# Introducing 2022 GIFCT Working Group Outputs



Dr. Erin Saltman Director of Programming, GIFCT In July 2020, GIFCT launched a series of Working Groups to bring together experts from across sectors, geographies, and disciplines to offer advice in specific thematic areas and deliver on targeted, substantive projects to enhance and evolve counterterrorism and counter-extremism efforts online. Participation in Working Groups is voluntary and individuals or NGOs leading Working Group projects and outputs receive funding from GIFCT to help further their group's aims. Participants work with GIFCT to prepare strategic work plans, outline objectives, set goals, identify strategies, produce deliverables, and meet timelines. Working Group outputs are made public on the GIFCT website to benefit the widest community. Each year, after GIFCT's Annual Summit in July, groups are refreshed to update themes, focus areas, and participants.

From August 2021 to July 2022, GIFCT Working Groups focused on the following themes:

- · Crisis Response & Incident Protocols
- Positive Interventions & Strategic Communications
- · Technical Approaches: Tooling, Algorithms & Artificial Intelligence
- Transparency: Best Practices & Implementation
- Legal Frameworks

A total of 178 participants from 35 countries across six continents were picked to participate in this year's Working Groups. Applications to join groups are open to the public and participants are chosen based on ensuring each group is populated with subject matter experts from across different sectors and geographies, with a range of perspectives to address the topic. Working Group participants in 2021–2022 came from civil society (57%), national and international government bodies (26%), and technology companies (17%).

Participant diversity does not mean that everyone always agrees on approaches. In many cases, the aim is not to force group unanimity, but to find value in highlighting differences of opinion and develop empathy and greater understanding about the various ways that each sector identifies problems and looks to build solutions. At the end of the day, everyone involved in addressing violent extremist exploitation of digital platforms is working toward the same goal: countering terrorism while respecting human rights. The projects presented from this year's Working Groups highlight the many perspectives and approaches necessary to understand and effectively address the ever-evolving counterterrorism and violent extremism efforts in the online space. The following summarizes the thirteen outputs produced by the five Working Groups.

#### **Crisis Response Working Group (CRWG):**

The GIFCT Working Group on Crisis Response feeds directly into improving and refining GIFCT's own Incident Response Framework, as well as posing broader questions about the role of law enforcement, tech companies, and wider civil society groups during and in the aftermath of a terrorist or violent extremist attack. CRWG produced three outputs. The largest of the three was an immersive virtual series of Crisis Response Tabletop Exercises, hosted by GIFCT's Director of Technology, Tom Thorley. The aim of the Tabletops was to build on previous Europol and Christchurch Call-led Crisis Response events, with a focus on human rights, internal communications, and external strategic communications in and around crisis scenarios. To share lessons learned and areas for

improvement and refinement, a summary of these cross-sector immersive events is included in the 2022 collection of Working Group papers.

The second output from the CRWG is a paper on the Human Rights Lifecycle of a Terrorist Incident, led by Dr. Farzaneh Badii. This paper discusses how best GIFCT and relevant stakeholders can apply human rights indicators and parameters into crisis response work based on the 2021 GIFCT Human Rights Impact Assessment and UN frameworks. To help practitioners integrate a human rights approach, the output highlights which and whose human rights are impacted during a terrorist incident and the ramifications involved.

The final CRWG output is on Crisis Response Protocols: Mapping & Gap Analysis, led by the New Zealand government in coordination with the wider Christchurch Call to Action. The paper maps crisis response protocols of GIFCT and partnered governments and outlines the role of tech companies and civil society within those protocols. Overall, the output identifies and analyzes the gaps and overlaps of protocols, and provides a set of recommendations for moving forward.

#### **Positive Interventions & Strategic Communications (PIWG):**

The Positive Interventions and Strategic Communications Working Group developed two outputs to focus on advancing the prevention and counter-extremism activist space. The first is a paper led by Munir Zamir on Active Strategic Communications: Measuring Impact and Audience Engagement. This analysis highlights tactics and methodologies for turning passive content consumption of campaigns into active engagement online. The analysis tracks a variety of methodologies for yielding more impact-focused measurement and evaluation.

The second paper, led by Kesa White, is on Good Practices, Tools, and Safety Measures for Researchers. This paper discusses approaches and safeguarding mechanisms to ensure best practices online for online researchers and activists in the counterterrorism and counter-extremism sector. Recognizing that researchers and practitioners often put themselves or their target audiences at risk, the paper discusses do-no-harm principles and online tools for safety-by-design methodologies within personal, research, and practitioner online habits.

#### Technical Approaches Working Group (TAWG):

As the dialogue on algorithms and the nexus with violent extremism has increased in recent years, the Technical Approaches Working Group worked to produce a longer report on Methodologies to Evaluate Content Sharing Algorithms & Processes led by GIFCT's Director of Technology Tom Thorley in collaboration with Emma Llanso and Dr. Chris Meserole. While Year 1 of Working Groups produced a paper identifying the types of algorithms that pose major concerns to the CVE and counterterrorism sector, Year 2 output explores research questions at the intersection of algorithms, users and TVEC, the feasibility of various methodologies and the challenges and debates facing research in this area.

To further this technical work into Year 3, TAWG has worked with GIFCT to release a Research Call

for Proposals funded by GIFCT. This Call for Proposals is on Machine Translation. Specifically, it will allow third parties to develop tooling based on the <u>gap analysis</u> from last year's TAWG Gap Analysis. Specifically, it seeks to develop a multilingual machine learning system addressing violent extremist contexts.

#### Transparency Working Group (TWG):

The Transparency Working Group produced two outputs to guide and evolve the conversation about transparency in relation to practitioners, governments, and tech companies. The first output, led by Dr. Joe Whittaker, focuses on researcher transparency in analyzing algorithmic systems. The paper on Recommendation Algorithms and Extremist Content: A Review of Empirical Evidence reviews how researchers have attempted to analyze content-sharing algorithms and indicates suggested best practices for researchers in terms of framing, methodologies, and transparency. It also contains recommendations for sustainable and replicable research.

The second output, led by Dr. Courtney Radsch, reports on Transparency Reporting: Good Practices and Lessons from Global Assessment Frameworks. The paper highlights broader framing for the questions around transparency reporting, the needs of various sectors for transparency, and questions around what meaningful transparency looks like.

#### The Legal Frameworks Working Group (LFWG):

The Legal Frameworks Working Group produced two complementary outputs.

The first LFWG output is about Privacy and Data Protection/Access led by Dia Kayyali. This White Paper reviews the implications and applications of the EU's Digital Services Act (DSA) and the General Data Protection Regulation (GDPR). This includes case studies on Yemen and Ukraine, a data taxonomy, and legal research on the Stored Communications Act.

The second LFWG output focuses on terrorist definitions and compliments GIFCT's wider Definitional Frameworks and Principles work. This output, led by Dr. Katy Vaughan, is on The Interoperability of Terrorism Definitions. This paper focuses on the interoperability, consistency, and coherence of terrorism definitions across a number of countries, international organizations, and tech platforms. Notably, it highlights legal issues around defining terrorism based largely on government lists and how they are applied online.

#### **Research on Algorithmic Amplification:**

Finally, due to the increased concern from governments and human rights networks about the potential link between algorithmic amplification and violent extremist radicalization, GIFCT commissioned Dr. Jazz Rowa to sit across three of GIFCT's Working Groups to develop an extensive paper providing an analytical framework through the lens of human security to better understand the relation between algorithms and processes of radicalization. Dr. Rowa participated in the Transparency, Technical Approaches, and Legal Frameworks Working Groups to gain insight into

the real and perceived threat from algorithmic amplification. This research looks at the contextuality of algorithms, the current public policy environment, and human rights as a cross-cutting issue. In reviewing technical and human processes, she also looks at the potential agency played by algorithms, governments, users, and platforms more broadly to better understand causality.

We at GIFCT hope that these fourteen outputs are of utility to the widest range of international stakeholders possible. While we are an organization that was founded by technology companies to aid the wider tech landscape in preventing terrorist and violent extremist exploitation online, we believe it is only through this multistakeholder approach that we can yield meaningful and long-lasting progress against a constantly evolving adversarial threat.

We look forward to the refreshed Working Groups commencing in September 2022 and remain grateful for all the time and energy given to these efforts by our Working Group participants.

# Participant Affiliations in the August 2021 - July 2022 Working Groups:

Tech Sector	Government Sector	Civil Society / Academia / Practitioners	Civil Society / Academia / Practitioners
ActiveFence	Aqaba Process	Access Now	Lowy Institute
Amazon	Association Rwandaise de Défense des Droits de l'Homme	Anti-Defamation League (ADL)	M&C Saatchi World Services Partner
Automattic	Australian Government - Department of Home Affairs	American University	Mnemonic
Checkstep Ltd.	BMI Germany	ARTICLE 19	Moonshot
Dailymotion	Canadian Government	Australian Muslim Advocacy Network (AMAN)	ModusIzad - Centre for applied research on deradicalisation
Discord	Classification Office, New Zealand	Biodiversity Hub International	New America's Open Technology Institute
Dropbox, Inc.	Commonwealth Secretariat	Bonding Beyond Borders	Oxford Internet Institute
ExTrac	Council of Europe, Committee on Counter- Terrorism	Brookings Institution	Partnership for Countering Influence Operations, Carnegie Endowment for International Peace
Facebook	Department of Justice - Ireland	Business for Social Responsibility	Peace Research Institute Frankfurt (PRIF); Germany
JustPaste.it	Department of State - Ireland	Centre for Analysis of the Radical Right (CARR)	PeaceGeeks
Mailchimp	Department of State - USA	Center for Democracy & Technology	Point72.com
MEGA	Department of the Prime Minister and Cabinet (DPMC), New Zealand Government	Center for Media, Data and Society	Polarization and Extremism Research and Innovation Lab (PERIL)
Microsoft	DHS Center for Prevention Programs and Partnerships (CP3)	Centre for Human Rights	Policy Center for the New South (senior fellow)
Pex	European Commission	Centre for International Governance Innovation	Public Safety Canada & Carleton University
Snap Inc.	Europol/EU IRU	Centre for Youth and Criminal Justice (CYCJ) at the University of Strathclyde, Scotland.	Queen's University
Tik Tok	Federal Bureau of Investigation (FBI)	Cognitive Security Information Sharing & Analysis Center	Sada Award, Athar NGO, International Youth Foundation
Tremau	HRH Prince Ghazi Bin Muhammad's Office	Cornell University	Shout Out UK
Twitter	Ministry of Culture, DGMIC - France	CyberPeace Institute	Strategic News Global
You Tube	Ministry of Foreign Affairs - France	Dare to be Grey	S. Rajaratnam School of International Studies, Singapore (RSIS)
	Ministry of Home Affairs (MHA) - Indian Government	Dept of Computer Science, University of Otago	Swansea University
	Ministry of Justice and Security, the Netherlands	Digital Medusa	Tech Against Terrorism
	National Counter Terrorism Authority (NACTA) Pakistan	Edinburgh Law School, The University of Edinburgh	The Alan Turing Institute

Organisation for Economic Co-operation and Development (OECD)	European Center for Not-for-Profit Law (ECNL)	The Electronic Frontier Foundation
Office of the Australian eSafety Commissioner (eSafety)	Gillberg Neuropsychiatry Centre, Gothenburg University, Sweden,	The National Consortium for the Study of Terrorism and Responses to Terrorism (START) / University of Maryland
Organization for Security and Co-operation in Europe (OSCE RFoM)	George Washington University, Program on Extremism	Unity is Strength
Pôle d'Expertise de la Régulation Numérique (French Government)	Georgetown University	Université de Bretagne occidentale (France)
North Atlantic Treaty Organization, also called the North Atlantic Alliance (NATO)	Georgia State University	University of Auckland
Secrétaire général du Comité Interministériel de prévention de la délinquance et de la radicalisation	Global Network on Extremism and Technology (GNET)	University of Groningen
State Security Service of Georgia	Global Disinformation Index	University of Massachusetts Lowell
The Royal Hashemite Court/ Jordanian Government	Global Network Initiative (GNI)	University of Oxford
The Office of Communications (Ofcom), UK	Global Partners Digital	University of Queensland
UK Home Office	Global Project Against Hate and Extremism	University of Salford, Manchester, England,
United Nations Counter-terrorism Committee Executive Directorate (CTED)	Groundscout/Resonant Voices Initiative	University of South Wales
UN, Analytical Support and Sanctions Monitoring Team (1267 Monitoring Team)	Hedayah	University of the West of Scotland
United Nations Major Group for Children and Youth (UNMGCY)	Human Cognition	Violence Prevention Network
United States Agency for International Development (USAID)	Institute for Strategic Dialogue	WeCan Africa Initiative & Inspire Africa For Global Impact
	International Centre for Counter-Terrorism	Wikimedia Foundation
	Internet Governance Project, Georgia Institute of Technology	World Jewish Congress
	Islamic Women's Council of New Zealand	XCyber Group
	JOS Project	Yale University, Jackson Institute
	JustPeace Labs	Zinc Network
	Khalifa Ihler Institute	
	KizBasina (Just-a-Girl)	
	Love Frankie	

Transparency Reporting: Good Practices and Lessons from Global Assessment Frameworks GIFCT Transparency Working Group





Dr. Courtney Radsch Center for Media, Data and Society

# Overview

The GIFCT seeks to inform its work on transparency by examining practices and approaches from an array of sectors and stakeholders to identify good practices, lessons learned, and approaches from beyond the technology sector. There is no singular definition of transparency reporting. Rather, the concept, principles, and need for transparency are embedded in assessment frameworks covering such areas as Corporate Social Responsibility (CSR), Environmental, Social, and Governance (ESG), Environmental Impact Assessments (EIA), Social Impact Assessment (SIA), Human rights Impact Assessments (HRIA), and corporate reporting in finance/accounting/taxation, extractive industries, and information and communication technologies (ICT). The following research scoping agenda identifies key themes, good practices, and lessons drawn from these frameworks along with academic, governmental, and civil society reporting and assessments of transparency and impact reporting more broadly.<sup>1</sup> Following this analysis, it turns briefly to transparency reporting in the tech sector and identifies common practices and recommendations.

Reporting takes place in response to mandatory requirements as well as occurring voluntarily. A wide range of sectors and stakeholders are increasingly conducting voluntary and/or mandatory reporting on various dimensions of their work, from governance and process to activities and impacts. They are also creating sets of principles and expectations through third-party assessments, rankings, and impact reporting. Such reporting is often aimed at increasing transparency, building trust, and enabling accountability; it may also be required by law or implemented in response to external pressure. Most corporate transparency reporting is voluntary, even in highly regulated industries like the financial sector, resulting in a wide range of quality and comprehensiveness.<sup>2</sup>

Transparency reporting by information and communications technologies is relatively new compared to other sectors. Nonetheless, in just over a decade, regular transparency reporting by application layer tech firms has become an emergent norm,<sup>3</sup> though the quality, comprehensiveness, and expectations of what should be included continue to evolve even as some standards coalesce while others are contested.<sup>4</sup> While telecom firms and businesses that have existed in more mature sectors may have engaged in transparency or impact reporting over a longer period of time, reporting on data requests, content moderation, terms of service issues, and other issues specific to the expressive and privacy dimensions of the tech sector are still nascent.

As a relatively new organization, the GIFCT must consider how it conceptualizes and addresses

1 Transparency reporting is distinct from data access or data sharing.

2 Barbara Kowalczyk-Hoyer, "Transparency in Corporate Reporting: Assessing the World's Largest Companies," Transparency International, (2012): 37.

3 Joan Donovan, "Navigating the Tech Stack: When, Where and How Should We Moderate Content?," Centre for International Governance Innovation, October 28, 2019, <a href="https://www.cigionline.org/articles/navigating-tech-stack-when-where-and-how-should-we-moderate-content/">https://www.cigionline.org/articles/navigating-tech-stack-when-where-and-how-should-we-moderate-content/</a>.

4 Access Now, "Transparency Reporting Index - Access Now's Global Database," 2022, <u>https://www.accessnow.org/transparency-reporting-index/</u>; Ranking Digital Rights, "2020 Ranking Digital Rights Corporate Accountability Index," Accessed February 3, 2022, <u>https://rankingdigitalrights.</u> <u>org/index2020/</u>; Priya Kumar, "Ranking Digital Rights Findings on Transparency Reporting and Companies' Terms of Service Enforcement," Ranking Digital Rights, March 2016, <u>https://rankingdigitalrights.org/wp-content/uploads/2016/03/RDR-Transparency-Findings.pdf</u>. transparency, and how then it translates this into its own transparency reporting and expectations for its member companies. This scoping paper is intended to guide a more detailed research paper that will examine in greater depth the specific themes and practices identified herein.

# Define objective(s)

The foundation of any good report requires clearly identifying the objective(s) of the information, the primary user(s) of the information, and the qualitative characteristics of useful information. Establishing the objective(s) of transparency reporting is the basis for effective reporting, and determining the audience(s) for such reports is a fundamental best practice that ensures the alignment of needs and expectations with what is covered by the report. Generally accepted government auditing standards, for example, require that the objectives of an audit be defined.<sup>5</sup>

Transparency reports can be used to "spark and grow the trust" of a company's user base,<sup>6</sup> signal to and inform policymakers, and force firms to build systems to enable them to capture and report on specific information. Other efforts are "narrowly designed to make government data more easily accessible to private sector and other stakeholders and do not attempt to consciously link these transparency mechanisms to accountability or participatory processes."<sup>7</sup> Defining objective(s) enables better design and assessment of whether transparency reporting is effective.

In voluntary transparency reporting, employing conceptual frameworks enables the development of standards that can be understood and correctly interpreted by all parties. Conceptual frameworks should be based on consistent concepts and the development of consistent reporting practices where no standard applies. Consider the following examples:

- The data-intensive Global Reporting Initiative (GRI) was designed to increase company transparency with respect to sustainability and improve decision making by companies and their stakeholders; and
- The Extractive Industries Transparency Initiatives (EITI), which implements the global standard for countries and companies to voluntarily disclose information on key aspects of the governance of oil, gas and mining revenues across the value chain, aims to encourage open and accountable management of those resources and "increase public and private sector responsiveness to citizen demands."<sup>8</sup> As a multi-stakeholder process operating at the country level, it provides data that can help inform reform efforts and strengthen public and corporate

8 Carothers and Brechenmacher, "Accountability, Transparency, Participation, and Inclusion."

<sup>5 &</sup>quot;San Francisco Police Department Use-of-Force Data Audit: Interim Key Issue Report: Best Practices in Reporting Use-of-Force Data," City and County of San Francisco: Audits Division, City Services Auditor, Office of the Controller, December 18, 2019, <u>https://sfgov.org/dpa/sites/default/</u> files/SFPD\_Key\_Issue\_Report\_Use\_of\_Force\_Data\_Reporting\_12\_18\_19\_FINAL.pdf.

<sup>6</sup> Peter Micek and Deniz Duru Aydin, "Non-Financial Disclosures in the Tech Sector: Furthering the Trend," in The Responsibilities of Online Service Providers, eds. Mariarosaria Taddeo and Luciano Floridi (Cham: Springer International Publishing, 2017), 241–61, <u>https://doi.org/10.1007/978-3-319-47852-4\_13</u>.

<sup>7</sup> Thomas Carothers and Saskia Brechenmacher, "Accountability, Transparency, Participation, and Inclusion: A New Development Consensus?," Carnegie Endowment for International Peace, October 20, 2014, <u>https://carnegieendowment.org/2014/10/20/accountability-transparency-participation-and-inclusion-new-development-consensus-pub-56968</u>.

#### governance.

### Good practice: Identify stakeholders and target audience(s)

Transparency reporting can be aimed at external audiences, but it also sends a signal to internal stakeholders about what is expected, acceptable, and important. There may be various audiences defined, but in most domains the general public or "users" are not typically an effective target audience. That said, in the case of consumer-facing products and services that are used on a regular basis, communicating in a way that service users can understand transparency standards may be important. An audience may also be the industry itself, with one objective of such reporting to raise standards, inculcate norms, and/or promote best practices. Transparency reporting that requires extensive and intensive data collection, and which typically serves a specialist audience, would be a relevant issue to examine further as the objective(s) and audience(s) of transparency reporting are interdependent but also dependent on resources. The diversity of approaches is reflected in the following examples:

- The primary audience for GRI reports is specialists rather than the general public or consumers.<sup>9</sup> A primary audience for ESG reports is investors. In the case of ESG reports, there are 77 industry-specific SASB Standards, which are specifically aimed at helping businesses convey financial material sustainability information related to ESG issues to investors.
- The audience for financial transparency reporting is government regulators, as well as internal personnel and those charged with maintaining due diligence and observing regulations.
- The audience for the Internet Commission's report is primarily tech companies and regulators, mainly based in the UK and the Global North.<sup>10</sup>
- The audience for Ranking Digital Right's annual Corporate Accountability Index is its key stakeholders, namely participating tech firms and the digital rights community. However, it was designed with investors in mind and the index indicators were developed within the ESG framework (which is used by investors), with a focus on the social and governance. A few years in, the index is now being used by more and more investors, who turn to it to supplement and embed this data in their own activities.<sup>11</sup>
- The audience for technology sector transparency reporting on content moderation is not well defined.
- √ Analyzing whether the stated objectives of reporting frameworks like GRI and EITI are in fact achieved and perceived as successful by key stakeholders and their intended audience can also inform whether ambitions for transparency reporting are realistic.<sup>12</sup>

10 Ioanna Noula, author interview with the Internet Commission, August 13, 2021.

11 Jan Rydzak and Amy Brouillette, author interview with Ranking Digital Rights, August 13, 2021.

<sup>9</sup> See <u>https://www.globalreporting.org/about-gri/;</u> Stephanie Watts, "Corporate Social Responsibility Reporting Platforms: Enabling Transparency for Accountability," Information Technology and Management 16, no. 1 (March 1, 2015): 19–35, <u>https://doi.org/10.1007/s10799-014-0192-2</u>; Klaus Dingwerth and Margot Eichinger, "Tamed Transparency: How Information Disclosure under the Global Reporting Initiative Fails to Empower," Global Environmental Politics 10, no. 3 (August 1, 2010): 74–96, <u>https://doi.org/10.1162/GLEP\_a\_00015</u>.

<sup>12</sup> Watts, "Corporate Social Responsibility Reporting Platforms"; Carothers and Brechenmacher, "Accountability, Transparency, Participation, and Inclusion."

√ Increased recognition that Global North/industrialized West companies have significant impact and influence on the Global South/developing countries underscores the need to consider various stakeholders and communities in disenfranchised localities. Similarly, specific communities affected by a company's operations, product, or services could be considered as a target audience.

# What information should be in a transparency report?

What should GIFCT and/or its member companies include in transparency reports? On what issues is it reporting? What counts as data, how it is collected, presented, disseminated, and the costs involved are all relevant considerations for designing a transparency report or broader industry standards. A review of the literature and existing reporting across industries and sectors indicates that transparency reports should typically include information on governance, policy, process, actions taken, impacts, results, and relations with government/authorities.

Good practices include providing quantitative and qualitative data and making sure that it is understandable,<sup>13</sup> including defining terms and abbreviations/relationships, explaining the methodology for collecting and analyzing the data, using data visualizations and examples, and explaining trends revealed by data.<sup>14</sup> Both **data sources** and **presentation** should include quantitative and qualitative information while contextualizing and explaining the information contained within.

A 2012 study of corporate transparency reporting by Transparency International recommended a series of baseline policies that all multi-nationals should adopt,<sup>15</sup> amounting to a set of best practices, some of which could be relevant for the GIFCT, including:

- · Data transparency at the organizational, country, and corporate-level
- · An informative website in at least one international language
- · Including a list of all subsidiaries, affiliates, and related entities

## Standardization and comparability

Standardization and consistency of data within a given industry or sector are widely recognized as a best practice. This enables comparison across entities, time, and data/issues. Taxonomies are important to allow data to be structured for sharing and comparison.<sup>16</sup> Verifiability and standardization of approaches for measuring are considered best practices across sectors. Minimum reporting requirements for an industry convey the basic expectations for transparency reports.

<sup>13</sup> The provision of data in transparency reporting is distinct from the issue of access to data sources and the raw data itself.

<sup>14</sup> As noted in several examples throughout. Also see, for example, Chloë Poynton, "Five Best Practices in Human Rights Reporting," BSR: Our Insights (blog), June 29, 2012, <u>https://www.bsr.org/en/our-insights/blog-view/five-best-practices-in-human-rights-reporting</u>.

<sup>15</sup> It is notable that tech companies scored worst among the nine sectors analyzed in terms of transparency into corporate governance; see Kowalczyk-Hoyer, "Transparency in Corporate Reporting."

<sup>16 &</sup>quot;Statement of Intent to Work Together Towards Comprehensive Corporate Reporting," Impact Management Project, World Economic Forum and Deloitte, September 2020, https://www.globalreporting.org/media/bixjklud/statement-of-intent-to-work-together-towards-comprehensive-corporate-reporting.pdf.

Interoperability between different reporting requirements also helps "ensure that companies can collect information about performance on a given matter once and can use that same information to serve different objectives when the information is suitable for the needs of those different objectives."<sup>17</sup> Examples of such an approach include:

- The OECD Voluntary Transparency Reporting Framework (VTRF) version 1.0 was adopted in late 2021. It contains a set of baseline transparency questions for companies and a glossary of key terms that were developed and negotiated by a multi-stakeholder group of experts<sup>18</sup> over two years and then approved by the OECD's Committee on Digital Economy Policy. These minimum reporting standards on terrorist and violent extremist content (TVEC) moderation transparency represent a consensus from member countries about what the private sector should be regularly reporting on publicly.<sup>19</sup> and will be further elaborated and refined after a pilot period.<sup>20</sup> The OECD has created such reporting standards in other domains.
- A process is underway to develop SASB Standards for content governance to help investors assess the scope and scale of content moderation practices.<sup>21</sup>

Independent audits are an important principle across industries, and in some cases are embedded as requirements in mandatory transparency reporting.<sup>22</sup> Financial reporting is seen as the gold standard in transparency reporting because of its maturity and "adherence to internationally recognized accounting standards that bring transparency, accountability, and efficiency to financial markets around the world."<sup>23</sup> However, the cost of data collection, analysis, and verification can be significant, and thus pose significant hurdles for small companies and can even create an uneven playing field. For example, the U.S. Securities and Exchange Commission (SEC) conflict mineral rule requires certification and standardized audits by independent bodies. However, this can be prohibitively expensive for smaller companies, and there is a lack of oversight and verification of auditors by the SEC. Additionally, data access is often restricted to auditors or specific agencies. The broader issue of data access for researchers, journalists, or the public is a separate topic from transparency reporting.

18 The author was a member of the expert group.

21 https://www.sasb.org/standards/process/active-projects/content-governance-in-the-internet-media-and-services-industry/.

<sup>17 &</sup>quot;Reporting on Enterprise Value Illustrated with a Prototype Climate-Related Financial Disclosure Standard," Impact Management Project, World Economic Forum and Deloitte, December 2020, <a href="https://29kjwb3armds2g3gi4lq2sxl-wpengine.netdna-ssl.com/wp-content/uploads/Report-ing-on-enterprise-value\_climate-prototype\_Dec20.pdf">https://29kjwb3armds2g3gi4lq2sxl-wpengine.netdna-ssl.com/wp-content/uploads/Report-ing-on-enterprise-value\_climate-prototype\_Dec20.pdf</a>.

<sup>19</sup> Scott Morrison, "Media Release: More Action to Prevent Online Terror," Prime Minster of Australia, August 26, 2019, <u>https://www.pm.gov.au/</u> media/more-action-prevent-online-terror.

<sup>20</sup> Jeremy West, "Why We Need More Transparency to Combat Terrorist and Violent Extremist Content Online," OECD Innovation Blog, September 15, 2020, <u>https://oecd-innovation-blog.com/2020/09/15/terrorist-violent-extremist-content-internet-social-media-transparency-tvec/</u>. The web portal is intended to be live in March 2022 at <u>http://www.oecd-vtrf-pilot.org/</u>

<sup>22</sup> One study found noteworthy "the importance of the national audit oversight bodies and the absence of professional bodies in the development of transparency reporting practice". Sakshi Girdhar and Kim K. Jeppesen, "Practice Variation in Big-4 Transparency Reports," Accounting, Auditing & Accountability Journal 31, no. 1 (January 1, 2018): 277, <u>https://doi.org/10.1108/AAAJ-11-2015-2311</u>; OECD, "OECD Best Practices for Budget Transparency," OECD Journal on Budgeting 1, no. 3 (May 16, 2002): 7–14, <u>https://doi.org/10.1787/budget-v1-art14-en</u>.

<sup>23</sup> IFRS Foundation, "The case for global accounting standards," 2021, <u>https://www.ifrs.org/use-around-the-world/why-global-accounting-stan-</u> <u>dards/</u>.

though it is relevant to issues of verifiability.24

# Corporate Social Responsibility (CSR) and the Global Reporting Initiative (GRI)

Transparency is a core pillar of CSR. Companies that engage in CSR also report on those practices (typically voluntarily), although some jurisdictions mandate CSR reporting. An entire industry has emerged around CSR reporting, making it one of the most mature areas for further investigations into best practices.

With 95 percent of the world's biggest companies reporting on their CSR efforts, the vast majority of them use the GRI as the dominant reporting standard.<sup>25</sup> The GRI is data-intensive and provides standardization and comparability. It is also costly.

The GRI suggests four dimensions that would constitute a **best practice for transparent reporting: accuracy, completeness, timeliness and relevance**. GRI is built around six categories (Economic, Environmental, Human Rights, Society, Labor Practices and Decent Work, and Product Responsibility), each with its own subcategories and indicators. There are a total of 41 standards focused on an organization's outward impacts, which represent a broad consensus on good practice for reporting on a range of economic, environmental, and social impacts with respect to sustainable development.

The GRI offers three levels of reporting detail that firms can adhere to, which underscores the need for capacity/capability and intent when it comes to transparency. Such levels also are an acknowledgment of the fact that some firms may not have the capacity (technical, financial, etc.) to generate or collect particular data, which is a relevant concern for tech companies as well. The exploration of specific sectoral standards and a review of the Human Rights Assessment standard to identify principles, approaches, and good practices would help align technology platform reporting.

#### **Determine level of analysis & granularity**

An organization undertaking transparency reporting must determine what it means by 'organization' and whether a corporation will report data at the corporate or country level, and whether it covers subsidiaries, affiliates, and related entities. This should flow from its objective(s) and audience target(s).

Reporting on both subsidiaries *and* at the country level is a good practice in a range of sectors because it (a) enables evaluation of a company's activities and impact in each jurisdiction and by each service, which may cross jurisdictional boundaries and (b) "sheds light on any special

<sup>24</sup> Policymakers are considering mandating access for independent and/or accredited researchers. See for example Article 30 and 31 of the EU's Digital Services Act and US proposals such as H.R. 3451 Social Media DATA Act.

https://ec.europa.eu/info/sites/default/files/proposal\_for\_a\_regulation\_on\_a\_single\_market\_for\_digital\_services.pdf.

<sup>25</sup> Watts, "Corporate Social Responsibility Reporting Platforms."

arrangements between governments and companies, resulting in greater accountability."<sup>26</sup> This second aspect is particularly important for technology platforms and content moderation (and discussed further below). For example, a corporation like Facebook is a parent company with various subsidiaries (Instagram, WhatsApp) and regional offices, any or all of which could be the subject of its transparency reporting.

Some companies claim that reporting on a specific issue – for example anti-corruption efforts at the country level – would put them at a competitive disadvantage, whereas others view this as an internal aspect of risk management. The existence of the GIFCT as an industry-funded, membership non-profit organization that also oversees a technology (the shared hash and URL databases) to which various private ICT companies contribute material that would itself be the subject of their own transparency reporting raises the importance of defining the objectives in order to determine the level of analysis.

More granular approaches to transparency reporting can provide a better incentive to provide accurate information and thus result in better compliance. A review of tax transparency reporting in Europe found that country-by-country mandated reporting as opposed to regional level reporting resulted in fewer discrepancies and evasions.<sup>27</sup> Another European study found that greater transparency incentivized banks to improve their credit practices following a disclosure initiative introduced by the European Central Bank that required greater loan level information collection by banks and "stronger market discipline" in transparency reporting.<sup>28</sup> (See Tech Reporting Below for what this might mean in terms of transparency reporting on content moderation practices or impacts.)

√ What good practices would provide a roadmap for thinking about which levels of analysis to apply to GIFCT's transparency reporting? Should these apply to member companies? How would this contribute to setting standards for transparency in the tech sector?

#### Governance

Some companies report governance data in a stand-alone report, but in many cases it is incorporated into transparency reporting. If it is not in a stand-alone report, high-level data about governance as it relates to the objective(s) of the report and to policies and impacts assessed should be included in the transparency report itself (which can link to more extensive governance information).

While governance reporting is an entire field unto itself, good practices for providing basic

<sup>26</sup> It is notable that tech companies scored worst among the nine sectors analyzed in terms of transparency into corporate governance; see Kowalczyk-Hoyer, "Transparency in Corporate Reporting."

<sup>27</sup> Niels Johannesen and Dan Thor Larsen, "The Power of Financial Transparency: An Event Study of Country-by-Country Reporting Standards," Economics Letters 145 (August 2016): 120–22, https://doi.org/10.1016/j.econlet.2016.05.029.

<sup>28</sup> Aytekin Ertan, Maria Loumioti, and Regina Wittenberg-Moerman, "Enhancing Loan Quality Through Transparency: Evidence from the European Central Bank Loan Level Reporting Initiative," Journal of Accounting Research 55, no. 4 (2017): 877–918.

information on governance should be addressed.<sup>29</sup> Furthermore, with Diversity, Equality, and Inclusion (DEI) becoming a more explicit goal for the private and public sectors alike, transparency reporting on governance and staffing should consider addressing this, particularly given concerns about censorship, definitions of problematic content, and geopolitical power dynamics inherent in the GIFCT's structure and mandate. The E.U.'s 2014 <u>Accounting Directive</u> added legally binding requirements for the disclosure of non-financial and diversity information by large companies and groups, in addition to existing financial disclosures, related to ESG issues.<sup>30</sup> Reports bound by its expectations must include information relating to "environmental matters, social and employee aspects, respect for human rights, anti-corruption and bribery issues, and diversity in their board of directors."

### Addressing relationship with government

Several approaches, including CSR, ESG, and taxation reporting, provide information and details about the relationship of the organization or management with the government and/or relevant authorities (e.g. Ministry of Interior, the tax authority), suggesting that such reporting constitutes a good practice. Such governmental entities are also stakeholders for such reports.<sup>31</sup> For example, Oxfam's <u>Behind the Brands Scorecard</u> assessed the agricultural sourcing policies of the world's ten largest food and beverage companies, including transparency about those policies as well as corporate governance and influence efforts.<sup>32</sup>

Most content moderation-related transparency reports from tech firms have focused on government requests of the company (see Transparency Reporting by tech companies). There is widespread agreement across stakeholder groups (private sector, government, civil society) that companies should report on government requests of ICT firms. In the 11 years since Google released the sector's first transparency report, more than 80 companies have started releasing transparency reports, indicating an emerging self-regulatory practice for companies to disclose privacy and freedom of expression threats (particularly with respect to government requests).<sup>33</sup> However, the mere existence of a transparency report does not mean that it is sufficient or fit for purpose, especially given that content moderation transparency reporting emerged from the tech industry in a seeming vacuum.

Transparency reporting on government requests should distinguish between content and noncontent information; however, making such a determination about whether information is content

31 Rachel F. Wang, Timothy C. Irwin, and Lewis K. Murara, "Trends in Fiscal Transparency: Evidence from a New Database of the Coverage of Fiscal Reporting," Proceedings, Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association 108 (2015): 1–37.

32 See https://www.behindthebrands.org/company-scorecard/.

<sup>29 &</sup>quot;Corporate Governance: Simple, Practical Proposals for Better Reporting of Corporate Governance," Report Leadership, n.d., <u>https://www.cimaglobal.com/Documents/Thought\_leadership\_docs/Governance/Report-Leadership-Corporate-Governance-Report.pdf</u>: See also Ranking Digital Rights Corporate Transparency Index, <u>https://rankingdigitalrights.org/index2020/explore-indicators</u>.

<sup>30</sup> See https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting\_en.

<sup>33</sup> Whereas some information related to national security can only be reported in bands, companies and civil society alike have pushed to be allowed to provide greater detail. See Access Now, "Transparency Reporting Index."

or non-content is not always straightforward.<sup>34</sup> Further review and assessment of how specific data is classified under relevant U.S. and E.U. legal frameworks governing government access to communications would be helpful.

- The OECD released a report on good practices and minimum standards for reporting on government requests for access to personal data held by the private sector which should inform principles and practices for transparency reporting in the tech sector more broadly.<sup>35</sup>
- Similarly, EFF's annual assessment of content moderation policies, <u>Who Has Your Back</u>, focuses on identifying whether companies report on government takedown requests and provide meaningful notice and appeals process to users.<sup>36</sup> The nine years of reports are built around assessing industry best practices (from a human rights perspective) with respect to content moderation and government requests.

Notably for the tech sector, there is **little transparency reporting by government actors** on requests made to platforms either with respect to personal data, content moderation, or other purposes. The entire onus is currently put on ICT companies, although there is recognition and advocacy for governments to also produce such reporting, which would be in line with other frameworks.<sup>37</sup> Just as reporting by government and industry can reduce corruption and build trust through comparison and verifiability, it also has the potential to address concerns about undue political influence or pressure on tech companies outside of legitimate legal processes. If governments also reported on their moderation requests to platforms, then it would enable comparability and add a level of verifiability that currently does not exist.

#### Availability and access to data sources

Most third-party transparency reports, impact assessments, and rankings make use of publicly available data, including information found on websites, annual reports, and corporate transparency reports.<sup>38</sup> Embedded in these assessments and indices are ideas about what constitutes good practices when it comes to reporting transparently on corporate governance, company policies and practices, government requests, and user notification and remedy. Several ICT-related transparency reporting initiatives have emerged that further research should analyze to identify common principles, expectations, and practices with respect to data sources and access.<sup>39</sup>

<sup>34</sup> Liz Woolery, Ryan Hal Budish, and Kevin Bankston, "The Transparency Reporting Toolkit: Best Practices for Reporting on U.S. Government Requests for User Information," 2016, https://dash.harvard.edu/handle/1/28552578.

<sup>35</sup> José Tomás Llanos, "Transparency Reporting: Considerations for the Review of the Privacy Guidelines," OECD Digital Economy Papers, April 23, 2021, <u>https://www.oecd-ilibrary.org/science-and-technology/transparency-reporting\_e90c1lb6-en</u>.

<sup>36</sup> See the 2015 edition: https://www.eff.org/who-has-your-back-government-data-requests-2015#best-practices,

<sup>37</sup> The Australian eSafety Commissioner is required to publish annual reports, and Australia's new Online Safety Act, which took effect in 2022, contains some specific requirements as to what our future reports must contain.

<sup>38</sup> This is due to a range of factors including access to relevant data, time and costs required to collect or generate data, and a commitment to transparency of their own processes.

<sup>39</sup> PCIO Baseline Datasets, "Transparency Reporting & Data Sharing," Partnership for Countering Influence Operations, accessed February 18, 2022, https://ceip.knack.com/pcio-baseline-datasets#transparency--data-sharing/?view\_69\_page=1&view\_69\_filters=%7B%22match %22%3A%22and%22%2C%22rules%22%3A%5B%7B%22field\_22%3A%22field\_448%22%2C%22operator%22%3A%22contains%22%2C%22value%22%3A%22government%22%2C%22field\_name%22%3A%22Stakeholder%20%22%7D%5D%7D.

Using publicly available data is a good practice because it enables verification and comparison, but often raw data emerges from proprietary business operations, so a better understanding of how data is generated and verified in other sectors is needed to inform the technology sector's approach. Some reporting is required by law and/or mandates disclosure of specific information. The U.S. SEC and OECD, for example, have reporting requirements for companies involved in conflict minerals that require disclosure about their uses and information about their due diligence processes.

- The OECD process has been deemed to be the most established one for compliance with conflict minerals reporting and merits closer review. Article 40 of the E.U.'s Eighth Directive, which sets out transparency reporting requirements for accounting and audit firms, is quite broad and leaves some discretion to member states in relation to the implementation of the directive into local law.<sup>40</sup>
- The Global Network Initiative (GNI) human rights assessment and the Internet Commission's responsibility evaluation framework both rely on confidential information obtained from companies under review, including data and interviews, and require researchers/auditors to sign non-disclosure agreements (NDAs).<sup>41</sup> This limits the replicability or verifiability of these assessments, but provides access to proprietary information and officials that would otherwise not be available.
- Many researchers, civil society groups, and journalists dislike and will not sign NDAs, and these are increasingly seen as problematic.<sup>42</sup>

# Dissemination

A critical part of a successful transparency report is ensuring that it gets to the people who need it. Using the correct, publicly available technology for access to reporting information is important.<sup>43</sup> A best practice from various domains is to make data contained within transparency reports publicly available in a machine-readable format, including explanatory charts, infographics, etc., that tell the story embedded in the data. Ensuring they are presented in a way that addresses the target audience/s is also an important factor.

- Embedded in the EITI is a commitment to open data, accessibility, and compatibility, making country-level data available through an **API** and via direct downloadable files along with the individual annual reports.
- Further research into how to best disseminate reports is needed, taking into consideration the

43 "Statement of Intent," Impact Management Project; Renata Avila et al., "Global Mapping of Technology for Transparency and Accountability." Transparency Accountability Initiative, 2010, <u>https://www.transparency-initiative.org/wp-content/uploads/2017/03/global\_mapping\_of\_technol-ogy\_finall.pdf.</u>

<sup>40</sup> S. Girdhar and K.K. Jeppesen, "Practice variation in Big-4 transparency reports," Accounting, Auditing & Accountability Journal 31, no. 1 (2018): 261–285, https://doi.org/10.1108/AAAJ-11-2015-2311.

<sup>41</sup> The author was a board member of the GNI; Noula, author interview.

<sup>42</sup> Based on the author's decade of experience working in the sector and specific discussions about the use of NDAs. For example, the author previously worked for the Committee to Protect Journalists which refused on principle to sign NDAs with technology platforms. This issue was also discussed in civil society groups including the Christchurch Call Advisory Network.

objective(s) and audience(s).44

#### Lesson: Content moderation

Transparency reporting in the tech sector is overwhelmingly focused on content moderation issues, though how these are addressed and assessed has evolved over the past decade. Access Now's Transparency Reporting Index collects links to transparency reports from major internet and telecom companies around the world by year, but does not assess the quality of those reports.<sup>45</sup> That said, common across all reports is reporting on user privacy data, specifically the number of requests for user data coming from governments, police, or other law enforcement agencies (and in some cases intelligence agencies).<sup>46</sup> This can mean distinguishing between criminal and national security requests, although the level of granularity varies widely, with some companies specifying the specific category of content while others lump them together.

Just as the 2013 Snowden revelations about NSA spying and access to U.S. tech companies' data propelled more ICT entities to report on government access and removal requests, the cumulative efforts to counter violent extremism and terrorism online appears to have propelled an increasing number of companies to engage in reporting on content moderation as it relates to terrorism/ CVE.<sup>47</sup> Similarly, since 2017 and amid the COVID-19 "Infodemic," reporting on information/influence operations has become a more regular practice among the largest social media firms, which are most often the target of such campaigns.<sup>48</sup> Facebook, Google/YouTube, and Twitter began reporting on these platform takedowns campaigns in 2017 but have significantly increased since then. However, there is no shared definition of misinformation or other problematic content categories, making comparison difficult. These reports are ad hoc and report on dimensions defined by the company at hand, sometimes in collaboration with civil society groups that identified the information operation.<sup>49</sup>

A comprehensive 2016 survey of U.S. internet and telecom transparency reporting identified eight principles and illustrates the importance of including key metadata (e.g. date) and a static URL for each report.<sup>50</sup> This study identified clear and granular categorization of specific legal processes as well as reporting on the subjects of requests and how users are impacted; comprehensive explanations of legal processes; and the need for standardization of definitions in order to achieve standardization in categories as best practices. It also recommends including a detailed and

44 Avila et al., "Global Mapping of Technology."

47 OECD, "Transparency Reporting on Terrorist and Violent Extremist Content Online: An Update on the Global Top 50 Content Sharing Services," OECD Digital Economy Papers, accessed August 12, 2021, <u>https://www.oecd.org/digital/transparency-reporting-on-terrorist-and-violent-extrem-ist-content-online-8af4ab29-en.htm</u>.

48 "Disinfodex," Partnership for Countering Influence Operations, accessed February 5, 2022, <u>https://disinfodex.org/</u>; Jon Bateman Smith and Victoria Natalie Thompson, "How Social Media Platforms' Community Standards Address Influence Operations," Carnegie Endowment for International Peace, April 1, 2021, <u>https://carnegieendowment.org/2021/04/01/how-social-media-platforms-community-standards-address-influence-operations-pub-84201</u>.

49 "Disinfodex," Partnership for Countering Influence Operations.

50 Woolery, Budish, and Bankston, "The Transparency Reporting Toolkit."

<sup>45</sup> Access Now, "Transparency Reporting Index."

<sup>46</sup> Micek and Aydin, "Non-Financial Disclosures in the Tech Sector."

illustrative (though non-exhaustive) list of how a provider can respond along with provider-specific examples in various categories and definitions. But the framework also suggests that in addition to best practices, good, standard, and notable practices may also provide useful information about transparency reporting principles and implementation.

There appears to be a trend toward providing greater granularity and reporting on a company's own content moderation policies and aggregate impacts on specific types of content (and not just at the behest of government/law enforcement). This could indicate an emerging best practice toward specificity and comprehensiveness with respect to content and account removals, with some advocating for greater details about content and account enforcement, actions taken, and rationale.<sup>51</sup> Amid ongoing advocacy from civil society and academia, policymakers in the E.U. and the U.S. appear poised to reinforce this emergent trend through mandates or regulation.<sup>52</sup> An emerging consensus on minimum expectations for reporting on content moderation by ICTs has emerged through various reports, recommendations, frameworks, and principles specific to the tech industry.<sup>53</sup> Further research should examine what expectations are embedded in such proposals, as they reflect normative expectations of key stakeholders and may indicate areas of consensus.<sup>54</sup>

For example, the Internet Commission's transparency reporting framework for social media content moderation proposes five categories to be assessed with qualitative and quantitative indicators: reporting, moderation, notice, process of appeal, resources, and governance.<sup>55</sup> These were built in part on the Santa Clara Principles on Transparency and Accountability in Content Moderation,<sup>56</sup> which spell out a set of minimum expectations for reporting, and are in line with the terms of service recommendations from the Internet Governance Forum Dynamic Coalition on Platform Regulation.<sup>57</sup> The 2.0 Principles outline general expectations for a broader range of content moderation actions and policies. The Global Disinformation Index recommends tech companies use unique error codes corresponding to the policy under which a piece of content was removed and developing a common notice and takedown regime, both of which would contribute to reporting standardization and comparison.<sup>58</sup>

One overarching challenge with the tech sector as compared to others is that policies and standards

•••••••••••••••••••••••

51 Based on the author's observations and engagement in a variety of venues where tech sector transparency reporting is being discussed.

52 See the EU's Digital Services Act and the US Senate's proposed Algorithmic Justice and Online Platform Transparency Act (S. 1896).

53 Spandana Singh and Leila Doty, "The Transparency Report TrackingTool: How Internet Platforms Are Reporting on the Enforcement of Their Content Rules," New America: Open Technology Institute, December 9, 2021, <u>http://newamerica.org/oti/reports/transparency-report-track-ing-tool/</u>; "Transparency Reporting & Data Sharing," Partnership for Countering Influence Operations.

54 As the 2021 OECD TVEC report notes, "the number of jurisdictions that have TVEC-related laws and regulations in force or under consideration is growing, but they are not consistent, either. That presents a risk of divergent reporting standards and requirements." It also poses an opportunity to identify commonalities and minimums.

55 See https://www.dropbox.com/s/fdzvwqeyosdezb9/The%20Internet%20Commission%20%E2%80%93%20transparency%20reporting%20 framework.pdf?dl=0.

56 See https://santaclaraprinciples.org/.

57 See Annex 12.5.1 Degree of Monitoring in Luca Belli et al., "Platform Regulations: How Platforms Are Regulated and How They Regulate Us," FGV Direito Rio, 2017, https://bibliotecadigital.fgv.br/dspace/handle/10438/19402.

58 Benjamin T. Decker and Tim Boucher, "Disrupting Online Harms: A New Approach," The Global Disinformation Index, July 2021, <u>https://disinformationindex.org/wp-content/uploads/2021/07/2021-07-23-Disrupting-Online-Harms-A-New-Approach.pdf</u>.

governing content and behavior on a specific platform are regularly revised based on new or emerging issues. This means that there is no required minimum or standard set of reporting criteria and that the internal criteria shift. Because there is typically no archive of previous policy iterations, it is difficult to assess how these evolve over time, much less to audit consistency with internal guidelines.

A more detailed review and analysis focused on multi-stakeholder and consensus-driven frameworks should be conducted to identify good practices, minimum transparency expectations, and technical feasibility for reporting on content moderation.

# **Beyond content moderation: Digital responsibility**

Ranking Digital Rights' Corporate Accountability Index evaluates the policies and practices of digital and telecom companies that affect human rights on an annual basis using publicly available information.<sup>59</sup> Companies are evaluated on a range of indicators that fall broadly into three buckets: governance, freedom of expression, and privacy. Each one contains a range of indicators related broadly to access to various types of information, including several indicators related to content and account moderation, as well as remedy and appeals and other related processes. These criteria mirror much of the same type of information that is available in tech company transparency reports.

The Internet Commission's Evaluation Framework for Digital Responsibility proposes a detailed set of qualitative and quantitative indicators related to organization, people, governance, content moderation, automation, and safety. It uses public data as well as proprietary information and interviews.<sup>60</sup>

# **Good practices**

The following summary of good practices apply to data and transparency and impact reporting in a range of sectors and across approaches:

- User-friendly and concise
- Accurate and Clear
  - Use of illustrative examples
  - · Use of tables, charts, infographics
  - · Complete
  - Glossary as needed
  - · Explains trends or data interpretation to avoid misinterpretation
  - · Considers which data may need additional context or interpretation
- Accessible
  - Downloadable
  - · Machine readable/API
- Timely

#### •••••••••••••••••••••••

59 Since policies and URLs change and information may be buried in a website, RDR maintains a database of snapshots of the corresponding content for each indicator and an explanation of its assessment, meaning that its assessment is both verifiable and replicable.

<sup>60</sup> See https://drive.google.com/file/d/13aaltNDoynvXHeNZLF2yj1liQsRcdbvb/view.

- Relevant
  - Verifiable
    - · Replicable/Auditable
    - Access to data/information sources

Figure 1: Examples of how an assessment of how San Francisco's police use-of-force transparency reports measured up to best practices<sup>61</sup>

Best Practices for Reporting Data			96A Report	
Context	Reports should provide context to assist users in interpreting data and facilitate informed decision making.	۲	0	
User needs	Reports should include data that is summarized, stratified, and provided in appropriate detail to meet the needs of stakeholders relying on the data.	۲	$\bigcirc$	
Key points	Reports should include a concise and organized executive summary to improve the structure of the report and ensure users can easily follow relevant points.	۲	Θ	
Visualization	Reports should represent data, especially more complex data, through graphics that accurately show trends, relationships, and the most significant information.	۲	$\Theta$	
Open Data	Data that supports reports should be available to increase public trust.	۲	۲	
Accuracy and completeness	Stakeholders should be able to rely on the accuracy and completeness of the data <i>underlying</i> reports to make informed decisions.	CSA will assess this in its full audit*		
Complies with best practice Partly complies with best practice ( Does not comply with best practice				

Exhibit 1: The Police Department Can Improve Its	Use-of-Force Reports by	Aligning Them With Best Practices
Exhibit i. The Fonce Department can improve its	b ose-or-rorce keports by	Angling mem with best riactices

\*As part of its full audit, CSA is assessing the accuracy and completeness of the data underlying the EIS and 96A reports. Source: Best practices from publications on writing statistics for governments; compliance with best practices assessed by CSA.

#### Next steps:

The references and organizations identified in this scoping paper provide a jumping-off point for identifying who to interview<sup>62</sup> and what reports to analyze. There are several databases and existing reports on content moderation and tech sector transparency that can form the basis for further study. Although there are published assessments and analyses of standard and best practices, there are many aspects of the process that can only be gleaned through interviews and analysis of specific reports themselves. These include the collection and analysis of data, the costs involved in different approaches and good practices and how tradeoffs between good practices and resource considerations are made, the internal mechanics in terms of staffing and process that goes into the production of the report, and what has worked for effective dissemination of the report. Interviews can provide further insight into how an organization decided what and when to measure and how it negotiated among its various stakeholders and the expectations of the field in which they are situated. Assessing whether a given set of transparency reporting objective(s) were achieved and whether they reached their target audience(s) - and what worked or didn't in terms of communicating information in a meaningful way (e.g., traffic to relevant webpages, efficacy of printed reports, and adherence to other best practices laid out above) - will be an important part of the research.

#### ••••••••••••••••••••••••

61 Screenshot from https://sfgov.org/dpa/sites/default/files/SFPD\_Key\_lssue\_Report\_Use\_of\_Force\_Data\_Reporting\_12\_18\_19\_FINAL.pdf.

62 This should include a range of stakeholders including government and civil society actors.

To learn more about the Global Internet Forum to Counter Terrorism (GIFCT), please visit our website or email outreach@gifct.org.