

The Governance Gap: A Measurable Framework for Assessing Institutional Readiness to Deliver the 2030 Agenda

Sam Polkar¹, Chidinma Adebayo¹

¹ Advocacy Unified Network

DOI: 10.62692/opuspublica.gppd.2023.01 Published: August 29, 2023

ABSTRACT

As the international community approaches the midpoint of the 2030 Agenda for Sustainable Development, available evidence indicates that progress on the Sustainable Development Goals (SDGs) is seriously off track. While financing deficits, pandemic disruption, and geopolitical instability are widely recognised obstacles, this paper argues that a further, less systematically examined condition also shapes implementation: the gap between formal SDG commitments and the governance arrangements required to deliver them. The paper defines the governance gap as the measurable distance between the institutional, procedural, and normative requirements for effective SDG implementation, as articulated in the 2030 Agenda, and the actual state of governance arrangements and practices in a given jurisdiction. It then develops the Governance Gap Assessment Matrix (GGAM), a structured framework comprising five dimensions—normative, institutional, operational, data and accountability, and participatory—assessed across three levels: formal, functional, and outcome-linked. For each of the resulting fifteen cells, the paper proposes indicators, operational definitions, evidence sources, and provisional scoring anchors, and it introduces two derived measures: a Governance Readiness Score (higher values indicate stronger readiness) and a Governance Gap Score (its transparent complement). The paper positions the GGAM against existing instruments, including the Worldwide Governance Indicators, the OECD Policy Coherence for Sustainable Development framework, the SDG 16 monitoring architecture, and the VNR Quality Index, and illustrates its potential application through a qualitative synthesis of evidence from the Voluntary National Review (VNR) process and related literature. The paper is explicitly positioned as a conceptual and foundational framework rather than a systematic empirical study; the illustrative evidence is used to demonstrate analytical utility, not to establish statistically representative findings. Limitations concerning measurement subjectivity, data availability, contextual generalisability, and causal inference are discussed. The paper concludes that the GGAM offers a transparent, replicable scaffold for subsequent empirical validation and for applied governance-gap assessment by governments, civil society, and international monitoring bodies, and that governance is a cross-cutting enabling condition that can substantially influence whether financial, technical, and institutional resources translate into effective SDG implementation.

TABLE OF CONTENTS

Abstract

1. Introduction

2. Conceptual Foundations and Literature Review

2.1 From Global Governance Gaps to SDG-Specific Governance Gaps

2.2 Existing Governance Measurement Frameworks

2.3 The Distinctive Contribution of the GGAM

3. Defining the Governance Gap: Conceptual Architecture

3.1 Definitional Boundaries and Scope

3.2 The Five Dimensions of the Governance Gap

3.2.1 Normative Governance Gap

3.2.2 Institutional Governance Gap

3.2.3 Operational Governance Gap

3.2.4 Data and Accountability Governance Gap

3.2.5 Participatory Governance Gap

4. Operationalising the Framework: The Governance Gap Assessment Matrix

4.1 Matrix Architecture

4.2 Indicators and Measurement Approach

4.3 Scoring Methodology

4.4 Governance Readiness Score (GRS) and Governance Gap Score (GGS)

4.5 Weighting

5. Illustrative Application: Evidence from the VNR Process

5.1 Methodological Approach

5.2 Recurring Governance Challenges in VNR-Related Evidence

5.3 Illustrative Findings and Their Limits

6. Counterarguments and Limitations

6.1 Governance versus Capacity

6.2 Measurement Challenges, Reliability, and Data Availability

6.3 Contextual Specificity, Generalisability, and Causal Inference

7. Implications and the Path Forward

7.1 For the SDG Summit 2023 and Beyond

7.2 For National Governments

7.3 For Civil Society and the Accountability Ecosystem

8. Conclusion

References

Appendix A. Detailed GGAM Measurement and Scoring Protocol

Appendix B. Hypothetical Worked Scoring Example (Illustrative)

Appendix C. Source Data and Verification Record for Figure 4: VNR Presenters at HLPF Sessions, 2016–2023

The Governance Gap:

A Measurable Framework for Assessing Institutional

Readiness to Deliver the 2030 Agenda

A Foundational Framework Paper

This framework paper is a conceptual contribution intended to structure subsequent empirical research and applied assessment. It is a scholarly publication and does not represent the official position of any government or intergovernmental organisation.

DOI: Pending registration. The intended identifier (10.62692/aun.govgap.2023) has not yet been registered and should not be presented or resolved as a live link until the publisher confirms official registration, correct resolution, and metadata match. The suggested citation therefore omits the DOI URL.

Abstract

As the international community approaches the midpoint of the 2030 Agenda for Sustainable Development, available evidence indicates that progress on the Sustainable Development Goals (SDGs) is seriously off track. While financing deficits, pandemic disruption, and geopolitical instability are widely recognised obstacles, this paper argues that a further, less systematically examined condition also shapes implementation: the gap between formal SDG commitments and the governance arrangements required to deliver them. The paper defines the governance gap as the measurable distance between the institutional, procedural, and normative requirements for effective SDG implementation, as articulated in the 2030 Agenda, and the actual state of governance arrangements and practices in a given jurisdiction. It then develops the Governance Gap Assessment Matrix (GGAM), a structured framework comprising five dimensions—normative, institutional, operational, data and accountability, and participatory—assessed across three levels: formal, functional, and outcome-linked. For each of the resulting fifteen cells, the paper proposes indicators, operational definitions, evidence sources, and provisional scoring anchors, and it introduces two derived measures: a Governance Readiness Score (higher values indicate stronger readiness) and a Governance Gap Score (its transparent complement). The paper positions the GGAM against existing instruments, including the Worldwide Governance Indicators, the OECD Policy Coherence for Sustainable Development framework, the SDG 16 monitoring architecture, and the VNR Quality Index, and illustrates its potential application through a qualitative synthesis of evidence from the Voluntary National Review (VNR) process and related literature. The paper is explicitly positioned as a conceptual and foundational framework rather than a systematic empirical study; the illustrative evidence is used to demonstrate analytical utility, not to establish statistically representative findings. Limitations concerning measurement subjectivity, data availability, contextual generalisability, and causal inference are discussed. The paper concludes that the GGAM offers a transparent, replicable scaffold for subsequent empirical validation and for applied governance-gap assessment by governments, civil society, and international monitoring bodies, and that governance is a cross-cutting enabling condition that can substantially influence whether financial, technical, and institutional resources translate into effective SDG implementation.

Keywords: governance gap; SDG implementation; institutional readiness; 2030 Agenda; Governance Gap Assessment Matrix; policy coherence; Voluntary National Reviews; governance measurement

1. Introduction

In September 2015, all 193 United Nations Member States adopted the 2030 Agenda for Sustainable Development, a plan of action for people, planet, and prosperity encompassing seventeen Sustainable

Development Goals (SDGs) and 169 associated targets [1]. The Agenda is explicitly universal in scope, applying to all countries regardless of development status, and is grounded in the recognition that eradicating poverty in all its forms remains the greatest global challenge and an indispensable requirement for sustainable development [1]. The ambition was unprecedented. As the midpoint of the implementation period is reached in 2023, however, the evidence assembled by the United Nations, the Sustainable Development Solutions Network (SDSN), and numerous independent assessments points to a sobering reality: the world is falling significantly short of meeting most of the Goals by 2030 [2, 3].

The conventional diagnostic has attributed this shortfall primarily to financing deficits, pandemic disruption, and geopolitical instability. The Sustainable Development Report 2023 estimates that the annual SDG financing gap for developing countries is on the order of USD 4 trillion [3], while the United Nations Special Edition report on SDG progress warns that only about 15 per cent of assessable targets are on track to be met, with close to half moderately or severely off track and roughly one third having stalled or regressed below the 2015 baseline [2]. These figures are alarming, and the financing and geopolitical dimensions are undeniably critical. This paper argues, however, that beneath these more visible deficits lies a deeper, more structural, and less systematically examined condition: the gap between formal SDG commitments and the governance arrangements required to translate those commitments into delivered outcomes. The paper does not claim that governance is the sole or sufficient explanation for implementation shortfalls; rather, it argues that governance is a cross-cutting enabling condition that can substantially influence the extent to which financial, technical, institutional, and social resources are converted into effective implementation.

The term governance gap has an established lineage in the scholarly literature. Thomas G. Weiss and Ramesh Thakur, in their work on global governance, identified multiple categories of gaps—including jurisdictional, operational, incentive, and normative gaps—that describe the distance between the scale of global challenges and the capacity of existing institutions to address them [4]. The OECD has developed an extensive body of work on Policy Coherence for Sustainable Development (PCSD), addressing the systemic misalignment between sectoral policies that undermines integrated action on the SDGs [5]. The World Bank's Worldwide Governance Indicators (WGI), since 1996, have provided six aggregate dimensions of governance quality across more than 200 countries and territories [6]. SDG 16, with its focus on peace, justice, and strong institutions, provides a normative anchor for governance concerns within the Agenda [7]. What is comparatively underdeveloped, however, is a unified, operationalisable framework that translates these disparate insights into a single, coherent, and measurable construct designed specifically for assessing the governance dimensions of SDG implementation readiness.

Research question. *How can the gap between formal SDG commitments and the governance arrangements required for their effective implementation be systematically conceptualised and assessed?*

To address this question, the paper pursues four objectives. First, it defines the governance gap as an analytically distinct concept relevant to SDG implementation, distinguishing it from related but different constructs such as the financing gap, the data gap, political will, and technical capacity. Second, it synthesises relevant insights from global governance theory, governance measurement, policy coherence, SDG monitoring, and institutional implementation literatures, identifying what existing approaches capture and what they leave under-specified. Third, it develops the Governance Gap Assessment Matrix (GGAM) as a structured framework—organised around five dimensions and three assessment levels—for diagnosing institutional readiness and governance-related implementation barriers, accompanied by a provisional measurement and scoring protocol. Fourth, it illustrates the potential application of the framework using an indicative synthesis of evidence from the Voluntary National Review (VNR) process and related literature, while being explicit that this illustration is not a systematic empirical study.

The paper is positioned as a conceptual and foundational framework paper with an illustrative synthesis of evidence, rather than as a systematic empirical cross-country study. It does not report statistically representative findings, definitive rankings, or causal estimates. Where the illustrative evidence suggests recurring patterns, the paper states that these patterns warrant systematic investigation through subsequent, dedicated research. The framework is designed for subsequent empirical validation. The remainder of the paper proceeds as follows: Section 2 reviews the conceptual and measurement literatures and positions the GGAM against existing instruments; Section 3 sets out the conceptual architecture of the governance gap and its five dimensions; Section 4 operationalises the framework, presenting the matrix, the measurement approach, and the scoring methodology, including the Governance Readiness Score and the Governance Gap Score; Section 5 presents an illustrative application to VNR-related evidence; Section 6 discusses counterarguments and limitations; Section 7 sets out implications for the upcoming SDG Summit, national governments, and the accountability ecosystem; and Section 8 concludes.

2. Conceptual Foundations and Literature Review

2.1 From Global Governance Gaps to SDG-Specific Governance Gaps

The concept of governance gaps has its intellectual roots in the global governance literature of the early twenty-first century. Weiss and Thakur's taxonomy identified several categories of gaps in global governance: jurisdictional gaps, arising when no institution has the authority to address a problem; operational gaps, reflecting the distance between institutional mandates and actual implementation capacity; incentive gaps, where actors lack motivation to act; normative gaps, where shared principles are absent or contested; and participatory and compliance gaps, denoting failures of inclusion and adherence to agreed rules [4, 8]. This taxonomy was developed primarily to analyse governance failures at the international system level, examining the capacity of the United Nations and other multilateral institutions to manage transnational challenges such as armed conflict, pandemics, and climate change.

The relevance of this taxonomy to SDG implementation is both direct and comparatively underexplored. The 2030 Agenda is, at its core, a governance framework. It establishes normative standards through its declaration and goals, creates policy expectations through its targets, requires institutional arrangements for implementation at national and subnational levels, and depends on compliance mechanisms—primarily the VNR process—to maintain accountability [1]. Each of these functions corresponds to elements of the global-governance-gap taxonomy. A normative governance gap manifests when national development plans and political discourse fail to internalise the integrated, indivisible vision of the 2030 Agenda. An operational gap appears where coordination mandates exist on paper but lack the staffing, budgetary authority, or political backing required to function effectively. A compliance gap is reflected in the voluntary, non-binding nature of the VNR process, which provides limited accountability pressure to drive transformative change [9]. The participatory gap captures the distance between the Agenda's commitment to multi-stakeholder partnership and the often consultative reality of stakeholder engagement.

The global-governance-gap concept has been further elaborated to emphasise the mismatch between the scale of challenges and the limitations of existing institutions, a mismatch structurally embedded in a decentralised international system built on state sovereignty [11]. Translating these system-level insights to the national and subnational setting requires care: the unit of analysis shifts from intergovernmental institutions to domestic governance arrangements, and the relevant gaps concern the internal capacity of states to honour universal commitments rather than the capacity of the international system to manage cross-border externalities. The GGAM proposed here makes this translation explicit, recasting the system-level taxonomy as a diagnostic instrument for national and subnational SDG implementation.

2.2 Existing Governance Measurement Frameworks

Several established frameworks measure aspects of governance, each with distinct strengths and limitations when applied to the SDG context. The World Bank's WGI represent the most widely used cross-national governance dataset, covering six dimensions: voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption [6]. These indicators have been influential in establishing the empirical association between governance quality and development outcomes. The methodology has, however, been subject to critique regarding measurement error and the assumption of independence among the underlying baseline indicators, with subsequent work showing that relaxing this assumption can alter the conclusions drawn from the indicators [12]. Moreover, the WGI captures general governance capacity but does not specifically address the institutional arrangements, policy-coherence mechanisms, or data systems most directly relevant to SDG implementation.

A second relevant framework is the OECD's PCSD analytical framework, which provides a comprehensive approach to integrating the economic, social, and environmental dimensions of sustainable development throughout the policy cycle [5]. The 2023 OECD report on driving policy coherence indicates that, while many countries engage with long-term visioning, significant challenges remain in translating vision into cross-sectoral action and in embedding impact assessment into routine decision-making [13]. The PCSD framework is particularly strong in its attention to the policy cycle, from planning through coordination and implementation to monitoring. It is, however, designed primarily for OECD member states and addresses policy coherence as one component of governance rather than as an integrated, multi-dimensional governance assessment spanning all seventeen Goals.

SDG 16 itself provides a third measurement framework, with twelve targets and twenty-three indicators addressing violence reduction, access to justice, and institutional effectiveness [7]. The SDG 16 Data Initiative, a consortium of civil society organisations coordinated by International IDEA, complements official statistics by tracking progress on these indicators and highlighting data availability challenges [14]. The Global Progress Report on SDG 16 indicators, jointly produced by UNODC, UNDP, and OHCHR and released in July 2023, assesses trends in violence, justice, and institutional trust and notes that progress is worryingly slow and that data coverage remains insufficient for evidence-informed policy in many areas [15]. While directly relevant to governance, SDG 16 indicators cover only a subset of the governance challenges that affect SDG implementation more broadly: a country may perform well on specific SDG 16 indicators such as homicide rates or birth registration while still exhibiting significant gaps in cross-ministerial coordination, data integration, or stakeholder participation.

A fourth instrument is the VNR Quality Index developed by CEPEI, which provides a methodology for qualitatively assessing the quality of VNR reports against principles of transparency, inclusiveness, and evidence-based analysis, and which also includes a governance-quality index component [21]. Broader scholarship on measuring governance for sustainable development emphasises both the promise and the pitfalls of governance indices, warning against treating composite scores as objective truths and underscoring the political economy of how such measures are produced and used [26]. Collectively, these frameworks provide valuable but partial lenses: each illuminates a dimension of governance relevant to the SDGs, but none offers an integrated, SDG-centric diagnosis of the gap between formal commitments and implementation-ready governance arrangements across all Goals.

2.3 The Distinctive Contribution of the GGAM

Rather than claiming to be the first framework of its kind, the GGAM seeks to integrate dimensions that are often assessed separately within existing governance and SDG implementation frameworks. Its distinctiveness lies in three respects. First, it is explicitly SDG-centric: each dimension is defined in terms of its direct

relevance to the 2030 Agenda's implementation requirements, not to governance in the abstract. Second, it is multi-dimensional yet integrated: the five dimensions are analytically distinct but empirically interrelated, recognising that governance shortfalls in one dimension tend to compound shortfalls in others. Third, it is operationally measurable: each dimension is associated with specific indicators that can be assessed using publicly available evidence—primarily VNRs, international organisation reports, and existing governance datasets—so that the framework is practicable without primary data collection, while remaining open to refinement through empirical validation.

Table 2 summarises how the GGAM relates to four established instruments. The comparison is not meant to imply superiority in all respects; each existing framework was designed for a purpose at which it excels. The point is to clarify what the GGAM adds: a unified, SDG-specific, multi-level lens that treats the gap between formal arrangements and implementation functionality as the central diagnostic object, and that links this gap transparently to implementation outcomes without claiming causal attribution.

Table 2. Comparative Positioning of the GGAM Against Existing Governance and SDG Frameworks

Framework	Primary purpose	Unit of analysis	Treatment of the formal-functional gap and SDG implementation
Worldwide Governance Indicators (WGI) [6]	Cross-national benchmarking of general governance quality.	Country (aggregate, perception-based).	Captures de facto governance outcomes broadly; not SDG-specific; does not trace the formal-to-functional translation for SDG delivery.
OECD Policy Coherence for Sustainable Development (PCSD) [5, 13]	Aligning sectoral policies with sustainable development across the policy cycle.	Country (mostly OECD members).	Strong on policy coherence and cross-sectoral coordination; does not separately diagnose normative, data, or participatory readiness for all SDGs.
SDG 16 monitoring framework / SDG 16 Data Initiative [7, 14, 15]	Tracking Goal 16 targets and indicators on peace, justice, institutions.	Country, target-level.	Goal-specific and indicator-rich; limited to Goal 16 scope; does not assess cross-Government implementation architecture for all SDGs.
VNR Quality Index (CEPEI) [21]	Assessing the quality of VNR reports and institutional reporting processes.	VNR report / reporting process.	Assesses reporting quality and includes a governance-quality component; focuses on the review document rather than the underlying implementation gap across all dimensions.
Governance Gap Assessment Matrix (GGAM) (this paper)	Diagnosing the gap between formal SDG commitments and implementation-ready governance arrangements across all SDGs.	Country (or subnational), across 5 dimensions x 3 levels.	Treats the formal-functional-outcome-linked gap as the central diagnostic object; SDG-specific and cross-Governmental; explicitly avoids causal attribution of outcomes to governance alone.

Note. The comparison is analytical and illustrative. Existing frameworks are summarised in terms of their primary published purpose; the GGAM is intended to complement, not replace, these instruments.

3. Defining the Governance Gap: Conceptual Architecture

3.1 Definitional Boundaries and Scope

Before proceeding to the dimensional analysis, it is essential to establish the definitional boundaries of the governance gap as used in this paper. The governance gap refers to the measurable distance between the institutional, procedural, and normative requirements for effective SDG implementation, as articulated in the

2030 Agenda and subsequent intergovernmental guidance, and the actual state of governance arrangements and practices in a given country or jurisdiction. This definition deliberately encompasses both structural dimensions, such as the existence and design of coordination mechanisms, and functional dimensions, such as the quality of policy coherence, data availability, and stakeholder engagement. It is deliberately framed as a diagnostic construct rather than as a causal explanation: the framework identifies where governance arrangements fall short of implementation requirements, but it does not by itself establish that closing a given gap will produce a specific SDG outcome.

Several delineations are important. The governance gap is distinct from the financing gap, although the two are related: weak institutional coordination can lead to inefficient allocation of available resources, thereby exacerbating financing shortfalls. It is distinct from the data gap as conventionally understood, although data deficiencies constitute one of its five dimensions. It is distinct from political will, although the absence of political commitment is a primary driver of governance gaps; the framework focuses on observable, measurable institutional and procedural attributes rather than on the underlying motivations of political actors. Finally, the governance gap is distinct from capacity in the narrow sense of technical skills, although capacity constraints interact with governance (see Section 6.1). A country may possess abundant technical expertise yet still exhibit a significant governance gap if institutional arrangements prevent that expertise from being effectively deployed for SDG delivery.

Figure 1 presents the conceptual architecture of the governance gap. It depicts the movement from formal SDG commitments, through institutional and functional governance arrangements and implementation processes, to outcome-linked results, and indicates the points at which a governance gap may emerge. The figure is a conceptual representation: it identifies where gaps can arise in the chain from commitment to outcome, but it does not assert a proven causal model, since outcomes are jointly shaped by financing, economic conditions, conflict, climate shocks, demographics, and other structural variables (see Section 6.3).

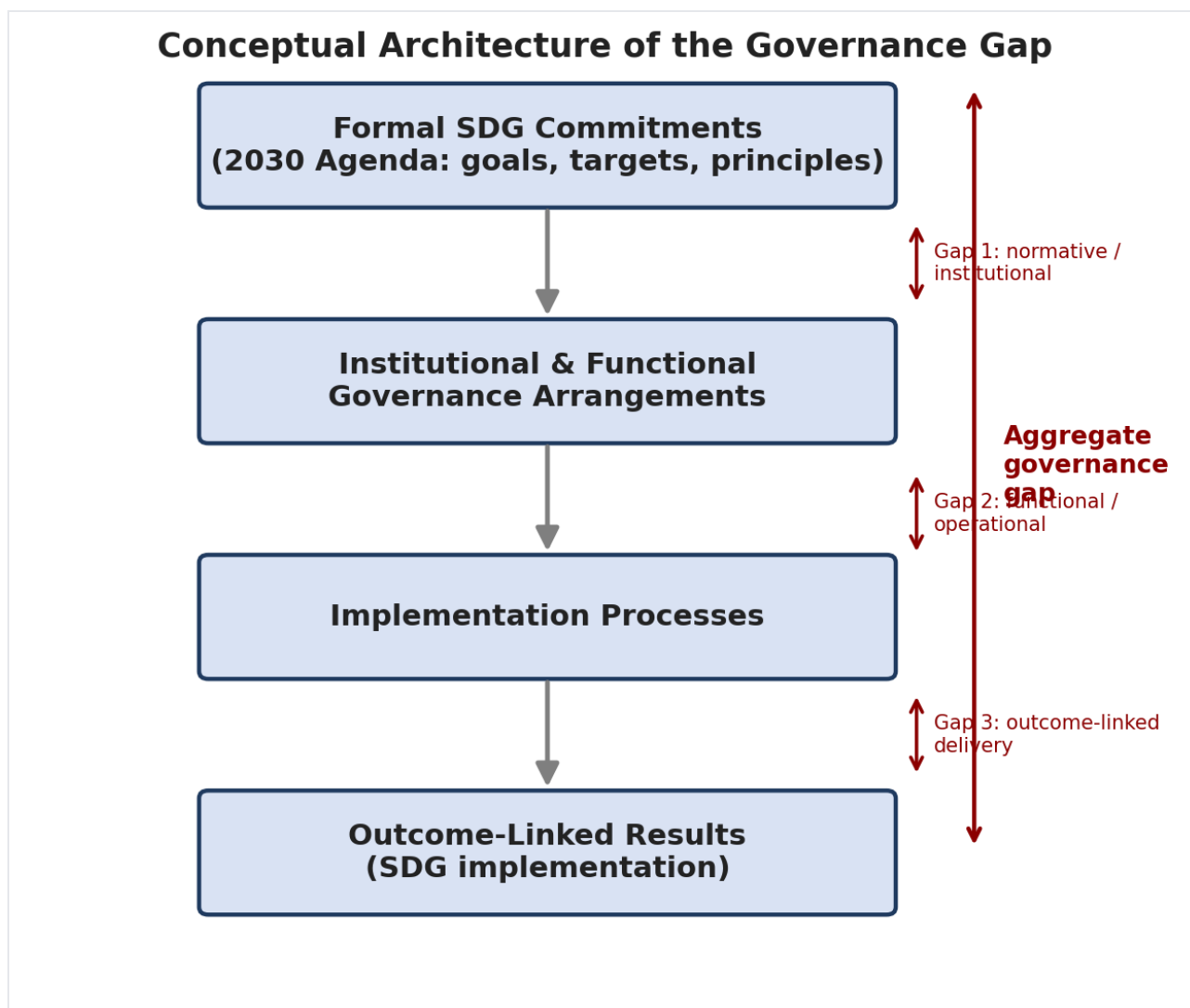


Figure 1. Conceptual architecture of the governance gap.

Source: Authors' conceptualisation. The figure is a conceptual representation and does not depict a proven causal model.

3.2 The Five Dimensions of the Governance Gap

3.2.1 Normative Governance Gap

The normative governance gap captures the extent to which the foundational principles of the 2030 Agenda—universality, integration, indivisibility, and the pledge to leave no one behind—have been internalised into national policy discourse, legal frameworks, and political culture [1]. A wide normative gap exists when the SDGs are treated as a peripheral concern grafted onto existing development plans without substantive reorientation of policy priorities, or when the integrated nature of the Goals is ignored in favour of selective adoption of politically convenient targets. The normative dimension is not merely rhetorical; it has practical consequences for how institutional resources are allocated and how policy trade-offs are negotiated. Research on integrated policymaking shows that countries adopt different institutional designs for coordinating SDG implementation, and that these designs are shaped by contextual factors including political regime type, administrative capacity, and exposure to international norms [17]. The normative gap is, in this sense, foundational: where the normative vision of the 2030 Agenda has not been internalised, investment in

the institutional, operational, data, or participatory infrastructure required for effective implementation is less likely to follow.

3.2.2 Institutional Governance Gap

The institutional governance gap refers to the absence, weakness, or inadequate design of the organisational structures and coordination mechanisms required for whole-of-government SDG delivery. The 2030 Agenda explicitly calls for integrated, cross-sectoral implementation, recognising that the Goals are interrelated and that progress in one area often depends on progress in others [1]. This integration requirement poses a fundamental challenge to the sectorally organised structures of most governments, where each ministry is responsible for a discrete portfolio and has limited incentives or mechanisms for coordination with others. The institutional gap manifests in several forms: the absence of a designated SDG coordination body; the existence of coordination bodies that lack political authority or budgetary control; the failure to align subnational governance structures with national SDG strategies; and the lack of institutional links between SDG implementation and national planning, budgeting, and audit processes.

Available evidence indicates considerable variation in how countries have addressed this challenge. Comparative work on integrated policymaking has developed typologies of national SDG coordination organised around political leadership, administrative coordination, stakeholder engagement, and localisation, and has found that contextual factors such as regime type and state capacity significantly shape institutional innovation and inertia [17]. The OECD's analysis of policy-coherence mechanisms indicates that while many countries have established cross-sectoral coordination bodies, the effectiveness of these bodies varies considerably, with some functioning as genuinely influential policy integrators and others existing in name only [13]. UNDP's review of VNR submissions identifies multi-level mechanisms engaging stakeholders from national to local levels as central to effective coordination and oversight of SDG implementation [18]. The institutional governance gap is thus not a binary condition but a spectrum, with countries occupying different positions along a continuum from integrated governance architecture to institutional absence.

3.2.3 Operational Governance Gap

The operational governance gap captures the distance between formal institutional mandates and actual implementation capacity on the ground. This dimension addresses a blind spot in much of the governance literature, which tends to focus on institutional design while paying insufficient attention to whether designed institutions actually function as intended. The operational gap encompasses several interrelated elements: the availability and adequacy of human resources dedicated to SDG coordination and implementation; the existence and enforcement of inter-ministerial working protocols; the integration of SDG considerations into budgetary allocation; the alignment of regulatory frameworks with SDG targets; and the capacity of subnational governments to translate national strategies into local action. The COVID-19 pandemic brought questions of public-sector readiness, responsiveness, and resilience to the fore, with uneven effects across countries, and reinforced the case for treating institutional transformation as a central component of SDG delivery [3, 13]. The operational gap is, in many respects, the dimension most directly actionable by governments, as it concerns the allocation of resources, the enforcement of protocols, and the strengthening of administrative systems rather than the more abstract normative or structural questions addressed by other dimensions.

3.2.4 Data and Accountability Governance Gap

The data and accountability governance gap refers to the inadequacy of data systems, monitoring frameworks, and accountability mechanisms for tracking SDG progress and ensuring that implementation shortfalls are identified and addressed. This dimension is critical because the 2030 Agenda's follow-up and review architecture depends on the availability of reliable, timely, and disaggregated data. The global

indicator framework comprises 231 indicators; although the formal tier classification that distinguished indicators without an established methodology was phased out around 2020, a substantial proportion of indicators continued to lack regularly produced data for a majority of countries as of 2023, constraining comprehensive progress assessment [20, 22]. The Global Progress Report on SDG 16 indicators, released in July 2023, underscores that progress is worryingly slow and that limited information availability hampers evidence-informed policy across Goal 16 and, by extension, the broader Agenda [15].

The accountability dimension of this gap is equally important. The VNR process, while valuable as a voluntary peer-learning mechanism, has been widely characterised as limited in its accountability force: countries select which Goals to highlight, may present an unduly favourable picture of progress, and face no formal consequences for inadequate performance [9]. The VNR Quality Index developed by CEPEI provides a methodology for qualitatively assessing VNR report quality against principles of transparency, inclusiveness, and evidence-based analysis, but it is a civil society instrument without formal standing in the UN system [21]. PARIS21, as a custodian agency for three statistical capacity-related indicators of SDG 17, plays a role in strengthening national statistical offices and the data-to-policy alignment that VNRs require [22]. The data and accountability gap thus represents both a technical challenge—stronger statistical systems—and a political challenge—more robust accountability mechanisms that complement the voluntary VNR process.

3.2.5 Participatory Governance Gap

The participatory governance gap measures the extent to which civil society, the private sector, local communities, and other non-state actors are meaningfully involved in SDG planning, implementation, monitoring, and review. The 2030 Agenda is explicit in its call for multi-stakeholder engagement, stating that all countries and all stakeholders, acting in collaborative partnership, will implement the plan [1]. Despite this rhetorical commitment, the reality of stakeholder engagement varies enormously across countries and often falls short of the transformative partnership envisioned by the Agenda. In many contexts, engagement is limited to consultation exercises that lack genuine influence over policy decisions, while in others civic space is constrained by legal, regulatory, or political barriers.

The participatory dimension is particularly relevant to the governance gap framework because it serves both as an independent indicator of governance quality and as an enabler of progress in other dimensions. Meaningful stakeholder engagement can strengthen the normative dimension by broadening societal ownership of the SDGs, improve the institutional dimension by providing external accountability pressure on coordination bodies, reduce the operational gap by mobilising non-state resources and expertise, and narrow the data gap by contributing alternative data sources and independent monitoring. Guidance for civil society identifies SDG coordination bodies as critical entry points for engagement and emphasises the importance of identifying and engaging with multiple ministries and departments given the breadth of SDG 16+ targets [10]. The participatory governance gap thus has a multiplier character: where civil society is effectively engaged, it can contribute to narrowing gaps across other dimensions simultaneously. Whether this potential is realised in practice is an empirical question that the framework is designed to help investigate.

4. Operationalising the Framework: The Governance Gap Assessment Matrix

4.1 Matrix Architecture

The Governance Gap Assessment Matrix is structured as a five-by-three analytical grid. The rows correspond to the five governance-gap dimensions: normative, institutional, operational, data and accountability, and participatory. The columns represent three assessment levels: the formal level, which examines the existence

of laws, policies, and institutional arrangements; the functional level, which assesses whether these formal arrangements operate in practice; and the outcome-linked level, which examines whether functioning arrangements are associated with improvements in implementation results. This three-level architecture draws on the well-established distinction in comparative institutional analysis between *de jure* and *de facto* governance, and extends it to include an outcome-linked dimension that connects governance processes to implementation results. The label 'outcome-linked' is used deliberately in place of 'impact' to avoid implying causal attribution: outcome-linked indicators are used to examine the relationship between governance arrangements and implementation results, but they do not by themselves establish that governance arrangements cause those results, since outcomes are also shaped by financing, economic conditions, conflict, climate shocks, demographics, global crises, technical capacity, and other structural variables (see Section 6.3).

The matrix is designed to be populated using a combination of qualitative and documentary evidence and, where available, quantitative data. Formal-level assessments rely primarily on documentary analysis of national strategies, legal frameworks, and institutional mandates. Functional-level assessments draw on VNR reports, OECD PCSD evaluations, civil society shadow reports, and institutional assessments. Outcome-linked assessments utilise SDG indicator data and contextual analysis where available. The scoring methodology, described below, is designed to produce both dimension-specific scores and an aggregate measure that enables comparison across dimensions and, with appropriate caution, across countries and over time.

Figure 2 sets out the GGAM assessment and diagnostic workflow. The assessment phase moves from evidence collection, through cell-level assessment of the fifteen cells, to dimension-level and aggregate readiness scores and the derived governance gap score. The diagnostic and response phase then interprets the dimensional gap profile to inform a targeted institutional response. The figure is explicit that assessment does not by itself establish causal attribution of outcomes to governance; the transition from assessment to response is an interpretive step that should be underpinned by contextual analysis.

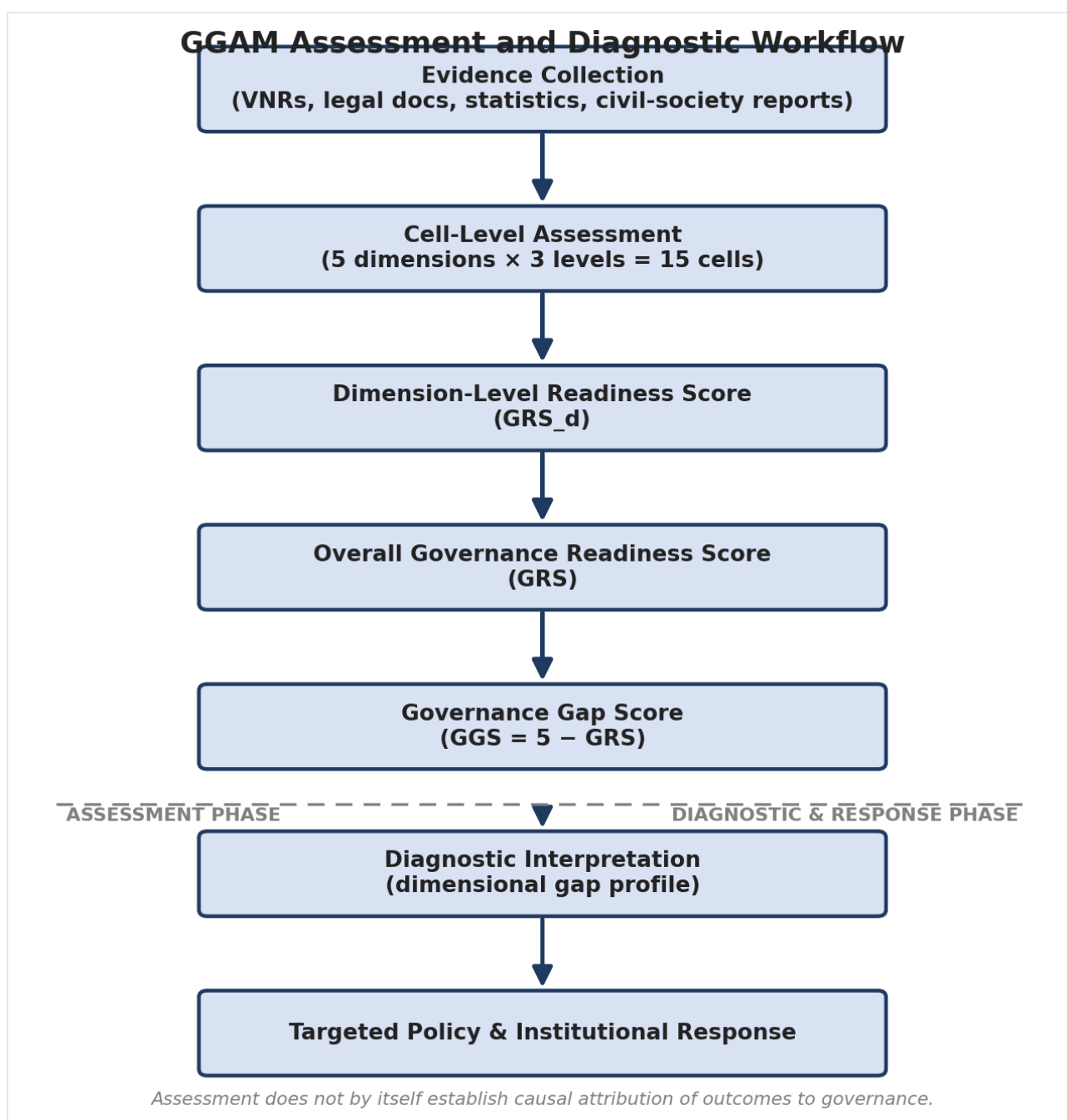


Figure 2. GGAM assessment and diagnostic workflow.

Source: Authors' conceptualisation. The workflow is a proposed application pathway; it does not imply that scoring automatically produces causal policy conclusions.

4.2 Indicators and Measurement Approach

Table 1 presents the indicator framework for the GGAM. For each dimension and assessment level, specific indicators are proposed that can be assessed using publicly available evidence. The indicators are deliberately designed to be actionable: they point to governance attributes that governments can realistically address through policy and institutional reform, rather than to structural factors such as regime type or income level that lie beyond the immediate scope of governance reform. The selection of indicators draws on existing measurement frameworks, including the WGI, the SDG 16 indicators, the OECD PCSD framework, and the VNR Quality Index, but adapts and extends these to the specific requirements of SDG governance analysis.

The full measurement and scoring protocol, including operational definitions, recommended evidence sources, and provisional score anchors for each of the fifteen cells, is set out in Appendix A.

Table 1. Governance Gap Assessment Matrix (GGAM): Indicator Framework

Dimension	Formal Level	Functional Level	Outcome-Linked Level
Normative	References to the 2030 Agenda and its principles in national strategy and law; legal alignment of SDG commitments.	Integration of SDGs in political discourse and in budgeting; cross-Government adoption beyond lead ministry.	Policy priorities and resource allocation reflect an integrated SDG approach rather than selective cherry-picking.
Institutional	Existence and mandate clarity of an SDG coordination body; subnational alignment provisions.	Cross-ministerial protocols enforced in practice; functioning subnational coordination mechanisms.	Coordination effectiveness visible in joint planning, budgeting, and review processes.
Operational	Staffing and budget for SDG coordination; regulatory review against SDG targets.	Implementation protocols followed; budget execution for SDG-related programmes.	Improvements in SDG-relevant service delivery and programme outputs.
Data & Accountability	National SDG indicator framework adopted; statistical capacity provisions.	Data timeliness and disaggregation; quality of VNR reporting.	Evidence-based policy adjustments made; independent review and scrutiny occur.
Participatory	Legal framework for stakeholder engagement; VNR consultation provisions.	Civil society involvement in monitoring; access to coordination platforms.	Stakeholder influence on policy decisions; quality of shadow reporting and independent monitoring.

Note. This table presents the indicator framework at a summary level. The detailed measurement and scoring protocol for each of the fifteen cells—operational definitions, recommended evidence sources, and provisional Score 1–5 anchors—is provided in Appendix A.

4.3 Scoring Methodology

Each cell in the GGAM is scored on a five-point scale designed to be as transparent and replicable as possible, while acknowledging that some interpretive judgement is inevitable, particularly at the functional and outcome-linked levels. The scale is defined in terms of governance readiness: a score of 1 indicates that the governance attribute is entirely absent or non-functional; 2 indicates that formal elements exist but are largely non-functional; 3 indicates partial functionality with significant weaknesses; 4 indicates substantial functionality with minor gaps; and 5 indicates that the attribute is fully present and functioning effectively. Detailed, cell-specific descriptors for each anchor are provided in Appendix A. The scoring anchors are presented as proposed criteria for evaluation, not as empirically validated thresholds; where exact

quantitative thresholds cannot yet be justified, the criteria are provisional and require calibration against country evidence. This distinction between proposed anchors and validated thresholds is maintained throughout the framework and is a key limitation discussed in Section 6.2.

4.4 Governance Readiness Score (GRS) and Governance Gap Score (GGS)

To avoid the conceptual inconsistency that arises when a measure labelled a 'gap' increases as governance improves, the paper defines two related but distinct measures. The Governance Readiness Score (GRS) is the primary scoring measure: higher values indicate stronger governance readiness and smaller governance gaps. The Governance Gap Score (GGS) is derived transparently from the readiness scale and is expressed so that higher values indicate a larger remaining governance gap. Keeping both measures explicit allows the framework to speak the language of 'gaps' for diagnostic communication while preserving a readiness scale that is internally consistent with the scoring direction.

Formally, let s denote the cell-level readiness score for dimension d ($d = 1, \dots, 5$) at assessment level l ($l \in \{\text{formal, functional, outcome-linked}\}$), with $s \in \{1, 2, 3, 4, 5\}$. The dimension-level Governance Readiness Score is the simple average of its three cell scores:

$$GRS_d = (1/3) \sum_l s_{d,l} \quad (1)$$

where GRS_d is the readiness score for dimension d ; the summation runs over the three assessment levels l (formal, functional, outcome-linked); and $s_{\{d,l\}}$ is the cell-level readiness score. GRS_d ranges from 1 (critical gap) to 5 (no significant gap).

The aggregate Governance Readiness Score is the equally weighted average of the five dimension scores, which is equivalent to the simple average of all fifteen cell scores under equal weighting:

$$GRS = (1/5) \sum_{d=1}^5 GRS_d = (1/15) \sum_{d=1}^5 \sum_l s_{d,l} \quad (2)$$

where GRS is the overall Governance Readiness Score and $GRS \in [1, 5]$. The equality of the two expressions holds because equal weighting is applied across both dimensions and assessment levels (see Section 4.5).

The corresponding Governance Gap Score is derived by a simple linear reflection of the readiness scale, so that it ranges from 0 (no gap) to 4 (maximum gap). The dimension-level and aggregate forms are:

$$GGS = 5 - GRS, GGS \in [0, 4] \quad (3)$$

$$GGS_d = 5 - GRS_d, GGS_d \in [0, 4] \quad (4)$$

where GGS is the aggregate Governance Gap Score and GGS_d is the dimension-level gap score. Higher GGS values denote larger remaining governance gaps. The two measures (GRS and GGS) carry identical information and are presented together so that readiness reporting and gap-framed diagnosis remain mutually consistent.

Under this transformation, a cell, dimension, or country scoring 5 on the readiness scale corresponds to a gap score of 0 (no significant gap), a score of 3 corresponds to a gap score of 2 (partial functionality with significant weaknesses), and a score of 1 corresponds to a gap score of 4 (the attribute entirely absent or non-functional). Figure 3 displays this inverse relationship compactly. Dimension scores can be reported alongside the aggregate to preserve the diagnostic granularity that is the framework's principal value.

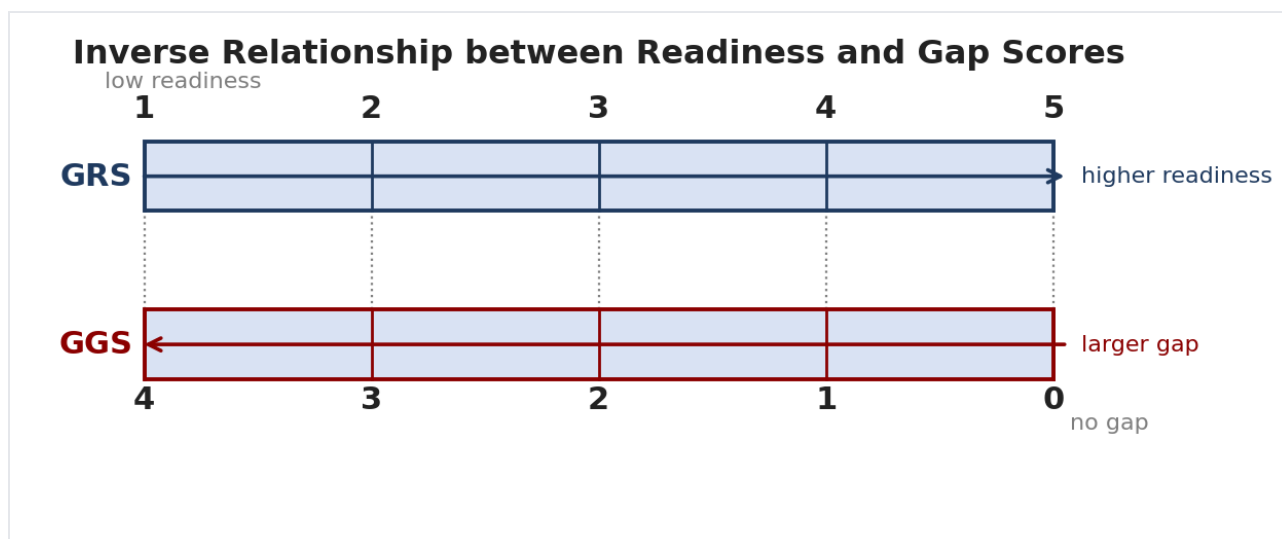


Figure 3. Inverse relationship between the Governance Readiness Score (GRS) and the Governance Gap Score (GGS).

Source: Authors' conceptualisation. $GRS \in [1, 5]$; $GGS = 5 - GRS \in [0, 4]$. The figure illustrates the scoring transformation and does not depict empirical data.

4.5 Weighting

The baseline specification applies equal weighting across the five dimensions. This choice is a transparent baseline assumption rather than a claim that the dimensions are of equal real-world importance. Equal weighting is adopted here for three reasons: it is transparent and easy to reproduce; it avoids imposing contested value judgements at the framework's founding stage; and it is appropriate for an initial framework whose empirical performance has not yet been validated. The paper does not assert that the operational dimension, for example, is in fact as consequential as the normative dimension; it asserts only that, without empirical grounds to weight them differently, treating them equally is the most defensible starting point.

Several future approaches to weighting are compatible with the framework and should be explored as the evidence base develops. These include expert elicitation, in which specialists assign weights based on structured judgement; stakeholder consultation, in which weights reflect the priorities of affected actors; empirical validation, in which weights are estimated from the observed association between dimension scores and implementation outcomes; and sensitivity analysis, in which results are reported across alternative weighting schemes to test robustness. Until such evidence is available, equal weighting is retained, and users are encouraged to treat aggregate scores as indicative and to examine dimension-level profiles wherever decisions depend on them.

5. Illustrative Application: Evidence from the VNR Process

5.1 Methodological Approach

To illustrate the analytical utility of the GGAM, this section presents a qualitative, indicative application of the framework to evidence from the VNR process and related literature. The VNRs, submitted by countries to the High-Level Political Forum on Sustainable Development (HLPF), represent the most comprehensive publicly available source of information on national SDG governance arrangements. By the 2023 HLPF, 188 countries and the European Union had carried out at least one Voluntary National Review, with the 2023 HLPF featuring thirty-eight countries and the European Union presenting Voluntary National Reviews (39 presenters in total),

with several presenting second- or third-cycle reviews [9, 16, 23]. The synthesis presented here draws on VNR synthesis reports, the UN DESA VNR database, the CEPEI VNR Quality Index methodology, the UN DESA Handbook for the preparation of VNRs, and secondary analyses published by UNDP, OECD, and civil society organisations [13, 16, 18, 21].

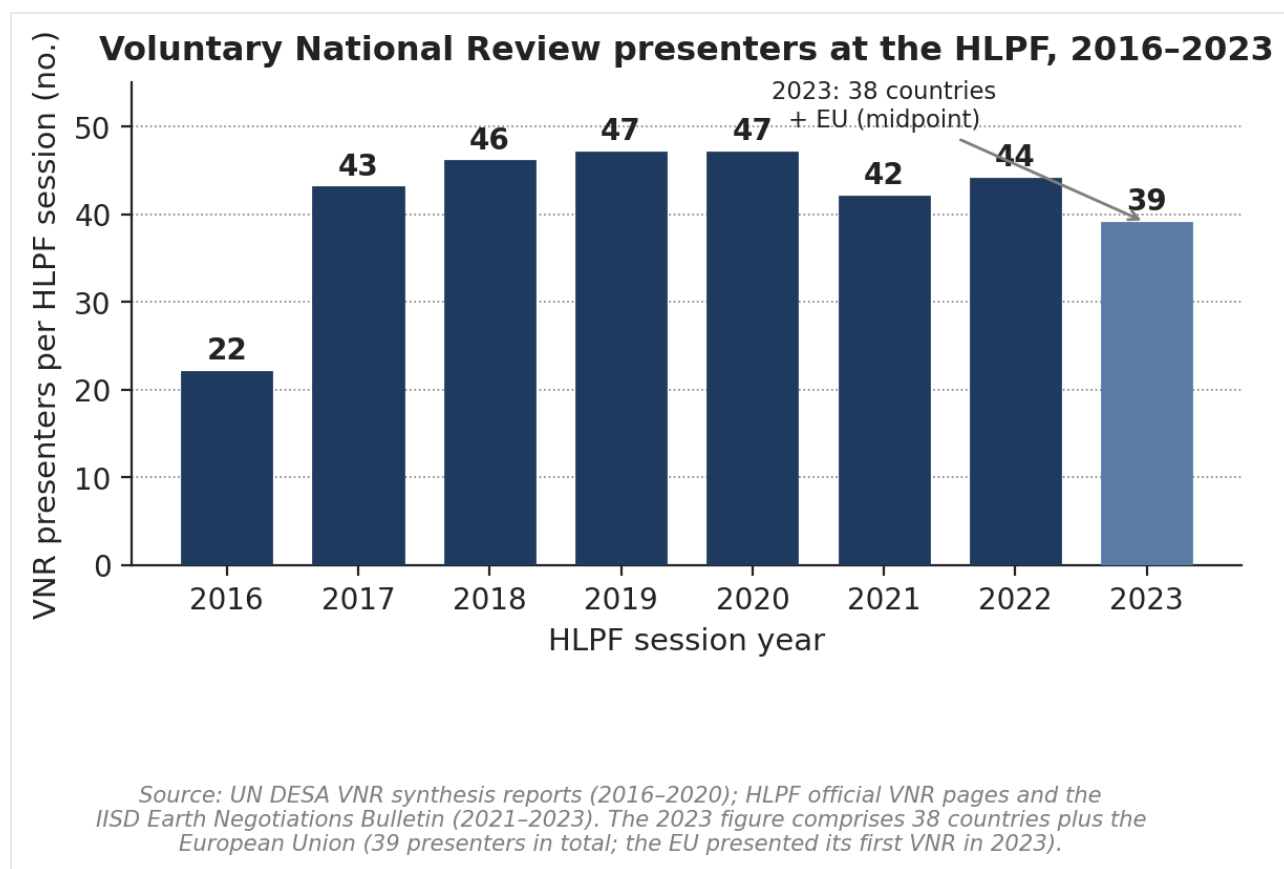


Figure 4. Voluntary National Review presenters at the HLPF, 2016–2023.

Source: Author compilation from the verified source record reproduced in Appendix C, based on authoritative HLPF and UN DESA documentation (UN DESA VNR synthesis reports, 2016–2020; HLPF official VNR pages and the IISD Earth Negotiations Bulletin, 2021–2023). The 2023 figure comprises 38 countries plus the European Union (39 presenters in total; the EU presented its first VNR in 2023). Plotted values are the number of VNR presenters per HLPF session, not unique participating countries (several countries presented more than once over the period).

The year-by-year source data and verification record underlying Figure 4 are provided in Appendix C to support transparency and reproducibility.

It is important to be precise about the status of this section. This is an illustrative synthesis of evidence, not a rigorous systematic application of the GGAM. No structured coding of VNR texts was undertaken, no inter-coder reliability was established, and no country-level scores were assigned. A full quantitative application, using the five-point scoring methodology set out in Section 4 and the detailed protocol in Appendix A, would require a dedicated research project with systematic coding of VNR texts against the cell-specific criteria, and is envisaged as a subsequent phase of this research programme. Accordingly, the observations below are framed as recurring governance challenges that the available evidence suggests, and as patterns that warrant systematic investigation rather than as established findings.

5.2 Recurring Governance Challenges in VNR-Related Evidence

Several broad observations recur across VNR-related evidence. First, there has been progressive evolution in the institutionalisation, alignment, and evaluation of the SDGs within national frameworks over successive review cycles. Earlier VNRs tended to focus on initial alignment and prioritisation of targets, while later reviews show greater evidence of multi-stakeholder coordination, data integration, and policy-coherence assessment [13, 16]. This evolution is consistent with the view that the VNR process itself can function as a governance-strengthening mechanism, providing a structured incentive for countries to develop more robust institutional arrangements over time.

Second, the available evidence suggests that the operational and data-and-accountability dimensions are where implementation difficulties most often recur across diverse country contexts. Many countries report that formal coordination arrangements exist but that translating these into effective cross-ministerial action is a persistent challenge, and that data availability—particularly disaggregated data—remains uneven [13, 15, 18]. The institutional dimension shows wide variation: most countries report some form of coordination body, but the authority and effectiveness of these bodies vary considerably. The normative dimension appears to have strengthened over successive cycles, with the SDGs increasingly reflected in national strategies, though full integration remains uneven. The participatory dimension varies most widely across countries, reflecting differences in civic space, civil society capacity, and legal frameworks for engagement.

Third, the evidence indicates that countries that have invested in formal institutional design have not necessarily seen commensurate improvements in implementation outcomes, suggesting that the relationship between governance architecture and governance effectiveness is more contingent than is often assumed. This observation is consistent with the framework's distinction between the formal, functional, and outcome-linked levels: a country can score well at the formal level while scoring substantially lower at the functional and outcome-linked levels. Table 3 summarises these recurring challenges in illustrative terms. The table is deliberately qualitative and does not assign country-level scores; its purpose is to show how the GGAM dimensions map onto patterns in the existing evidence, not to establish prevalence.

Table 3. Illustrative Governance Challenges Identified in VNR-Related Evidence

GGAM Dimension	Illustrative pattern in the evidence	Indicative prominence across reports reviewed
Normative	Progressive integration of SDG language into national strategies over successive cycles, with uneven internalisation of integration and leave-no-one-behind principles.	Moderate
Institutional	Wide variation: most countries report coordination bodies, but their authority, resourcing, and effectiveness differ markedly.	Moderate to wide
Operational	Recurring difficulties translating coordination mandates into cross-ministerial action, adequate staffing, and budget-aligned implementation.	Wide (recurring theme)
Data & Accountability	Uneven data availability and disaggregation; variable VNR reporting quality; limited formal accountability beyond voluntary review.	Wide (recurring theme)
Participatory	Wide variation by country context; engagement often consultative rather than substantive; civic-space constraints in some settings.	Moderate to wide (highly variable)

Note. The patterns are qualitative and illustrative, drawn from a non-systematic synthesis of VNR-related literature; they are not prevalence estimates and should not be read as statistically representative findings. The third column summarises the relative salience of these themes in the reports reviewed, not their measured frequency.

5.3 Illustrative Findings and Their Limits

The most salient observation from this illustrative synthesis is the recurring prominence of operational and data-and-accountability challenges across diverse country contexts. Even countries that appear to perform relatively well on normative and institutional indicators—having established high-level political commitment and formal coordination structures—frequently report difficulties in translating these into effective implementation and robust monitoring. This pattern suggests that governance shortfalls are not simply a matter of political will or institutional design but reflect deeper, recurring challenges in the relationship between governance architecture and implementation functionality. The implication for reform is that strengthening governance requires attention not only to establishing the right institutions but also to ensuring that those institutions are adequately resourced, empowered, and embedded within systems of public administration and accountability.

These observations are offered as illustrative findings, not as established empirical results. They are consistent with the broader literature on SDG governance challenges, including UN DESA guidance on VNR preparation and the Sustainable Development Report 2023, which both note that despite some progress in institutional setup, the operationalisation of integrated governance remains a critical bottleneck [3, 16]. Whether these patterns hold systematically across the full population of VNRs, and how they vary by region, income group,

and governance context, are questions that the framework is designed to help answer through subsequent, systematic empirical work.

6. Counterarguments and Limitations

6.1 Governance versus Capacity

A potential critique of the governance gap framework is that it may conflate governance with administrative or technical capacity. If governance gaps were simply a manifestation of broader capacity deficits, the framework would offer little analytical value beyond existing capacity assessment tools. This critique has merit and warrants careful engagement. To address it, the framework draws a deliberate distinction among several related but distinct constructs. Governance arrangements refer to the formal and informal rules, structures, and processes by which authority is exercised for SDG delivery. Institutional design refers to the architecture of those arrangements—the existence, mandate, and configuration of coordination bodies and mechanisms. Administrative capacity refers to the human resources, procedures, and organisational competence available to operate these arrangements. Technical capacity refers to specialist skills, such as statistical methodology or sectoral expertise. Financial and resource constraints refer to the budgetary and fiscal envelope within which all of the above operate.

These constructs interact but are not identical. A country may have sound governance arrangements and adequate institutional design yet lack the administrative capacity to operate them—manifesting as a functional-level gap rather than a formal-level one. Conversely, a country may possess strong technical capacity and adequate financing yet still exhibit a significant governance gap if institutional arrangements prevent that capacity from being deployed for SDG delivery. The three-level assessment architecture captures this distinction: a country that has established a well-designed coordination body but cannot staff it adequately would score relatively well at the formal level but poorly at the functional level, clearly separating governance design from the capacity constraint. The framework does not seek fully to disentangle these causal pathways, which would require analytical precision that is neither feasible with available evidence nor necessary for the framework's diagnostic purpose. It does, however, remain a governance-focused framework: capacity is treated as a contributor to governance effectiveness, a constraint on functional performance, and an explanatory factor for outcome-linked results—not as a substitute for governance. The framework is not reconstructed as a generic governance-capacity composite.

6.2 Measurement Challenges, Reliability, and Data Availability

A second limitation concerns the measurement challenges inherent in assessing governance, particularly at the functional and outcome-linked levels. Formal-level indicators, such as the existence of a coordination body or the inclusion of SDG references in a national strategy, are relatively straightforward to assess using documentary evidence. Functional and outcome-linked indicators require more interpretive judgement: assessing whether a coordination body is effective or whether stakeholder engagement is meaningful involves evaluation that may vary across evaluators, introducing a risk of inconsistency that could undermine the comparability of scores across countries or over time.

The framework mitigates this risk in several ways. Indicators are designed to be as specific and observable as possible, anchored in concrete evidence such as the frequency of inter-ministerial meetings, the number and breadth of stakeholders consulted during VNR preparation, or the timeliness of SDG data publication. The five-point scale is defined with detailed, cell-specific descriptors in Appendix A, providing a common reference frame for evaluators. The framework recommends the use of multiple evidence sources and, where feasible, multiple evaluators, with explicit reporting of assessor identity and any disagreements, to enhance

reliability. These measures cannot eliminate subjectivity but can reduce it to a level acceptable for a diagnostic framework intended to guide policy discussion rather than to produce definitive rankings. The WGI, despite similar measurement challenges and documented sensitivity to assumptions about error structure [12], has proven analytically valuable for cross-national research, suggesting that the benefits of systematisation can outweigh the costs of imprecision when limitations are openly acknowledged. A related limitation is data availability: outcome-linked assessment depends on SDG indicator data that remain uneven across countries and goals [15, 20], so outcome-linked scores will be more reliable for some dimensions and jurisdictions than for others.

6.3 Contextual Specificity, Generalisability, and Causal Inference

A third concern relates to the contextual specificity of governance arrangements and to the limits of causal inference. The framework's indicators assume a relatively standard model of governance, with a central government coordinating SDG implementation through ministerial structures, engaging with civil society through consultation mechanisms, and monitoring progress through national statistical systems. This model may not accurately reflect governance realities in all contexts. In federal systems, subnational arrangements may be more important than national-level coordination; in conflict-affected states, the formal architecture may be largely irrelevant to actual decision-making; and in small island developing states, governance challenges may be qualitatively different from those in large bureaucracies. The framework is designed to be adaptable: the five-dimensional structure is a universal analytical scaffold, but the specific indicators can and should be adapted to national context, and cross-country comparison is most meaningful among countries with similar structures and development contexts.

The most important inferential limitation concerns causality. The framework deliberately uses outcome-linked rather than impact-level indicators because SDG outcomes are jointly determined by governance and by many other factors—financing, economic conditions, conflict, climate shocks, demographic structure, global crises, and technical capacity, among others. An observed association between stronger governance arrangements and better implementation results does not establish that governance caused the results, nor that strengthening governance alone will produce comparable results elsewhere. The framework therefore supports diagnosis and hypothesis generation; it does not, by itself, support causal claims. Establishing causation requires complementary research designs—natural experiments, longitudinal analysis, and process tracing—that lie beyond the scope of a single assessment instrument. Users should treat the GRS and GGS as structured summaries of governance readiness, not as evidence of the governance contribution to particular outcomes.

7. Implications and the Path Forward

7.1 For the SDG Summit 2023 and Beyond

The SDG Summit scheduled for 18–19 September 2023 represents a pivotal moment for the 2030 Agenda. Convened under the auspices of the UN General Assembly at the level of Heads of State and Government, the Summit is expected to adopt a political declaration that reaffirms global commitment to the Goals and outlines acceleration actions; negotiations on the declaration were under way in the first half of 2023, with the declaration expected to be a consensual, action-oriented, and transformative document [24, 25]. The governance gap framework is offered as one structured contribution to that agenda. A political declaration that explicitly recognises the gap between formal commitments and implementation-ready governance arrangements—and that encourages Member States to undertake systematic governance-gap assessments—would help translate the Summit's acceleration rhetoric into a concrete diagnostic agenda. The SDG Summit

Acceleration and Accountability Platform, which registers initiatives and commitments, provides an existing infrastructure through which governance-reform commitments could be monitored [27].

The Summit could also usefully address the structural accountability deficit in the SDG follow-up and review architecture. The current reliance on voluntary, self-reported VNRs, while valuable for peer learning, is widely considered insufficient to generate the accountability pressure needed to drive governance reform at the pace required [9]. Complementary mechanisms that could be explored include independent civil society assessments, parliamentary oversight of SDG implementation, and an optional governance-gap module within the VNR guidelines that would provide a standardised framework for countries to report on their governance arrangements. None of these measures requires renegotiating the voluntary character of the VNR; each would complement it by widening the evidence base on which assessment depends.

7.2 For National Governments

For national governments, the governance gap framework provides a structured diagnostic tool that can guide reform in a systematic and prioritised manner. A first practical step is to undertake a self-assessment using the GGAM, either as a standalone exercise or integrated into VNR preparation. Such a self-assessment would identify which governance dimensions exhibit the widest gaps and where reform effort should be concentrated. The framework's three-level architecture would help governments distinguish between situations in which the problem is the absence of governance arrangements and situations in which existing arrangements fail to function effectively, thereby pointing toward different types of intervention—institutional reform in the first case, capacity and operational strengthening in the second.

A second practical step is to develop a governance-gap closure plan, with specific, time-bound targets for reducing gaps in each dimension, integrated into national SDG implementation strategies and subject to the same monitoring and reporting mechanisms as other commitments. The OECD PCSD Recommendation provides a useful template for institutionalising policy coherence, and the framework proposed here can be understood as extending the PCSD approach into a broader governance assessment that encompasses institutional design, operational functionality, data systems, and stakeholder engagement as well as policy coherence [5, 13]. Governments should also invest in the participatory dimension, recognising that civil society engagement is not only a normative commitment but a practical strategy for enhancing governance effectiveness across all other dimensions.

7.3 For Civil Society and the Accountability Ecosystem

For civil society organisations and the broader accountability ecosystem, the framework offers a common analytical language and a structured methodology for assessing governance. Civil society organisations can apply the GGAM to produce independent governance-gap assessments that complement official VNRs and provide a more critical, evidence-based account of governance strengths and weaknesses. The VNR Quality Index developed by CEPEI provides an existing model for civil-society-led assessment, and the GGAM extends this approach to a broader set of governance dimensions [21].

The framework can also inform advocacy by identifying the specific dimensions where gaps are widest and where reform is most urgent, allowing organisations to move from general calls for governance reform to targeted, evidence-based demands—such as the establishment of parliamentary oversight committees for SDG implementation, the enactment of legal frameworks for stakeholder participation, or the allocation of dedicated budgetary resources for SDG data systems. Finally, the framework can facilitate cross-country learning by enabling organisations in different countries to compare governance-gap profiles and share reform experience. The intended trajectory of this work is that the GGAM be subjected to independent

empirical validation and open methodological scrutiny, so that any subsequent applied tools build on a publicly testable foundation.

8. Conclusion

This paper has argued that the gap between formal SDG commitments and the governance arrangements required to deliver them is a distinct, analytically tractable, and operationally measurable condition that deserves systematic attention alongside the financing, technical, and geopolitical deficits that dominate current discourse. It has defined the governance gap as the measurable distance between the institutional, procedural, and normative requirements for effective SDG implementation and the actual state of governance arrangements and practices in a given jurisdiction, and it has developed the Governance Gap Assessment Matrix as a structured framework comprising five dimensions—normative, institutional, operational, data and accountability, and participatory—assessed across three levels: formal, functional, and outcome-linked. The framework is accompanied by a provisional measurement and scoring protocol, two derived measures (the Governance Readiness Score and the Governance Gap Score), and an explicit equal-weighting baseline.

What the paper establishes is a conceptual and methodological foundation: a defensible definition, a comparative positioning against existing instruments, a transparent scoring logic, and a detailed measurement protocol suitable for subsequent application. What remains to be empirically tested is whether the framework, applied systematically to a coded corpus of VNRs and supporting evidence, produces reliable, comparable, and policy-relevant assessments; whether the proposed scoring anchors can be calibrated against country evidence; whether alternative weighting schemes materially change the conclusions; and whether the outcome-linked level can be connected to implementation results without overstating causation. The illustrative synthesis presented in Section 5 suggests that operational and data-and-accountability challenges recur across diverse contexts, but this observation is offered as a hypothesis for systematic investigation, not as an established finding.

The practical relevance of the GGAM is that it gives governments, civil society, and international monitoring bodies a common, transparent scaffold for diagnosing where governance arrangements fall short of implementation requirements and for tracking whether reforms narrow those shortfalls over time. It does not promise that closing governance gaps is sufficient for SDG achievement, nor that governance is the sole determinant of implementation success. It offers the more modest but, it is hoped, more durable claim that governance is a cross-cutting enabling condition that can substantially influence whether financial, technical, institutional, and social resources are translated into effective SDG implementation, and that this condition can and should be assessed with the same seriousness accorded to financing and other deficits. The next steps are empirical: a systematic, coded application of the protocol to VNR evidence; calibration of the scoring anchors; testing of weighting and reliability; and open publication of the data and code so that the framework's performance can be independently assessed. The paper is offered as the conceptual foundation for that work.

References

- [1] United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development (A/RES/70/1). United Nations. <https://sdgs.un.org/2030agenda>
- [2] United Nations. (2023). The Sustainable Development Goals Report 2023: Special Edition. United Nations Department of Economic and Social Affairs. <https://unstats.un.org/sdgs/report/2023>

- [3] Sachs, J. D., Lafortune, G., Fuller, G., & Drumm, E. (2023). Sustainable Development Report 2023: Implementing the SDG Stimulus. Dublin University Press.
- [4] Weiss, T. G., & Thakur, R. (2010). Global Governance and the UN: An Unfinished Journey. Indiana University Press.
- [5] OECD. (2019). Recommendation of the Council on Policy Coherence for Sustainable Development (OECD/LEGAL/0381). OECD Publishing. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0381>
- [6] Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). The Worldwide Governance Indicators: Methodology and Analytical Issues (Policy Research Working Paper No. 5430). World Bank.
- [7] United Nations. (2015). Sustainable Development Goal 16: Peace, Justice and Strong Institutions. <https://www.un.org/sustainabledevelopment/peace-justice>
- [8] Weiss, T. G., & Wilkinson, R. (2019). Rethinking Global Governance. Polity Press.
- [9] Persson, Å., & Ringius, L. (2016). Follow-up and review of the SDGs: The politics of assessing progress. In N. Choukroune (Ed.), Sustainable Development Goals and Human Rights (pp. 231–247). Springer.
- [10] TAP Network. (2021). SDG16+ Civil Society Toolkit and Resource Guide: Identifying and Engaging with SDG Coordination Bodies. <https://www.sdg16toolkit.org>
- [11] Centre for International Research on Sustainability. (2023). Global governance gaps. <https://www.ciris.info/learningcenter/global-governance-gaps>
- [12] Magnusson, L. M., & Tarverdi, Y. (2020). Measuring governance: Why do errors matter? World Development, 136, 105061. <https://doi.org/10.1016/j.worlddev.2020.105061>
- [13] OECD. (2023). Driving Policy Coherence for Sustainable Development: Accelerating Progress on the SDGs. OECD Publishing.
- [14] SDG16 Data Initiative. (n.d.). About the SDG 16 Data Initiative [Consortium coordinated by International IDEA]. International IDEA.
- [15] United Nations Office on Drugs and Crime (UNODC), United Nations Development Programme (UNDP), & Office of the United Nations High Commissioner for Human Rights (OHCHR). (2023). Global Progress Report on Sustainable Development Goal 16 Indicators 2023. UNODC.
- [16] United Nations Department of Economic and Social Affairs (UN DESA). (2023). Handbook for the Preparation of Voluntary National Reviews (2023 Edition). United Nations. <https://hlpf.un.org>
- [17] Breuer, A., Leininger, J., Malerba, D., & Tosun, J. (2023). Integrated policymaking: Institutional designs for implementing the sustainable development goals (SDGs). World Development, 170, 106317. <https://doi.org/10.1016/j.worlddev.2023.106317>
- [18] UNDP. (2022). Institutional mechanisms and thematic best practices from the Voluntary National Reviews of various countries. UNDP SDG Knowledge Hub.
- [19] United Nations Statistics Division. (2022). Practical Guide for Country-Led Evidence-Based Voluntary National Reviews. United Nations.
- [20] United Nations Statistics Division. (2023). Global indicator framework for the Sustainable Development Goals and targets (updated). <https://unstats.un.org/sdgs/indicators/indicators-list>

- [21] CEPEI. (2023). SDG VNR Quality Tracker: VNR Quality Index methodology. CEPEI. <https://cepei.org/en/sdg-vnr-quality-tracker>
- [22] PARIS21. (2023). Voluntary National Reviews: When Data Meets Policy. OECD/PARIS21. <https://www.paris21.org>
- [23] United Nations Department of Economic and Social Affairs (UN DESA). (2023). Voluntary National Reviews database. <https://hlpf.un.org/countries>
- [24] United Nations. (2023). The 2023 SDG Summit [Official webpage]. <https://sdgs.un.org/SDGSummit>
- [25] International Institute for Sustainable Development (IISD). (2023, March 23). Countries exchange views on elements to include in SDG Summit declaration. SDG Knowledge Hub. <https://sdg.iisd.org>
- [26] Malito, D. V., Umbach, G., Savoia, A., & Hulme, D. (2021). Measuring governance to achieve sustainable development: Promises and challenges. In W. Leal Filho et al. (Eds.), *Encyclopedia of the UN Sustainable Development Goals*. Springer. https://doi.org/10.1007/978-3-319-95960-3_32
- [27] United Nations. (2023). SDG Summit Acceleration and Accountability Platform. <https://sdgs.un.org/SDGSummitActions>

Appendix A. Detailed GGAM Measurement and Scoring Protocol

This appendix sets out the detailed measurement and scoring protocol for each of the fifteen cells of the GGAM (five dimensions × three levels). For each cell, it specifies the indicator, an operational definition, a recommended evidence source, the unit or basis of assessment, and proposed anchors for Scores 1–5, together with important limitations. The anchors are proposed criteria for evaluation; they are not empirically validated thresholds. Where the criteria invoke a quantitative benchmark, that benchmark is provisional and intended to be calibrated against country evidence in subsequent empirical work. The anchors are designed so that two competent evaluators, working independently from the same evidence base, could apply them with reasonable consistency; where residual judgement is unavoidable, evaluators are asked to record the evidence on which their score rests. Scoring direction follows Section 4: higher readiness scores (5) indicate stronger governance readiness and smaller gaps; the corresponding Governance Gap Score is $GGS = 5 - \text{readiness score}$.

Cell 1. Normative × Formal

Indicator	Existence and legal force of 2030 Agenda references in national strategy and law.
Operational definition	Whether a national development plan or equivalent strategy explicitly references the 2030 Agenda and its guiding principles (universality, integration, indivisibility, leave no one behind), and whether SDG commitments are reflected in binding legal or regulatory instruments.
Recommended evidence source	National development strategy; primary legislation and regulations; VNR formal-section text.
Unit / basis of assessment	Country (de jure documents).
Score 1 (critical gap)	No national strategy references the 2030 Agenda; SDGs absent from legal/regulatory frameworks.
Score 2 (formal only)	2030 Agenda referenced in a strategy document but not in binding instruments; references generic and non-operational.
Score 3 (partial)	2030 Agenda referenced in the principal national strategy and at least one binding instrument; references cover several Goals but remain partial.
Score 4 (substantial)	2030 Agenda embedded across the principal strategy and multiple binding instruments; integration and leave-no-one-behind principles explicitly operationalised.
Score 5 (no significant gap)	2030 Agenda fully integrated into national strategy and binding instruments; principles operationalised with clear legal force across Government.
Important limitations	Documentary existence does not guarantee political salience; legal force varies by legal system; provisions may be aspirational.

Cell 2. Normative × Functional

Indicator	Integration of SDGs in political discourse and in budgeting beyond the lead ministry.
Operational definition	Whether SDG language and principles appear in senior political discourse and in budget documents across ministries, not only in the lead ministry's outputs.
Recommended evidence source	Budget statements, ministerial policy speeches, sector plans, VNR substantive sections.
Unit / basis of assessment	Country (de facto practice).
Score 1 (critical gap)	SDGs absent from political discourse and from sectoral budgeting.
Score 2 (formal only)	SDGs appear in lead-ministry communications but not in cross-Government discourse or budgeting.
Score 3 (partial)	SDGs appear in some sectoral budgets and ministerial discourse; integration uneven across portfolios.
Score 4 (substantial)	SDGs consistently reflected in political discourse and in most sectoral budgets; trade-offs acknowledged.
Score 5 (no significant gap)	SDGs fully internalised across political discourse and budgeting; integrated priorities visible across all portfolios.
Important limitations	Discourse indicators are interpretive; budget alignment is hard to verify from external documents.

Cell 3. Normative × Outcome-linked

Indicator	Policy priorities and resource allocation reflecting an integrated SDG approach.
Operational definition	Whether observable policy priorities and allocations reflect integration and leave-no-one-behind rather than selective cherry-picking of targets.
Recommended evidence source	Budget allocations by Goal; policy- coherence assessments; SDG leave-no-one-behind indicators.
Unit / basis of assessment	Country (implementation results).
Score 1 (critical gap)	No evidence of integrated prioritisation; resources allocated without reference to SDG integration.
Score 2 (formal only)	Limited evidence; a few integrated initiatives but allocations largely SDG-blind.
Score 3 (partial)	Some evidence of integrated prioritisation in priority sectors; leave-no-one-behind partially addressed.
Score 4 (substantial)	Substantial evidence across sectors; allocations increasingly aligned with integration and leave-no-one-behind.
Score 5 (no significant gap)	Strong, systematic alignment of priorities and allocations with integrated, leave-no-one-behind SDG approach.
Important limitations	Outcome-linked only; cannot establish causation; allocations are influenced by many factors beyond governance.

Cell 4. Institutional × Formal

Indicator	Existence and mandate clarity of an SDG coordination body and subnational alignment provisions.
Operational definition	Whether a designated cross-Government SDG coordination body exists with a clear mandate, and whether subnational alignment is formally provided for.
Recommended evidence source	Government organisation charts; coordination-body mandate documents; VNR institutional-section text.
Unit / basis of assessment	Country (de jure architecture).
Score 1 (critical gap)	No designated SDG coordination body; no formal subnational alignment provisions.
Score 2 (formal only)	Coordination body exists but mandate unclear, or subnational alignment absent.
Score 3 (partial)	Coordination body with a clear mandate exists; subnational alignment provisions partial.
Score 4 (substantial)	Coordination body with a clear, resourced mandate; subnational alignment provisions present for most levels.
Score 5 (no significant gap)	Comprehensive coordination architecture with clear mandates at all levels and explicit subnational alignment.
Important limitations	Formal existence does not ensure functionality; mandates may be nominal.

Cell 5. Institutional × Functional

Indicator	Cross-ministerial protocols enforced and functioning subnational coordination.
Operational definition	Whether coordination-body protocols operate in practice (e.g., regular inter-ministerial meetings, joint decisions) and subnational coordination is active.
Recommended evidence source	VNR functional sections; meeting minutes where public; OECD PCSD assessments; civil society reports.
Unit / basis of assessment	Country (de facto operation).
Score 1 (critical gap)	Coordination body inactive; no functioning inter-ministerial protocol; subnational coordination absent.
Score 2 (formal only)	Body meets infrequently; protocols largely unenforced; subnational coordination nominal.
Score 3 (partial)	Body meets regularly; some joint decisions; subnational coordination uneven.
Score 4 (substantial)	Body meets regularly with substantive outputs; protocols largely enforced; subnational coordination active in most regions.
Score 5 (no significant gap)	Robust, routine cross-ministerial coordination with documented outputs; strong subnational coordination throughout.
Important limitations	Functionality is interpretive and often under-reported; minutes may be unavailable.

Cell 6. Institutional × Outcome-linked

Indicator	Coordination effectiveness visible in joint planning, budgeting, and review.
Operational definition	Whether coordination produces visible joint planning, budgeting, and review outputs across Government.
Recommended evidence source	Joint planning documents; programme-based budgets; integrated review reports.
Unit / basis of assessment	Country (implementation results).
Score 1 (critical gap)	No evidence of joint planning, budgeting, or review attributable to coordination.
Score 2 (formal only)	Limited joint outputs; coordination rarely translates into shared plans or budgets.
Score 3 (partial)	Some joint planning or budgeting in priority sectors; review uneven.
Score 4 (substantial)	Regular joint planning and budgeting across most sectors; integrated review in place.
Score 5 (no significant gap)	Systematic joint planning, budgeting, and review across Government with demonstrable integration.
Important limitations	Outcome-linked; effectiveness is influenced by capacity and political factors, not governance alone.

Cell 7. Operational × Formal

Indicator	Staffing and budget for SDG coordination; regulatory review against SDG targets.
Operational definition	Whether dedicated staffing and budget are formally assigned to SDG coordination, and whether regulatory frameworks are scheduled for review against SDG targets.
Recommended evidence source	Coordination-body staffing and budget documents; regulatory review schedules; VNR operational-section text.
Unit / basis of assessment	Country (de jure resourcing).
Score 1 (critical gap)	No dedicated staffing or budget; no regulatory review against SDG targets.
Score 2 (formal only)	Minimal staffing/budget; regulatory review absent or ad hoc.
Score 3 (partial)	Modest dedicated staffing/budget; regulatory review planned for some sectors.
Score 4 (substantial)	Adequate dedicated staffing/budget; systematic regulatory review across most sectors.
Score 5 (no significant gap)	Ample dedicated staffing and budget; comprehensive, scheduled regulatory review across Government.
Important limitations	Formal allocation does not guarantee execution; budgets may be under-spent.

Cell 8. Operational × Functional

Indicator	Implementation protocols followed and budget execution for SDG-related programmes.
Operational definition	Whether implementation protocols are followed in practice and SDG-related budget allocations are actually executed.
Recommended evidence source	Budget execution reports; audit reports; VNR functional sections; implementation reviews.
Unit / basis of assessment	Country (de facto execution).
Score 1 (critical gap)	Protocols largely unenforced; budget execution for SDG programmes very low.
Score 2 (formal only)	Protocols partially followed; execution uneven and often below allocation.
Score 3 (partial)	Protocols generally followed; execution moderate and improving.
Score 4 (substantial)	Protocols consistently followed; execution high for most SDG programmes.
Score 5 (no significant gap)	Protocols fully enforced; high, on-time budget execution across SDG programmes.
Important limitations	Execution data may be aggregated and hard to attribute to SDGs specifically.

Cell 9. Operational × Outcome-linked

Indicator	Improvements in SDG-relevant service delivery and programme outputs.
Operational definition	Whether there are measurable improvements in SDG-relevant service delivery and programme outputs attributable to operational functioning.
Recommended evidence source	SDG indicator data for service-delivery targets; programme output reports.
Unit / basis of assessment	Country (implementation results).
Score 1 (critical gap)	No measurable improvements; outputs declining or stagnant.
Score 2 (formal only)	Limited improvements in a few areas; most outputs stagnant.
Score 3 (partial)	Visible improvements in several priority service areas.
Score 4 (substantial)	Broad improvements across most SDG-relevant service areas.
Score 5 (no significant gap)	Sustained, broad-based improvements in SDG-relevant service delivery and outputs.
Important limitations	Outcome-linked only; improvements are driven by many factors beyond operational governance.

Cell 10. Data & Accountability × Formal

Indicator	National SDG indicator framework adopted and statistical capacity provisions.
Operational definition	Whether a national SDG indicator framework is formally adopted and whether statistical-capacity provisions are in place.
Recommended evidence source	National indicator framework documents; statistical-system legislation; VNR data sections.
Unit / basis of assessment	Country (de jure framework).
Score 1 (critical gap)	No national SDG indicator framework; no statistical-capacity provisions.
Score 2 (formal only)	Framework partially adopted; capacity provisions minimal.
Score 3 (partial)	Framework adopted; capacity provisions present but limited.
Score 4 (substantial)	Comprehensive framework adopted; capacity provisions adequate and funded.
Score 5 (no significant gap)	Comprehensive framework with strong, funded statistical-capacity provisions.
Important limitations	Formal adoption does not ensure data production; capacity provisions may be unfunded.

Cell 11. Data & Accountability × Functional

Indicator	Data timeliness and disaggregation; quality of VNR reporting.
Operational definition	Whether SDG data are produced timely and disaggregated, and whether VNR reporting meets evidence-based quality standards.
Recommended evidence source	National statistical office outputs; CEPEI VNR Quality Index scores; data-availability metadata.
Unit / basis of assessment	Country (de facto data practice).
Score 1 (critical gap)	Data largely absent, outdated, or non-disaggregated; VNR reporting weak.
Score 2 (formal only)	Some timely data; limited disaggregation; VNR reporting uneven.
Score 3 (partial)	Timely data for many indicators; partial disaggregation; VNR reporting moderate.
Score 4 (substantial)	Timely, disaggregated data for most indicators; VNR reporting strong and evidence-based.
Score 5 (no significant gap)	Timely, disaggregated data across indicators; VNR reporting consistently evidence-based and transparent.
Important limitations	Disaggregation and timeliness vary by indicator; VNR quality is self-reported and partly self-assessed.

Cell 12. Data & Accountability × Outcome-linked

Indicator	Evidence-based policy adjustments and independent review.
Operational definition	Whether data and evidence are used to adjust policy, and whether independent review or scrutiny occurs.
Recommended evidence source	Policy-revision records; parliamentary/audit scrutiny reports; civil society shadow reports.
Unit / basis of assessment	Country (implementation results).
Score 1 (critical gap)	No evidence of data-driven adjustment; no independent review.
Score 2 (formal only)	Occasional adjustments; little independent scrutiny.
Score 3 (partial)	Some data-driven adjustments in priority areas; periodic independent review.
Score 4 (substantial)	Regular data-driven adjustments; routine independent review.
Score 5 (no significant gap)	Systematic, data-driven policy adjustment with robust independent review.
Important limitations	Outcome-linked; policy change is influenced by many factors beyond the data system.

Cell 13. Participatory × Formal

Indicator	Legal framework for stakeholder engagement and VNR consultation provisions.
Operational definition	Whether a legal framework for stakeholder engagement exists and whether VNR consultation is formally required.
Recommended evidence source	Constitutional/legal provisions; VNR preparation guidelines; stakeholder-engagement statutes.
Unit / basis of assessment	Country (de jure framework).
Score 1 (critical gap)	No legal framework for engagement; VNR consultation not required.
Score 2 (formal only)	Consultation required but no broader legal framework; provisions weak.
Score 3 (partial)	Legal framework for engagement present; VNR consultation required.
Score 4 (substantial)	Comprehensive legal framework; structured VNR consultation required.
Score 5 (no significant gap)	Strong legal framework protecting civic space; structured, inclusive VNR consultation required.
Important limitations	Legal provisions may not reflect practice; civic space can narrow despite formal rights.

Cell 14. Participatory × Functional

Indicator	Civil society involvement in monitoring and access to coordination platforms.
Operational definition	Whether civil society and non-state actors are substantively involved in monitoring and have access to coordination platforms.
Recommended evidence source	VNR stakeholder sections; civil society reports; coordination-body access records.
Unit / basis of assessment	Country (de facto practice).
Score 1 (critical gap)	Civil society excluded from monitoring; no platform access.
Score 2 (formal only)	Consultation occurs but without substantive influence or platform access.
Score 3 (partial)	Civil society involved in some monitoring; selective platform access.
Score 4 (substantial)	Civil society substantively involved in monitoring; broad platform access.
Score 5 (no significant gap)	Strong, routine civil-society involvement in monitoring with full platform access.
Important limitations	Involvement is interpretive; access records are often unavailable.

Cell 15. Participatory × Outcome-linked

Indicator	Stakeholder influence on policy and quality of shadow reporting.
Operational definition	Whether stakeholder engagement demonstrably influences policy decisions, and whether independent shadow reporting is robust.
Recommended evidence source	Policy documents referencing stakeholder inputs; shadow reports; accountability-ecosystem assessments.
Unit / basis of assessment	Country (implementation results).
Score 1 (critical gap)	No evidence of influence; no robust shadow reporting.
Score 2 (formal only)	Occasional influence in isolated cases; weak shadow reporting.
Score 3 (partial)	Visible influence in some policy areas; periodic shadow reporting.
Score 4 (substantial)	Regular influence across most policy areas; robust shadow reporting.
Score 5 (no significant gap)	Systematic stakeholder influence across policy; strong, independent shadow reporting.
Important limitations	Outcome-linked; influence is hard to attribute and often contested.

Note. All quantitative thresholds implied above (e.g., "high execution", "most indicators", "most regions") are provisional and require empirical calibration. The protocol is intended to be applied by at least two independent evaluators with documented evidence and inter-rater reliability reporting. The anchors are proposed criteria, not validated thresholds.

Appendix B. Hypothetical Worked Scoring Example (Illustrative)

The following worked example demonstrates how the proposed scoring architecture may be applied. The values below are hypothetical and are included solely to demonstrate the calculation of dimension-level readiness scores, the overall Governance Readiness Score, and the derived Governance Gap Score. They do not represent empirical findings, country-level assessments, or validated governance scores for any real jurisdiction.

Consider an illustrative jurisdiction, "Country X," assessed against the fifteen GGAM cells using the provisional anchors in Appendix A. Suppose two independent evaluators, working from the same evidence base, assign the cell-level readiness scores shown in Table B1. The scores reflect a common illustrative pattern: formal arrangements score relatively well, while functional and outcome-linked levels score lower—the pattern the framework is designed to surface.

Table B1. Hypothetical cell-level readiness scores for an illustrative jurisdiction (Country X).

Dimension	Formal	Functional	Outcome-linked	GRS _d	GGG _d = 5 – GRS _d
Normative	4	3	2	3.00	2.00
Institutional	4	3	3	3.33	1.67
Operational	3	2	2	2.33	2.67
Data & Accountability	3	2	2	2.33	2.67
Participatory	3	3	2	2.67	2.33

Note. Cell values (Formal, Functional, Outcome-linked) are hypothetical readiness scores on the 1–5 scale. GRS_d is the mean of the three level scores (Equation 1); GGS_d = 5 – GRS_d (Equation 4). Values are illustrative and do not represent empirical findings.

The overall Governance Readiness Score is the equally weighted average of the five dimension scores (Equation 2):

Interpretation. The illustrative jurisdiction shows the widest governance gaps in the operational and data-and-accountability dimensions (GGG_d ≈ 2.67), moderate gaps in the normative and participatory dimensions (GGG_d ≈ 2.0–2.33), and the smallest gap in the institutional dimension (GGG_d ≈ 1.67). The aggregate GGS of approximately 2.27 indicates a moderate overall governance gap. The pattern—formal arrangements scoring higher than functional and outcome-linked levels—illustrates the framework’s capacity to distinguish between the existence of governance arrangements and their effective operation, which is the central diagnostic value of the three-level architecture. In an applied setting, this profile would direct reform attention toward operational resourcing and data-and-accountability strengthening rather than toward the creation of new formal structures.

Note. This appendix is a hypothetical worked example for pedagogical purposes only. The scores are not derived from any real assessment and must not be cited as evidence about “Country X” or any actual jurisdiction. The example demonstrates the scoring architecture; it does not validate it. Validation requires systematic application with documented evidence and inter-rater reliability reporting, as noted in Sections 4 and 6.

Appendix C. Source Data and Verification Record for Figure 4: VNR Presenters at HLPF Sessions, 2016–2023

This appendix reproduces the year-by-year source dataset underlying Figure 4 to support transparency and reproducibility. The dataset uses one consistent measurement unit throughout: the number of Voluntary National Review (VNR) presenters per HLPF session. This is distinct from the number of VNR presentations, the number of unique participating countries, and the cumulative number of countries that have ever undertaken a VNR. For 2016–2022, the presenters are countries (the European Union did not present a VNR in those years). For 2023, the figure comprises 38 countries plus the European Union (39 presenters in total;

the EU presented its first VNR in 2023). All values are drawn from authoritative United Nations and IISD records and were checked against the primary source for each year; no values were estimated, interpolated, or fabricated.

Table C1. Year-by-year source data and verification record for Figure 4.

Year	VNR presenters	Measurement unit	Authoritative source	Verification status
2016	22	VNR presenters per HLPF session	UN DESA, 2016 Voluntary National Reviews Synthesis Report (New York: United Nations). Cross-checked in UN DESA, 2017 VNR Synthesis Report: “up from 22 countries that presented inaugural reviews at the 2016 HLPF”.	Verified
2017	43	VNR presenters per HLPF session	UN DESA, 2017 Voluntary National Reviews Synthesis Report (New York: United Nations): “A total of 43 countries presented voluntary national reviews.”	Verified
2018	46	VNR presenters per HLPF session	UN DESA, 2018 Voluntary National Reviews Synthesis Report (New York: United Nations): “A total of 46 countries presented voluntary national reviews, up from 43.”	Verified
2019	47	VNR presenters per HLPF session	UN DESA, 2019 Voluntary National Reviews Synthesis Report (New York: United Nations): “47 countries presented their Voluntary National Reviews (VNR) at the HLPF.”	Verified
2020	47	VNR presenters per HLPF session	UN DESA, 2020 Voluntary National Reviews Synthesis Report (New York: United Nations): “47 countries presented their Voluntary National Reviews (VNRs) during the first virtual high-level political forum.”	Verified
2021	42	VNR presenters per HLPF session	HLPF 2021 Voluntary National Reviews page (hlpf.un.org/2021/voluntary-national-reviews); UN Deputy Secretary-General remarks, 12 July 2021: “This year, 42 countries will be presenting their VNRs.”	Verified
2022	44	VNR presenters per HLPF session	IISD Earth Negotiations Bulletin, HLPF 2022 Summary Report (enb.iisd.org): “44 countries presented VNRs.”	Verified
2023	39	VNR presenters per HLPF session	HLPF 2023 Voluntary National Reviews page (hlpf.un.org/2023/vnrs): “Following 39 presenters carried out VNRs at the HLPF 2023” — list explicitly includes the European Union; counted: 38 countries + EU = 39 presenters. Cross-checked against the HLPF 2023 Information Note.	Verified (38 countries + European Union)

Note. Unit throughout: number of VNR presenters per HLPF session. For 2016–2022, presenters = countries (the European Union did not present a VNR in those years). For 2023, presenters = 38 countries + the European Union (39 in total; the EU presented its first VNR in 2023). Plotted values are presenters per session, not unique participating countries (several countries presented more than once over the period). The pre-event announcement of “42 presenters” for 2023 (IISD, October 2022) was superseded by the actual post-event count of 39. The addition of this source-data table improves transparency, reproducibility, and evidence traceability; it does not itself constitute empirical validation of the GGAM framework.

Reproducibility. A future researcher or editor can verify each plotted value by consulting the authoritative source listed in the fourth column. The UN DESA VNR Synthesis Reports for 2016–2020 are available via the UN DESA Publications portal (desapublications.un.org). The HLPF official VNR pages for 2021 and 2023 are available at hlpf.un.org. The IISD Earth Negotiations Bulletin summary reports are available at enb.iisd.org. All eight plotted values in Figure 4 exactly match the corresponding entries in Table C1.