

Audit Tool Survey

Argyro Kartsonaki & Stefan Wolff
Institute for Conflict, Cooperation and Security, University of Birmingham

A. Introduction: Audit tools as instruments for complex social change

Audit tools have significantly grown in popularity over the past decade as part of a global drive to measure the performance of various institutions against a set of indicators and to score performance in a range of indices. As such, they constitute a form of power in international and transnational governance (Lebaron and Lister 2015, Sending and Lie 2015). Because of the resultant institutional and behavioural effects that the increasing reliance on auditing and benchmarking has, significant responsibility rests with those involved in the design and conduct of audits and in follow-up actions, including advocacy and policy advice (Power 2003). In this section, we therefore, first, address two issues that are central to the GCP's considerations for developing a pluralism audit tool—legitimacy and credibility.

Legitimacy and credibility of audit tools are closely intertwined with each other and derive, in part, from the same components underpinning the design of audit tools and the conduct of audits. Regardless of the methodology by which an audit is to be conducted, audits require data gathering and analysis according to pre-defined indicators and, more often than not, involve the scoring of measurements against equally pre-defined benchmarks. This can create a normative trap: as benchmarking represents a normative vision of what things should be like (Broome and Quirk 2015), audit results tend to attribute responsibility directly or indirectly, either by initiating a blame game about who is responsible for an unsatisfactory state of affairs (Clegg 2015), or by conferring authority to effect changes to an existing status quo (Sending and Lie 2015). This is not in itself problematic, but the underlying logic of what is being measured in most audits is not value free and often differs across different cultures. This is the case with 'popular' audit topics like democracy (Giannone 2010) and human rights (Raworth 2001), and undoubtedly also the case for pluralism. This suggests the need for a context-sensitive design of a pluralism audit tool, with a degree of in-built flexibility in terms of its underpinning methodology (including indicators and benchmarks) that allows for locally nuanced application.

Closely related to this normative dimension of the development of a pluralism audit tool is, therefore, the robustness of the underlying science (Freistein 2015) from the validity of the constructs used (Thomas 2010), to the definition of indicators (Langbein and Knack 2008) and the qualitative and quantitative, perceptual and non-perceptual data that inform their measurement (Heywood and Rose 2014).

Data gathered in the course of an audit need to be consistent and coherent, and as comprehensive as possible, across the countries and/or sectors on which an audit is conducted. In addition, data should be able to capture progress over time and space, while being simple and easy to understand and offering straightforward measuring (Rosga and Satterthwaite 2008). The degree to which they fulfil these criteria will have a clear impact on policy implications and recommendations that can be derived from an audit (Freistein 2015, Giannone 2010).

Critical in this respect is the formulation of benchmarks: the process of developing them needs to be embedded in relevant networks and public so that these networks can grow (Porter 2015). These networks and the public nature of benchmarking thus underpin the development of a credible and legitimate audit tool: they ensure that benchmarks address salient issues, reflect

robust thresholds for assessing performance against specific indicators, and have traction among relevant change agents.¹

While the debate on the effectiveness of audit tools, benchmarks, and indices as instruments for complex social change is inconclusive (Clegg 2010, Dominique et al. 2013), there is some consensus that legitimacy is critical to effective advocacy (Guterman 2014). Yet, it would be a mistake to assume that legitimacy alone is sufficient. Rather, as noted above, scientific, and especially methodological, credibility of the tool is equally important, and in part effects legitimacy.²

In addition, linking a pluralism audit tool to an actual pluralism index (i.e., demonstrating its application) will further increase the likelihood of inducing positive change. Wide and instant distribution of a pluralism index would produce awareness and, through ‘social pressure’, action effects (Kelley and Simmons 2015). Combined with transparency on data gathering and measurement, as well as display and use (Rosga and Satterthwaite 2008), a pluralism audit tool that can be adapted to specific local needs (and capacities) would recommend itself as an instrument for generating practical insights and applicable lessons (Dominique et al. 2013) that can effect social change.

For such approach to work, it would be important to get the balance right between data and insights gained from the audit, their scoring against benchmarks, and the actions (and responsibilities for taking them) that flow from that (Fukuda-Parr 2006). Equally importantly, there needs to be consideration of how the choice of data and data sources for and audit and its related benchmarks may pre-determine responsibilities for action, e.g., by ‘pre-selecting’ the state as the main change agent (Homolar 2015), while simultaneously—positively or negatively—constraining the policy options available for change (Kuzemko 2015).³

The use of benchmarks in auditing tends to lead to a quantification of performance. This has advantages in terms ‘easy’ conclusions being drawn from audits and ‘clear’ messages being sent, but its drawbacks include limitations of substantive outcomes of, and accountability for, policies aimed at addressing identified shortcomings (Harrison and Sekalala 2015). The challenge for a pluralism audit tool in this respect would appear to be in developing an instrument that can generate easily understood insights into existing shortcomings that combine quantifiable performance indicators with qualitative and/or perception-based measures that would enable benchmarking based on objective/de-jure criteria (e.g., laws on the statute book) and subjective/de-facto criteria (e.g., experienced policy implementation and application of law).

Against this background, the remainder of this paper gives an overview of the range of existing audit tools in Section II and offers a number of recommendations to GCP for the further development of a pluralism audit tool.

B. Audit tool profiles

This report presents a compilation of resources that could serve as models for the creation of a pluralism audit tool to be developed by the Global Centre for Pluralism. Three types of resources are considered as relevant: audit tools in a narrow sense of the meaning of the term, i.e., practical guides on how to audit a particular issue; audit tools in the broader sense of indices that present

¹ Seabrooke and Wigan (2015) use the terms salience, expertise, and will in this context to describe three essential requirements for developing benchmarks.

² From a GCP perspective, we thus have to consider a two-way effect: the credibility of the pluralism audit affects GCP’s legitimacy, and the legitimacy of GCP affects the effectiveness of a pluralism audit tool as an instrument for social change.

³ By way of, admittedly crude, illustration: a positive constraint would be the use of a benchmark for respect of religious diversity which would militate against a response to religious exclusion that would seek to forcibly convert members of the excluded religious group; while a negative constraint would be one in which a benchmark of equality of opportunity provides a pretext to ban affirmative action policies to foster inclusion.

the results of an audit and are transparent about their underpinning methodology; and similar indices that lack a transparent discussion of their methodology. Each of these different tools offer some relevant insights into the critical components of audit tools as outlined in the preceding section, including on aspects data sources, gathering and analysis. Surveying different audit tools in this section, thus, prepares the ground for the assessment and recommendations offered in the final section with a view to the development by GCP of a Pluralism Audit Tool.

Such a pluralism audit tool as such does not exist, and the need to create one derives from the assumption that it could serve as a useful instrument to support complex social change towards more inclusive societies. With some of relevant considerations in this respect outlined in the Introduction to this report, the following section focuses on an overview of existing audit tools and relevant indices generated by audit tool-like methodologies that are available online. Constructed predominantly either by international organisations or think tanks or by a collaboration of the two, they target wide-ranging and diverse audiences, including national and international stakeholders, such as politicians, policy-makers, government officials, analysts, international organisations officials, civil society activists, journalists and academics. While often linked to particular agendas (e.g., development, good governance), their utility reaches beyond specific audiences targeted as change agents (e.g., governments, international organisations, professional associations). They disseminate their findings online; the vast majority of audit tool allows open, free of charge access to their reports. They research and measure several state and governance characteristics, focusing mainly on freedom and democracy, development, state stability and corruption. Methodologies vary from purely descriptive, i.e., sole use of questionnaires and public statistics, to purely qualitative, i.e., finding based only on interviews. Most of them however, adopt mixed methodologies combining research on the ground, experts' opinions and public statistics.

This section consists of three parts. The first provides a general overview of audit tools and indices. The second part is a more focused presentation of specific indices, presenting their utility and methodology, which might serve as examples for the Centre's Pluralism audit tool. Finally, the third section comprises two widely used indices that provide information that could inform the construction of the GCP audit tool, but that lack a detailed and transparent discussion of their underpinning methodology.

1. Overview of audit tools

There are numerous audit tools assessing different elements of successful governance, standards of living, state capacity and transparency, as well as audit tools which deal with characteristics of specific regions. This section has divided existing audit tools in five wider themes:

- a) Good governance, including audit tools on governance, democracy, freedom(s) and human rights
- b) Development, referring not only to standards of living and wealth/poverty, but also to the environment of economy, competitiveness and entrepreneurship
- c) State capacities, referring to peace and state stability or fragility
- d) Corruption and transparency audits
- e) Audits conducted according to region, including surveys as well.

These themes are indicative of the main focus of each audit tool, while acknowledging that there is significant overlap, especially regarding indicators and data sources used.⁴

⁴ Note, in this context, also the analysis by Langbein and Knack (2010) of six governance indicators which found that despite claims of measuring separate, distinct dimensions of governance quality, the same broad perception of a more general notion of governance quality was measured.

a) Governance

Audit tools assessing the quality of governance cover a wide conceptual range, including measurements of democracy in its various manifestations, as well as measurement of the rights and freedoms it gives rise to.

The *International Institute for Democracy and Electoral Assistance (IDEA)* publishes a practical guide for “*Assessing the Quality of Democracy*” (<http://www.idea.int/sod-assessments/>). In light of the different forms democracy can have and their worldwide proliferation over the last decades, IDEA has developed a framework for democracy assessment that combines a commitment to the fundamental principles of democracy, the mediating values that are related to these principles, and a range of questions about democratic performance.

The IDEA audit tool employs a “self-assessment” methodology. One of the main characteristics of their approach is that only citizens and others who live in the country being assessed should carry out a democracy assessment, since only they can know from experience how their country’s history and culture shape its approach to democratic principles. If a democracy assessment is conducted by citizens and residents of another country or external agents, this may happen only under strict safeguards of the impartiality of the assessment.

The aim of the audit is to contribute to public debate and consciousness raising and should also assist in identifying priorities for reform and monitoring their progress. Moreover, the exercise ought to allow for the expression of popular understanding as well as any elite consensus.

In terms of methods, the assessments should be qualitative judgements of strengths and weaknesses in the areas of civil and political rights, electoral democracy and accountable governments, strengthened by quantitative measures where appropriate. The assessors should choose benchmarks or standards for assessment, based on the country’s history, regional practice and international norms, as they think appropriate, while the assessment process should involve wide public consultation, including a national workshop to validate the findings. Finally, both old and new democracies can and should be subject to a similar framework of assessment.

In putting together this guide, IDEA aimed to provide a user-friendly knowledge resource for those seeking to improve the quality of their democracies. For those engaged in democracy assistance, IDEA’s self-assessment methodology provides an opportunity for such assistance to be informed by locally defined priorities for democratic reform.

The United Nations Development Programme (UNDP) has introduced so-called “*Governance Assessments*” (http://www.undp.org/content/undp/en/home/ourwork/democraticgovernance/oslo_governance_centre/governance_assessments.html), as well as a “*User’s Guide to Measuring Local Governance*” (http://www.undp.org/content/undp/en/home/librarypage/democratic-governance/oslo_governance_centre/governance_assessments/a-users-guide-to-measuring-local-governance-.html).

Through the Global Programme on Democratic Governance Assessments, UNDP seeks to assist developing countries in producing disaggregated and unranked governance indicators with the aim of developing the capacities of governments, national statistical offices and civil society in the collection, maintenance and analysis of governance-related data and assisting the development of an inclusive, consultative framework for systematic assessment and monitoring of democratic governance goals and targets expressed in national development plans. UNDP currently provides financial and technical support to governance assessments projects across the world, including in Angola, Barbados, Bhutan, Chile, Djibouti, Egypt, FYR Macedonia, Indonesia, Malawi, Mexico, Mongolia, Nicaragua, Nigeria, Senegal and Tajikistan.

The User’s Guide to Local Governance is intended to respond to an increasing demand from a wide range of national stakeholders for guidance on the multiplicity of tools and methods that

are being used to measure, assess and monitor governance at the local level. The guide contains an extensive source guide with more than 20 ready-made tools for assessing local governance. In terms of measuring local governance, there are four broad focus areas which an assessment might address: local governance; decentralisation processes; local democracy; and local government.

Indicators can yield both quantitative and qualitative data. Governance indicators can be classified in many ways. Most researchers and practitioners work with national governance indicators that measure dimensions such as inputs, outputs, processes, outcomes, and impact. This framework is applied to local governance indicators as well and encourages combining reliance on multiple indicators in any assessment.⁵ Such integrated assessment strategies always provide a stronger factual basis and a much sounder foundation for development planning, and policy making than an assessment with a strict emphasis on only one type of indicator or only one measurement approach.

The *Center for Systemic Peace* publishes the “*Polity IV*” series (<http://www.systemicpeace.org/polityproject.html>). The Polity IV Country Report project provides key regime data, regime trend graph, and a brief narrative description of executive recruitment, executive constraints, and political competition for each country covered by the Polity data series. The Polity conceptual approach examines concomitant qualities of democratic and autocratic authority in governing institutions, rather than discreet and mutually exclusive forms of governance. This perspective envisions a spectrum of governing authority that spans from fully institutionalized autocracies through mixed, or incoherent, authority regimes (termed “anocracies”) to fully institutionalized democracies.

The Polity Score captures this regime authority spectrum on a 21-point scale ranging from -10 (hereditary monarchy) to +10 (consolidated democracy). The Polity scores can also be converted into regime categories in a suggested three-part categorization of “autocracies” (-10 to -6), “anocracies” (-5 to +5 and three special values: -66, -77 and -88), and “democracies” (+6 to +10). The Polity scheme consists of measures that record key qualities of executive recruitment, constraints on executive authority and political competition. They assess how institutionalised the procedures for open, competitive and deliberative political participation are, to what extent the choice and the replacement of chief executives takes place in open, competitive elections and to what extent substantial checks and balances on the discretionary powers of the chief executive are disposed.⁶ It also records changes in the institutionalized qualities of governing authority.

The *World Governance Assessment* was produced by the *Overseas Development Institute* from 2000 to 2007 (<https://www.odi.org/projects/1286-world-governance-assessment>). The WGA treats governance as a political process, assuming that the extent to which rules-in-use are perceived as legitimate matters as a determinant of policymaking and regime stability. The WGA is a survey of ‘well informed persons’ (WIPs) in a country, systematically covering a number of relevant groups and led by a local researcher. The survey captures a wide range of indicators/specific questions related to the overall concept of governance, covering key arenas and issues. The advantage of asking about perceptions is that this can capture how rules are *perceived to work*; this is influenced both by formal rules in place and by the actual (informal) practices of applying them. Furthermore, opinions of WIPs are valuable since they represent wide groups of agents who are crucial in shaping the governance of a country. The key parameters in the WGA approach are the six “arenas of governance” i.e., civil society; political society; government; bureaucracy; economic society; judiciary and the six “principles of good governance” being participation; fairness; decency; accountability; transparency; efficiency.

⁵ UNDP, “A Users’ Guide to Measuring Local Governance,” p. 14.

⁶ Centre for Systemic Peace, Global Report 2014, p.20, <http://www.systemicpeace.org/vlibrary/GlobalReport2014.pdf>

The WGA was originally developed as an approach to assessing governance in 1999/2000 and was first field-tested in 2001. Based on the initial experience, changes were made for the second round, which was implemented in 2006. The assessments have been initiated as 'rounds', i.e., taking place in different countries during roughly the same time period. In the second round of the WGA, several changes were introduced. The sample size of WIPs was doubled to around 70 interviewed per country. An internet-based data-warehouse system has been developed in order to manage real-time information and communication with country coordinators. As a result, study management and the overall quality of the sample and data were greatly enhanced. An online training course on technical issues was developed, generating specific information and learning services for all members of country teams.

The *Research and Expertise on World Economy*, in collaboration with the *French Ministry of Economy and Finance*, has launched the "*Institutional Profiles Database*" (<http://www.cepii.fr/institutions/EN/ipd.asp>). The Database provides an original measure of countries' institutional characteristics through composite indicators built from perception data. It was designed in order to facilitate and stimulate research on the relationship between institutions, long-term economic growth, and development. The 2012 edition of the database follows on from the 2001, 2006 and 2009 editions. It covers 143 countries and contains 130 indicators, derived from 330 variables describing a broad range of institutional characteristics, structured in nine functions: political institutions; security, law and order, control of violence; functioning of public administrations; free operation of markets; coordination of stakeholders, strategic vision and innovation; security of transactions and contracts; market regulations, social dialogue; openness; social cohesion and social mobility. The perception data needed to build the indicators were gathered through a survey completed by the country/regional Economic Services (Services Économiques) of the Ministry for the Economy and Finance (MEF) and the French Development Agency (AFD).

Finally, the *International Institute of Social Sciences* has produced an audit tools based on six *Indices of Social Development* (<http://www.indsocdev.org/home.html>). The Indices of Social Development brings together 200 indicators, synthesising them into a usable set of measures to track how different societies perform along six dimensions of social development. These six dimensions are: civic activism, measuring use of media and protest behaviour; clubs and associations, defined as membership in local voluntary associations; intergroup cohesion, which measures ethnic and sectarian tensions, and discrimination; interpersonal safety and trust, focusing on perceptions and incidences of crime and personal transgressions; gender equality, reflecting gender discrimination in home, work and public life; inclusion of minorities, measuring levels of discrimination against vulnerable groups such as indigenous peoples, migrants, refugees, or lower caste groups. The indices are composed from 25 reputable data sources for 193 countries, over the period from 1990 to 2010, and are updated as new data become available. The indices are aggregated using the innovative method of "matching percentiles."⁷

The results reveal the achievements and challenges facing societies across space (from the richness of community life in Sub-Saharan Africa, to the high levels of personal safety and security in the Persian Gulf, to violence in the Caribbean) and time (the growth of civic engagement in Eastern Europe, gender empowerment in the Middle East, or inclusion of minorities in Southern Africa). The indicators show that while economic and social development are closely correlated, many high income societies continue to face problems of discrimination and exclusion, while some developing countries have overcome these challenges. The indices allow estimating the effects of social development for a large range of countries on indicators like economic growth, human development, and governance.

⁷ Indices of Social Development Website, <http://www.indsocdev.org/home.html>

b) Development

Development audit tools consider mainly the economic and social characteristics of the countries they assess, the standards of living and the investment environment. Numerous studies are conducted by the **World Bank** in the framework of “**Country Policy Institutional Assessment**” project (<http://data.worldbank.org/data-catalog/CPIA>). This is a cross-sectional rating of countries against a set of 16 criteria grouped in four clusters: economic management, structural policies, policies for social inclusion and equity, and public sector management and institutions. It assesses the economy and growth and the public sector of 95 economies in East Asia, Pacific, Europe, Central Asia, Latin America and Caribbean, Middle East and North Africa, South Asia, Sub Saharan Africa. The CPIA is a diagnostic tool that is intended to capture the quality of a country’s policies and institutional arrangements—i.e., its focus is on the key elements that are within the country’s control, rather than on outcomes (such as growth rates) that are influenced by elements outside the country’s control. More specifically, the CPIA measures the extent to which a country’s policy and institutional framework supports sustainable growth and poverty reduction, and consequently the effective use of development assistance. The outcome of the exercise produces an overall score and scores for all of the sixteen criteria that compose the CPIA. The CPIA tool was developed and first employed in the mid-1970s and over the years the World Bank has periodically updated and improved it to reflect the lessons of experience and the evolution of thinking about development.

In collaboration with the **European Bank for Reconstruction and Development**, the **World Bank Group** also launched the “**Business Environment and Enterprise Performance Survey**” (BEEPS) (<http://ebrd-beeps.com/>). BEEPS is a firm-level survey of a representative sample of an economy’s private sector with the aim of gaining an understanding of firms’ perceptions of the environment in which they operate. BEEPS covers a broad range of business environment topics including access to finance, corruption, infrastructure, crime, competition, and performance measures. Its findings can be used to help policy makers to understand better how the businesses experience the business environment and to identify, prioritize and implement reforms of policies and institutions that support efficient private economic activity.

The survey was first undertaken on behalf of the EBRD and the World Bank in 1999–2000, when it included some 4,100 enterprises in 25 countries of Eastern Europe and Central Asia (including Turkey) to assess the environment for private enterprise and business development. The fifth round of BEEPS in 2011-2014 covered almost 16,000 enterprises in 30 countries, including 4,220 enterprises in 37 regions in Russia.⁸

The **International Country Risk Guide** is a commercial source of country risk analysis and ratings (<https://www.prsgroup.com/about-us/our-two-methodologies/icrg>). Updated monthly, *ICRG* monitors 140 countries. Each issue provides financial, political, and economic risk information and forecasts. *ICRG* assigns values to 22 indicators indicating *ICRG*’s business-oriented model for quantifying risk, examining such country-specific elements as currency risk, political leadership, the military and religion in politics, and corruption.

The International Fund for Agricultural Development (IFAD) audit tool “**Rural Sector Performance Assessments**” (<http://www.gaportal.org/resources/detail/ifad-land-tenure-indicators>) is a rules-based system that uses a formula that incorporates measures of country need and country performance in the rural sector. The goal of RSPA is to increase the effectiveness of the use of IFAD’s resources and to establish a more transparent basis and predictable level of future resource flows. Equitable access to land and tenure security for IFAD’s target groups are

⁸ BEEPS website, <http://ebrd-beeps.com/about/>

essential for rural development and poverty eradication. Tenure security influences the extent to which farmers are prepared to invest in improvements in production and land management.

The tool has a clear aspiration to achieve change, namely to:

- strengthen the capacity of the rural poor and their organisations providing policy and legal framework for rural organisations and facilitating dialogue between government and rural organisations;
- improve equitable access to productive natural resources and technology, i.e., access to land, access to water for agriculture and access to agricultural research and extension services;
- increase access to financial services and markets, including enabling conditions for rural financial services development, improving the investment climate for rural businesses and facilitate access to agricultural input and produce markets;
- improve gender issues including access to education in rural areas and support female representation; enhance public resources management and accountability promoting accountability and combating transparency and corruption in rural areas.

c) State capacity

State capacity audit tools measure the stability or fragility of states and the sustainability of peace. While there are several indices that measure state stability and peace that will be presented in the next section, there are very limited comprehensive audit tools on this issue. Perhaps the only comprehensive audit tool on fragility is the UNDP's *User's Guide on Measuring Fragility* (http://www.undp.org/content/dam/undp/library/Democratic%20Governance/OGC/usersguide_measure_fragility_ogc.pdf). However, even the UNDP audit tool for fragility actually presents and compares already existing fragility indices, instead of providing an individual tool for the assessment of peace and stability.

The UNDP Guide presents a comparative analysis of cross-country fragility indices. It assesses their conceptual premises, methodological approach and possible uses. Through the compilation, review and evaluation of several already existing indices on fragility, UNDP has created a guide that provides answers to questions, such as what fragility indices are there, what concepts do they intend to measure, how well do they measure these concepts, and how should fragility indices be applied. The UNDP Fragility tool includes the following indices: Bertelsmann Transformation Index: State Weakness Index, Carleton University: Country Indicators for Foreign Policy Fragility Index, The World Bank: Resource Allocation Index, Institute for Economics and Peace: Global Peace Index, Brookings Institution: Index of State Weakness in the Developing World, University of Maryland: Peace and Conflict Instability Ledger, Economist Intelligence Unit: Political Instability Index, George Mason University: State Fragility Index, The World Bank: World Governance Indicators-Political Stability and Absence of Violence.⁹

d) Corruption

Relevant tools for measuring corruption and transparency include the UNDP User Guide to Measuring Corruption; as well as the Global Integrity Report and various sector-specific and local derivatives.

The UNDP *User Guide to Measuring Corruption*" (http://www.undp.org/content/dam/aplaws/publication/en/publications/democratic-governance/dg-publications-for-website/a-users-guide-to-measuring-corruption/users_guide_measuring_corruption.pdf) is targeted at national stakeholders, donors and international actors involved in corruption measurement and anti-

⁹ These indices are discussed in more detail in the following section of this report focusing specifically on indices that could inform the construction of the GCP's audit tool.

corruption programming. It explains the strengths and limitations of different measurement approaches, and provides practical guidance on how to use the indicators and data generated by corruption measurement tools to identify entry points for anti-corruption programming. The Guide presents a series of methodologies, tools and practices that have been used and validated by the anti-corruption community over the last few years. It uses perceptions, experiences and external assessment as data, collecting them through surveys, expert surveys, monitoring and evaluation systems, crowd-sourcing, compliance/field testing and indicator-driven cases. In addition to classifying corruption indicators according to scale, what is being measured, methodology and the role that internal and external stakeholders play, this guide also examines corruption metrics based on the various types of indicators. This approach should help the user better understand the data presented. The main types of corruption indicators are: perception-based indicators and experience-based indicators, indicators based on a single data source and composite indicators and proxy indicators.

The **“Global Integrity Report”** (<http://www.globalintegrity.org/research/reports/global-integrity-report/>) is a guide to anti-corruption institutions and mechanisms around the world, intended to help policymakers, advocates, journalists and citizens identify and anticipate the areas where corruption is more likely to occur within the public sector.

The Report evaluates both anti-corruption legal frameworks and the practical implementation and enforcement of those frameworks, and takes a close look at whether citizen can effectively access and use anti-corruption safeguards.

The Report is prepared by local researchers, lawyers, journalists and academics using a double-blind peer review process. Each country assessment contained in the Global Integrity Report comprises two core elements: a qualitative Reporter’s Notebook and a quantitative Integrity Indicators scorecard. An Integrity Indicators scorecard assesses the existence, effectiveness, and citizen access to key governance and anti-corruption mechanisms through more than 300 actionable indicators. They are scored by a lead in-country researcher and blindly reviewed by a panel of peer reviewers, a mix of other in-country experts as well as outside experts. Reporter’s Notebooks are reported and written by in-country journalists and blindly reviewed by the same peer review panel.

Another Global Integrity audit is the **“Money, Politics, and Transparency (MPT) Campaign Finance Indicators”** (<https://data.moneypoliticaltransparency.org/>) mobilising a network of more than 110 political finance experts from academia, journalism, and civil society to generate comparative, country-level data on the transparency and effectiveness of political finance regimes across the world. A cross-national survey, MPT examines both the de jure legislation regulating political finance and the de facto implementation of that legislation.

In an iterative process, Global Integrity, the Sunlight Foundation and the Electoral Integrity Project worked in consultation with a reference group of political finance experts to develop a concise set of 50 indicator questions, which were compiled into a comparative country scorecard. The project partners also selected an economically, politically and regionally diverse sample of 54 countries in which to apply the scorecard. The selection process, though not randomized, ensures that MPT reflects the exceptional variety characterizing the range of political finance systems across the world.¹⁰

The MPT scorecard evaluates the key components of effective political finance regimes, including the regulation of direct and indirect public funding, limits on contributions and expenditure,

¹⁰ Global Integrity, “Money, Politics and Transparency Campaign Finance Indicators, Assessing Regulation and Practice in 54 Countries across the World in 2014,” p.61, <https://www.globalintegrity.org/wp-content/uploads/2015/07/MPT-CFI-2014-Key-Findings.pdf>

reporting and public disclosure, the regulation of third-party actors, and monitoring and enforcement. Researched scorecards account for both the existing legal particulars of each of these issue areas and the de facto realities of practical implementation in each country.

MPT delves into critical aspects of political finance by examining not only what laws are on the books, but also whether and how those laws are effectively enforced. The combination of quantitative scores and detailed, evidence-based explanations supporting those scores — in addition to the inclusion of a number of non scored, open-text questions that provide additional, context-specific detail — make the MPT indicators a source of information for interested stakeholders, policy makers and reformers.

The third Global Integrity project is the **“Local Integrity Initiatives”** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/>). The Local Integrity Initiatives are a collection of projects assessing anti-corruption and governance at the sub-national and at sector levels. As an extension of the nationally-focused Global Integrity Report, these joint projects are carried out with local partners -directly relevant to the diverse governance challenges found at the local level. They often involve follow-on outreach and advocacy. Local Integrity Initiatives were concluded in Liberia, Argentina, Ecuador, Peru, the Philippines, Kenya and Papua New Guinea. They will be presented in more detail in the following section on Regional Audit Tools.

e) Regional Audit Tools

Under this category are the audit tools that target a specific region, whether it refers to a continent, subcontinent or groups of countries. This is a more general category that includes audits, indices and surveys.

As mentioned in the end of the previous section, the Local Integrity Initiatives are projects assessing anti-corruption and governance at the sub-national and at sector levels. They were carried out in Liberia, Argentina, Ecuador, Peru, the Philippines, Kenya and Papua New Guinea.

The **“Sub-National Integrity Indicators”** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/local-integrity-liberia-2008/>) scorecards for Liberia assess the existence, effectiveness, and citizen access to key governance and anti-corruption mechanisms in each of Liberia’s 15 counties. Each county scorecard comprises more than 200 individual Integrity Indicators questions that are guided by consistent scoring criteria across all counties and supported by original document research and interviews with experts. The indicators explore issues such as the transparency of the local budget process, media freedom, asset disclosure requirements for political leaders, and conflicts of interest regulations at the county level. Scorecards take into account both existing legal measures on the books as well as de facto realities of practical implementation in each county. They were scored by a lead in-country researcher and blindly reviewed by a panel of peer reviewers. The Integrity Indicators for the Liberia Local Governance Toolkit were jointly developed by Global Integrity and the Center for Transparency and Accountability in Liberia (CENTAL).

The **“Local Integrity Argentina”** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/local-integrity-argentina-2009/>) assessments focus on the existence and effectiveness of public sector transparency and anti-corruption mechanisms in Argentina’s provinces. The methodology and fieldwork were carried out in 2008 and 2009 jointly with CIPPEC, an Argentinian think tank that worked with researchers and analysts at the provincial level across the country to score the project’s indicators. These assessments were part of a larger regional project that involved similar but customized governance assessments in Peru’s regions and Ecuador’s largest municipalities.

The ***“Ecuador Municipal Governance”*** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/local-integrity-ecuador-2009/>) assessments focus on the existence and effectiveness of public sector transparency and anti-corruption mechanisms in Ecuador’s largest municipalities. The methodology and fieldwork were carried out in 2008 and 2009 jointly with Grupo FARO, an Ecuadorian non-governmental organisation that worked with researchers and analysts at the municipal level across the country to score the project’s indicators. These assessments were part of a larger regional project that involved similar but customized governance assessments in Peru’s regions and Argentina’s provinces.

The ***“Peru Regional Governance”*** (<https://www.globalintegrity.org/research/reports/local-integrity-initiative/local-integrity-peru-2009/>) assessments focus on the existence and effectiveness of public sector transparency and anti-corruption mechanisms in Peru’s regions. The methodology and fieldwork were carried out in 2008 and 2009 jointly with Cuidadanos al Dia, a Peruvian non-governmental organisation that worked with journalists and analysts at the regional level across the country to score the project’s indicators. These assessments were part of a larger regional project that involved similar but customized governance assessments in Argentina’s provinces and Ecuador’s largest municipalities.

The ***“Philippines Local Governance Transparency and Accountability Indicators”*** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/philippines-local-governance-2011/>) are a set of measures jointly developed by the La Salle Institute of Governance (LSIG) and Global Integrity to assess the strengths and weaknesses of mechanisms for transparency and accountability in local governance.

With the city/municipal level as the basic unit of analysis, the indicators systematically identify best practices and areas for improvement, thereby empowering local stakeholders to plan and implement evidence-based policy and institutional reforms towards improvements in transparency and accountability in local governance. Ten municipalities were assessed: Balanga City, Carmen, Lapu Lapu City, Lawann-Eastern Samar, Miag ao-Iloilo, Quezon City, Santa Maria-Laguna, Tacurong City-Sultan Kudarat, Taytay-Rizal, and Zamboanga City.

The indicators are based on an extensive review of the transparency and accountability literature, and an examination of what are relevant in the Philippine local context. These mechanisms comprise “best practices” for improving transparency and accountability that have been found to work well in various parts of the Philippines and internationally. The indicators, taken in their entirety, aim to capture the extent to which these mechanisms for improving transparency and accountability are in place (existence indicators), whether the design of these mechanisms indicate that they are likely to work (effectiveness indicators), and whether citizens are able to adequately utilize these mechanisms (access indicators). The Philippines local indicators therefore contain a combination of de jure and de facto indicators that look at laws and institutions “on the books,” as well as their implementation and enforcement in practice.

In all, the Philippines local indicators include a total of 205 specific indicator questions. These indicators are spread over six categories representing various aspects of transparency and accountability in local governance. The categories are as follows: civil society, public information and media; local elections; local government accountability; local fiscal processes; local civil service; and local regulatory functions.

The ***“Kenya City Integrity Report”*** (<https://www.globalintegrity.org/research/reports/local-integrity-initiative/kenya-city-integrity-2011/>) is a tool for understanding the existence, effectiveness, and citizen access to key accountability and anti-corruption mechanisms at the city level in Kenya. Rather than measure corruption, the city assessments seek to understand the ways applied to cure it: the public policies, institutions, and practices that deter, prevent, or punish corruption. By understanding where those mechanisms are stronger or weaker in each of the

three cities, they seek to predict where corruption is more or less likely to manifest itself within a city's public sector. Rather than act as a "name and shame" tool, the Report focuses on concrete reforms that city governments, citizens and the private sector can implement to build systems of integrity within their own cities.

The Kenya City Integrity Report was generated by teams of local researchers who gathered original information through document research and interviews with key informants and local stakeholders from the public, private, and civic sectors in Nairobi, Mombasa and Kisumu. This bottom-up approach yields a trusted assessment of each city's situation as told by local citizens themselves – not external or foreign experts. In the production of the Kenya City Integrity Report, Global Integrity and Center for International Private Enterprise (CIPE) worked together with local researchers from the Kenya Association of Manufacturers (KAM) in Nairobi, the Economic and Social Rights Centre (HAKIJAMII/HAKI YETU) in Mombasa, and the Civil Society Organisation Network in Kisumu, as well as two independent peer reviewers.

The ***"Papua New Guinea Provincial Level Access to Information"*** (<http://www.globalintegrity.org/research/reports/local-integrity-initiative/papua-new-guinea-healthcare-2011/>) is a tool for understanding the existence and effectiveness of access to information mechanisms at the provincial level in Papua New Guinea's healthcare system. It seeks to address, if not fully answer, a simple question: does increased access to information empower patients in PNG to demand greater healthcare service delivery in the provinces? The question is important for accountability and transparency reform because information has been theorized to be one of the key pillars of good governance in health care service delivery.

To unpack this claim, Global Integrity collaborated with the Consultative Implementation and Monitoring Council (CIMC) to identify and carry out fieldwork on four key dimensions of information access in the health sector that are theorised to have a significant impact on citizen empowerment and participation in lower-income countries: 1) Basic Issues around the Existence and Usability of Information in Health Care; 2) Redress Mechanisms that Enforce Accountability in the Health Contexts; 3) Availability of Fiscal/Budget Information with which to Conduct Citizen Audits of Local Clinics; and 4) Citizen Participation in Local Decision-Making as Influenced by Availability of Information.

Working in close partnership with CIMC in PNG, this research was generated by teams of local analysts who traveled to each of the five provinces covered. They gathered original information through document research and interviews with key informants and local stakeholders from the public, private, and civic sectors. This bottom-up approach yields a trusted assessment of each province's situation as told by local citizens themselves – not external or foreign experts.

The scorecards then underwent a two-step review process: after being examined by Global Integrity staff for the quality of references, comments and scoring consistency, the data were evaluated by an independent peer reviewer with expertise in PNG health care. After adjustments were made in light of the reviewer's comments, the final scorecards were then published.

The ***"Africa Integrity Indicators"*** (All) (http://aii.globalintegrity.org/scores-map?stringId=transparency_accountability&year=2016) is produced in cooperation of ***"Global Integrity"*** with the ***"Mo Ibrahim Foundation."*** The All assesses key social, economic, political and anti-corruption mechanisms at the national level across the continent. Global Integrity staff recruit and manage teams of in-country contributors in 54 countries to generate original governance data on an annual basis. Using a blend of social science and journalism, in-country teams of independent researchers, academics and journalists report on the de jure as well as de facto reality of corruption and governance. Measuring both the existing legal framework and the "in practice" implementation is key in our effort to produce actionable governance data that help

governments, citizens and civil society understand and evaluate the status quo and identify intervention points for subsequent reform efforts.

The questionnaire has 114 indicators and is divided in two main categories: transparency and accountability.

- Drawing from the indicators of the Global Integrity Report (GIR), the Transparency and Accountability category consists of 59 indicators examining issues divided in the thematic areas of rule of law, accountability, elections, public management integrity, civil service integrity, and access to information. The indicators look into transparency of the public procurement process, media freedom, asset disclosure requirements, independence of the judiciary, and conflict of interest laws, among others.
- The Social Development category consists of 51 indicators about gender, rights, welfare, rural sector, business environment, health and education. It is important to note that this category of the questionnaire was designed to feed the Ibrahim Index of African Governance (IIAG) in areas not covered by the secondary data sources it utilizes. Therefore, it does not attempt to be a comprehensive assessment by itself. Because the Social Development portion of the questionnaire only includes a small number of questions per each topic area, Global Integrity only provides the scores per each individual indicator and does not provide category or subcategory scores. For example, there are only two questions about Health and users can access the score for each of those two indicators, but they cannot obtain an overall Health score or an overall Social Development score.

The Africa Integrity Indicators are scored by in-country researchers following an evidence-based investigation methodology. The resulting data points are then reviewed blindly by a panel of peer reviewers, drawing on the expertise of a mix of in-country experts as well as outside experts. Rather than relying on experiences or pre-existing perceptions by experts, the strength of Global Integrity's methodology is that it requires a variety of sources of information to be reviewed and documented (including legal and scholarly reviews, interviews with experts, and reviews of media stories) to substantiate the score choice.

The research rounds are dated from the completion of the research process. The period of study for each research cycle is 12 months, and the research is completed approximately 4-6 months after the close of the period of study. The research covers all 54 African countries. The pilot phase covered 50 out of the 54 African countries, excluding the Republic of Congo, Guinea-Bissau, Niger and Lesotho. Beginning with the second round research, all African countries are covered.

Each indicator is presented for the user with three elements - score, explanatory comment, and sources. These components mean that a given scorecard presents a wealth of information. Scores allow for general comparisons across countries, while sources and comments provide a unique window into the realities of regulation and enforcement in each country.

The "**Mo Ibrahim Foundation**" also produces the "**Ibrahim Index of African Governance**" (IIAG) (<http://mo.ibrahim.foundation/iiag/>), which is an annual statistical assessment of the quality of governance in every African country. The Mo Ibrahim Foundation defines governance as the provision of the political, social and economic goods that any citizen has the right to expect from his or her state, and that any state has the responsibility to deliver to its citizens.

The IIAG assesses governance provision within four distinct conceptual categories: safety and rule of law, participation and human rights, sustainable economic opportunity and human development. The 2015 IIAG was made up of 14 sub-categories, being rule of law, accountability, personal safety, national security, participation, rights, gender, public management, business environment, infrastructure, rural sector, welfare, education and health. It consisted 93

indicators that were calculated using data from 33 independent, external data sources, including official data, expert assessments and opinion surveys.¹¹

Another index for Africa is the “*Harvard Kennedy School Index of African Governance*” (<http://belfercenter.ksg.harvard.edu/files/Strengthening%20African%20Governance%202008%20Index%20of%20African%20Governance.pdf>), prepared under the auspices of the Mo Ibrahim Foundation, the Kennedy School of Government, Harvard University and the World Peace Foundation.

The Index measures the degree to which each of the “essential political goods” is provided within the forty-eight African countries south of the Sahara. The index defines and summarises the essential political goods in five categories: safety and security; rule of law, transparency, and corruption; participation and human rights; sustainable economic opportunity; and human development.¹² By measuring the performance of government in this manner, that is, by measuring governance, it is able to offer a report card on the accomplishments of each government for the years being investigated. The index uses 55 indicators, combining expert data and public statistics. They retrieve the data through the African Development Bank, African Economic Outlook, CIRI Human Rights Data Project, Environmental Performance Index, Freedom House, Global Peace Index, Heritage Foundation, Institute for Health Metrics and Evaluation, International Centre for Prison Studies, International Monetary Fund, International Road Federation, International Telecommunication Union, Norwegian Refugee Council, Reporters without Borders, Transparency International, UN, The UN Office for the Coordination of Humanitarian Affairs, UN Security Council, UNESCO, UNHCR, UNICEF, Uppsala Conflict Database, U.S. Committee for Refugees and Immigrants, U.S. Energy Information Administration, WHO, World Bank.

The Indicators are standardised each year with reference to the extreme values across all years. This means that each year, index scores for all years need to be updated, with the advantage of maintaining both a closed scale of 0-100 while producing time invariant scores. The index scale is 0.0-100.0 (worst to best; 1 digit displayed, ~13 digits reported in data file) and the overall score is the arithmetic average of the five categories employed.¹³ Each category consists of 2-4 sub-categories calculated by arithmetic average from 1-11 indicators each, with the exception of safety and security, whose sub-categories are weighted. All categories, sub-categories and indicators are weighted equally, with the exception of national security which is weighted by a factor of two in the category of safety and security to account for insufficient data in the second sub-category, public safety.

The “*Afrobarometer Surveys*” project is conducted by the *African Development Bank* under the “*Country Policy and Institutional Assessments*” initiative (<http://www.afrobarometer.org/>). The Afrobarometer is a pan-African, non-partisan research network that conducts public attitude surveys on democracy, governance, economic conditions, and related issues in more than 35 countries in Africa.

The “*Latinbarometro Surveys*” (<http://www.latinbarometro.org/lat.jsp>) is an annual public opinion survey that involves some 20,000 interviews in 18 Latin American countries, representing more than 600 million inhabitants. Latinobarómetro Corporation researches the development of democracy and economies as well as societies, using indicators of opinion, attitudes, behaviour

¹¹ Ibrahim Index of African Governance, “Methodology 2015,” p. 3, <http://static.moibrahimfoundation.org/u/2015/09/03150715/2015-iiag-methodology.pdf>

¹² Ibrahim Index of African Governance, “Strengthening African Governance,” p.7, <http://belfercenter.ksg.harvard.edu/files/Strengthening%20African%20Governance%202008%20Index%20of%20African%20Governance.pdf>

¹³ UNDP, “User’s Guide on Measuring Fragility,” p. 62, http://www.undp.org/content/dam/undp/library/Democratic%20Governance/OGC/usersguide_measure_fragility_ogc.pdf

and values. Its results are used by social and political actors, international organizations, governments and the media.

The **“East Asia Barometer”** (<http://www.jdsurvey.net/eab/eab.jsp>) is a comparative survey for democratisation and value changes. This project represents a systematic comparative survey of attitudes and values toward politics, power, reform, democracy and citizens' political actions in East Asia. Cross-national comparative surveys have been implemented in nine East Asian political systems -namely Hong Kong, Indonesia, Japan, Mainland China, Mongolia, the Philippines, South Korea, Taiwan and Thailand. In each of the nine countries or regions, a national research team administers a country-wide face-to-face survey using standardized survey instruments to compile the required micro-level data under a common research framework and research methodology.

The **“Asia Barometer”** (<http://www.asianbarometer.org/intro/program-objectives>) has as the three main objectives: 1. to generate a region-wide base of scientifically reliable and comparable data. Public opinion data on issues such as political values, democracy, governance, human security, and economic reforms were gathered through face-to-face interviews with randomly-selected pools of respondents that represent the adult population of the country in question. By insisting on strict research standards, we seek to ensure that are data is trustworthy and accurate. 2. To strengthen intellectual and institutional capacity for research on democracy. The network fosters mutual exchange of expertise by bringing partners and scholars together for planning and analysis at the national, continental, and global levels. 3. To disseminate survey results to academics and policy circles. The Asia barometer uses the Globalbarometer Survey (GBS) report as model and distributes its data to a wide variety of individuals and organisations, including decision makers in legislative and executive branches of government, policy advocates and civic educators, journalists, academic researchers, international investors, and NGOs concerned with democratic development.

The **“Eurobarometer”** (http://ec.europa.eu/public_opinion/index_en.htm) produced since 1973 by the European Commission, has been monitoring the evolution of public opinion in the Member States, thus, helping the preparation of texts, decision-making and the evaluation of its work. Surveys and studies address major topics concerning European citizenship such as enlargement, social situation, health, culture, information technology, environment, the Euro, defence. The Standard Eurobarometer survey consists of approximately 1000 face-to-face interviews per country with reports being published twice yearly. The Special Eurobarometer reports are based on in-depth thematic studies carried out for various services of the European Commission or other EU Institutions and integrated in the Standard Eurobarometer's polling waves. Flash Eurobarometers are ad hoc thematic telephone interviews conducted at the request of any service of the European Commission. Flash surveys enable the Commission to obtain results relatively quickly and to focus on specific target groups, as and when required. The Eurobarometer strongly support the conduct of qualitative studies as they investigate in-depth the motivations, feelings and reactions of selected social groups towards a given subject or concept, by listening to and analysing their way of expressing themselves in discussion groups or with non-directive interviews.

The **“European Social Survey”** (ESS) (<http://www.esrc.ac.uk/research/our-research/european-social-survey-ess/>) is an academically-driven cross-nationally comparative social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. The ESS was established in 2001 and covers over 30 nations.

Finally, the **“EU Minorities and Discrimination Survey”** (<http://fra.europa.eu/en/project/2011/eu-midis-european-union-minorities-and-discrimination-survey>) provides an extensive data set on discrimination and victimisation faced by ethnic minorities and immigrants in the EU. The one-off survey of 500 randomly selected immigrants

and members of ethnic minorities in each of the then 27 member states of the EU was the first of its kind to systematically survey minority groups across the EU through face-to-face interviews using a standardised questionnaire.

2. Selected index profiles

This section presents specific indices that could inform the construction of the Pluralism Audit Tool. They are assembled in the same categories as the previous section, excluding regional audit tools. Thus, they are divided in four main themes, being governance, development, state capacity and corruption indices. Each category provides one to three examples of well-constructed instruments, presenting their profiles, their purpose and target audience, what change is sought through the index and how is it constructed and disseminated.

a) Governance

This section begins with an index created by the *Quality of Government Institute*, University of Gothenburg. The main objective of their research is to address the theoretical and empirical problem of how political institutions of high quality can be created and maintained. A second objective is to study the effects of Quality of Government on a number of policy areas, such as health, the environment, social policy, and poverty, approaching these problems from a variety of different theoretical and methodological perspectives. For the purposes of their research they created the “*Quality of Government Datasets*” (QoG Standard dataset) aiming to make publicly available cross-national comparative data on quality of governance and its correlates (<http://www.gaportal.org/global-indicators/quality-of-government-datasets-qog-standard-dataset>).

The datasets contain a wide range of different types of data: expert-coded indicators and classifications, various demographic measures, national accounts data and aggregated individual-level survey data. The data are compiled from numerous freely available and well-known data sources, including datasets produced by independent research projects, international research initiatives, NGOs and IGOs.

QOG produces two datasets:

- A cross-section dataset examining all countries in the world recognised by the United Nations in 2002, plus Taiwan, and
- A cross-section time-series dataset assessing all countries in the world recognised by the United Nations in 2002, plus Taiwan and 13 historical countries that no longer exist.

Each of the two datasets is divided into three data types or topics:¹⁴

- WII (What It Is), i.e., variables related to the core of the institute’s research area (e.g., corruption, bureaucratic quality, human rights and democracy);
- HTG (How To Get It), thus, variables believed to promote the quality of government (e.g., electoral rules, political institutions, legal and colonial origin, religion and social fractionalisation); and
- WYG (What You Get), being variables pertaining to some of the possible consequences of government quality (e.g., economic and human development, international and domestic peace, environmental sustainability, gender equality, and satisfied and trusting citizens).

A major strength of the datasets is that they often provide several measures of the same concept, allowing users to choose between several indicators and to test whether different ways of measuring a phenomenon has any effect on the analyses. Furthermore, the indicators in this

¹⁴ Quality of Government Datasets Website, <http://www.gaportal.org/global-indicators/quality-of-government-datasets-qog-standard-dataset>

dataset are not useful for only studying the quality of government; these variables can be used to inspire a research agenda focused on making claims regarding levels or quality of democracy.

The datasets are described in an extensive codebook, which is freely available for download. They are widely used in academic research.

The “*Media Sustainability Index*” (MSI) (<https://www.irex.org/resource/media-sustainability-index-msi-methodology>) produced by *International Research and Exchanges* (IREX) provides in-depth analyses of the conditions for independent media in 80 countries across the world. IREX developed the MSI to provide a complete picture of the development of sustainable, independent media. Looking beyond issues of free speech, the MSI aims to understand the degree to which the journalist corps is emerging as objective professionals, whether media firms can sustain robust news operations, and whether civil society supports the fourth estate.

The MSI seeks to make a difference in the lives of citizens in each country, by measuring a number of contributing factors of a well-functioning media system and considering both traditional media types and new media platforms. This level of investigation allows policymakers and implementers to analyse the diverse aspects of media systems and determine the areas in which media development assistance can improve citizens’ access to news and information. This way, citizens can help improve the quality of governance through participatory and democratic mechanisms, and help government and civil society actors devise solutions to pervasive issues such as poverty, healthcare, conflict, and education. The MSI also provides useable information for the media and media advocates in each country and region. By reflecting the expert opinions of media professionals in each country, its results inform the media community, civil society, and governments of the strengths and weaknesses of the sector.

The MSI, therefore, assesses how sustainable a media sector is in the context of providing the public with useful, timely, and objective information and how well it serves as a facilitator of public discussion. By “sustainability” IREX refers to the ability of media to play its vital role as the “fourth estate.” To measure this, the MSI assesses five objectives that shape a media system: freedom of speech, professional journalism, plurality of news, business management, and supporting institutions. A score is attained for each objective by rating between seven and nine indicators, which determine how well a country meets that objective.

The primary source of information is a panel of local experts that IREX assembles in each country to serve as panellists. These experts are drawn from the country’s media outlets, NGOs, professional associations, and academic institutions. Panellists may be editors, reporters, media managers or owners, advertising and marketing specialists, lawyers, professors or teachers, or human rights observers. Additionally, panels comprise the various types of media represented in a country. The panels also include representatives from the capital city and other geographic regions, and they reflect gender, ethnic, and religious diversity as appropriate. For consistency from year to year, at least half of the previous year’s participants are included on the following year’s panel. IREX identifies and works with a local or regional organization or individual to oversee the process.

The scoring is completed in two parts. First, panel participants are provided with a questionnaire and explanations of the indicators and scoring system. Descriptions of each indicator clarify their meanings and help organize the panellist’s thoughts. For example, the questionnaire asks the panellist to consider not only the letter of the legal framework, but its practical implementation, too. A country without a formal freedom-of-information law that enjoys customary government openness may well outperform a country that has a strong law on the books that is frequently ignored. Furthermore, the questionnaire does not single out any one type of media as more important than another; rather it directs the panellist to consider the salient types of media and

to determine if an underrepresentation, if applicable, of one media type impacts the sustainability of the media sector as a whole.

In this way, MSI captures the influence of public, private, national, local, community, and new media. Each panellist reviews the questionnaire individually and scores each indicator. The panellists then assemble to analyse and discuss the objectives and indicators. While panellists may choose to change their scores based upon discussions, IREX does not promote consensus on scores among panellists. The panel moderator (in most cases a representative of the host-country institutional partner or a local individual) prepares a written analysis of the discussion, which IREX staff members edit subsequently. Names of the individual panellists and the partner organization or individual appear at the end of each country chapter.

IREX editorial staff members review the panellists' scores, and then provide a set of scores for the country, independently of the panel. This score carries the same weight as an individual panellist. The average of all individual indicator scores within the objective determines the objective score. The overall country score is an average of all five objectives.

In some cases, where conditions on the ground are such that panellists might suffer legal retribution or physical threats as a result of their participation, IREX will opt to allow some or all of the panellists and the moderator/author to remain anonymous. In severe situations, IREX does not engage panellists as such; rather the study is conducted through research and interviews with those knowledgeable of the media situation in that country.

Panellists are directed to score each indicator from 0 to 4, using whole or half points. The average scores of all the indicators are averaged to obtain a single, overall score for each objective. Objective scores are averaged to provide an overall score for the country. Guidance on how to score each indicator is as follows:

- 0 = Country does not meet the indicator; government or social forces may actively oppose its implementation.
- 1 = Country minimally meets aspects of the indicator; Forces may not actively oppose its implementation, but business environment may not support it and government or profession do not fully and actively support change.
- 2 = Country has begun to meet many aspects of the indicator, but progress may be too recent to judge or still dependent on current government or political forces.
3 = Country meets most aspects of the indicator; implementation of the indicator has occurred over several years and/or through changes in government, indicating likely sustainability.
- 4 = Country meets the aspects of the indicator; implementation has remained intact over multiple changes in government, economic fluctuations, changes in public opinion, and/or changing social conventions.

The average scores of all the indicators are averaged to obtain a single, overall score for each objective. Objective scores are averaged to provide an overall score for the country. IREX interprets the overall scores as follows:

- *Unsustainable, Anti-Free Press* (0-1): Country does not meet or only minimally meets objectives. Government and laws actively hinder free media development, professionalism is low, and media-industry activity is minimal.
- *Unsustainable Mixed System* (1-2): Country minimally meets objectives, with segments of the legal system and government opposed to a free media system. Evident progress in free-press advocacy, increased professionalism, and new media businesses may be too recent to judge sustainability.
- *Near Sustainability* (2-3): Country has progressed in meeting multiple objectives, with legal norms, professionalism, and the business environment supportive of independent

media. Advances have survived changes in government and have been codified in law and practice. However, more time may be needed to ensure that change is enduring and that increased professionalism and the media business environment are sustainable.

- *Sustainable* (3–4): Country has media that are considered generally professional, free, and sustainable, or to be approaching these objectives. Systems supporting independent media have survived multiple governments, economic fluctuations, and changes in public opinion or social conventions.

Reporters without Borders (RWB) compile an index on “**Press Freedom**” (<https://rsf.org/en/ranking>) which measures the level of freedom available to journalists. The Index ranks 180 countries according to the level of freedom available to journalists. It is a snapshot of the media freedom situation based on an evaluation of pluralism, independence of the media, quality of legislative framework and safety of journalists in each country. It does not rank public policies even if governments obviously have a major impact on their country’s ranking, nor is it an indicator of the quality of journalism in each country.

The RWB calculates a global indicator and regional indicators that evaluate the overall performance of countries (in the world and in each region). It is an absolute measure that complements the Index’s comparative rankings. The global indicator is the average of the regional indicators, each of which is obtained by averaging the scores of all the countries in the region, weighted according to their population as given by the World Bank.

The degree of freedom available to journalists in 180 countries is determined by pooling the responses of experts to a questionnaire devised by RWB. This qualitative analysis is combined with quantitative data on abuses and acts of violence against journalists during the period evaluated. The criteria evaluated in the questionnaire are pluralism, media independence, media environment and self-censorship, legislative framework, transparency, and the quality of the infrastructure that supports the production of news and information.

To compile the Index, RWB has developed an online questionnaire with 87 questions focused on these criteria. Translated into 20 languages including English, Arabic, Chinese, Russian, Indonesian and Korean, the questionnaire is targeted at the media professionals, lawyers and sociologists who are asked to complete it. Scores are calculated on the basis of the responses of the experts selected by RSF combined with the data on abuses and violence against journalists during the period evaluated.

A team of specialists, each assigned to a different geographical region, keeps a detailed tally of abuses and violence against journalists and media outlets. These researchers also rely on a network of correspondents in 130 countries. The Abuses indicator for each country is calculated on the basis of the data about the intensity of abuses and violence against media actors during the period evaluated. This quantitative indicator is then used to weight the qualitative analysis of the situation in the country based on the replies to the questionnaires.

The questionnaire focuses on such criteria categories as the country’s performance as regards pluralism, media independence and respect for the safety and freedom of journalists. Each question in the questionnaire is linked to one of the six following indicators:

- *Pluralism* measures the degree to which opinions are represented in the media.
- *Media independence* measures the degree to which the media are able to function independently of sources of political, governmental, business and religious power and influence.
- *Environment and self-censorship* analyses the environment in which news and information providers operate.

- *Legislative framework* measures the impact of the legislative framework governing news and information activities.
- *Transparency* measures the transparency of the institutions and procedures that affect the production of news and information.
- *Infrastructure* measures the quality of the infrastructure that supports the production of news and information.

With each being given a score between 0 and 100, a cumulative score emerges as follows:

- 0 to 15 points: Good (white)
- 15.01 to 25 points: Fairly good (yellow)
- 25.01 to 35 points: Problematic (orange)
- 35.01 to 55 points: Bad (red)
- 55.01 to 100 points: Very bad (black)

Other indices measuring different aspects of freedom, good governance and democracy include:

- Democracy Index, by the Economist Intelligence Unit, http://www.eiu.com/public/topical_report.aspx?campaignid=DemocracyIndex2015
- Human Rights Database Index, <http://www.ohchr.org/EN/HRBodies/Pages/UniversalHumanRightsIndexDatabase.aspx>
- Multiculturalism Policy Index for Immigrant Minorities, <http://www.queensu.ca/mcp/sites/webpublish.queensu.ca.mcpwww/files/files/immigrantminorities/evidence/ImmigrantMinoritiesApr12.pdf>
- World Justice Project: Rule of Law Index, http://worldjusticeproject.org/sites/default/files/roli_2015_0.pdf
- Migrant Interaction Policy Index (MIPEX), <http://www.mipex.eu/>
- International Institute of Social Sciences: Indices of Social Development, including indices on Civic Activism, Clubs and Associations, Intergroup Cohesion, Interpersonal Safety and Trust, Gender Equality and Inclusion of Minorities, <http://www.indsocdev.org/home.html>
- Vanhanen Index of Democracy, <https://www.prio.org/Data/Governance/Vanhanen-index-of-democracy/>
- Democratic Audit UK, http://www.democraticaudit.com/?page_id=18787
- Sustainable Governance Indicators Survey, <http://www.sgi-network.org/2015/>
- The Minorities at Risk Project, <http://www.cidcm.umd.edu/mar/>
- Democracy Barometer, http://www.democracybarometer.org/concept_en.html

b) Development

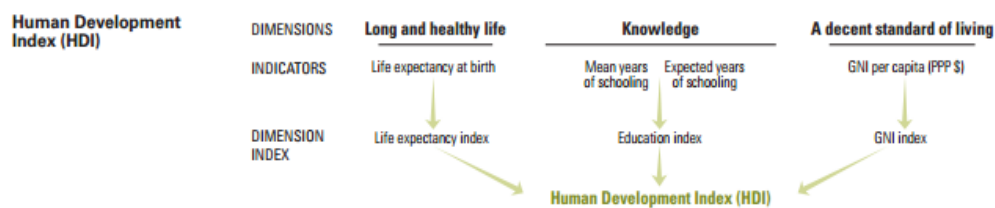
The **UNDP “Human Development Index”** (HDI) (<http://hdr.undp.org/en/content/human-development-index-hdi>) was created to emphasize that expanding human choices should be the ultimate criteria for assessing development results. Economic growth is a mean to that process, but is not an end by itself. The HDI can also be used to question national policy choices, asking how two countries with the same level of GNI per capita can end up with different human development outcomes. For example, Malaysia has GNI per capita higher than Chile, but in Malaysia, life expectancy at birth is about 7 years shorter and expected years of schooling is 2.5 years shorter than Chile, resulting in Chile having a much higher HDI value than Malaysia. These striking contrasts can stimulate debate about government policy priorities.

Thus, the HDI stresses that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone. The Index is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. The HDI is the geometric mean of normalised indices for each of the three dimensions.

The health dimension is assessed by life expectancy at birth, the education dimension is measured by mean of years of schooling for adults aged 25 years and more and expected years of schooling for children of school entering age. The standard of living dimension is measured by gross national income per capita. The HDI uses the logarithm of income, to reflect the diminishing importance of income with increasing GNI. The scores for the three HDI dimension indices are then aggregated into a composite index using geometric mean.

The HDI simplifies and captures only part of what human development entails. It does not reflect on inequalities, poverty, human security, empowerment, etc. The HDRO offers the other composite indices as broader proxy on some of the key issues of human development, inequality, gender disparity and human poverty.

A simplified graphic representation of the Index and its components is presented below.



The 2014 HDI covers 188 countries. The wide coverage is the result of efforts by the Human Development Report Office (HDRO) to work with UN agencies and the World Bank, who provide internationally standardised data and also with national statistical agencies to obtain required development indicators for the HDI. Since 2010, the Human Development Report data has been available on Google Public Data Explorer, in an initiative aimed at increasing its accessibility.

Further indices that measure economic and social development include

- Global Competitiveness Index, <http://reports.weforum.org/global-competitiveness-report-2015-2016/>
- Wall Street Journal & Heritage Foundation: Index of Economic Freedom, <http://www.heritage.org/index/>
- Commonwealth Youth Development Index, <http://youthdevelopmentindex.org/views/index.php#OVER>
- Commitment to Development Index, <http://www.cgdev.org/cdi-2015>

c) State capacity

The **“Fragile States Index”** (<http://fsi.fundforpeace.org/>) produced by the **Fund for Peace** is an annual ranking of 178 countries based on their levels of stability and the pressures they face. Focusing on the problems of weak and failing states, purpose of this research is to promote sustainable security, through training and education, engagement of civil society, building bridges across diverse sectors, and developing innovative technologies and tools for policy makers. The Fund’s objective is to create practical tools and approaches for conflict mitigation that are useful to decision-makers.

The Fragile States Index is a tool that highlights not only the normal pressures that all states experience, but also in identifying when those pressures are pushing a state towards the brink of failure. By highlighting pertinent issues in weak and failing states, the Index -and the social science framework and the data analysis tools upon which it is built- makes political risk assessment and early warning of conflict accessible to policy-makers and the public at large.

The Index adopts a “mixed methods” approach using qualitative and quantitative techniques to establish patterns and trends. The Fund for Peace collects thousands of reports and information from around the world, detailing the existing social, economic and political pressures faced by

each of the 178 countries it analyses. The Index is based on the Fund for Peace's proprietary Conflict Assessment System Tool (CAST) analytical platform. Based on comprehensive social science methodology, data from three primary sources is triangulated and subjected to critical review to obtain final scores for the Index.

Millions of documents are analysed every year. By applying highly specialized search parameters, scores are apportioned for every country based on twelve key political, social and economic indicators (demographic pressures, refugees and IDPs, group grievance, human flight and brain drain, uneven economic development, poverty and economic decline, state legitimacy, public services, human rights and rule of law, security apparatus, factionalised elites, and external intervention) that are the result of years of painstaking expert social science research.

Through integration and triangulation techniques, the CAST platform separates the relevant data from the irrelevant. Guided by the twelve primary social, economic and political indicators (each split into an average of 14 sub-indicators), the researchers first use content analysis using specialised search terms that flag relevant items. This analysis is then converted into a score representing the significance of each of the various pressures for a given country. The content analysis is further juxtaposed against two other key aspects of the overall assessment process: quantitative analysis and qualitative inputs based on major events in the countries examined.

This mixed-methods approach also helps to ensure that inherent weaknesses, gaps, or biases in one source are checked by the others. Though the basic data underpinning of the Index is already freely and widely available electronically, the strength of the analysis is in the methodological rigor and the systematic integration of a wide range of data sources.

The **"Global Peace Index"** (http://economicsandpeace.org/wp-content/uploads/2015/06/Global-Peace-Index-Report-2015_0.pdf) created by **The Institute for Economics and Peace** (IEP) provides metrics for measuring peace and uncovering the relationships between business, peace and prosperity. It also seeks to promote a better understanding of the cultural, economic and political factors that create peace. The IEP works with a wide range of partners internationally and collaborates with intergovernmental organisations on measuring and communicating the economic value of peace.

As peace is difficult to define, the Index approaches peace in terms of the harmony achieved by the absence of violence or the fear of violence, which has been described as Negative Peace. Negative Peace is a compliment to Positive Peace which is defined as the attitudes, institutions and structures which create and sustain peaceful societies.

The Global Peace Index measures a country's level of Negative Peace using three domains of peacefulness. The first domain, ongoing domestic and international conflict, investigates the extent to which countries are involved in internal and external conflicts, as well as their role and duration of involvement in conflicts. The second domain evaluates the level of harmony or discord within a nation; ten indicators broadly assess what might be described as societal safety and security. The assertion is that low crime rates, minimal terrorist activity and violent demonstrations, harmonious relations with neighbouring countries, a stable political scene and a small proportion of the population being internally displaced or made refugees can be equated with peacefulness. Seven further indicators are related to a country's militarisation, reflecting the link between a country's level of military build-up and access to weapons and its level of peacefulness, both domestically and internationally. Comparable data on military expenditure as a percentage of GDP and the number of armed service officers per head are gauged, as are financial contributions to UN peacekeeping missions.

The GPI comprises 23 indicators of the existence of violence or fear of violence. The indicators were originally selected with the assistance of an international panel of independent experts in

2007 and have been reviewed by the expert panel on an annual basis. In the 2015 edition of the index two new indicators have been introduced: number and duration of internal conflicts and number, duration and role in external conflicts (which replace number of external and internal conflicts fought). These indicators, produced by the Institute for Economics and Peace, improve the methodological structure of the index by distinguishing between domestic and international conflict and adding a degree of complexity to the scoring.

All scores for each indicator are normalised on a scale of 1-5, whereby qualitative indicators are banded into five groupings and quantitative ones are either banded into ten groupings or rounded to the first decimal point. The Economist Intelligence Unit's team of country analysts has scored seven of the eight qualitative indicators and also provided estimates where there have been gaps in the quantitative data. The indicators and the sources where data are retrieved from are the following:

1. Number and duration of internal conflicts, data by Uppsala Conflict Data Program (UCDP), Battle-Related Deaths Dataset, Non-State Conflict Dataset and One-sided Violence Dataset; Institute for Economics and Peace (IEP)
2. Number of deaths from organised conflict (external), UCDP Armed Conflict Dataset
3. Number of deaths from organised conflict (internal), International Institute for Strategic Studies (IISS) Armed Conflict Database (ACD)
4. Number, duration and role in external conflicts, UCDP Battle-Related Deaths Dataset; IEP
5. Intensity of organised internal conflict, qualitative assessment by EIU analysts
6. Relations with neighbouring countries, qualitative assessment by EIU analysts
7. Level of perceived criminality in society, qualitative assessment by EIU analysts
8. Number of refugees and internally displaced people as a percentage of the population, Office of the High Commissioner for Refugees (UNHCR) Mid-Year Trends; Internal Displacement Monitoring Centre (IDMC)
9. Political instability, qualitative assessment by EIU analysts
10. Political Terror Scale, qualitative assessment of Amnesty International and US State Department yearly reports
11. Impact of terrorism, IEP Global Terrorism Index (GTI)
12. Number of homicides per 100,000 people, United Nations Office on Drugs and Crime (UNODC) Surveys on Crime Trends and the Operations of Criminal Justice Systems (CTS); EIU estimates
13. Level of violent crime, qualitative assessment by EIU analysts
14. Likelihood of violent demonstrations, qualitative assessment by EIU analysts
15. Number of jailed population per 100,000 people, World Prison Brief, International Centre for Prison Studies, University of Essex
16. Number of internal security officers and police per 100,000 people, UNODC; EIU estimates
17. Military expenditure as a percentage of GDP, The Military Balance, IISS
18. Number of armed services personnel per 100,000 people, The Military Balance, IISS
19. Volume of transfers of major conventional weapons as recipient (imports) per 100,000 people, Stockholm International Peace, Research Institute (SIPRI) Arms Transfers Database
20. Volume of transfers of major conventional weapons as supplier (exports) per 100,000 people SIPRI Arms Transfers Database
21. Financial contribution to UN peacekeeping missions, United Nations Committee on Contributions; IEP
22. Nuclear and heavy weapons capabilities, The Military Balance, IISS; SIPRI; UN Register of Conventional Arms; IEP

23. Ease of access to small arms and light weapons, qualitative assessment by EIU analysts.¹⁵

When the GPI was launched in 2007 the advisory panel of independent experts apportioned scores based on the relative importance of each of the indicators on a scale 1-5. Two sub-component weighted indices were then calculated from the GPI group of indicators:

- A measure of how at peace internally a country is;
- A measure of how at peace externally a country is (its state of peace beyond its borders).

The overall composite score and index is then formulated by applying a weight of 60 percent to the measure of internal peace and 40 percent for external peace. This decision was based on the innovative notion that a greater level of internal peace is likely to lead to, or at least correlate with, lower external conflict. The weights have been reviewed by the advisory panel prior to the compilation of each edition of the GPI.

The Economists Intelligence Unit's Country Analysis team plays an important role in producing the GPI by scoring seven qualitative indicators and filling in data gaps on quantitative indicators when official data is missing. The EIU employs more than 100 full-time country experts and economists, supported by 650 in-country contributors. Analysts generally focus on two or three countries and, in conjunction with local contributors, develop a deep knowledge of a nation's political scene, the performance of its economy and the society in general. Scoring follows a process to ensure reliability, consistency and comparability:

1. Individual country analysts score qualitative indicators based on a scoring methodology and using a digital platform;
2. Regional directors use the digital platform to check scores across the region; through the platform they can see how individual countries fare against each other and evaluate qualitative assessments behind proposed score revisions;
3. Indicator scores are checked by the EIU's Custom Research team (which has responsibility for the GPI) to ensure global comparability;
4. If an indicator score is found to be questionable, the Custom Research team, and the appropriate regional director and country analyst discuss and make a judgment on the score;
5. Scores are assessed by the external advisory panel before finalising the GPI;
6. If the advisory panel finds an indicator score to be questionable, the Custom Research team, and the appropriate regional director and country analyst discuss and make a final judgment on the score, which is then discussed in turn with the advisory panel.¹⁶

Because of the large scope of the GPI, occasionally data for quantitative indicators do not extend to all nations. In this case, country analysts are asked to suggest an alternative data source or provide an estimate to fill any gap. This score is checked by Regional Directors to ensure reliability and consistency within the region, and by the Custom Research team to ensure global comparability. Again, indicators are assessed by the external advisory panel before finalisation. The findings are available online.

Other indices that measure state capacity, fragility and peace are the following:

- Bertelsmann Transformation Index: State Weakness Index, Transformation Index, <http://www.bti-project.org/de/index/>

¹⁵ Global Peace Index Report, p.101, http://economicsandpeace.org/wp-content/uploads/2015/06/Global-Peace-Index-Report-2015_0.pdf

¹⁶ Global Peace Index Report, p.103, http://economicsandpeace.org/wp-content/uploads/2015/06/Global-Peace-Index-Report-2015_0.pdf

- Carleton University: Country Indicators for Foreign Policy Fragility Index, <http://www4.carleton.ca/cifp/ffs.htm>
- The World Bank: Resource Allocation Index, <http://ida.worldbank.org/financing/ida-resource-allocation-index>
- The World Bank: World Governance Indicators-Political Stability and Absence of Violence, <http://info.worldbank.org/governance/wgi/index.aspx#home>
- Brookings Institution: Index of State Weakness in the Developing World, <http://www.brookings.edu/research/reports/2008/02/weak-states-index>
- University of Maryland: Peace and Conflict Instability Ledger, <http://www.cidcm.umd.edu/>
- Economist Intelligence Unit: Political Instability Index, http://viewswire.eiu.com/site_info.asp?info_name=social_unrest_table&page=noads&rf=0
- George Mason University: State Fragility Index, <http://www.systemicpeace.org/globalreport.html>

d) Corruption

The *International Budget Partnership* promotes public access to budget information and the adoption of accountable budget systems. Through the *Open Budget Initiative* the “*Open Budget Index*” (OBI) (<http://www.internationalbudget.org/opening-budgets/open-budget-initiative/open-budget-survey/>) was created that evaluates whether governments give the public access to budget information and opportunities to participate in the budget process at the national level and assigns a score to each country based on the information it makes available to the public throughout the budget process. The findings address government officials, watchdogs of public finances and supreme audit institutions, donors, the media as well as citizens and members of the public.

Public budgets translate a government’s policies, political commitments, and goals into decisions on how much revenue to raise, how it plans to raise it, and how to use these funds to meet the country’s competing needs, from bolstering security to improving health care to alleviating poverty. While a government’s budget directly or indirectly affects the lives of every one of its citizens, it can have the greatest impact on the well-being and prospects of certain groups, such as the elderly, children, the poor, rural residents, and minorities. Budget cuts tend to have the greatest impact on programs that benefit the poor and vulnerable, whereas items like the interest owed on government debt, wages for public employees, or military expenditures are more likely to have first claim on scarce funds. Moreover, even when funds have been allocated to specific programs—whether for minorities, children, or the disabled—poor management and misuse can keep these funds from reaching the intended beneficiaries. Marginalized people lack political power, so it is hard for them to hold their government to account—another factor in poor budget execution

Therefore, the IBP believes that open budgets are empowering because they allow people to be the judge of whether or not their government officials are good stewards of public funds. While providing the public with comprehensive and timely information on the government’s budget and financial activities and opportunities to participate in decision making can strengthen oversight and improve policy choices, keeping the process closed can have the opposite effect. Restricting access to information creates opportunities for governments to hide unpopular, wasteful, and corrupt spending, ultimately reducing the resources available to fight poverty.

Thus, the Open Budget Index is a measure of central government budget transparency. The OBI assigns countries covered by the Open Budget Survey a transparency score on a 100-point scale using 109 of the 140 questions on the Survey. These questions focus specifically on whether the government provides the public with timely access to comprehensive information contained in eight key budget documents in accordance with international good practice standards.

The Open Budget Survey assesses the three components of a budget accountability system: public availability of budget information; opportunities for the public to participate in the budget process; and the strength of formal oversight institutions, including the legislature and the national audit office (referred to here as the supreme audit institution). The majority of the Survey questions assess what occurs in practice, rather than what is required by law.

The Survey assesses the public availability of budget information by considering the timely release and contents of eight key budget documents that all countries should issue at different points in the budget process, according to generally accepted good practice criteria for public financial management. Many of these criteria are drawn from those developed by multilateral organizations, such as the International Monetary Fund's (IMF) Code of Good Practices on Fiscal Transparency, the Public Expenditure and Finance Accountability initiative (whose secretariat is hosted by the World Bank), the Organisation for Economic Co-operation and Development's (OECD) Best Practices for Fiscal Transparency, and the International Organisation of Supreme Audit Institutions' Lima Declaration of Guidelines of Supreme Audit Precepts. The strength of such guidelines lies in their universal applicability to different budget systems around the world, including countries with different income levels.

The Open Budget Survey 2015 is a collaborative research process in which IBP worked with civil society partners in 102 countries over the past 18 months. The 102 countries cover all regions of the world and all income levels.

The results for each country in the 2015 Survey are based on a questionnaire, comprising 140 questions, that is completed by researchers typically based in the country surveyed. Almost all of the researchers responsible for completing the questionnaire are from academic institutions or civil society organisations. Most of the researchers belong to organisations with a significant focus on budget issues.

Most of the Survey questions require researchers to choose from five responses. Responses "a" or "b" describe best or good practice, with "a" indicating that the full standard is met or exceeded, and "b" indicating the basic elements of the standard have been met. Response "c" corresponds to minimal efforts to attain the relevant standard, while "d" indicates that the standard is not met at all. An "e" response indicates that the standard is not applicable, for example, when an OECD country is asked about the foreign aid it receives. Certain other questions, however, have only three possible responses: "a" (standard met), "b" (standard not met), or "c" (not applicable).

Once completed, the questionnaire responses are quantified. For the questions with five response options: "a" receives a score of 100, "b" receives 67, "c" receives 33, and "d" receives 0. Questions receiving "e" are not included in the country's aggregated scores. For the questions with three response options: "a" receives 100, "b" receives 0, and "c" responses are not included in the aggregated score.

For the 2015 Survey, researchers began collecting data for the Survey in May 2014 and completed the questionnaire for their country by the end of June 2014. The Open Budget Survey 2015 thus assesses only events, activities, or developments that occurred up to 30 June 2014; any actions occurring after this date are not accounted for in the 2015 Survey results.¹⁷

All responses to the Survey questions are supported by evidence. This includes citations from budget documents; the country's laws; or interviews with government officials, legislators, or experts on the country's budget process. Throughout the research process, IBP staff assisted the researchers in following the Survey methodology, particularly the guidelines for answering Survey questions.

¹⁷ Open Budget Survey 2015 Website, Methodology, <http://www.internationalbudget.org/opening-budgets/open-budget-initiative/open-budget-survey/research-resources/methodology/>

Upon completion, IBP staff members analysed and discussed each questionnaire with the individual researchers over a three- to six-month period. IBP sought to ensure that all questions were answered in a manner that was internally consistent within each country, and consistent across all survey countries. The answers were also cross checked against published budget documents and reports on fiscal transparency issued by international institutions, such as the IMF, World Bank, and the OECD.

Each questionnaire was then reviewed by an anonymous peer reviewer who has substantial working knowledge of the budget systems in the relevant country. The peer reviewers, who were not associated with the government of the country they reviewed, were identified through professional contacts and variety of other channels.

IBP also invited the governments of nearly all survey countries to comment on the draft Survey results. The decision to invite a government to comment on the draft results was made after consulting with the relevant research organization responsible for the Survey. Of the 98 governments that IBP contacted, 53 commented on the Survey results for their country.

Hence, each country's OBI 2015 score is calculated from a subset of 109 Survey questions. Though each of the eight key budget documents assessed may have a different number of questions related to it, the OBI score is a simple average of all 109 questions which had responses a, b, c, or d. In calculating the OBI scores, no method of explicit weighting was used.

Though using a simple average is clear, it implicitly gives more weight to certain budget documents than others. In particular, 54 of the 109 OBI questions assess the public availability and comprehensiveness of the Executive's Budget Proposal, and thus are key determinants of a country's overall OBI score. In contrast, the Citizens Budget and the Enacted Budget are the focus of only four and six questions, respectively.

This implicit weighing is justified. From a civil society perspective, the Executive's Budget Proposal is the most important budget document, as it lays out the government's budget policy objectives and plans for the upcoming year. It typically provides details on government fiscal policies not available in any other document. Access to this information is critical for civil society to understand and influence the budget prior its approval, and to have as a resource throughout the year. That said, one of the changes IBP introduced to the questionnaire was to increase the emphasis on the other seven budget documents, reflecting their role in ensuring sufficient information is provided throughout the entire budget cycle.

3. Indices without a clearly transparent methodology

This section presents indices that provide data without publishing their methodology in detail. Nevertheless, they are widely used and cited as informed and accurate sources of information.

The "*Corruption Perception Index*" (<https://www.transparency.org/cpi2015/>) produced by *Transparency International* scores and ranks countries/territories based on how corrupt a country's public sector is perceived to be. The aim of the tool is to raise awareness about the consequences of corruption and help to fight it. It is based on perceptions rather than hard empirical data, because corruption generally comprises illegal activities, which are deliberately hidden and only come to light through scandals, investigations or prosecutions. Thus, there is no meaningful way to assess absolute levels of corruption in countries or territories on the basis of hard empirical data. Possible attempts to do so, such as by comparing bribes reported, the number of prosecutions brought or studying court cases directly linked to corruption, cannot be taken as definitive indicators of corruption levels. Instead, they show how effective prosecutors, the courts or the media are in investigating and exposing corruption. Capturing perceptions of corruption of those in a position to offer assessments of public sector corruption is the most reliable method of comparing relative corruption levels across countries. As such, however,

although their methodology is publicly available, it has a certain level of vagueness. Nevertheless, the Index is one of the most cited indicators worldwide.

The Freedom House annual “*Freedom in the World*” report (https://freedomhouse.org/sites/default/files/FH_FITW_Report_2016.pdf) assesses the condition of political rights and civil liberties around the world. Being published since 1973, it has become one of the most widely read and cited report of its kind, used on a regular basis among others by policymakers, journalists, academics, activists practitioners interested in the state of democracy and human freedom.

Freedom in the World 2016 evaluates the state of freedom in 195 countries and 15 territories during calendar year 2015. Each country and territory is assigned between 0 and 4 points on a series of 25 indicators, for an aggregate score of up to 100. These scores are used to determine two numerical ratings, for political rights and civil liberties, with a rating of 1 representing the most free conditions and 7 the least free. A country or territory’s political rights and civil liberties ratings then determine whether it has an overall status of Free, Partly Free, or Not Free.

The note available about their methodology states that indicators are derived from the Universal Declaration of Human Rights and applied to all countries and territories, irrespective of geographic location, ethnic or religious composition, or level of economic development. *Freedom in the World* assesses the real-world rights and freedoms enjoyed by individuals, rather than governments or government performance per se. Political rights and civil liberties can be affected by both state and non-state actors, including insurgents and other armed groups.

External analysts assess the 210 countries and territories, using a combination of on-the-ground research, consultations with local contacts, and information from news articles, nongovernmental organizations, governments, and a variety of other sources. Expert advisers and regional specialists then vet the analysts’ conclusions. The final product represents the consensus of the analysts, advisers, and Freedom House staff. For each country and territory, Freedom in the World analyses the electoral process, political pluralism and participation, the functioning of the government, freedom of expression and of belief, associational and organisational rights, the rule of law, and personal autonomy and individual rights. Their findings are easily accessible online.

Similarly, indices that are widely cited but also lack a transparent methodology include:

- Freedom House: Countries at Crossroads, <https://freedomhouse.org/report-types/countries-crossroads>
- Transparency International: Bribe Payers Index, <http://www.transparency.org/research/bpi/overview>
- Milken Institute: Opacity Index, <http://www.milkeninstitute.org/publications/view/384>
- Peoples Under Threat Index, <http://peoplesunderthreat.org/about-the-peoples-under-threat-index/>
- Varieties of Democracy, <https://www.v-dem.net/en/>

C. Assessment and Recommendations

In the final section of this report, we offer an overall assessment of the feasibility of a pluralism audit tool and make some recommendations on likely candidate models and suitable next steps. We begin with some broader preliminary observations that place the audit tool in the context of GCP’s more general work and then provide a more detailed set of assessments and recommendations for the further development of a pluralism audit tool. We then conclude by briefly pointing out some possible candidates on which GCP’s pluralism audit tool could be modelled.

1. Preliminary observations

Audit tools in both the narrower and broader sense in which we understood them in this report can serve a variety of different purposes, one of which tends to be facilitation of change in relation to a status quo perceived to fall short of expectations. These expectations, in turn, are commonly expressed in the form of scores against an established benchmark. In addition, there is also an assumption that it is not only absolute scores against benchmarks that matter but also performance relative to other entities (countries, institutions, organisations) that are subjected to the same audit. Hence, one frequent use of audit tools is their application for the purpose of global or regional rankings that establish performance in relation to certain benchmarks and simultaneously rank entities in terms of their performance. This would suggest that GCP should consider the development of an audit tool in the context of establishing pluralism benchmarks and a pluralism index: well-defined benchmarks will be critical to give practical meaning to an audit tool; an audit tool, while in some sense an independent instrument to be used by others to assess the state of pluralism, for example, in their country, region, or city, can simultaneously provide the methodology underpinning the construction of a national, regional, or global pluralism index; and a pluralism index thus constructed can serve as one form of validation for the utility of an audit tool and the relevance of its benchmarks. **A necessary decision for GCP to make, therefore, is the extent to which the development of an audit tool should proceed closely linked to the development of a pluralism index.**

Audit tools, benchmarks, and indices, while conceptually closely connected, nonetheless can serve distinct purposes. From GCP's perspective, a logical assumption would be that their primary purpose relates to achieve, and sustain, high levels of pluralism. This would involve

- Establishing present levels of pluralism (i.e., conducting an initial audit);
- Identifying dimensions in which there are shortfalls against set benchmarks (i.e., analysing initial audit findings and scoring);
- Determining actions aimed at mitigating these shortfalls (i.e., drawing on comparative knowledge and understanding, based for example on insights from a global pluralism index, in order to improve institutional performance to achieve higher levels of pluralism);
- Making commitments about timelines in which such mitigation policies are implemented and what levels of resource will be made available to achieve this; and
- Monitoring the fulfilment of these commitments either on an ongoing basis or by way of a follow-up audit.

Related to the possible combination of work on an audit tool, benchmarks, and index, GCP thus needs to decide which of these second-order purposes are priorities and why, which in turn is related to identifying a suitable theory of change, related to the audit tool (as well as possibly a pluralism index).

2. Issues in designing an audit tool

Focusing on the practicalities of the audit tool,¹⁸ the overview of existing instruments and the more detailed profiles of particular such tools in the preceding section has highlighted three critical questions for the design of a feasible and viable audit tool, i.e., a tool that can be implemented and that will produce relevant findings and insights:

- a) Who conducts the audit?
- b) How is the audit conducted?
- c) What happens with the audit results?

¹⁸ Note that audits are potentially very resource and labour intensive. While we do not assess this particular dimension here any audit tool to be developed should include some guidance on costs.

We discuss these three issues in turn with a focus on what they imply for the design of a pluralism audit tool, before concluding with a few reflections on possible candidates on which a pluralism audit tool could be modelled in the following, and final, section of this report.

a) Who conducts the audit?

This question is primarily about how local the audit team should be, and the range of options extends from a purely local team to a purely international (external) team, with most existing tools relying on a mix of auditors.¹⁹ Different rationales underpin these choices: on the one hand, one could argue along the lines of the International IDEA democracy that really only people who have direct experience of living in a specific country should audit the quality of its democracy. On the other hand, there may be circumstances, explicitly acknowledged by IREX, where local auditors would be exposed to undue risks if they conducted an audit, and the audit team, therefore, is composed of purely external members.

Beyond such clearly context-dependent considerations, the overarching consensus appears to be that what matters most is that the audit team as a whole and all of its individual members are possessed of a high degree of integrity and do not have a conflict of interest in relation to the audit.

Assuming that GCP's audit tool will be freely available for anyone to use, the Center would have limited control over the composition of teams that choose to conduct a pluralism audit using GCP's tool. That said, it would be useful to provide guidance on the composition of an audit team, as well as to offer training and capacity building to would-be auditors. Given the specific nature of pluralism as a focal point for an audit, such **guidance offered by the GCP should extend, at a minimum, to encourage appropriate levels of diversity among auditors and pinpoint particular areas of (subject and methodological) expertise that should be represented among auditors.**

b) How is the audit conducted?

Here the question relates primarily to issues of data gathering and analysis, as well as a review and validation of findings.

In terms of data gathering, and bearing in mind the specific nature of a pluralism audit tool, a mix of existing and original data and of perceptual and non-perceptual data would seem the most appropriate. This is, for the most part, the approach adopted by the majority of existing audit tools.²⁰ A balance between existing and original data is critical for the utility of the audit tool: as it should be as widely applicable as possible, guidance needs to be provided to would-be users about sources of existing data (such as national statistics, global and regional indices, etc.), as well as about potential sources of original data and the extent to which original can compensate for the absence, and/or add a corrective to, existing data.

Regarding original, (i.e., audit-specific) data, in their majority, existing tools primarily rely on surveys (using standardised questionnaires) and interviews and focus groups (key informants, experts, well-informed persons). Data gathered through surveys and interviews would, for the most part, be perception-based, i.e., reflect views and experiences of interlocutors. Such perception-based data also underpins a large number of indices. In other words, it is important to understand the methodologies by which existing data to be used in an audit was generated in order to ensure the right balance between perceptual and non-perceptual data.

¹⁹ Note that a purely local audit team could still gather data through interviews and focus groups from external experts or mixed panels of experts, while a purely external audit team could solely rely on local interlocutors. In other words, an important distinction needs to be drawn between auditors and their information sources.

²⁰ Among the exceptions, the Harvard Kennedy School Index of African Governance is the closest to an index of indices among the tools surveyed in this report.

Another source of audit-specific original data could be gathered regarding specific indicators that may be relevant for a pluralism audit. This could involve reports on crimes against particular communities, inflammatory media reports, cross-community development projects, etc. This might be useful as complement to, for example, existing national crime statistics or donor reports on aid spending. Critical in all of these cases would be a process of data review and validation in order to ensure that any subsequent analysis happens on the basis of credible data. Appropriate methods in this regard would be triangulation, fact-checking across multiple sources, and expert review.

Data analysis, similarly, would involve at a minimum a two-stage process of analysis and review of the analysis. The analysis stage would depend on the kind of data that has been gathered, and could involve a variety of statistical methods, discourse analysis, and (comparative) case studies. Crucial at this stage of the audit would be the application of a consistent scoring methodology for each indicator relative to a set benchmark, including illustrative descriptors for each score. This is normally performed at an individual auditor level first before it is reviewed by experts or panels of experts or in a process of discussion among auditors. A choice then also needs to be made about when the score for a particular indicator is a weighted or unweighted compound of individual scores or whether there is a 'hierarchy' of scorers where earlier scores can be changed later in the process. Scores in one dimension or area of the audit should then be combined into a single score for this particular dimension on the basis of a consistently applied formula.

As the GCP develops its audit tool, it would be useful to offer some sample methodologies for data gathering, data validation, and data analysis, as well as a methodology for scoring. This could be drawn from a more extensive methodology underpinning a potential GCP pluralism index.

c) What happens with the audit results?

Once an audit has been conducted, it should clearly establish how well a particular entity scores against set pluralism benchmarks and enable the identification of shortfalls, as well as of area in which levels of pluralism have been achieved that require sustaining, rather than increasing. Put differently, a pluralism audit would be equivalent to a 'needs assessment' for the entity that has been audited and can potentially also help in establishing priority areas for change. As such, it can inform government policies and donor programming with a view to achieve change, as well as serve as an advocacy tool to push governments (and donors) to commit to change and to follow through on these commitments. From a donor perspective, a pluralism audit would also serve to focus capacity-building in relevant areas identified as priorities for change.

From a GCP perspective, its pluralism audit tool might therefore also include suggestions and guidelines for follow-up that auditors can usefully undertake, depending on the purpose of the audit they have conducted. A potentially critical element in this would be to indicate pathways towards achieving higher levels of pluralism—both from the perspective of a process leading there (e.g., building effective and sustainable coalitions for change) and the actual substance of required policy and institutional changes. Additionally, it could be useful to indicate potentially multiple pathways resulting from particular policy and institutional changes and offering, to the extent possible, contingent impact and risk assessments.

The latter two points, and the question of audit follow-up more generally, raises the issue of the link of the audit tool with the wider work of GCP. Pluralism case studies and possibly a pluralism index would provide crucial knowledge and understanding of pathways to higher levels of pluralism and of the impact of particular policy and institutional choices. They would also serve as factor in risk mitigation for the GCP by creating a solid foundation for the use of insights

gained from a pluralism audit, thus enhancing the relevance of the audit and increasing the likelihood of it supporting positive change.

3. Towards a GCP pluralism audit tool

The obvious candidate on which a pluralism audit tool could be modelled is the International IDEA guide to *Assessing the Quality of Democracy*. This is an exceptionally detailed guide that offers a broad menu of options on what to assess, how to do it, and what to do with the assessment. It is clearly driven by a desire to support change towards a higher quality of democracy. The guide also offers a vast range of detailed questions to be asked, potential sources of data, and suggestions on the structure of the process of a democracy assessment from the decision to conduct it through to achieving the desired change.

This strength is, however, at the same time a potential weakness from the perspective of a pluralism audit. It offers a vast, almost confusing range of options. Critically, it also does not provide a clear set of benchmarks and is very thin on data analysis methodology, including a scoring methodology. As a result, it is useful to consider several other audit tools, including methodologies underpinning existing indices.

In terms of the detail provided on methodology, and the utility of the approaches adopted, IREX's *Media Sustainability Index* and Reporters without Borders' *Press Freedom* index offer a useful guide of how to approach data gathering and analysis, including scoring. Being focused on a specific issue, both offer a manageable number of audit dimensions: IREX has five 'objectives' with between seven and nine indicators each that are assessed by country-based expert panels, Reporters without Borders uses an online questionnaire with 87 questions across six indicators. Both then use a multi-stage process of data validation and analysis, and scoring.

IREX's focus on sustainability is additionally useful from the perspective that a pluralism audit would not only need to be able to identify levels of pluralism at or above a set benchmark but also offer recommendations for how to sustain them. Again, this should be seen in the context of IREX's index and the possibility of a pluralism index: monitoring levels of pluralism over time would enable the identification/forecasting of trends, including in relation to performance against benchmarks and in relation to policies. In other words, IREX offers a useful methodology to determine how sustainable a certain level of press freedom is, thus enabling the identification of priority areas for follow-up actions.

The other useful complement to IDEA's democracy audit concerns the way in which the Quality of Government Institute presents its *Quality of Government Datasets*, distinguishing three different 'topics': What it is; How to get it; and What you get. As noted earlier, the first set of data focuses on what might be considered the core of an audit tool, i.e., a range of critical dimensions and their indicators which could then be measured and scored as part of the audit. The other two types of data could usefully inform follow-up actions:

- By helping to establish appropriate institutional and policy changes to achieve higher levels of pluralism in under-performing areas (or policies and institutions that can sustain pluralism levels at or above set benchmarks;
- By assisting in the identification of the likely impact of changing or maintaining particular policies and/or institutions linked to aspects of pluralism.

This would require a close linking of the audit tool with GCP's pluralism case studies and a potential pluralism index. The clear, tangible benefit of such a connection would be the greater likelihood of the audit tool serving as an effective instrument for positive change, while providing greater coherence to GCP's work as a whole—from the conceptual work on benchmarks through to the deep engagement of its policy work.

References

- Broome, A. and J. Quirk (2015). "Governing the world at a distance: the practice of global benchmarking." Review of International Studies 41(Special Issue 05): 819-41.
- Clegg, L. (2010). "Our Dream is a World Full of Poverty Indicators: The US, the World Bank, and the Power of Numbers." New Political Economy 15(4): 473-92.
- Clegg, L. (2015). "Benchmarking and blame games: Exploring the contestation of the Millennium Development Goals." Review of International Studies 41(Special Issue 05): 947-67.
- Dominique, K. C., A. A. Malik and V. Remoquillo-Jenni (2013). "International benchmarking: Politics and policy." Science and Public Policy 40(4): 504-13.
- Freistein, K. (2015). "Effects of Indicator Use: A Comparison of Poverty Measuring Instruments at the World Bank." Journal of Comparative Policy Analysis: Research and Practice, 10.1080/13876988.2015.1023053: 1-16.
- Fukuda-Parr, S. (2006). "Millennium Development Goal 8: Indicators for International Human Rights Obligations?" Human Rights Quarterly 28(4): 966-97.
- Giannone, D. (2010). "Political and ideological aspects in the measurement of democracy: the Freedom House case." Democratization 17(1): 68-97.
- Guterman, E. (2014). "The legitimacy of transnational NGOs: lessons from the experience of Transparency International in Germany and France." Review of International Studies 40(02): 391-418.
- Harrison, J. and S. Sekalala (2015). "Addressing the compliance gap? UN initiatives to benchmark the human rights performance of states and corporations." Review of International Studies 41(5): 925-45.
- Heywood, P. M. and J. Rose (2014). "'Close but no Cigar': the measurement of corruption." Journal of Public Policy 34(03): 507-29.
- Homolar, A. (2015). "Human security benchmarks: Governing human wellbeing at a distance." Review of International Studies 41(Special Issue 05): 843-63.
- Kelley, J. G. and B. A. Simmons (2015). "Politics by Number: Indicators as Social Pressure in International Relations." American Journal of Political Science 59(1): 55-70.
- Kuzemko, C. (2015). "Climate change benchmarking: Constructing a sustainable future?" Review of International Studies 41(Special Issue 05): 969-92.
- Langbein, L. and S. Knack (2008). "The worldwide governance indicators and tautology: causally related separable concepts, indicators of a common cause, or both?" World Bank Policy Research Working Paper Series, Vol.
- Langbein, L. and S. Knack (2010). "The Worldwide Governance Indicators: Six, One, or None?" The Journal of Development Studies 46(2): 350-70.
- Lebaron, G. and J. Lister (2015). "Benchmarking global supply chains: the power of the 'ethical audit' regime." Review of International Studies 41(05): 905-24.
- Porter, T. (2015). "Global benchmarking networks: the cases of disaster risk reduction and supply chains." Review of International Studies 41(Special Issue 05): 865-86.
- Power, M. (2003). "Evaluating the Audit Explosion." Law & Policy 25(3): 185-202.
- Raworth, K. (2001). "Measuring Human Rights." Ethics & International Affairs 15(1): 111-31.
- Rosga, A. J. and M. L. Satterthwaite (2008). "The trust in indicators: Measuring human rights." Berkley Journal of International Law (BJIL), Forthcoming: 08-59.
- Seabrooke, L. and D. Wigan (2015). "How activists use benchmarks: Reformist and revolutionary benchmarks for global economic justice." Review of International Studies 41(Special Issue 05): 887-904.
- Sending, O. J. and J. H. S. Lie (2015). "The limits of global authority: World Bank benchmarks in Ethiopia and Malawi." Review of International Studies 41(Special Issue 05): 993-1010.
- Thomas, M. A. (2010). "What do the worldwide governance indicators measure?" European Journal of Development Research 22(1): 31-54.