakeholders stated that GIFCT should play a role in addressing higher-up-the-chain issues. However, they also noted the complex and varied ecosystem of these efforts around the world and the role of localized, on-the-ground responses, and the importance of GIFCT identifying its value add to these spaces, rather than duplicating existing efforts.

To address potential adverse human rights impacts, GIFCT has a responsibility to help address (rather than entrench) this bias, such as through research, advocacy, or by engaging with affected stakeholders throughout its work.

6.2 Recommendations

01 RECOMMENDATION

Develop a common understanding of terrorist and violent extremist content.

The process to create the common understanding should commence with terrorist content, and then move on to violent extremist content, where the more challenging definitional issues will exist.

BSR proposes that the common understanding be created in a collaborative and multi-stakeholder manner, including the new IAC taxonomy subgroup, relevant working groups, and Global Network on Extremism and Technology (GNET), the academic research arm of GIFCT GNET research. There should be consultation with relevant UN Special Rapporteurs, dialogue with affected stakeholders, and transparency about how final resolutions are reached.

These common understandings should be written to a level of granularity that is actionable and practical for companies to use, including, but not limited to, requirements for content added to the hash-sharing database.

While each company will likely retain its own definition, these common understandings can shape and inform the work of GIFCT. They should be reviewed on a regular basis to ensure they accurately reflect reality and adequately limit overreach.

01 EXPLANATION

Principle 16 of the UNGPs states that company human rights policy should be "informed by relevant internal and/or external expertise."

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

The lack of shared definitions of terrorist and violent extremist content is a human rights risk when governments and other entities deploy vague or overbroad definitions to attack legitimate speech and target rightsholders—an authoritative and multi-stakeholder defined common understanding can be used as a collaborative bulwark against these efforts.

Implementing this recommendation will improve the capability of smaller companies without extensive policy teams and improve shared awareness of the relationship between human rights and terrorist and violent extremist content. During this assessment BSR encountered significant interest from stakeholders and experts inside and outside GIFCT to participate in this effort.

02 RECOMMENDATION

Build the common understanding of terrorist and violent extremist content on "behavior" rather than "groups."

GIFCT should align its approach with the definitions proposed by the UN Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism.

The common understanding should be as narrow as possible, so as to limit the adverse human rights impacts that would result from "scope creep" into other objectionable speech areas. For example, they could include clear red lines and thresholds, and reference to relevant legal standards. They should also include "guardrails" to prevent slippage into other extremist content and hate speech.

The common understanding can be used to inform specific GIFCT activities, such as labels, descriptions, and criteria for inclusion in the hash-sharing database.

02 EXPLANATION

Principle 16 of the UNGPs states that company human rights policy should be "informed by relevant internal and/or external expertise."

During this assessment there was near consensus among experts and stakeholders to define terrorist and violent extremist content on "behavior" rather than "groups."

Focusing on groups enables the criminalization and stigmatization of group membership by associating entire communities with terrorist groups, which in turn can adversely impact rights to freedom

of association, freedom of religion, and non-discrimination, among others.

Group-based definitions have been used in ways that have contributed to anti-Islamic bias and discrimination against Muslim communities.

Group-based definitions are not suited to the full range of terrorist threats, such as right-wing terrorism and white supremacy, where group membership is not always a feature.

03 RECOMMENDATION

Participate in efforts to pursue counterterrorism and violent extremism priorities from a holistic and strategic perspective.

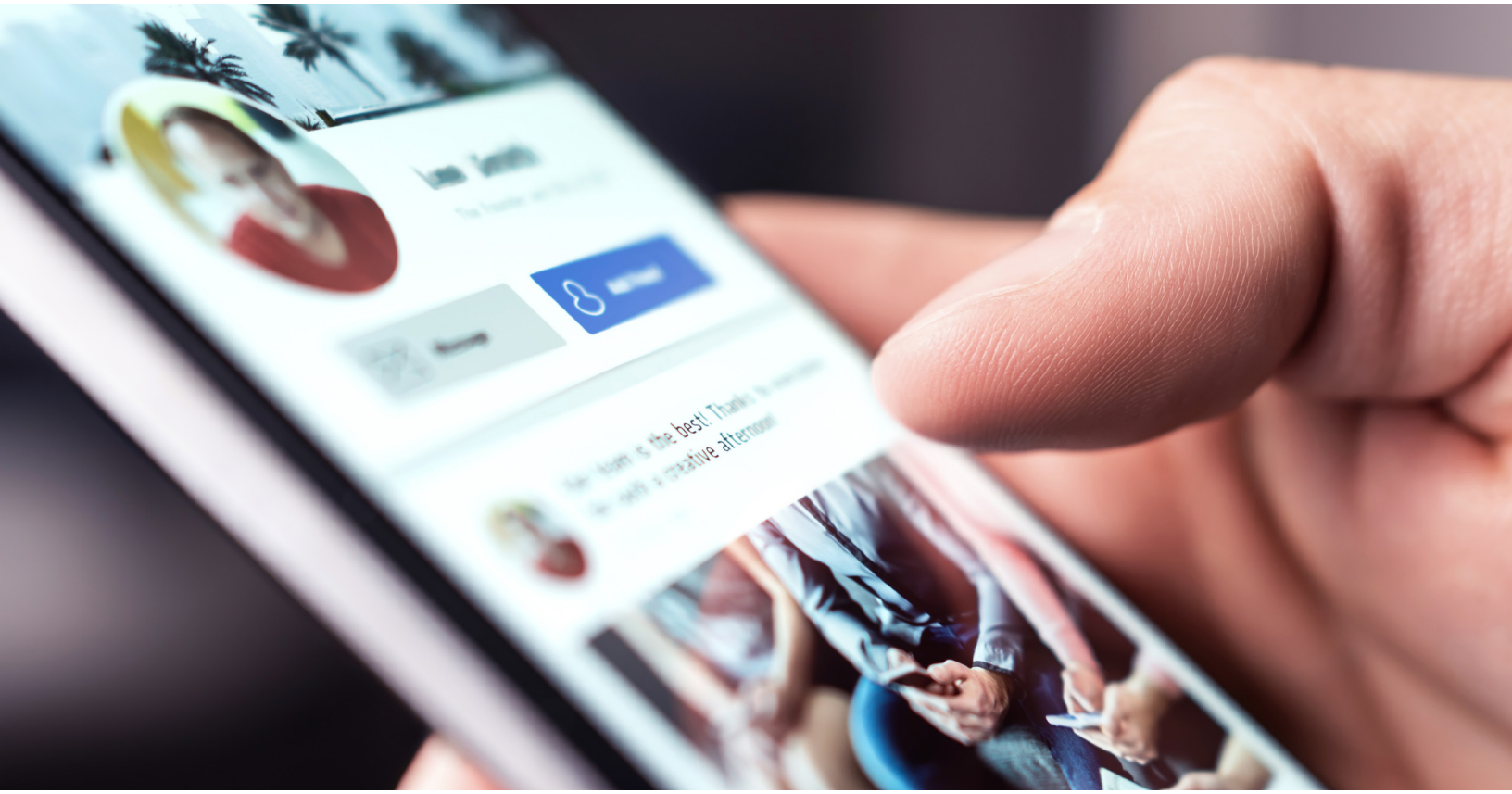
The online environment is only one element of a broader counterterrorism and violent extremism agenda, and GIFCT can usefully participate in efforts to examine the root causes of terrorism and violent extremism, and shape an agenda for how this can be addressed online.

GIFCT should be thoughtful and cautious when entering areas such as prevention, counter-speech, and de-radicalization—for example, focusing on areas of greatest complementarity with other efforts. This might include, for example, convening in-country practitioners about the online sphere so that they can apply lessons learned in their own work.

03 EXPLANATION

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

Experts interviewed for this assessment emphasized that the online and offline environments cannot be neatly divided, and that interaction between the online and offline goes both ways—offline harm can appear in the form of online content, and online content can be used to inspire, organize, and incite offline harm.



7. Content Removal And Preservation

IN THIS SECTION

7.1 Analysis

7.2 Recommendations

7.1 Analysis

Several of the human rights impacts identified in section 5 concern the removal and preservation of content, and these raise very specific questions relating to the hash-sharing database and the preservation of removed content that may have value as evidence in legal and access-to-remedy processes at a later date.

In addition to the responsibility to take action to address potentially adverse human rights impacts, this issue also raises questions of transparency. Here the UNGPs state that companies should communicate externally when concerns are raised by or on behalf of affected stakeholders in a form and frequency that reflects human rights impacts and that are accessible to intended audiences.[17]

How should the freedom of expression risks associated with the hash-sharing database be addressed?

During this assessment, affected stakeholders raised with BSR the concern that overbroad inclusion of hashes in the hash-sharing database may lead to the removal of legal and legitimate content, including legitimate political speech, journalistic reporting, and content that is intended to expose, analyze, and draw attention to human rights violations. This concern becomes more salient given GIFCT's expansion to addressing violent extremism as well as terrorism, which is even more challenging to define and risks a slippery slope into other content areas, such as hate speech and nonviolent extremism.

Several stakeholders also raised the concern that overbroad inclusion could arise via government pressure to include certain content in the hash-sharing database, which companies may feel obligated to acquiesce to given the trend of increased regulation in this area. Further, concerns were raised that bias in the counterterrorism field (see section 6) may result in disproportionate and discriminatory impacts on Muslim rightsholders, who are especially vulnerable to overbroad content removals.

Another key concern of affected stakeholders was that, because a few large companies contribute the majority of hashes to the database, the database could export their models of content governance across the internet. This was especially worrying for stakeholders given the risk that smaller companies may not have the resources to conduct human review of the content identified via hashes on their own platforms. BSR explored this concern during the assessment, and it is noteworthy that stakeholders familiar with the processes of small companies stated that the opposite is often true—smaller companies often do not have the technical capacity for automated moderation, and therefore conduct human reviews out of necessity.

Although the hash-sharing database receives a disproportionate amount of external attention, it is important to note that similar concerns exist with URL sharing. Additionally, several stakeholders indicated that the hash-sharing database is just one possible technical solution to address terrorist and violent extremist exploitation of the internet, and that it will be important for GIFCT to carefully consider and address the potential adverse human rights impacts of any new solutions it explores in the future.

BSR's recommendations draw upon a range of suggestions we received to address freedom of expression risks, including restrictions and controls on who should be able to add hashes to the database, the ability to challenge hashes, conducting third party reviews of the hash-sharing database, governance and oversight of the hash-sharing database, and transparency (see below).

What is the right model for transparency for the hash-sharing database?

There is significant interest among affected stakeholders and relevant experts to gain greater insights into the content and use of the hash-sharing database. However, this also presents challenges owing to both the format of the database (a hash is not itself content, but a string of numbers and letters, and only the individual platforms have access to the content) and the nature of the associated content (hosting terrorist content is illegal in many jurisdictions).

BSR's assessment surfaced near-consensus on the need for a third-party review of the hash-sharing database. The key questions are about "how" rather than "whether" one is undertaken. BSR also observes that aligning third-party reviews of the hash-sharing database with an annual cycle of transparency by GIFCT would help fulfil the UNGPs expectation that disclosures happen in a form and frequency that meets the needs of intended audiences.

BSR's assessment also identified very different assumptions about the significance of the hash-sharing database for actions taken by companies to remove content. Larger companies viewed the database as predominantly a "source of tips and signals" to supplement their existing and far more expansive efforts to address terrorist and violent extremist abuse of their platforms, whereas smaller companies viewed the database as playing a very significant role in expanding the volume of content removed. Further, it was emphasized that companies will choose to use the database differently. For example, some platforms may allow certain content that others do not, and therefore choose to use only select categories of hashes.

In reality the significance of the database is likely to vary from company to company, and be more significant for the growing membership of smaller companies than the larger founding companies, which are able to use their own extensive resources, technical capabilities, and infrastructure.[18] However, this divergence illustrates that transparency

shouldn't be limited to the database itself, but also cover the relative significance of the database in broader company actions; this will require transparency from GIFCT member companies, in addition to GIFCT itself.

How should GIFCT address the need to preserve removed content that may be used as evidence later?

Article 8 of the UDHR states that "everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the constitution or by law," while the UNGPs state that companies should "provide for or cooperate in their remediation through legitimate processes" when associated with adverse human rights impacts.[19] Remedy can take a variety of forms, including satisfaction, restitution, guarantees of non-repetition, rehabilitation, and compensation.[20]

While GIFCT member companies are right to rapidly remove terrorist and violent extremist content, there is a risk that the ability of affected rightsholders to realize their right to effective remedy may be compromised if that content is deleted and no longer available for use in relevant processes, including to investigate serious crimes under international humanitarian law.

During this assessment three main processes were emphasized with BSR: (1) legal processes, including both national law enforcement processes and international tribunals trying those responsible for serious crimes under international humanitarian law; (2) transitional justice processes in countries emerging from periods of conflict and repression; and (3) social science research into terrorist and violent extremist content characteristics and how to prevent it. All three represent potential pathways to remedy for victims of terrorism and violent extremism, and all three may require their own unique approach.

Many stakeholders emphasized the removal, deletion, and lack of archiving of content as the main challenge in this context, especially if—as increasingly occurs—companies remove larger amounts of content before a relevant law enforcement agency becomes aware of its existence or is able to store its own copies. Even if a company does retain rather than delete removed content, it faces legal obstacles to holding on to or disclosing this content if it does not receive a law enforcement preservation request or valid warrant—for example, most content can only be retained for 90 days under the US Stored Communications Act. As companies face increasing legal requirements to remove terrorist content as swiftly as possible, the loss of valuable evidence is likely to increase.[21]

However, other interviewees emphasized the challenge of too much removed-but-not-deleted content for investigators to draw upon, and the challenge of identifying content most suitable for documentation and specific use cases. With such a large volume of removed content, there is a shared interest in identifying the most urgent, significant, and focused needs.

The link between these issues and GIFCT is two-fold: first, there is concern that use of the hash-sharing database may result in the more voluminous and rapid deletion of relevant content; second, there is interest in the role GIFCT may be able to play in addressing this challenge.

During this assessment BSR did not find many stakeholders recommending that GIFCT itself create and manage a repository of removed content for evidence retention. However, BSR encountered significant interest in GIFCT playing a constructive role in addressing this challenge, such as advocating for "carve-outs" for the legal restrictions companies currently face for storing removed content that may have evidential value later, or facilitating the identification of specific needs or cases.

How should GIFCT address the risk that the Content Incident Protocol (CIP) results in over-removal of content?

The CIP is a process by which GIFCT member companies become aware of, quickly assess, and

act on potential content circulating online resulting from a real-world terrorism or violent extremist event. It is based on the existence of content online relating to the real-world terrorism or violent extremism event and potential distribution of that content, including a live streaming.

For many stakeholders, particularly in government, the CIP is the most impactful part of GIFCT. While the CIP has received less public attention, given its role in coordinating company responses to identifying and removing content from real-world terrorist events, it does have significant implications for human rights. Several stakeholders expressed concern over the lack of transparency about the CIP and how it works, including when and how governments are able to contribute to the process.

7.2 Recommendations

Content Preservation

01 RECOMMENDATION

Convene multi-stakeholder discussions to advance acceptance and adoption of legal carve-outs for evidentiary content.

There is widespread interest across all stakeholder groups, including government and civil society, in addressing the challenge of content preservation. As a multi-stakeholder effort, GIFCT can play a convening role to bring together all major stakeholders to discuss this challenge and advance possible solutions.

These discussions could include best practices and guidance for how companies should retain evidence for the three potential different processes: law enforcement and international tribunal processes; transitional justice processes; and social science research.

These discussions could also (for example) result in a GIFCT point of view in favor of a "legal carve-out" or "safe harbor" for companies storing removed terrorist and violent extremist content that may be used as evidence later.

01 EXPLANATION

Principle 3 of the UNGPs states that governments should put in place laws that enable rather than constrain business respect for human rights.

Principle 7 of the UNGPs emphasizes the role of governments in supporting business respect for human rights in conflict-affected areas.

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

Article 8 of the UDHR states that "everyone has the right to an effective remedy by the competent national tribunals for acts violating the fundamental rights granted him by the while the UNGPs state that companies should "provide for or cooperate in their remediation through legitimate processes" when associated with adverse human rights impacts.

There was a shared reluctance among interviewees for GIFCT to act as a host for this evidentiary content. By contrast, many emphasized the importance of the company-to-law enforcement agency relationship—that, ultimately, processes and systems for the retention of content for access by law enforcement officials should be a matter for each company.

Content Incident Protocol

02 RECOMMENDATION

Conduct a review after the CIP has been initiated / completed to identify lessons learned, human rights impacts, content for use later, and opportunities for transparency.

Company action in response to a crisis must be swift, so it will be difficult to adequately review whether or not responses are effective and what adverse human rights impacts may have occurred in the process.

However, it is important to ensure that adverse human rights impacts are identified, lessons learned are shared, and (where appropriate) removed content is preserved for evidentiary purposes.

GIFCT should publish summaries of these reports publicly, with due consideration for operational effectiveness, stakeholder safety, and national security.

BSR notes that company debriefs already take place in the aftermath of a CIP, and that the Crisis Response Working Group is developing an expanded threshold for debriefs. In this context, BSR recommends that reviews (including human rights impacts) of the two CIPs declared to date be undertaken and published.

02 EXPLANATION

Principle 13 of the UNGPs states that companies should "seek to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts."

Principle 19 of the UNGPs states that "If the business enterprise has leverage to prevent or mitigate the adverse impact, it should exercise it. And if it lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example, offering capacity-building or other incentives to the related entity, or collaborating with other actors."

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

In feedback on drafts of this human rights assessment, some reviewers expressed an interest in human rights assessments of the two CIPs declared to date. BSR agrees with this feedback, but for reasons of scope and timing was unable to undertake CIP reviews as part of this assessment.

03 RECOMMENDATION

Publicly disclose information about the CIP process and relevant metrics when activated.

GIFCT should continue to enhance its explainer about the CIP, including how it works, what the thresholds are, and which actors are involved. It is particularly important to explain how government actors contribute to the process.

In addition, GIFCT should publish metrics as part of its transparency report. This might include:

- Number of proposed incidents
- Number of approved incidents
- Descriptions of each incident and why they were / were not approved
- Number of company content takedowns as a result / percentage removal

BSR notes that exploration of expanded transparency for the CIP is already underway.

BSR also notes that any public disclosure of information about the CIP process and relevant metrics should be reviewed to take into account implications for operational effectiveness and the need to protect security interests and investigations.

03 EXPLANATION

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

However, Principle 21 also notes that public communications should not "pose risks to affected stakeholders [or] personnel."

04 RECOMMENDATION

Enhance disclosure about collaboration that occurs for incidents not triggering the CIP.

The "bar" to declaring a CIP is high, and only two of over 100 potential cases have been declared a CIP since it was established in 2019.

However, it would be misleading to suggest that collaboration doesn't occur when an incident doesn't trigger the CIP. For this reason, BSR recommends enhanced disclosure—for example, on the types of incidents, the nature of collaboration that does take place, relevant human rights impacts, and lessons learned.

BSR notes that an exploration of a "second tier" of the CIP is already underway, and the notion of increased transparency and disclosure can be considered in this context.

04 EXPLANATION

Principle 13 of the UNGPs states that companies should "seek to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products, or services by their business relationships, even if they have not contributed to those impacts."

Principle 19 of the UNGPs states that "If the business enterprise has leverage to prevent or mitigate the adverse impact, it should exercise it. And if it lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example, offering capacity-building or other incentives to the related entity, or collaborating with other actors."

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

Hash-sharing Database

05 RECOMMENDATION

Use the GIFCT "common understanding" of terrorist and violent extremist content to determine inclusion in the hash-sharing database.

GIFCT's recently announced a multi-stakeholder effort to create an expanded and more clearly defined taxonomy for the hash-sharing database is an important step toward more clearly bounding the types of content that may be included in the database. The effort will also help GIFCT move away from the previous approach that limited content to UN-designated terrorist organizations and therefore had both (1) potentially discriminatory impacts on Muslim communities and (2) failed to adequately address the much broader ideological scope of terrorism today.

Throughout this process, GIFCT should ensure the taxonomy is consistent with international human rights standards (see BSR's recommendation on a "common understanding" for terrorist and violent extremist content definitions in section 6). These common understandings should be as clearly and narrowly scoped as possible to prevent scope creep and improper additions to the database and subsequent removals, and should avoid over-reliance on group membership rather than behavior / actions.

05 EXPLANATION

The purpose of this recommendation is to establish an unambiguous rule for content submitted to the database, thereby protecting rights to freedom of expression, association, and assembly, and nondiscrimination.

Principle 19 of the UNGPs states that companies should "integrate the findings from their impact assessments across relevant internal functions and processes."

06 RECOMMENDATION

Introduce and expand transparency and oversight mechanisms alongside the extension of content in the hash-sharing database.

Given the gray area around violent extremist content, and the risk of a slippery slope into other objectionable forms of content, appropriate transparency and oversight mechanisms should be in place to ensure that risk is adequately addressed.

06 EXPLANATION

The purpose of this recommendation is to protect the rights to freedom of expression, association, and assembly, and nondiscrimination by ensuring the inclusion of violent extremist content in the hash-sharing database does not result in the removal of borderline and/or legitimate content.

07 RECOMMENDATION

Require contributing companies to conduct human review and approval prior to adding hashes to the database.

GIFCT member companies who contribute hashes to the hash-sharing database should be required to undertake manual review of any contributions to ensure they meet the standards outlined in the forthcoming taxonomy. Ideally this review would involve two staff members (to reduce error rates), with final sign-off from a senior staff member—though this may not always be possible in smaller companies.

07 EXPLANATION

The purpose of this recommendation is to protect the rights to freedom of expression, association, and assembly, and nondiscrimination by ensuring hashes added to the database are limited to clearly defined terrorist and violent extremist content and do not result in the removal of borderline and/or legitimate content.

This is an example of a recommendation where implementation by smaller companies may take a different form (e.g., fewer reviewers) than implementation by larger companies.

08 RECOMMENDATION

Do not allow governments to directly add hashes to the database, and ensure any content that governments recommend for inclusion undergoes a review against the standards outlined in the taxonomy.

In order to ensure the integrity of the hash-sharing database and to prevent any political weaponization, government actors and/or entities working on behalf of governments should not be allowed to directly contribute hashes (as is the case today).

Government actors may recommend content be included either through GIFCT or individual member companies. However, any such recommendations should undergo a review against the standards outlined in the forthcoming taxonomy. If it is deemed to meet these standards, it should include a label that indicates it was recommended by X government.

Additionally, GIFCT should publicly disclose the number of submissions by governments, by which governments, the details of that content, and whether or not the submissions were approved or not. This should be included as part of the recommended expansions to GIFCT's transparency report (see below).

08 EXPLANATION

The purpose of this recommendation is to protect the rights to freedom of expression, association, and assembly, and nondiscrimination by ensuring hashes added to the database are limited to clearly defined terrorist and violent extremist content and does not result in the removal of borderline and/or legitimate content.

Principle 3 of the UNGPs states that governments should ensure that laws and policies "do not constrain but enable business respect for human rights" and "provide effective guidance to business enterprises on how to respect human rights throughout their operations."

Principle 23 of the UNGPs states that companies should "seek ways to honor the principles of internationally recognized human rights when faced with conflicting requirements."

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

09 RECOMMENDATION

Require companies utilizing the hash-sharing database for content moderation decisions to have an adequate appeals mechanism, and build in relevant remedy capabilities into the functioning of the database itself.

GIFCT should require all companies that utilize the hash-sharing database to have an adequate appeals process, consistent with the UNGPs requirements for effective operational grievance mechanisms. This requirement should include appropriate flexibility for resource constrained smaller companies.

In addition, mechanisms can be built into the hash-sharing database itself to facilitate remedy. For example, the recent label / dispute functionality can

be built out to enable companies to explain why they dispute a hash and include a label to mark hashes that an individual company has issued an appeal on. The database could also include an automatic block on hashes that receive a certain number of disputes, triggering review by appropriate stakeholders within GIFCT to verify whether the content meets the taxonomy requirements.

09 EXPLANATION

According to the UNGPs, companies must provide remedy to affected rightsholders for adverse human rights impacts that they cause or contribute to. One key aspect of this is establishing operational grievance mechanisms. In the context over erroneously removed content, this means establishing effective appeals mechanisms.

Principle 31 of the UNGPs lays out the effectiveness criteria for operational grievance mechanism. These include legitimacy, accessibility, predictability, equity, transparency, rights-compatible, and a source of continuous learning.

GIFCT member companies choose how to utilize the hash-sharing database and removal of content is done by the companies themselves, so primary responsibility for the remedy of erroneously removed content lies with each company.

However, enabling "blocks/holds" on highly disputed hashes is an important integrity mechanism to consider as GIFCT membership expands, because currently only the contributing company can remove a hash from the database.

This is an example of a recommendation where implementation by smaller companies may take a different form than implementation by larger companies.

10 RECOMMENDATION

Investigate how to enable third-party reviews of the hash-sharing database to assess whether hashes are consistent with the GIFCT taxonomy.

This review should include the ability to challenge and remove content that is inconsistent with the taxonomy. It should take place annually, and the results should be publicly disclosed.

Despite widespread agreement that third-party reviews of the hash-sharing database are necessary, it not yet clear how such reviews can be carried out given the hash-sharing database itself has no content. A key goal of GIFCT's transparency working group should be to work with the GIFCT Director of Technology to assess and recommend the best approach to these reviews.

Stakeholders consulted for this assessment proposed several possible approaches to third-party reviews, including:

- Requiring all companies contributing hashes to host their own databases of content for reviewers to access.
- GIFCT creating and hosting a central repository of content.
- Reviewing hashes against a body of known terrorist content, such as the Tech Against Terrorism database (TCAP). This approach is most feasible in the short term because it would work with the way the hash-sharing database is currently structured and managed.

BSR notes that one key barrier to third-party audits are legal barriers to holding removed content. Comprehensive third-party reviews of the hash-sharing database may therefore not be possible until legal carve-outs are achieved.

10 EXPLANATION

This recommendation is intended to ensure effective and transparent oversight of the hash-sharing database.

There is a debate about whether a repository of content included in the hash-sharing database should be created—either a single centralized repository, or repositories maintained by each company containing the content that corresponds to the hashes they have submitted.

For some, a repository would be essential to enable third-party review and the preservation of content that may have value as evidence later; for others, a "honeypot" repository would present too many legal, privacy, and security risks to be practical.

11 RECOMMENDATION

Develop a process for enabling researcher access to the hash-sharing database and associated content.

Defining how to enable access by researchers could take place in tandem with the development of a third-party review process.

Given the sensitivity of the content, it will be important to establish an approval process and adequate controls to prevent privacy violations or other adverse consequences. For this reason, BSR recommends that the granting of researcher access initially only take place through the GNET workstream.

11 EXPLANATION

This recommendation is intended to improve social science research into how to most effectively prevent terrorists and violent extremists from exploiting digital platforms.

Several reviewers recommended that the only third-party access to the hash-sharing database should be for transparency and accountability purposes, not social science research. The main concerns were: (a) lack of control over "who" would have access, and (b) unclear "purpose" and "outcome" of researcher access. A more nuanced version of this feedback is that researcher access should be allowed, but only with very clear controls, parameters, and desired outcomes in place. By limiting access to GNET participants, we believe that our recommendations address these concerns, while still enabling the intended benefit of researcher access—i.e., research into the most effective methods to counter terrorist and violent extremist activity online in a rights-respecting manner.

12 RECOMMENDATION

Publish a detailed explanation of how the hash-sharing database functions and conduct webinars for interested stakeholders.

External stakeholders do not adequately understand how the hash-sharing database works. This has resulted in the proliferation of numerous myths and misunderstandings, and also limits the ability of stakeholders to make constructive recommendations for improved management and transparency.

To address this, GIFCT (with significant assistance from Facebook ThreatExchange, which hosts the database) should create a detailed written explanation of how the hash-sharing database works and publish it on the GIFCT website.

The explanation should include a start-to-finish description of the process for adding and classifying hashes, how companies can customize their use of the database, labeling / dispute mechanisms, how hashes are removed, and technical challenges relating to hash quality.

The explanation should include a description of how the database is governed, including the role of the different actors involved. It should also include an explanation of the taxonomy, with illustrative examples, and a list of terrorist groups included.

This document should be updated when changes to the functioning and management of the hash-sharing database are made.

To augment the release of this document, GIFCT should also release a document with a more technical explanation of hash-sharing database functionality, and host webinars for interested parties to learn more about how the database functions.

12 EXPLANATION

This recommendation is meant to ensure effective, transparent oversight of the hash-sharing database.

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

13 RECOMMENDATION

Establish a multi-stakeholder process to develop metrics about how the hash-sharing database is used.

Rather than define specific metrics ourselves, we believe there is benefit from a process involving GIFCT members, participants, and other stakeholders

(but not governments) to agree on the most decision-useful metrics for disclosures.

That said, BSR suggests the following metrics for consideration:

- # hashes added, by whom, about what groups, and with what tag(s)
- # hashes disputed, by type, group, reason
- # hashes removed following appeal, by reason
- # hashes of disputed legal standing
- # hashes broken down by type / taxonomy / terrorist organization (including those triggered by the CIP)

These metrics would be published annually by GIFCT, with due consideration given to the risk that disclosure may lead to adverse human rights impacts (e.g., demands by governments for a significant / overbroad increase in hashes, if they conclude that volumes are too low).

These metrics would be accompanied by a narrative explaining how the database works in practice (see previous recommendation), provide insight into trends, and explain why numbers increased or decreased in certain areas. For example, the narrative would explain the role of governments and the protections in place to mitigate the risk that governments exert pressure toward an over-inclusion of hashes.

13 EXPLANATION

Principle 21 of the UNGPs states that "in order to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

The hash-sharing database and its use are widely viewed by stakeholders as opaque, and many assumptions—some true, some false—are made about how it works and how it is used. This recommendation is intended to address these concerns.

There are also concerns that the hash-sharing database may augment the "bias" of major social media platforms across the entire internet.

Some of the questions raised with BSR included: How is AI used? Do governments play a role? Can governments add hashes to the database? What are the rates of appeal? How often are the classifications rejected?

It is important to involve external stakeholders in this process to co-create greater transparency around the database.

14 RECOMMENDATION

Require companies that contribute to and utilize the hash-sharing database to commit to specific disclosures.

Most of the information external stakeholders seek to know about the hash-sharing database relates to how the companies use it. To facilitate this, GIFCT should establish minimum transparency requirements for member companies that contribute to and/or utilize the hash-sharing database. This information can be published as part of companies own transparency reporting processes, and be linked to in the GIFCT transparency page.

Because companies utilize the hash-sharing database in different ways, it may not be possible to have standards reported across all members. However, GIFCT should utilize its Transparency Working Group, liaise with member companies, and consult with external stakeholders to establish minimum requirements and suggested areas for disclosure.

Areas of interest and specific metrics that have arisen during stakeholder interviews include:

- How companies utilize the hash-sharing database, including whether they conduct human review of content flagged on their platforms via database utilization.
- # of pieces of content flagged and removed as a result, segmented by taxonomy.
- # of appeals as a result of removal, along with acceptance rate, segmented by taxonomy.
- # of appeals as a result of content posted for documentary, journalistic, or artistic reasons.
- Time intervals to removal of content.
- Whether or not hashes are used on their own or go into a broader terrorist and violent extremist content review stream.
- How companies judge the effectiveness of terrorist and violent extremist content removal.

This requirement should include appropriate flexibility for resource-constrained smaller companies, and the BSR recommendation in section 9 below that GIFCT provide technical assistance to smaller companies to address human rights risk is relevant here.

14 EXPLANATION

Principle 21 of the UNGPs states that "to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

This is an example of a recommendation where implementation by smaller companies may take a different form than implementation by larger companies.

15 RECOMMENDATION

Enable multi-stakeholder governance of the hash-sharing database to the extent possible under the current management model (i.e., hosted by Facebook Threat Exchange), and develop a plan for long-term governance and oversight.

GIFCT should hold multi-stakeholder discussions to plan for the long-term management and oversight of the hash-sharing database.

GIFCT should work with Facebook Threat Exchange to increase its independent oversight of the hash-sharing database. This should include regular meetings to discuss trends, challenges, and new features, and regular report-outs to the IAC and operating board. GIFCT should consider creating a hash-sharing database subcommittee with both company and IAC members to facilitate oversight. Governments should not form part of the multi-stakeholder governance of the hash-sharing database.

15 EXPLANATION

Many have suggested that GIFCT itself should host and manage the hash-sharing database, in addition to having oversight over it. Although independent management makes logical sense, there are challenges, such as cost relative to GIFCT's overall budget and access to consistent engineering resources. Adequate resourcing has been a problem for other similar hash databases in the past, such as the NCMEC database for child exploitation material.

8. Theory of Change and Programmatic Priorities

IN THIS SECTION

8.1 Analysis

8.2 Recommendations

8.1 Analysis

The UNGPs establish the expectation that companies integrate the findings from their impact assessments across relevant internal functions and processes, and take appropriate action.[22] In the case of GIFCT, BSR interprets this principle as meaning that GIFCT should develop a theory of change and a mix of programs that enables effective action on human rights.

At the time of writing, GIFCT's three main programmatic areas are technical innovation (i.e., the hash-sharing Consortium, Content Incident Protocol, and URL Sharing), research (i.e., GNET, acting as the academic research arm of GIFCT) and expertise development (i.e., the various GIFCT working groups). This section reviews these programmatic priorities in the context of GIFCT's human rights impacts, and considers whether additional and/or revised programmatic priorities are merited.

What is the relative role of different GIFCT workstreams as it relates to protecting and respecting human rights?

While recognizing that there will always be resource constraints, BSR's instinct is toward a strategic, holistic, and systematic approach to preventing terrorists and violent extremists from exploiting digital platforms. This integrated approach would encompass research, expertise development, proactive engagement with relevant policymakers, and services (such as the hash-sharing database) provided to members.

Terrorism and violent extremism are "whole of society" problems that require whole of society solutions, of which GIFCT is one part. However, as a multi-stakeholder effort, GIFCT has the opportunity to convene stakeholders for constructive and solutions-oriented dialogue in areas that have long been characterized by exclusivity and opaqueness.

In addition to focusing on the rapid removal of terrorist content, GIFCT can provide an important space for reflection, learning, and expertise development. It was frequently noted during this assessment that while most public attention has been focused on the hash-sharing database, the opportunities provided by GIFCT for field building can be one of its most valuable contributions.

Should GIFCT have an independent and expert point of view on rights-based approaches to addressing the exploitation of digital platforms by terrorists and violent extremists, such as what policies, actions, and strategies governments and companies should deploy?

The UNGPs emphasize the importance of exercising leverage to address human rights impacts, where leverage is defined as the ability to affect change in the wrongful practices of an entity that causes a harm.[23] In BSR's experience, seeking to influence the policies, actions, and strategies of governments can form an important element of leverage, given how significant government action is to defining the human rights impacts of technology.

In addition, the first pillar of the UNGPs emphasizes the role of governments in creating and enforcing laws that are aimed at companies respecting human rights, ensuring that laws do not constrain but enable business respect for human rights, and supporting business respect for human rights in conflict-affected areas.[24]

During interviews and dialogue to inform this assessment, BSR sought views on whether GIFCT should develop an expert point of view on the rights-based laws, policies, regulations, and strategies needed to more effectively address the exploitation of digital platforms by terrorists and violent extremists, and proactively express this point of view with relevant governments, policymakers, and regulators.

BSR believes that the most important role GIFCT can play in public policy dialogue will be to combine (1) evidence based, nuanced, specialist, substantive, and technical insights into how to prevent terrorists and violent extremists from exploiting digital platforms, with (2) a strong point of view that counterterrorist activities should be undertaken in ways that respect human rights, and (3) the perspective that taking action online is only one element of a more holistic approach.

BSR surfaced considerable interest in the benefits of an independent multi-stakeholder effort expressing this view, combined with a need to take a step-by-step and cautious approach to engaging in policy dialogue. The emphasis should be on increasing the quality of the terrorist and violent extremist content dialogue, raising awareness of the technical feasibility of different policy solutions, and commentating on trends and real world events. The emphasis would not be lobbying on specific laws and regulations.

Where, when, and how should this point of view be conveyed, and who should decide what this point of view is?

As a multi-stakeholder effort, it will be important for GIFCT to have a clearly defined process for defining and developing this point of view. A human rights-based approach implies that this process should involve understanding the concerns of potentially affected stakeholders by consulting with them directly[25] and drawing upon feedback from internal and external sources.[26]

However, GIFCT's multi-stakeholder structure also presents a paradox: GIFCT's range of participants is an obstacle for reaching consensus on every potential topic upon which a point of view could be developed; at the same time, the process of dialogue also represents one of the most notable benefits of a multi-stakeholder approach, and significantly strengthens the credibility of any point of view that is reached.

Our recommendations are intended to combine the role of the IAC and GIFCT working groups in securing stakeholder insights with the need for Operating Board accountability and staff responsiveness. We believe that developing position papers and "fact sheets" on broad topics (for example, a common understanding of terrorist and

violent extremist content or best practices in the use of technology to address terrorist and violent extremist content) will enable a more nimble, rapid, and widely understood positioning when the need arises to craft position statements or provide more specific points of view, and help break down silos in ways that emphasize areas of common agreement. BSR notes that comfort with position-taking will improve over time and with practice.

What research agenda would support the development of a point of view on rights-based approaches to addressing the exploitation of digital platforms by terrorists and violent extremists?

GIFCT's research program already refers to human rights in the list of categories for research outputs. However, it is BSR's instinct that an opportunity exists to develop a deliberate research and expertise development agenda that includes a focus on the intersection between terrorist and violent extremist content and human rights—in other words, exploring in detail elements of a rights-based approach to addressing the exploitation of digital platforms by terrorists and violent extremists. Several participants raised with BSR the significant potential for GIFCT to be a thought leader on how counterterrorism and human rights intersect online.

What other human rights-focused activities might GIFCT undertake to achieve its mission?

Aside from a more deliberate engagement on public policy, the other significant innovation proposed during this assessment is training and capacity building. Specifically, several interviewees surfaced the potential for learning, training, and development programs that bring together participants from different backgrounds—such as companies, governments, and civil society organizations—to learn about rights-based approaches to preventing terrorists and violent extremists from exploiting digital platforms.

8.2 Recommendations

01 RECOMMENDATION

Develop position statements on the rights-based laws, policies, regulations, and strategies needed to more effectively address the exploitation of digital platforms by terrorists and violent extremists.

These position statements could take the form of briefing documents, fact sheets, and position papers that provide evidence-based, nuanced, specialist, substantive, and technical insights into rights-based approaches to countering terrorist and violent extremist content.

These position statements would seek to connect the theory with the practical—in other words, how counterterrorism and violent extremism priorities can be achieved in ways that both fulfill and respect international human rights instruments.

The process to create these statements should involve understanding the concerns of potentially affected stakeholders by consulting with them directly and drawing upon feedback from internal and external sources.

Specifically, BSR recommends that the IAC and working groups are actively involved in creating these public policy positions, with the Operating Board having final sign off.

BSR notes that these position statements would represent the view of GIFCT, and not necessarily the

consensus position of all its member companies and participants. The position statements are not intended to constrain GIFCT member companies and participants from undertaking their own public policy activities.

01 EXPLANATION

Principle 3 of the UNGPs states that governments should put in place laws that enable rather than constrain business respect for human rights.

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

Principle 16 of the UNGPs states that company human rights policy should be "informed by relevant internal and/or external expertise."

During the interviews to inform this assessment there was significant interest in the benefits of an independent multi-stakeholder organization expressing a constructive point of view that increased the quality of public dialogue.

02 RECOMMENDATION

Proactively express this point of view with relevant governments, policymakers, and regulators.

The GIFCT point of view can be shared during dialogue focused on the long-term development of sustainable and rights-based approaches—for example, during the development of regional counterterrorism strategies.

BSR also recommends that GIFCT staff refer to the relevant briefing papers, fact sheets, and position documents when undertaking "rapid response" to political events. In other words, using pre-existing content to enable a more nimble response in ways that maintain a shared point of view.

There should be an emphasis on increasing the quality of the dialogue, raising awareness of the technical feasibility of different policy solutions, commentating on trends and real world events, and highlighting both the potential and limitations of tackling terrorist and violent extremist content for the broader terrorism and violent extremism priorities.

The emphasis should not be lobbying on specific laws and regulations.

02 EXPLANATION

Public policy advocacy is a key way for GIFCT to address adverse human rights impacts in the counterterrorism and violent extremism field. It is important for GIFCT staff to have the "room" to express a point of view, and be confident that this point of view is consistent with that of the GIFCT membership.

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

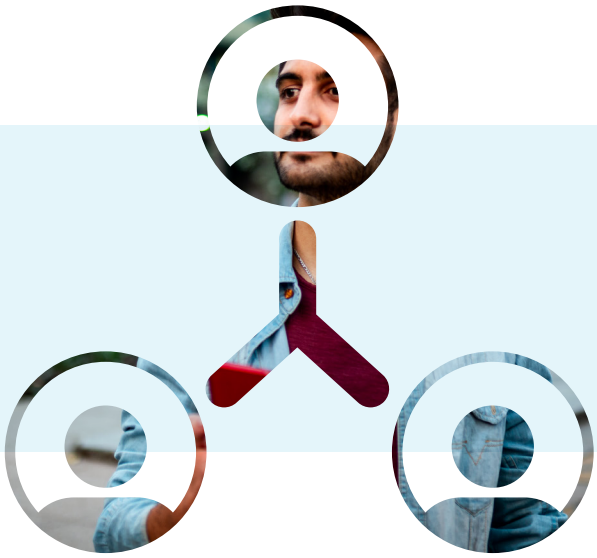
BSR notes that comfort with position-taking will improve over time and with practice.

9. GIFCT Membership

IN THIS SECTION

9.1 Analysis

9.2 Recommendations



9.1 Analysis

During the course of this assessment BSR encountered considerable debate around whether GIFCT should actively increase its company membership, especially (1) with companies headquartered outside the US and (2) with companies from elsewhere in "the technology stack." [27]

From a human rights perspective there are two dimensions to this question: (1) would increasing membership outside the US or elsewhere in "the technology stack" serve to advance or detract from GIFCT's mission, and (2) how can potential adverse human rights impacts arising from membership expansion be addressed?

How important is increasing company membership outside the US to the achievement of GIFCT's mission?

The UNGPs emphasize the importance of prioritizing the most severe impacts, stating that companies should "first seek to prevent and mitigate those that are most severe or where delayed response would make them irremediable."[28] In the context of GIFCT's global mission and remit, this suggests that a human rights-based approach should focus on the locations where impacts are most severe, rather than where they have the highest profile or media coverage.

For this reason, BSR concludes that GIFCT will be better positioned to prevent terrorists and violent extremists from exploiting digital platforms through more engagement with companies (and organizations) outside the US and Europe, rather than less. This is consistent with the views of many expressed during this assessment, that GIFCT will constrain its impact if it works with only a subset of the overall landscape of social media platforms, and that over time terrorist and violent extremist content may shift away from the "mainstream" platforms to other platforms outside of GIFCT's sphere of influence.

Further, we note that a key element of GIFCT's theory of change is to increase the expertise,

capability, and capacity of smaller companies to prevent terrorists and violent extremists from exploiting digital platforms, and that this also indicates a need to engage rather than avoid companies from outside the US and Western Europe. Increasing the capability of smaller companies would be a manifestation of GIFCT's leverage to affect change and have the potential to address adverse human rights impacts that might otherwise arise from counterterrorism efforts.

What minimum human rights-based criteria should be established for company membership? Under what human rights-based circumstances should companies be removed from membership? Are there options other than "full membership" that can help advance GIFCT's mission and grow membership while addressing human rights risk?

While BSR's instincts are toward a "big tent" on human rights grounds, expansion of membership would need to be undertaken in a manner that addresses potential adverse human rights impacts and maintains the integrity of GIFCT.

Here BSR's recommendations address five considerations: (1) the policies and commitments of applicant companies; (2) contextual factors, such as local laws, government pressure, and ownership that may impact the ability of a company to fulfill those commitments; (3) the potential for different membership categories that maximize benefits while minimizing harms; (4) the potential for enhanced transparency requirements; and (5) the ability to expel companies from membership in certain circumstances.

We are also cognizant of the need not to fall into the trap of viewing non-US companies and locations as having inherently greater risk; many stakeholders we engaged with for this assessment emphasized that the reverse can be equally true.

GIFCT already applies membership criteria, including content standards that prohibit the promotion of terrorism, the ability to receive and act on reports of illegal activity or activity violating terms of service, a desire to explore new technical solutions to content and conduct challenges, regular data transparency (i.e., a transparency report), a public commitment to respecting human rights when implementing content removal policies, support for expanding the capacity of civil society organizations to challenge violent extremism, and agreeing that governments will not be able to remove terrorist content directly from company platforms.

It is especially important to highlight the complexity of any additional human rights-based criteria for GIFCT membership. At first glance, it is reasonable to require member companies to achieve certain business and human rights benchmarks, such as having a human rights policy in place, providing evidence of human rights due diligence, and being transparent about its approach. However, in reality these criteria can be subject to local realities outside of the companies' own control—some companies may, for example, be under local legal expectations to provide direct access to law enforcement agencies or may be partially owned or controlled by a government associated with human rights harms or complicit in terrorist and violent extremist content activities.

Here it is BSR's instinct that the right approach for GIFCT is not to avoid these companies but rather to go in "eyes wide open" with appropriate measures in place to avoid, prevent, or mitigate adverse human rights impacts that may arise. This view is addressed in our recommendations and includes, for example, a company membership tier that allows for companies to participate in GIFCT activities (e.g., knowledge and expertise sharing) but not contribute to the hash-sharing database.

Further, it is important for GIFCT to consider the broader ecosystem of companies to influence regardless of membership—in other words, GIFCT's mission isn't going to be achieved by working with members only, but instead by engaging with a

broader range of actors, including nonmembers.

Are there companies elsewhere in "the stack" that can contribute toward the achievement of GIFCT's mission?

GIFCT's membership is currently limited to "companies operating internet platforms and services."[29] During the course of this assessment the January 6 insurrection in the US shed light on the role of companies elsewhere in "the stack" (such as cloud services providers, content delivery networks, and ISPs) in addressing the exploitation of digital platforms by terrorists and violent extremists, and therefore the potential to include them in GIFCT's membership.

BSR's instinct is that engaging companies elsewhere in "the stack" would enhance GIFCT's ability to achieve its human rights goals via more holistic approaches. However, as an initial step, this growth should be limited to companies that engage with content issues (e.g., cloud services companies, content delivery networks) rather than non-content challenges (e.g., surveillance).

9.2 Recommendations

01 RECOMMENDATION

Refine and publish human rights-based membership criteria.

GIFCT membership criteria on human rights should combine two elements—company human rights commitments and contextual factors.

BSR's proposed company human rights commitments are a moderate strengthening of existing GIFCT membership criteria.

Specifically, BSR recommends requiring (1) a public commitment to the International Bill of Human Rights and UNGPs; (2) a public commitment to the mission of GIFCT, including the view that combatting terrorism and violent extremism and ensuring respect for human rights are complementary and mutually reinforcing goals; (3) an ability to provide access to remedy via appeals channels for content decisions; (4) an annual public disclosure from the company about how they have participated in GIFCT, including how they have made use of GIFCT's technology innovation workstream (which includes the hash-sharing database); and (5) an annual private report to GIFCT staff on how the company is living up to membership commitments.

BSR's proposed contextual factors are also a moderate strengthening of existing GIFCT membership criteria, and we note that these can only be applied in a case-by-case fashion. Specifically, BSR recommends that GIFCT consider the extent to which (1) there is potential for home-government pressure and (2) the company ownership structure may have an adverse impact on GIFCT's mission. BSR does not propose a list of excluded countries; rather, each case would be considered on its merits.

Our recommendation on membership tiers (below) will help address these risks.

BSR recommends that the GIFCT membership criteria be published.

01 EXPLANATION

Principle 17 of the UNGPs states that "human rights due diligence should be initiated as early as possible in the development of a new activity or relationship."

Interviews undertaken to inform this assessment emphasized that the membership criteria represent the main source of leverage that GIFCT has over member companies, and that therefore it should be used to set out a structured view on how member

companies should fulfill their responsibility to respect human rights while undertaking efforts to prevent terrorists and violent extremists from exploiting digital platforms.

BSR notes that GIFCT may retain other membership criteria (such as not allowing direct government access to their platforms) not listed here—this recommendation covers the main human rights elements that can exist within broader membership criteria.

02 RECOMMENDATION

Refine the due diligence process for new company membership applications.

The existing GIFCT membership due diligence process includes a dialogue with GIFCT staff, a review by the IAC, and a final decision by the GIFCT Operating Board.

BSR recommends maintaining a very similar process, with the following key elements: (1) direct engagement by a designated member of the IAC and the GIFCT Executive Director with a senior executive leader from the company making the application to assess commitment to the GIFCT mission, (2) the option for the IAC to make a recommendation for or against membership, and (3) a final decision by the GIFCT Operating Board.

BSR notes that binary in/out membership criteria will not exist, and that GIFCT should develop comfort in the use of professional judgement on a case-by-case basis, especially as it relates to contextual factors.

As described in BSR's recommendation below, a decision to admit a company to the GIFCT membership has three potential pathways: (1) full membership, (2) associate membership, (3) observer status. The company applying for membership may request one of these three pathways; the IAC may advise one of these three; and/or the Operating Board may decide one of these three.

02 EXPLANATION

Involving a range of GIFCT stakeholders in membership applications where there could be human rights risks and implications for the work of GIFCT will be essential to maintaining the organization's integrity.

BSR recommends that the Operating Board rather than the IAC make final membership decisions for reasons of accountability and consistency. The IAC is an advisory body, and so should advise on membership; the Operating Board is a decision-making body, and so should make decisions on membership.

Further, given time and resource constraints, BSR proposes that the IAC be informed of all membership applications and given the option to advise on them; it may choose to engage with some membership applications and "pass" on others.

Principle 17 of the UNGPs states that "human rights due diligence should be initiated as early as possible in the development of a new activity or relationship."

03 RECOMMENDATION

Establish a tiered membership structure for GIFCT.

GIFCT membership should include four tiers: (1) founding member companies, (2) full member companies, (3) associate member companies, (4) observers.

The associate member companies would have access to the full benefits of GIFCT membership (e.g., access to use the hash-sharing database; participation in the crisis incident protocol; participation in working groups), but would not be able to add to the hash-sharing database or other technology-based solutions that may be developed over time.

The observer member category would apply to all new members and last for one year, followed by the option to apply for full or associate membership. BSR recommends that observer member companies be able to participate in working groups and learning activities, but not the technology innovation workstream—rather, observers would receive mentorship by Tech Against Terrorism.

It was noted during discussions about this recommendation that the relevance of GIFCT member benefits (e.g., the hash-sharing database) will vary according to business model, and that a rights-respecting approach should require that GIFCT member companies are only able to access the benefits that are relevant for them.

For all membership categories, there should be a probationary period of up to six months.

03 EXPLANATION

The observer member category is intended to provide potential new members and GIFCT / existing members the opportunity to get to know each other—as a result, the ambition to increase company membership could be achieved in a manner that minimizes human rights risks, particularly with regard to the hash-sharing database.

Principle 17 of the UNGPs states that "human rights due diligence should be initiated as early as possible in the development of a new activity or relationship."

04 RECOMMENDATION

Establish a process to (1) expel companies not living up to their membership commitments and / or (2) alter a company's membership tier

BSR recommends five criteria for removal or membership tier alteration: (1) abusing the hash-sharing database by persistently adding nonqualifying content, (2) no longer meeting GIFCT membership criteria, (3) changes in company ownership structure that present a material threat, (4) changes in local laws and regulations that present a material threat, (5) other factors with a material adverse impact on the achievement of GIFCT's mission.

The process should be based on a "strike system" where companies have the opportunity for course correction.

The process should include direct engagement between a designated member of the IAC and the GIFCT Executive Director with senior executive leadership from the company potentially being removed.

The final decision for removal should involve (1) a recommendation for or against expulsion by the IAC, and (2) a final decision by the GIFCT Operating Board.

04 EXPLANATION

It is an important principle of any membership organization that if there is a route in, there should also be a route out if commitments are not being met.

BSR recommends that the Operating Board rather than the IAC make final membership decisions for reasons of accountability and consistency—the IAC is an advisory body, and so should advise on membership; the Operating Board is a decision-making body, and so should make decisions on membership.

Principle 19 of the UNGPs states that "there are situations in which the enterprise lacks the leverage to prevent or mitigate adverse impacts and is unable to increase its leverage...here, the enterprise should consider ending the relationship, taking into account credible assessments of potential adverse human rights impacts of doing so."

05 RECOMMENDATION

Actively recruit new member companies, especially from non-US locations.

At the time of writing nearly all the members of GIFCT are US-based companies. BSR recommends a proactive effort to recruit companies from outside the US, especially companies with a presence in markets facing severe terrorist and violent extremist content challenges.

In the context of GIFCT's global mission and remit, we believe that a human rights-based approach should focus on the locations where impacts are most severe, rather than where they have the highest profile or media coverage. For this reason, we believe GIFCT will be better positioned to prevent terrorists and violent extremists from exploiting digital platforms through more engagement with companies outside the US and Europe, rather than less.

This recruitment process should include an element of capacity building to prepare companies for

membership in GIFCT, including the use of observer status.

05 EXPLANATION

Principle 24 of the UNGPs emphasizes the importance of prioritizing the most severe impacts, stating that companies should "first seek to prevent and mitigate those that are most severe or where delayed response would make them irremediable."

We note that a key element of GIFCT's theory of change is to increase the expertise, capability, and capacity of smaller companies to prevent terrorists and violent extremists from exploiting digital platforms, and that this also indicates a need to engage rather than avoid companies from outside the US and western Europe.

Increasing the capability of smaller companies would be a manifestation of GIFCT's leverage to affect change and have the potential to address adverse human rights impacts that might otherwise arise from counterterrorism efforts.

BSR is cognizant of the need not to fall into the trap of viewing non-US companies and locations as having inherently greater risk; many stakeholders we engaged with for this assessment emphasized that the reverse can be equally true.

06 RECOMMENDATION

Actively recruit members from elsewhere in the technology "stack."

BSR recommends that GIFCT seek membership from companies elsewhere in the technology "stack" that engage with content issues, such as cloud services companies, content delivery networks, and domain registrars.

Over time, GIFCT should study the merits of expanding this reach to companies with predominantly non-content challenges (e.g., surveillance), such as ISPs / telecommunications companies.

06 EXPLANATION

Engaging companies elsewhere in "the stack" would enhance GIFCT's ability to achieve its human rights goals via more holistic approaches.

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

07 RECOMMENDATION

Provide technical assistance to smaller companies to address human rights risks

Providing technical assistance to small companies to help them prevent and respond to abuse of their platforms by terrorists and violent extremists (e.g., technology tools to assist with content removal) forms a core part of GIFCT's mission. BSR recommends that this technical assistance extend to include additional elements relevant to human rights risks, such as the ability to receive and act upon user appeals about content decisions, or to publish transparency reports.

07 EXPLANATION

Principle 14 of the UNGPs states that "The means through which a business enterprise meets its responsibility to respect human rights will be proportional to, among other factors, its size. Small and medium-sized enterprises may have less capacity as well as more informal processes and management structures than larger companies, so their respective policies and processes will take on different forms. But some small and medium-sized enterprises can have severe human rights impacts, which will require corresponding measures regardless of their size."

Principle 19 of the UNGPs states that "if the business enterprise...lacks leverage there may be ways for the enterprise to increase it. Leverage may be increased by, for example...collaborating with other actors."

10. Stakeholder Engagement

IN THIS SECTION

10.1 Analysis

10.2 Recommendations

10.1 Analysis

GIFCT contains some features of a multi-stakeholder initiative (i.e., non-companies actively participate in the work of GIFCT) but lacks others (i.e., decision-making power rests solely with companies). For this reason, it is especially important to implement a robust approach to stakeholder engagement where mutual expectations are widely shared.

Stakeholder engagement plays a central role in a human rights-based approach. For example, when assessing human rights impacts, the UNGPs expect companies to "seek to understand the concerns of potentially affected stakeholders by consulting them directly," and where this isn't possible, engaging with "credible, independent expert resources, including human rights defenders and others from civil society."[30] Similar references to stakeholder engagement arise in relation to tracking progress, communicating action, and providing access to remedy.[31]

It is BSR's conclusion that GIFCT would benefit from a more deliberate and meaningful integration of affected stakeholders into its work, including (1) establishing greater clarity around the channels for stakeholder input, (2) broadening the range of groups engaged, and (3) clarifying the role of governments in GIFCT.

What is the right stakeholder engagement model for GIFCT—for example, how should the IAC and working groups evolve?

During the course of this assessment BSR encountered a strong desire for much greater clarity on the role of the IAC and working groups as a channel for stakeholder input into the work of GIFCT. For example, interviewees were often unclear about whether their roles should or should not encompass elements such as making formal

recommendations to GIFCT, co-creating solutions with other participants, or issuing public statements.

BSR makes several recommendations to address the desire for increased clarity, with a focus on stakeholder engagement as an essential element of a human rights-based approach. We propose a two-part model whereby the IAC and the working groups play distinctly different roles.

In short, we propose that the IAC deliberates on difficult dilemmas and challenges, makes formal recommendations to the Operating Board, reviews the GIFCT annual transparency report, and issues an annual independent statement. It would focus on review and accountability.

By contrast, we propose that the working groups focus on the collaborative development of solutions, including by engaging with a wide range of affected stakeholders and other experts. It would focus on outputs of practical value for member companies and other participants.

Which civil society organizations and academic experts should participate to ensure coverage of the most relevant human rights impacts?

The UNGPs are very clear in emphasizing the importance of engaging with "affected stakeholders" and paying special attention to populations that may be at heightened risk of vulnerability or marginalization. In the case of GIFCT this includes both the actual and potential victims of terrorist and violent extremist content (e.g., those whose rights will be more effectively realized and fulfilled when GIFCT is successful) and the actual and potential victims of efforts to address terrorist and violent extremist content (e.g., those whose rights may be impacted by overbroad restrictions on speech or the deletion of content.) With this breadth in mind, it will be important that GIFCT maintain a map of affected stakeholders and other relevant organizations, taking into consideration the need to pay special attention to vulnerable groups. BSR makes several recommendations to fulfil this role.

BSR notes that several interviewees emphasized that GIFCT participants would benefit from hearing from victims—both the victims of terrorism and the victims of efforts to address terrorist and violent extremist content—more frequently than they do today.

What risks and opportunities exist with government involvement in GIFCT? How can they be addressed? What are the human rights risks and opportunities associated with GIFCT becoming "the go-to place" for governments during a crisis, and how should they be addressed?

Engagement with governments presents enormous opportunities to develop more holistic, preventative, and rights-respecting approaches to terrorist and violent extremist content, but also presents human rights risks, both real and perceived. During this assessment, BSR encountered four different perspectives on government involvement in GIFCT as it relates to human rights.

First, aside from Ghana, the absence of government participation from outside the US, Western Europe, Australasia, and Japan was noted as problematic given the prevalence of terrorist and violent extremist content in other geographies. This curtails the opportunity for interaction between companies, civil society, and governments on terrorist and violent extremist content in key parts of the world that would benefit from it, and limits the ability to co-create rights-based approaches to terrorist and violent extremist content in regions with few human rights protections in place.

Second, the opportunity for enhanced dialogue between companies, civil society, and governments on rights-respecting approaches to terrorist and violent extremist content were emphasized, with the GIFCT working groups in particular offering the potential for greater shared understanding. While there may be differences of emphasis—and a fear that some governments may "use" GIFCT to pressure companies into overbroad restrictions on speech—the majority view leaned toward more dialogue, learning, and engagement as a positive

contribution to the field.

Third, several stakeholders raised with BSR the concern that some governments may use their participation in GIFCT to pursue objectives that may not be consistent with rights-based approaches to terrorist and violent extremist content, such as overbroad additions to the hash-sharing database. BSR makes several recommendations to address this risk.

Finally, while there is concern that some governments may "use" GIFCT to pressure companies into overbroad restrictions on speech, there is also an acknowledgment that there are plenty of other channels and venues for interaction between companies and governments on terrorist and violent extremist content (e.g., internet referral units), and that the importance of GIFCT should not be exaggerated. Here the need for greater interaction with the UN Special Procedures system and UN Office of Counter-Terrorism, and for greater support for multilateralism generally, was emphasized by some interviewees.

Is there a size limit for civil society and government participation—what is the right balance between size and securing diverse voices?

While increased participation by civil society, victims' organizations, and "non-Western" governments and companies was of interest to many during this assessment, there was also a recognition of time and resource constraints that exist for GIFCT and its members. Here many interviewees emphasized the importance of GIFCT establishing an influence and engagement strategy beyond its immediate membership—for example, forums to engage and influence the actions of nonmember companies, or platforms to engage and receive the advice of stakeholders in a range of geographies. There was also interest in rotation models for non-company participants.

10.2 Recommendations

01 RECOMMENDATION

Continue mapping stakeholders to further identify organizations and experts that would increase the diversity of rightsholders whose voices are heard in GIFCT activities—and create plans for their involvement.

In particular, GIFCT would benefit from increased participation from groups representing both the victims of terrorism and the victims of efforts to address terrorism, organizations from underrepresented regions (notably Africa, the Middle East, and South Asia), interfaith organizations, and those able to share a gender-perspective on terrorist and violent extremist content.

These stakeholders could join GIFCT in a formal capacity (e.g., IAC) or in a more informal capacity via participation in working groups or other activities. For example, GIFCT could establish a workstream dedicated to engaging with and obtaining a variety of perspectives from a range of affected stakeholders outside GIFCT membership / participants.

01 EXPLANATION

Principle 18 of the UNGPs states that when assessing human rights impacts, business enterprises should "involve meaningful consultation with potentially affected groups and other relevant stakeholders."

Principle 20 of the UNGPs states that tracking progress on human rights should "draw on feedback from both internal and external sources, including affected stakeholders."

02 RECOMMENDATION

Establish and maintain closer relationships with the United Nations system.

This should include structured interaction with relevant UN Special Rapporteurs (e.g., promotion and protection of the right to freedom of opinion and expression; right to privacy; promotion and protection of human rights and fundamental freedoms while countering terrorism; discrimination against women and girls; minority issues; human rights defenders; racism, racial discrimination, xenophobia and related intolerance) as well as the Office of the United Nations High Commissioner for Human Rights (OHCHR).

It should also include relations with the UN Office of Counter-Terrorism.

02 EXPLANATION

GIFCT is one part of more holistic approaches to counterterrorism and violent extremism; it is important that multilateral approaches are cultivated and institutions of international human rights law are engaged.

03 RECOMMENDATION

Train GIFCT participants in principles of good stakeholder engagement.

BSR recommends that GIFCT staff, the four founding companies, and member companies receive training on the principles of effective stakeholder engagement. For example, this might include: (1) stakeholder identification and prioritization, (2) two-way dialogue, (3) proactive and timely communications, and (4) recognition of power and informational asymmetry.

This training will enable the improved integration of stakeholder inclusive approaches across GIFCT's workstreams.

03 EXPLANATION

Principle 20 of the UNGPs states that "In order to verify whether adverse human rights impacts are being addressed, business enterprises should track the effectiveness of their response. Tracking should...draw on feedback from both internal and external sources, including affected stakeholders."

During this assessment interviewees from all backgrounds (e.g., companies, IAC members, external stakeholders) welcomed the increasingly ambitious approach to stakeholder engagement being taken by GIFCT. At the same time, interviewees expressed a desire for this ambition to be realized through the application of stakeholder engagement best practices as a normal way of conducting business—for example, that stakeholder engagement should be viewed as an opportunity to learn, grow, and improve.

04 RECOMMENDATION

Consider geographical diversity when rotating government membership of IAC.

The advancement of GIFCT mission would benefit from participation by a more diverse range of government participants, with a focus on governments facing the most significant challenges with terrorism and violent extremism.

This will likely need to involve moving beyond governments currently participating in the Freedom Online Coalition (a current requirement for government participation). However, this expansion should take place with due consideration of government commitment to human rights and the rule of law. BSR recommends that the IAC be consulted about new government members.

04 EXPLANATION

During this assessment the lack of diversity in government representation was noted.

The first pillar of the UNGPs emphasizes the role of governments in providing advice, guidance, and support for companies to respect human rights throughout their operations.

11. Governance, Accountability, And Transparency

IN THIS SECTION

11.1 Analysis

11.2 Recommendations

11.1 Analysis

The concepts of governance, accountability, and transparency feature prominently in the UNGPs.

On governance, the UNGPs emphasize effective integration, including decision-making, budget allocations, and oversight processes,[32] and tracking the effectiveness of response to adverse human rights impacts.[33]

On transparency, the UNGPs state that, in order to account for how they address their human rights impacts, companies should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders. The UNGPs go on to state that communications should "be of a form and frequency that reflect an enterprise's human rights impacts and that are accessible to its intended audiences" and "provide information that is sufficient to evaluate the adequacy of…response to the particular human rights impacts involved."[34]

It is BSR's instinct that a strategic and deliberate approach to transparency should form an important part of GIFCT's approach to human rights, and that this will (1) enable enhanced accountability; (2) spread expertise, insight, and learning on how to prevent terrorists and violent extremists from exploiting digital platforms; and (3) address various myths and misunderstandings that exist about GIFCT.

Should GIFCT transition to a multi-stakeholder board?

The term "multi-stakeholder" is used to convey different meanings. In BSR's view, a "multi-stakeholder initiative" (MSI) is characterized by a decision-making structure where no single constituency (e.g., companies, civil society organizations, investors, governments) has a majority of the votes. By contrast, a "multi-

stakeholder approach" implies the formal involvement of different constituencies, and "multi-stakeholder engagement" implies informal engagement with different constituencies; in the latter two cases, a single constituency (typically companies) retains exclusive or majority decision-making power.

Using these definitions, GIFCT takes a multi-stakeholder approach (e.g., the IAC) and undertakes multi-stakeholder engagement (e.g., the working groups), but is not a multi-stakeholder initiative.

During the assessment BSR surfaced a range of perspectives on whether GIFCT should move from a multi-stakeholder approach to become a multi-stakeholder initiative. Those making the case in favor argued that providing non-companies with a greater stake in the decision making and operations of GIFCT (e.g., Operating Board seats) will enhance their commitment to collaborative approaches and better enable GIFCT to foster collaboration in pursuit of its mission. Accountability and transparency would also be enhanced. By contrast, others argued that civil society voices would be weakened by participation on the GIFCT Operating Board given the risk of "capture"; these respondents felt that a multi-stakeholder approach is better pursued through a strengthened IAC with clear strategic direction. Some argued that GIFCT should be primarily "by industry, for industry," and should focus solely on effective multi-stakeholder engagement, short of formal governance. Notably, these different views for and against different multi-stakeholder models varied within constituencies, as well as between them.

It is BSR's view that human rights goals requiring collaborative approaches are best achieved through multi-stakeholder initiatives with multi-stakeholder boards. However, we also believe that governance shifts of this magnitude should be undertaken deliberately (e.g., with extensive consultation and planning), and recognize that a transition to a multi-stakeholder board would be a significant shift for GIFCT given its young age. We're also cognizant that our scope of work is limited to a human rights assessment, not a full governance review of GIFCT. Our recommendation—that GIFCT's multi-stakeholder approach is strengthened now (e.g., via enhancing the role and effectiveness of the IAC) and that a more complete governance review be undertaken in two years—seeks to strike this balance.

What qualitative and quantitative disclosures from GIFCT would be useful for those with an interest in the human rights impacts of GIFCT?

GIFCT's disclosures should provide the information necessary for an evaluation of GIFCT's approach to human rights. This disclosure can take two forms. First, qualitative disclosure of GIFCT's activities—such as the nonconfidential elements of Board meetings, IAC meetings, and working group activities—will provide external stakeholders with much improved insight into the work of GIFCT. Second, quantitative disclosures about the scale of GIFCT's work—such as data relating to the hash-sharing database, URL sharing, and content incident protocol—will provide external stakeholders insight into the extent of potential human rights risks. On quantitative data, it is extremely important to note the importance of accompanying narrative interpreting the data; for the issues GIFCT works on, data going "up" or "down" isn't necessarily "good" or "bad."

In both cases, these disclosures would need to be balanced against reasonable restrictions on external communications described in the UNGPs, such as stakeholder safety, mitigating the risk of further human rights harms, and legitimate concerns of commercial confidentiality.[35]

Should there be any transparency requirements for member companies and other participants? If yes, what would they be?

At present GIFCT member companies are required to have a commitment to public data transparency, interpreted as a regular transparency report. However, GIFCT could

choose to implement stricter transparency requirements for member companies, and here there are two main options: (1) requirements that GIFCT member companies publish reports about their approach to terrorist and violent extremist content and how they prevent terrorists and violent extremists from exploiting digital platforms; and/or (2) requirements that GIFCT member companies disclose how they make use of GIFCT resources, such as the hash-sharing database.

11.2 Recommendations

01 RECOMMENDATION

Institute a system of formal recommendations from the IAC to the Operating Board, and formal responses from the Operating Board to the IAC.

BSR recommends that the IAC be tasked with making formal recommendations to the Operating Board. These recommendations could be requested of the IAC by the Operating Board, or initiated proactively by the IAC. In either case, the Operating Board would make a formal response to the IAC recommendation.

This exchange between the IAC and Operating Board (with appropriate redactions where full disclosure would compromise stakeholder safety or risk adverse human rights impacts) would be published on a quarterly and/or annual basis.

01 EXPLANATION

During this assessment many interviewees raised with BSR the need for a clearer "role and mandate" for the IAC, and that this mandate should include a much stronger role for the IAC in GIFCT governance, accountability, and transparency. This recommendation is intended to help fulfill this desire.

This recommendation is modeled on a similar "recommend and respond" process being used by

the Facebook Oversight Board.

Principle 20 of the UNGPs states that "In order to verify whether adverse human rights impacts are being addressed, business enterprises should track the effectiveness of their response. Tracking should...draw on feedback from both internal and external sources, including affected stakeholders."

02 RECOMMENDATION

Task the IAC with publishing an annual statement about GIFCT's work.

This annual statement would be published alongside the GIFCT annual report. The content, focus, and length of the report would be at the IAC's discretion.

02 EXPLANATION

During this assessment many interviewees raised with BSR the need for a clearer "role and mandate" for the IAC. and that this mandate should include a much stronger role for the IAC in GIFCT governance, accountability, and transparency. This recommendation is intended to help fulfil this desire.

Concerns were also raised with BSR about "consultation theater" (i.e., consultation for show but not impact), which it is important to address substantively and practically. This recommendation is intended to provide a focus point for this consultation, demonstrate the independence of IAC members, and help build confidence in the work of the IAC and GIFCT.

Principle 20 of the UNGPs states that "In order to verify whether adverse human rights impacts are being addressed, business enterprises should track the effectiveness of their response. Tracking should... draw on feedback from both internal and external sources, including affected stakeholders."

03 RECOMMENDATION

Publish minutes of Operating Board and IAC meetings.

These minutes should be published in a timely manner, but redacted where needed for reasons of confidentiality and stakeholder safety.

The "spirit" of this recommendation is that the public minutes provide those outside GIFCT with improved insight into GIFCT's work and deliberations.

03 EXPLANATION

Many stakeholders interviewed during this assessment expressed a lack of clarity about what GIFCT is doing and how it works.

Principle 21 of the UNGPs states that "in order to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders." This communication should be in "a form and frequency that reflect an enterprise's human rights impacts and that are accessible to its intended audiences" and "in turn not pose risks to affected stakeholders, personnel or to legitimate requirements of commercial confidentiality."

04 RECOMMENDATION

In two years, review the merits of transitioning to a multi-stakeholder Operating Board.

Over the medium-to-long term an Operating Board consisting only of the four founding member companies is not a sustainable model. However, GIFCT is also a young organization, and governance transitions should be entered into cautiously. For this reason, BSR recommends a gradual transition to a multi-stakeholder decision-making model, beginning with a governance review in two years.

Specifically, consideration should be given to a fully multi-stakeholder Operating Board, with 50 percent votes held by companies and 50 percent votes held by civil society organizations and academics / experts. BSR does not advise that governments join the Operating Board.

GIFCT can also consider adding non-founding member companies to the Operating Board.

Further, GIFCT can consider segmenting the IAC civil society constituency into (1) academics / experts and (2) civil society organizations.

04 EXPLANATION

The importance of engaging with affected stakeholders is a fundamental principle in rights-based approaches, and a theme throughout the UNGPs, including impact identification (Principle 18), taking action (Principle 19), and tracking progress (Principle 20).

The UNGPs highlight the role of collaboration and multi-stakeholder initiatives in addressing system wide challenges of shared concern (Principles 19 and 23).

Providing civil society organizations with a greater stake in the decision-making and operations of GIFCT will enhance their commitment to collaborative approaches and is key for GIFCT to achieve its ambition to "bring together the technology industry, government, civil society, and academia to foster collaboration and information-sharing to counter terrorist and violent extremist activity online."

However, significant changes to organizational governance should be entered into deliberately, following a more substantial and focused review process than undertaken for this assessment.

05 RECOMMENDATION

Require GIFCT member companies to publish insights into their use of the hash-sharing database as part of their transparency reports, or similar disclosures.

These reports should provide insights into the absolute number and portion of company content removals that are (a) derived from and (b) contained in the hash-sharing database, as well as (c) appeals made against these content removals and (d) content the company has contributed to the hash-sharing database.

These reports would describe how companies make use of the hash-sharing database in practice, such as whether each hash is subject to human review, how quickly new hashes are actioned, and how many hashes are not actioned. This data should be segmented according to the tags / definitions used in the hash-sharing database, and be accompanied by an explanatory narrative that "demystifies" their use of the hash-sharing database.

Finally, these disclosures should contain a narrative from GIFCT member companies about where the hash-sharing database "fits" as part of broader efforts to prevent terrorists and violent extremists from exploiting digital platforms.

05 EXPLANATION

Principle 21 of the UNGPs states that "in order to account for how they address their human rights impacts, business enterprises should be prepared to communicate this externally, particularly when concerns are raised by or on behalf of affected stakeholders."

During this assessment, many interviewees raised with BSR questions about the hash-sharing database that can only be addressed by GIFCT member companies, not by GIFCT itself. For example, questions included issues such as whether companies accept what is in the database or independently review material.

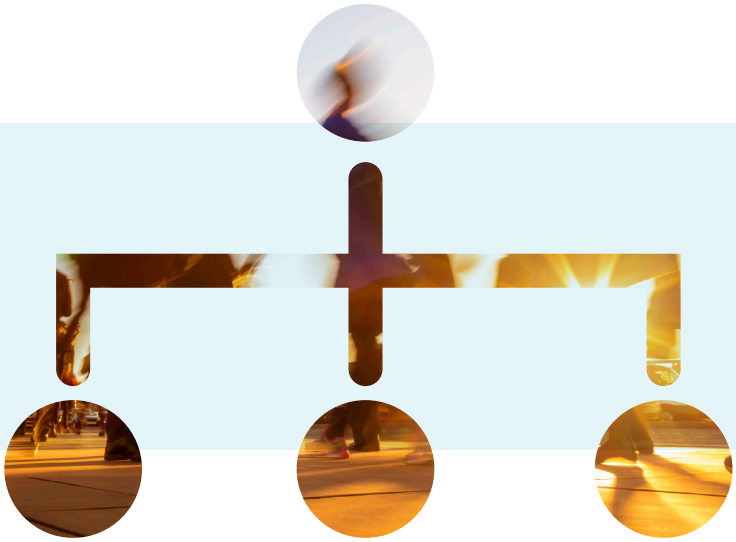
This is an example of a recommendation where implementation by smaller companies may take a different form than implementation by larger companies—and where smaller companies will benefit from BSR's recommendation of additional technical assistance from GIFCT.

12. Organizational Issues

IN THIS SECTION

12.1 Analysis

12.2 Recommendations



12.1 Analysis

Principle 19 of the UNGPs emphasizes the importance of integrating the findings from impact assessments across relevant internal functions and processes, and ensuring that assessment findings are properly understood, given due weight, and acted upon.

However, Principle 14 of the UNGPs states that the "size, sector, operational context, ownership and structure" of an entity should influence the means through which an organization meets its responsibility to respect human rights, and that processes, structures, and means may vary.

What level of resources does GIFCT need to effectively address human rights? Which working group (if any) should carry forward this work?

GIFCT is a small, young, and newly-independent organization, and BSR is very cognizant of the challenges associated with the combination of high expectations and limited resources. It is a challenge we have seen repeated across many multi-company and multi-stakeholder collaborations that we have been involved in over the past three decades, and throughout this assessment we've been attentive to the need to make actionable and practical recommendations. In this final section we make recommendations relating to the GIFCT organization itself.

12.2 Recommendations

01 RECOMMENDATION

Create a diversity, equity, and inclusion ambition for (1) GIFCT staff and (2) GIFCT participants.

This should include a deliberate hiring and stakeholder outreach strategy and an annual review of progress.

Diversity of GIFCT participants can be achieved both through diversifying membership and diversifying the roles and locations of member representatives participating in GIFCT. However, it is also important that GIFCT culture is inclusive, and that opportunities for participation are equitable (e.g., language, time zones, resources).

01 EXPLANATION

There is a direct link between the lived experiences of those engaged in GIFCT's work—both staff and participants—and the ability of GIFCT to achieve its mission of preventing terrorists and violent extremists from exploiting digital platforms.

The UNGPs state that organizations should pay "particular attention to the rights and needs of, as well as the challenges faced by, individuals from groups or populations that may be at heightened risk of becoming vulnerable or marginalized, and with due regard to the different risks that may be faced by women and men."

02 RECOMMENDATION

Establish a mechanism to provide stipends for non-company / non-government participants in GIFCT.

This stipend would be offered to participants in the IAC and working groups who might otherwise face barriers to full engagement.

This mechanism should include a dedicated budget line controlled by GIFCT staff, and diverse sources of funding (e.g., companies, foundations, governments) should be sought.

02 EXPLANATION

Principles 18-21 of the UNGPs all emphasize the importance of engagement with affected stakeholders, and the UNGPs overall prioritize impacts on individuals from groups or populations that may be at heightened risk of vulnerability or marginalization. This recommendation is intended to support the enhanced involvement of affected stakeholders in the work of GIFCT.

03 RECOMMENDATION

Continue enhancing GIFCT staff support for the IAC and working groups.

GIFCT staff doubled in size during the course of this assessment, significantly increasing capacity to support stakeholder engagement and enhancing the effectiveness of the organization. This trend should continue.

03 EXPLANATION

Effective engagement of affected stakeholders and working group effectiveness will require sufficient staff support.

During the assessment BSR encountered a view that more consistency of effectiveness across working groups would be welcome—that while it is early days, some groups are more effective than others.

04 RECOMMENDATION

Hire a "Human Rights & Stakeholder Engagement Director."

The Director would be responsible for creating and maintaining (1) a human rights strategy and (2) a strategy for engaging with affected stakeholders and their legitimate representatives, such as civil society organizations, governments, academics, and human rights defenders.

The Director should play a full and substantial role in the pursuit of GIFCT's mission and strategy; BSR does not envision this to be a "watchdog" role.

04 EXPLANATION

BSR sees significant synergy between a human rights strategy and a stakeholder engagement strategy, so we are proposing that they are combined into a single role.

05 RECOMMENDATION

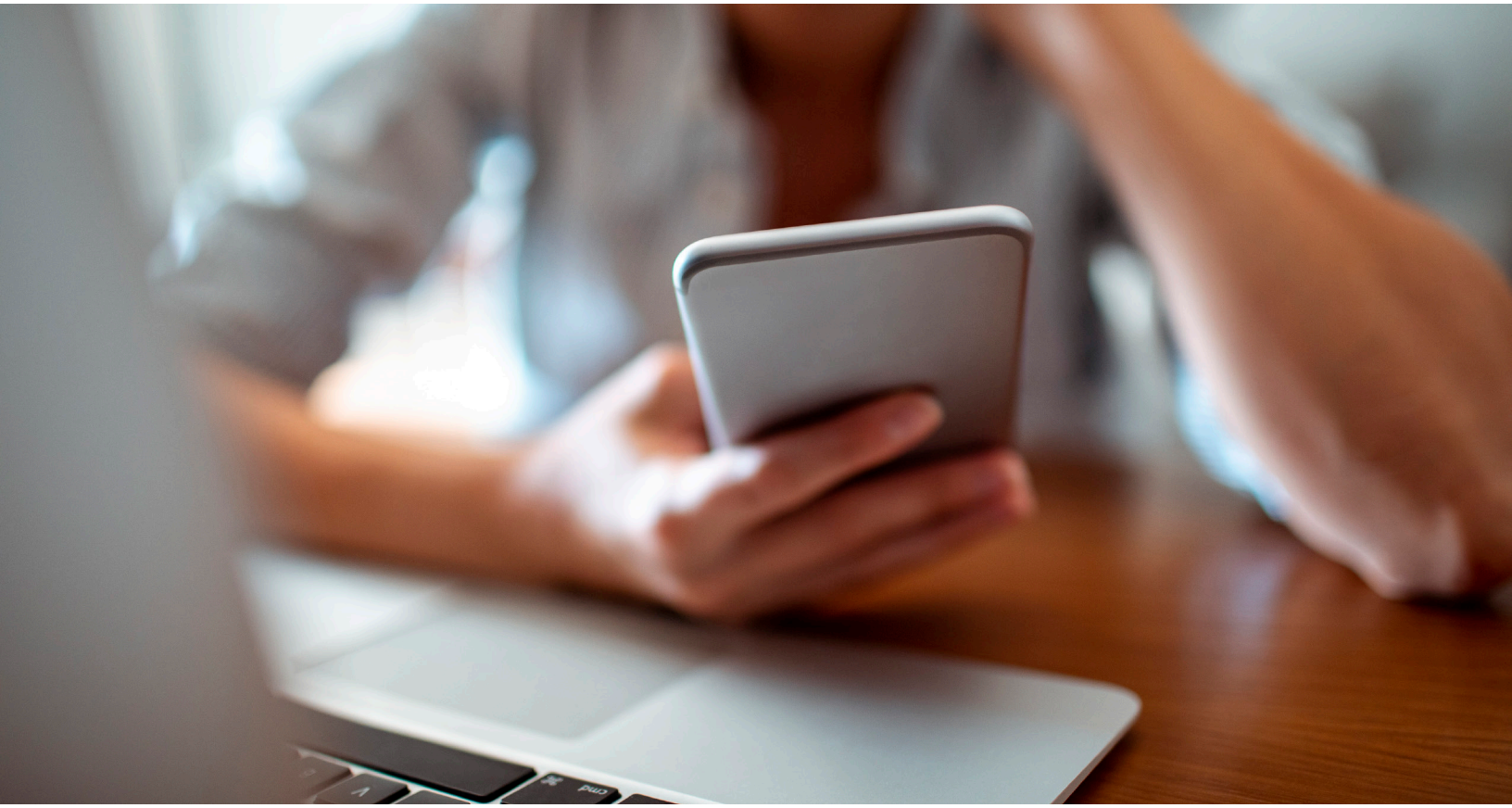
Create a GIFCT governance and management chart.

This simple chart would describe decision-making authority, advisory roles, and the respective roles and responsibilities of staff and participants. It would be available on the GIFCT website.

BSR recommends that this chart include a description of the role of governments in GIFCT, as that appears to be where there is most distance between perception and reality.

05 EXPLANATION

During the assessment BSR encountered a strong desire to better understand "how GIFCT works" and "who does what."



13. Conclusion

We commended GIFCT for commissioning an assessment of its human rights impacts less than a year after becoming an independent entity. The thoughtful integration of human rights into its mission, goals, and activities will both help address GIFCT's potential human rights impacts and enhance GIFCT's legitimacy and sustainability as a multi-stakeholder effort. Developing a plan to implement the recommendations we make in this assessment will be a significant first step.

It is important to emphasize that the recommendations in this report are not designed to be implemented at the same time; rather, they are a broad set of actions GIFCT can take over the coming years. Five areas in particular stand out as having the greatest potential to positively impact the work of GIFCT and its connection to human rights:

01. Better integration of human rights into the governing documents of GIFCT, the creation of a human rights policy, and the development of a framework for ongoing human rights due diligence.

02. Increased transparency around the Crisis Incident Protocol and the hash-sharing database to enhance accountability, address myths, and broaden understanding of how these important tools both function and are used.

03. Ensuring a stronger voice and role for stakeholders in GIFCT's work by broadening the diversity of stakeholders included and formalizing the advisory role of the IAC in relation to the GIFCT Operating Board. It is also important to clarify the role of governments within the IAC and more broadly.

04. Using the breadth of expertise within GIFCT to contribute to the well-documented challenges of bias in the counterterrorism field, such as the recently started work on developing a common understanding of terrorist and violent extremist content.

05. Prioritizing growth in the membership of GIFCT beyond the current US-centric company membership, but with checks in place to manage the human rights risks of expansion, such as a tiered membership structure.

Although this assessment focused on identifying and making recommendations to address GIFCT's adverse human rights impacts, it is also important to emphasize the potential GIFCT has to facilitate the realization of human rights. By pursuing rights-respecting approaches to combating the abuse of online platforms by terrorists and violent extremists, and implementing the recommendations in this assessment, GIFCT can defend the right to life, liberty, and security, protect freedom of expression, and enhance the realization of many other rights online.

Endnotes

1 https://gifct.org/about/.

2 BSR uses the term "multi-stakeholder effort" to describe GIFCT throughout. We note that GIFCT exhibits some important characteristics of multi-stakeholder initiatives (e.g., non-companies are formal participants in the work and governance of GIFCT), but it also lacks others (e.g., companies have 100 percent of Board votes).

3 UNGPs Principle 18.

4 UNGPs Principle 16.

5 Official letter to Facebook and A/HRC/60/288.

6 UNGPs Principle 12.

7 For example, see the June 2018 Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression on the topic of online content regulation.

8 UNGPs Principle 18.

9 BSR hasn't investigated or identified specific instances where the hash-sharing database has led to overbroad restrictions on content; rather, this is being presented here as a risk. Overbroad content removal by companies has occurred; it is unclear at present whether or not this has resulted from using the hash-sharing database. Later in the assessment we make recommendations to investigate this risk further.

10 BSR notes that, if they were to occur, GIFCT would likely lack visibility into specific cases.

11 https://undocs.org/A/HRC/46/36.

12 UNGPs Principle 17.

13 UNGPs Principle 19.

14 This states that terrorist content means one or more of the following: (a) inciting or advocating, including by glorifying, the commission of terrorist offenses, thereby causing a danger that such acts be committed; (b) encouraging the contribution to terrorist offenses; (c) promoting the activities of a terrorist group, in particular by encouraging the participation in or support of a terrorist group; (d) instructing on methods or techniques for the purpose of committing terrorist offenses.

15 A/HRC/16/51, practice 8 "Model offense of incitement to terrorism": It is an offense to intentionally and unlawfully distribute or otherwise make available a message to the public with the intent to incite the commission of a terroristoffense, where such conduct, whether or not expressly advocating terrorist offenses, causes a danger that one or more such offenses may be committed.

16 Letters to the European Commission from Human Rights Watch and the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; the Special Rapporteur on the right to privacy and the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism.

17 UNGPs Principle 21.

18 BSR's interviews with smaller companies did not always support the hypothesis that smaller companies would use the hash-sharing database without human review.

19 UNGPs Principle 22.

20 UNGPs Principle 25.

21 For example, the forthcoming EU terrorism content law requires removal within one hour of its appearance. See https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2372.

22 UNGPs Principle 19.

23 UNGPs Principle 19.

24 UNGPs Principles 3 and 7.

25 UNGPs Principle 18.

26 UNGPs Principle 20.

27 The "tech stack" refers to different layers of the technology industry, such as app stores, cloud services, content delivery networks, domain registrars, and internet service providers. For further analysis about the tech stack and content moderation, see Navigating the Tech Stack: When, Where and How Should We Moderate Content? by Joan Donovan, and A Framework for Moderation, by Ben Thompson.

28 UNGPs Principle 24.

29 GIFCT Membership Criteria.

30 UNGPs Principle 18.

31 UNGPs Principles 20, 21, and 31

32 UNGPs Principle 19.

33 UNGPs Principle 20.

34 UNGPs Principle 21.

35 UNGPs Principle 21.

BSR®

About BSR

BSR™ is an organization of sustainable business experts that works with its global network of the world's leading companies to build a just and sustainable world. With offices in Asia, Europe, and North America, BSR™ provides insight, advice, and collaborative initiatives to help you see a changing world more clearly, create long-term business value, and scale impact.