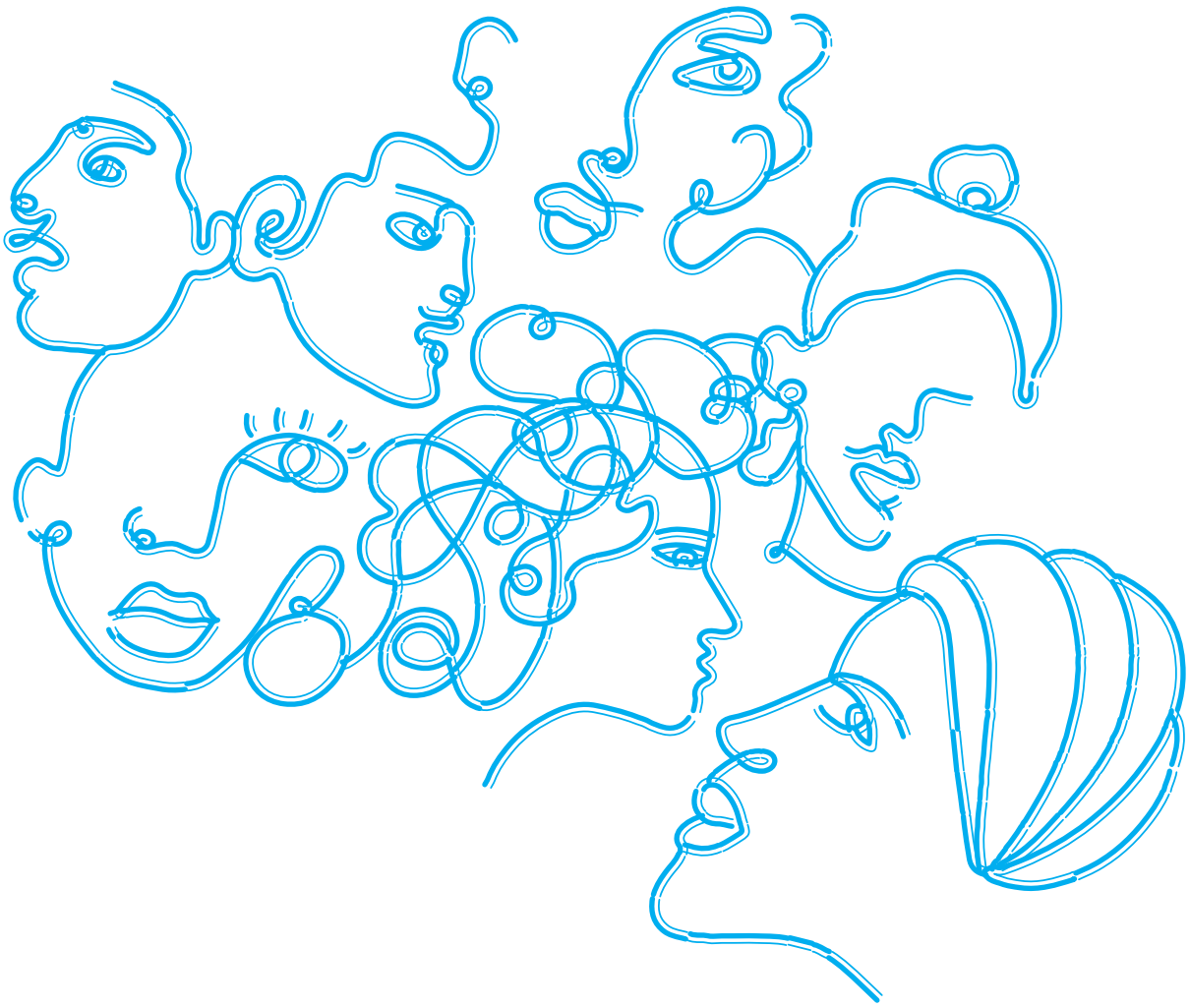
REPORT
FEBRUARY 2023

A DECOLONIAL, FEMINIST APPROACH TO DIGITAL ETHICS AND GOVERNANCE

Aarathi Krishnan

#GWL | VOICES FOR CHANGE & INCLUSION
www.gwlvoices.com



Global Governance Innovation Network

COALITION for the UN WE NEED

Bahá'í International Community

MESSAGE TO READERS

New technologies and the increasing sophistication of information technologies can have two conflicting impacts. On the one hand, it can open access and democratize information and communication in real time. Still, it can exert violence and harassment, especially against women.

Social media platforms have become indispensable after a pandemic accelerated the virtualization of political debates. Unfortunately, though these platforms might seem powerful tools to amplify women's participation in politics and public life, social media has transformed into a toxic space where politically active individuals, especially women, face online harassment and abuse.

Female politicians and, in general, women that have a public life face online violence daily, including insults and hate speech, embarrassment and reputational risk, physical threats, and sexualized misrepresentation.

Online gender-based violence is a scary manifestation affecting women's empowerment and represents a serious menace to the core of our political systems. On the one hand, women self-censor, minimizing their political speech and online activity to avoid attacks. But on the other hand, oppressive and authoritarian actors use this tainted tactic to intimidate women and girls who are politically active.

A global survey conducted with over 26,000 adolescent girls and young women across 33 countries showed that one in 4 participants feels less confident to share their views online, and one in 5 has stopped engaging in political or current affairs on online platforms. These statistics show that violence on social media effectively deprives women and girls of their right to express their political opinion freely, get involved in decision-making, and ultimately become leaders of their communities.

Under the current circumstances, the struggle to promote women's political participation is incrementally long and tortuous. According to UN Women, after the pandemic, we will have to wait at least 140 years to achieve gender parity in positions of power and leadership in the workplace and at least 40 years to achieve equal representation in national parliaments.

There are already too many battlefields in the different levels of government and decision-making to add one more, such as the digital environment, to the spaces where women are forced to claim their place and make their voices heard.

GWL Voices works to raise awareness of gender-based violence in the digital space, where it appears as if the law of the jungle prevails. Faced with online gender-based violence,

we cannot remain silent because our political participation is at risk solely because we are women, a prejudice rooted in the dominant political structure and replicated as if it were natural on social media platforms and all over the internet.

This report represents our collective stance on an issue needing immediate action. It provides an opportunity for reflection and serves as a call to action to elevate the protection of women at all levels of society.

We want to thank our partners, GGIN, C4UN, Bahá'í International, and the International Alliance of Women, whose foresight and support have enabled us to carry out this critical work.

Here's to elevating our voices for change and inclusion while advancing the issues we care about most.

#GWL VOICES FOR CHANGE & INCLUSION

AUTHOR

Aarathi Krishnan

Affiliate - Berkman Klein Centre for Internet and Society, Harvard University. 2020-2022 Tech and Human Rights Fellow - Carr Center for Human Rights Policy, Harvard Kennedy School

Like the technical architecture of classic colonialism, digital colonialism is rooted in the design of a tech ecosystem for the purposes of profit and plunder. If the railways and maritime trade routes were the "open veins" of the Global South back then, today, digital infrastructure takes on the same role: extraction of data gleaned from the streams of information given up as residents of all countries go online, register for state benefits, and connect with one another through applications whose terms of service demand they give up their personal and private information.

Digital futures are not neutral—nor are those that advocate for its expansion or its mitigation, neutral either. There are complex reasons behind this. Digital systems exist within complex imperial formations. As communications studies scholar Paula Chakravartty (Gallatin School of Individualized Study, n.d.) suggests in her studies of new media and racial capitalism, these are all interlocking formations, built on imperial rivalries and a tech worldview that imagines some figures—especially the migrant working classes of the Global South—as outside the world of tech itself. Like the technical architecture of classic colonialism, digital colonialism is rooted in the design of a tech ecosystem for the purposes of profit and plunder. If the railways and maritime trade routes were the "open veins" of the Global South back then, today, digital infrastructure takes on the same role: extraction of data gleaned from the streams of information given up as residents of all countries go online, register for state benefits, and connect with one another through applications whose terms of service demand they give up their personal and private information. As Mohamed, Png and Isaac argue, "the coloniality of power can be observed in digital structures in the form of socio-cultural imaginations, knowledge systems, and way of developing and using technology which are based on systems, institutions and values that persist from the past and remain unquestioned in the present"[1].

The act of colonialism, or colonization, removes power from the colonized, dispossesses and transfers economic resources, and removes culture in the name of 'civility.' Coloniality presents itself in a matrix of power that operates through control or hegemony over the economy,

1 Mohamed, Png & Isaac (2020), *Decolonial AI: Decolonial Theory as Socio-Technical Foresight in Artificial Intelligence, Philosophy and Technology* (405)

including land, labour, and natural resources; authority; gender and sexuality; and subjectivity and knowledge. Coloniality was presented to the world as 'modernization' but this was at grave expense to freedom, justice, equality, and a homogenous world view. The ideas of modernity that resulted come because of this domination. How we then conceive and transfer knowledge, as well as what knowledge we see as credible and valid are also based on these colonial beliefs. How we exist within these structures and how we interpret reality is deeply influenced by colonization as well. Digital colonialism isn't the only privileging force that digital systems sometimes uphold. It can also uphold forces like *patriarchy, race and ethnicity bias, paternalism, hetero/cis normativity, classism and class privilege, caste privilege, ableism and ageism.*

The concerns of the biases of AI and digital systems on populations are slowly getting traction. Some governments are moving to simultaneously limit the power of tech companies with an 'urgency and breadth that no single industry has experienced before'[2]. Technology firms are working to embed tech ethics into their ways of working, though this hasn't been without controversy or notions of ethics washing. Campaigns to ban specific technologies that disproportionately harm marginalised and historically oppressed communities are getting traction (i.e. facial recognition technologies in policing), **however this hasn't quite in equal ways globally**.

How we exist within these structures and how we interpret reality is deeply influenced by colonization as well. Digital colonialism isn't the only privileging force that digital systems sometimes uphold. It can also uphold forces *like patriarchy, race and ethnicity bias, paternalism, hetero/cis normativity, classism and class privilege, caste privilege, ableism and ageism*.

As it stands today, digital democratisation is fundamentally flawed as it "predicated on a future where soon everyone will have a personal internet connection, a social media account, and therefore will be able to create content and collectively engage online" (Kaurin, 2021). Digital tools for diverse communities are designed based on the assumption that local communities ought to meet on the platforms that international organisations are familiar with, rather than localising technology tools to be more easily accessible – i.e., translated into local languages, designing UX for diversity and for use in places with slower WiFI and poorer telecom

2 https://www.nytimes.com/2021/04/20/technology/global-tipping-point-tech.html

infrastructures[3]. Supporting this is the argument that scholars and civil society organisations tend to present – that the use of data technologies as harm-inducing techno-solutionism or techno-colonialism"[4].

In addition, the **values assigned to what determines complex principles such as rights, fairness, and privacy do not necessarily consider the cultural contexts** that those values and principles are being applied to. The right to privacy for example, has been one that has "historically been most difficult to define in a legal framework, due to not only its roots in cultural rituals, but also changing societal and political norms"[5]. Salil Shetty, former Secretary General of Amnesty International, said this of human rights: "human rights often mean different things to different people. And they don't mean anything at all for a good number of people in the developing world". Shetty further argues that "colonialism, and early, modern-day human rights fed upon each other…the development and flourishing of the institution of international law itself – with its definition and consolidation of the notions of sovereignty, statehood, trusteeship, and protection – become inextricably linked to the colonial project". By only continuing to draw on normative frameworks and definitions of these values and principles, it ensures that Western homogeneous definitions of privacy or the rights to be left alone, end up being imposed down on other cultures.

The fast evolution of AI technology amplifies the risks it brings with it. The confinement and risk to vulnerable people is happening at speed, meaning it is harder to seek justice. Digital technologies and AI mask ideologies of power and are wed to a market ideology of dominance. To intentionally carve a different type of ideology would require governance systems that are informed by different knowledge sources that tangibly influence decision making and that prioritise a focus not just on the firefighting of today but rather the implications on future generations.

3 Kaurin, D., *Tech Localisation: Why the localisation of aid requires the localisation of technology*, CDAC Network Annual Forum, 2021

4 Weitzberg, Cheeseman, & Martin (2021), *Between surveillance and recognition, Rethinking digital identity in aid*, Big Data & Society, Sage Pub: https://doi.org/10.1177/20539517211006744

5 Kaurin, D., *Data Protection and Digital Protection for Refugees*, World Refugee Council Research Paper Series, 2019: https://www.cigionline.org/publications/data-protection-and-digital-agency-refugees/

GOVERNING DIGITAL SYSTEMS

An incredible amount of work has gone into establishing data governance protocols that have just shifted how organisations think about data. This ranges from the work being led by the OCHA's Centre of Humanitarian Data on Data Responsibility, UNDGs Guidance Note on Data Privacy, Ethics and Protection, and numerous others. Emerging areas of research on AI and Human Rights are calling for a human rights centred approach to AI Governance . The UN Report on Roadmap for Digital Cooperation goes further into issues of digital public goods, digital cooperation, and digital human rights and points out that "developing countries are largely absent from or not well-represented in most prominent forums on artificial intelligence..." and that "current artificial intelligence-related initiatives lack overall coordination in a way that is easily accessible to other countries outside the existing groupings, other United Nations entities and other stakeholders...and without a broader, more systematic attempt to harness the potential and mitigate the risk of artificial intelligence, opportunities to use it for the public good are being missed"[6]. UNESCO's AI Decision Maker's Toolkit (2019) aims to elaborate a standard setting instrument on ethics of AI drawing on elements of trends, recommendations, implementation guides and capability building resources for the development of a human rights-based and ethical AI[7].

In addition to this, are questions of immunity (traditionally, institutions – multilateral or otherwise – do not go beyond the individual institutional governance mechanisms, that often are bordered by the institution's immunity); and appropriateness of the technology innovations we deploy.

The commonality across all these protocols as it exists today, is that it **narrows** the focus of governance on issues such as data and AI, rather than the broader governance of the use of digital systems **within the systems it intersects with.** Pizzi, Romanoff and Engelhardt (2021) argue that current *codes* of ethics are limited in that they are "not binding, like law and hence

6 Roadmap for Digital Cooperation (2020), United Nations: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf

7 https://en.unesco.org/artificial-intelligence/decision-makers-toolkit

do not promulgate compliance; they often reflect the values of the organisation that created them, rather than the diversity of those potentially impacted by AI systems; and they are not automatically operationalized by those designing and applying AI tools on a daily basis"[8]. In addition to this, are questions of **immunity** (traditionally, institutions - multilateral or otherwise - do not go beyond the individual institutional governance mechanisms, that often are bordered by the institution's immunity); and **appropriateness** of the technology innovations we deploy. More urgently, what weaves these protocols together is its focus on the *now* and its *homogeneity* - from how we understand harm, protection, human rights, and ethics solutions, and then **what the potential implications might be in the future on impacted populations.**

Operationalising this approach requires clear pathways to how it influences decision making and accountability — as without these, it is merely side-lined to tokenistic gestures of 'inclusion' that continue to affirm power as it looks like today in a status quo that arguably is not fit for the future.

The fast evolution of AI technology amplifies the risks it brings with it. The confinement and risk to vulnerable people is happening at speed, meaning it is harder to seek justice. Digital technologies and AI mask ideologies of power and are wed to a market ideology of dominance. To intentionally carve a different type of ideology would require governance systems that are informed by different knowledge sources that tangibly **influence** decision making and that **prioritise a focus not just on the firefighting of today but rather the implications on future generations.** Kaurin (2021) argues, "digital spaces are representative of ontic spaces; the same challenges regarding gender, age, ableism and discrimination that prevent diverse and representative engagement on the ground are replicated online"[9].

8 Pizzi, Romanoff and Engelhardt (2021), *AI for humanitarian action: Human rights and ethics,* Cambridge University Press.

9 Kaurin, ibid.

REIMAGINED DIGITAL ETHICS AND GOVERNANCE

This brief posits a digital ethics and governance framework that specifically aims at interrogating and analysing *context, motives, and impact* of use over a long-term time frame. This focus is to enable safeguards to be built-in to ensure a broader accountability to public interest as well as ensure that practices and frameworks do not advertently or inadvertently obstruct peoples current *and new* rights (new rights draw on the argument that rights are neither static nor immutable, and any efforts to design equitable, flourishing futures must consider entirely new forms of harms and rights[10] to be effective.

An incredible amount of work has gone into establishing data governance protocols that have just shifted how organisations think about data.

Ethics and governance models that are baselined in decolonial theory interrogates patterns of power that shape our intellectual, political, economic, and social world. Our models of governance and strategic design, require a broader evolution to consider the "sexual, gender, spiritual, epistemic, economic, political, linguistic, aesthetic, pedagogical and racial hierarchies of the "modern/colonial western-centric Christian-centric capitalist/patriarchal world-system"[11]. By embedding a decolonial critical approach within the technical practice of ethical and governance practice, digital systems can ensure that the impacts of these systems can amplify impacted populations' ability to flourish in the long term, rather than just to survive in the short term.

The commonality across all these protocols as it exists today, is that it narrows the focus of governance on issues such as data and AI, rather than the broader governance of the use of digital systems within the systems it intersects with.

10 Raman & Schultz, The Coming Good Society: Why New Realities Demand New Rights, Harvard University Press, 2020

11 Grosfoguel, Ramón; https://dialogoglobal.com/texts/grosfoguel/Grosfoguel-Decolonizing-Pol-Econ-and-Postcolonial.pdf

What the framework aims, is to ensure the following principles are imbued into the use and deployment of digital systems:

- **Positionality:** That *who* holds the responsibility for developing governance, ethics, protocols, and standards of use and deployment consider the positionality of the authorship and decision making within their wider metropolis, and assess the impacts any biases or privileges that that gives rise to
- **Future Impacts and Harm Assessment:** The organisations deploying these systems include an expansion on a range of criteria for assessment including plausible, possible, and probable future harms and impacts that might arise on impacted populations and on their future generations. This involves going beyond assessments to ensure that something is used only within its prescribed intent, but rather to assess function creep[12] current and into the future.
- **Plurality in Legitimacy:** Utilising a wider range of knowledge sources and experiences to legitimize a multiplicity of conceptual models related to 'ethics' and 'fairness' to ensure assessments and conclusions drawn, intentionally do not replicate an echo-chamber worldview via a limited homogenous perspective that do not account for normative and cultural realities.
- **Reverse Accountability against Meaningless Consent:** Ensuring that accountability mechanisms are designed to hold institutions to account so that the burden of harm is not continuously placed on impacted peoples. Specifically, this applies to issues of meaningless consent[13] where reverse accountability can provide redress and ensures institutions bear the costs when ethical principles are violated.
- **Beyond Empowerment and Inclusion:** Moving beyond just 'diversity and inclusion' as a metric for mitigating bias, but ensuring clear, transparent mechanisms that ensure the inclusion of historically excluded, impacted populations can *influence* decision making and provide a re-weighting of decision-making priorities.
- **Relational Ethics[14]:** Assessing patterns across a wider range of contextual social, technical, economic, and historical systems, norms, and structures to understand *why* rather than merely designing technical solutions and systems blindly.
- **Transparent Privilege and Dispossession:** Assessing whose rights are privileged and whose are dispossessed in decision making, and the risk of that weighting in the short term and long-term time horizon.
- **Objective Truth v Constructed Representation:** Ensuring governance systems are utilised to recognise the *context* digital systems will exist in as opposed to a singular representation at a specific point in time[15], to design malleability and adaptability in changing contexts over time

Operationalising this approach requires clear pathways to how it influences *decision making and accountability* - as without these, it is merely side-lined to tokenistic gestures of 'inclusion' that continue to affirm power as it looks like today in a status quo that arguably is not fit for

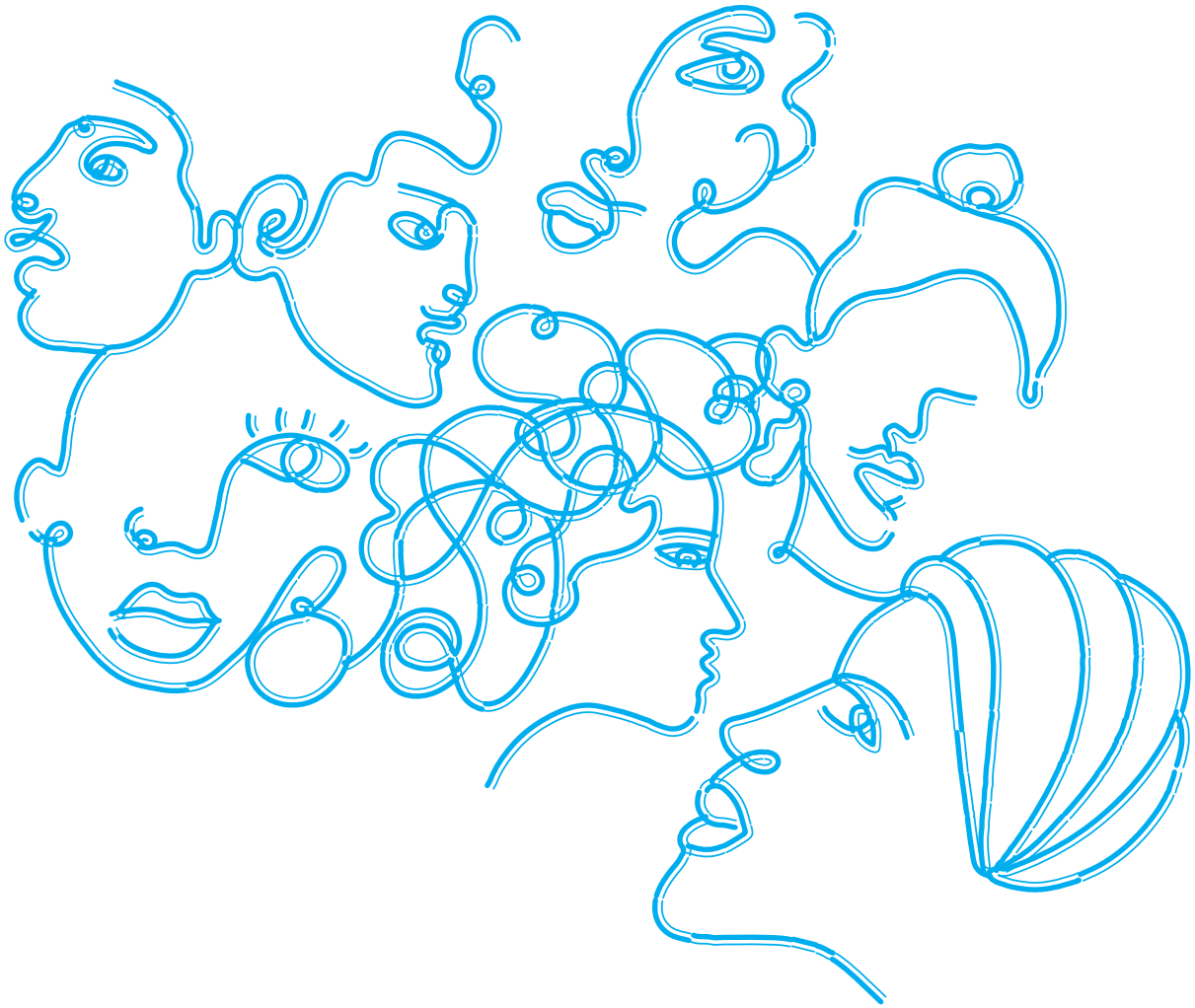
12 Wright and Verity, ibid
13 Wright and Verity, ibid.
14 Birhane, ibid.
15 McDonald, S; ibid.

the future. As Ahmed Ansari argues "the whole project of democratically-decided AI futures is a cover for essentially continued Anglo-European coloniality and re-asserting global white supremacy because mere representation doesn't necessarily equate to radical alterity"[16].

By embedding a decolonial critical approach within the technical practice of ethical and governance practice, digital systems can ensure that the impacts of these systems can amplify impacted populations' ability to flourish in the long term, rather than just to survive in the short term.

Radical alterity in this sense, is the collective responsibility of all humanitarian actors to not just expect dehumanizing resilience as a coping mechanisms for impacted populations to just deal with whatever might come from digital systems in their futures, but rather to radically work towards mitigating and designing systems that don't just say 'we leave no one behind' but rather very intentionally designing through justice, equity and resistance to never do so.

16 https://medium.com/a-new-ai-lexicon/a-new-ai-lexicon-modernity-coloniality-7f6979ffbe82



WWW.GWLVOICES.COM

Global Governance Innovation Network

COALITION for the UN WE NEED

Bahá'í International Community

#GWL | VOICES FOR CHANGE & INCLUSION