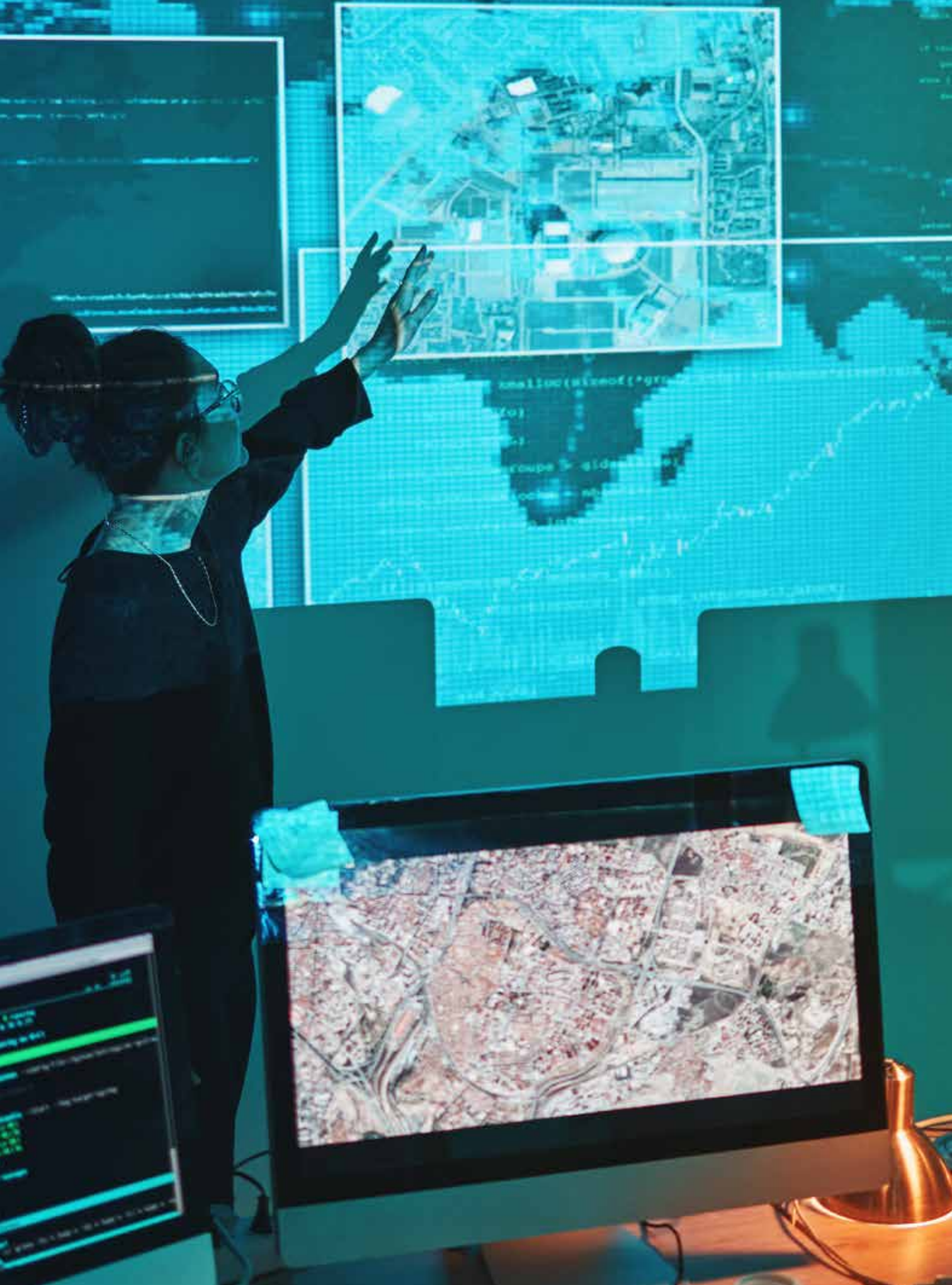

Public–private partnerships as the scaffolding for safer digital spaces

by **Dr. Nagham El Karhili**

The screens we scroll and the streets we walk are part of the same ecosystem: harms that begin online can and do spill into the physical world and vice versa. As features, platforms, and communities proliferate, so do the vectors through which terrorist and violent extremist (TVE) messages travel — from recruitment and radicalization to coordination and the circulation of traumatic imagery. That reality makes clear that keeping online spaces safe is not solely an engineering problem for companies, nor only a law-enforcement task for States; it is a collective challenge that demands durable public–private partnerships, multistakeholderism, and cross-sector cooperation.

Public–private partnerships are one way this idea takes shape in practice. No single sector can prevent, mitigate, or respond to the misuse of digital spaces on its own. Industry can move quickly on technical fixes, but it needs government alignment, civil society insight, and academic expertise to ensure those fixes are responsible, rights-based, and contextually informed. The Global Internet Forum to Counter Terrorism (GIFCT) is one prominent example of this approach.[1] Originally an industry-led partnership, it evolved into a multistakeholder organization and was formalized in 2017 when major platforms agreed to share signals and research to reduce the online reach of violent extremists.[2] Those early cooperative steps — including the establishment of a cross-platform hash-sharing database (HSDB) — underscored that technical progress gains strength when supported by multistakeholder guidance.[3]

The March 15, 2019, Christchurch attacks made this reality tragically visible. What began as an offline act of mass violence was rapidly amplified online: the attacker livestreamed the massacre

[1] "Global Internet Forum to Counter Terrorism".

[2] Microsoft Corporate Blogs, "Facebook, Microsoft, Twitter and YouTube provide update on Global Internet Forum to Counter Terrorism", December 2017.

[3] GIFCT, "GIFCT's Hash-Sharing Database".

and posted a manifesto, both of which spread virally across platforms. Within 24 hours, companies reported removing millions of re-shares, a crisis that underscored two truths: viral harm can outrun single-company defenses, and coordinated, cross-platform response is essential. The attack spurred greater collaboration, pushing platforms not only to scale up incident response together, but also to bring governments and civil society more directly into the process.[4] Post-incident multistakeholder debriefs — including those convened by GIFCT — became a way to reflect, share lessons, and collectively strengthen preparedness for future incidents.

From these early evolutions, a practical lesson emerges: technical fixes are necessary but not sufficient. A reliable response architecture needs three mutually reinforcing elements: rapid, privacy-respecting technical signals; consultative governance to set rules of engagement; and external, rights-focused scrutiny to ensure fairness and legitimacy. GIFCT's trajectory illustrates this lesson well. Its HSDB, for example, provides perceptual "fingerprints" of known terrorist content so re-uploads can be detected and removed more efficiently. But what gets hashed and why is not, and should not be, a purely technical choice: inclusion criteria, taxonomy, and review processes are set through consultative working groups, regular taxonomy reviews,[5] and cross-sector dialogue so that operational tools reflect context and rights considerations as they scale.

Governance is equally important in public–private partnership models. GIFCT's Operating Board — composed of member technology companies — sits alongside an Independent Advisory Committee (IAC) made up of representatives from government, civil society, and intergovernmental organizations. The IAC provides external guidance on priorities, human rights concerns, regional dynamics, and transparency — a model that balances industry agility with independent scrutiny, while amplifying the voices of practitioners and experts who witness harms in different contexts around the world.

In practice, multistakeholder systems can deliver value in three key ways:

First: reducing information asymmetries through convening and knowledge exchange. Threat intelligence is often fragmented: platforms, governments, local civil society, and researchers each hold pieces of the same puzzle. Regular working groups, regional workshops, and threat-specific briefings knit these pieces together so that a public-health-style response can replace ad-hoc firefighting. Routine convenings also help small or regional platforms adopt best practices and prevent gaps that malicious actors may exploit. GIFCT's regional workshops are one example of how consistent convening builds shared situational awareness.

Second: translating operational signals into interoperable technical tools — responsibly. Hash-sharing is a clear example of how industry cooperation reduces re-uploads without requiring every company to build duplicate detection systems from scratch. But the utility of a shared technical tool depends on governance: who decides what is included, what criteria are used, how content is reviewed, and how to handle edge cases where context matters. Iterative taxonomies, inclusion criteria, and incident-response protocols — developed through consultative working groups and subject to independent review — are what make technical interoperability credible and defensible. GIFCT's taxonomy work and transparency reporting illustrate this balancing act.

[4] GIFCT, "The Incident Response Framework".

[5] GIFCT Working Group, "Broadening the GIFCT Hash-Sharing Database Taxonomy: An Assessment and Recommended Next Step", July 2021.

Third: anchoring responses in human rights and independent oversight. Rapid technical responses can reduce harm, but they can also chill legitimate expression or unevenly burden marginalized communities if safeguards are not embedded. Independent advisory mechanisms, human-rights due diligence, and external impact assessments are therefore central to effective public–private partnership. GIFCT's engagement with external experts, its IAC, and its human-rights assessment work demonstrate how accountability and operational speed can be reconciled.

Looking ahead, the tech and threat landscapes continue to shift. Generative AI, new content formats, and a diversified tech stack — from gaming and marketplaces to encrypted services and decentralized platforms — expand both the vectors of exploitation and the tools for mitigation. That duality means multistakeholder systems must accelerate knowledge exchange to ensure policy and engineering are informed by context, ramp capacity building for smaller platforms, and invest in interdisciplinary research that connects technical signals with sociopolitical dynamics. GIFCT's academic research arm, the Global Network on Extremism and Technology (GNET), is one example of how industry and academia can be bridged to advance this work.[6]

There are hard choices ahead: how to preserve open discourse while reducing the reach of violent narratives; how to equip smaller companies without encouraging automated over-removal; how to ensure incident response remains rapid without centralizing control. The only realistic way to nav-

[6] "Global Network on Extremism and Technology (GNET)".

igate those tradeoffs is together. Multistakeholderism — not as a slogan but as a practice of shared rules, independent scrutiny, regional engagement, and transparent feedback loops — is the best path we have to design online spaces that are resilient, rights-respecting, and responsive.

In short: terrorists and violent extremists will keep adapting; so must we. If public safety, human dignity, and democratic values matter, then industry, governments, civil society, and academia must keep designing and testing solutions together. The task is ongoing, but the lesson is timeless: the screens we scroll and the streets we walk are inseparable — and when they intersect, our responses should too.

About the Author

Dr. Nagham El Karhili is the Membership and Programs Senior Lead at the Global Internet Forum to Counter Terrorism (GIFCT). She oversees the full membership cycle, recruiting and mentoring new tech platforms and facilitating ongoing engagement with industry, government, and civil society partners. Her work ensures GIFCT's membership and programming strategies advance its mission to combat terrorism and violent extremism online while upholding human rights and addressing diverse multistakeholder needs. Before joining GIFCT, Dr. El Karhili served as Program and Research Manager at the Horizon Forum, a think-and-do tank focused on hate-funding in philanthropy, where she led investigative projects and convenings that informed policy recommendations and sector best practices. She holds a PhD in Communication from Georgia State University, where she was a Presidential Fellow at the Transcultural Conflict and Violence Initiative, examining violent extremism, organizational religious identity, and civil society resilience. She has taught media, religion, and peacebuilding at Georgia State University and Purdue University, and has also earned a BS and an MS from the University of Louisiana at Lafayette.

FREEDOM FROM FEAR
M A G A Z I N E

The New Criminal Code

Deciphering Emerging Threats in Cyberspace

DOWNLOAD ISSUE 20

"Cyberspace has become a defining arena for contemporary crime, conflict and security"