

# Sustainability Benchmarking Good Practice Guide

Version 1.0 | July 2019





# Contents

<b>1. Preamble</b>	2
<b>2. Scope</b>	3
<b>3. Audience</b>	3
<b>4. Definitions</b>	4
<b>5. Principles</b>	5
<b>6. How Benchmarking is Applied</b>	7
<b>7. Benchmarking Models</b>	9
<b>8. Benchmarking Development Process</b>	12
8.1 Determine the purpose of the benchmark and who it is for	13
8.2 Determine who or what is being benchmarked	14
8.2.1 Type of entity	14
8.2.2 Supply chain scope	15
8.2.3 Market segmentation	15
8.3 Determine who will manage the benchmarking process	15
8.4 Determine the content of the benchmark	16
8.4.1 Things to consider when developing the benchmark	16
8.4.2 Core elements for benchmarking sustainability standards	18
8.4.3 Alignment with international norms and guides	19
8.4.4 Accommodating diverse approaches in a benchmark	20
8.5 Determine the evaluation structure of the benchmark	22
8.6 Determine the benchmarking methodology	23
8.6.1 Steps in implementing the benchmarking process	23
8.7 Determine how the results will be communicated	26
8.8 Further considerations	27
8.8.1 Competence	27
8.8.2 Cost and complexity	27
8.8.3 Improving the benchmark and process	28
8.8.4 Government use of benchmarks	28
<b>Annex 1: Summary of Sustainability Benchmarking Models</b>	29
<b>Annex 2: Benchmarking Good Practice Checklist</b>	30
<b>Annex 3: Benchmark Criteria for Evaluating Sustainability Standards</b>	32

The following symbols are used throughout this guidance document:



**ACTION**



**TAKE NOTE**



**IDEA**



# 1. Preamble

The rising prominence of environmental crises and social upheaval, coupled with ambitious global commitments like the UN Sustainable Development Goals (SDGs), has pushed sustainability into the mainstream.

Companies, governments and civil society are seeking tools they can use to evaluate sustainability performance and to recognise and reward good practice. For stakeholders wondering which tools to use and approaches to take, the landscape can seem both bewildering and challenging to navigate. One response to this challenge has been the development of benchmarks to evaluate, compare and qualify sustainability tools and company performance.

A benchmark is a reference point against which something is evaluated. Sustainability benchmarks can evaluate a wide variety of entities, from the sustainability performance of companies to the rigour of sustainability standards and certification. Benchmarking programmes define a specific reference point and carry out evaluations of sustainability policies, practices and tools against it. In this way, benchmarking programmes chart a path through the wilderness, providing users with comparable information about the benchmarked entities that then allows those users to choose between them.

The challenge with benchmarking is that there has been little guidance on how to develop and implement a credible benchmarking programme, leading to a proliferation of efforts of varying rigour, transparency and effectiveness. This is significant because it means that these programmes have the potential to recognise and reward lower performers, potentially limiting the effectiveness of our collective response to today's sustainability challenges.

Through this guidance, ISEAL aims to contribute a framework and practical set of good practices and considerations for those organisations and initiatives considering whether to carry out a benchmarking exercise or develop a benchmarking programme. We also aim to publicise existing benchmarking programmes as a means to limit their proliferation. Our goal is to support increased consistency in, and strengthening of, sustainability benchmarking programmes so that they can effectively support better practices and a faster transition to a more sustainable world.

## 2. Scope

This guidance applies to any ongoing benchmarking programme for analysing or evaluating sustainability initiatives or performance. It can also be applied to one-off benchmarking exercises.

The guidance covers benchmarks developed by any type of organisation, including companies, governments, NGOs, sustainability standards and others. A few sections, identified in the guidance, have been developed to support benchmarking of sustainability standards and certification specifically and do not apply to benchmarking of other entities.

The guidance does not propose criteria or requirements to be used in a benchmark but sets out considerations for developing those requirements.

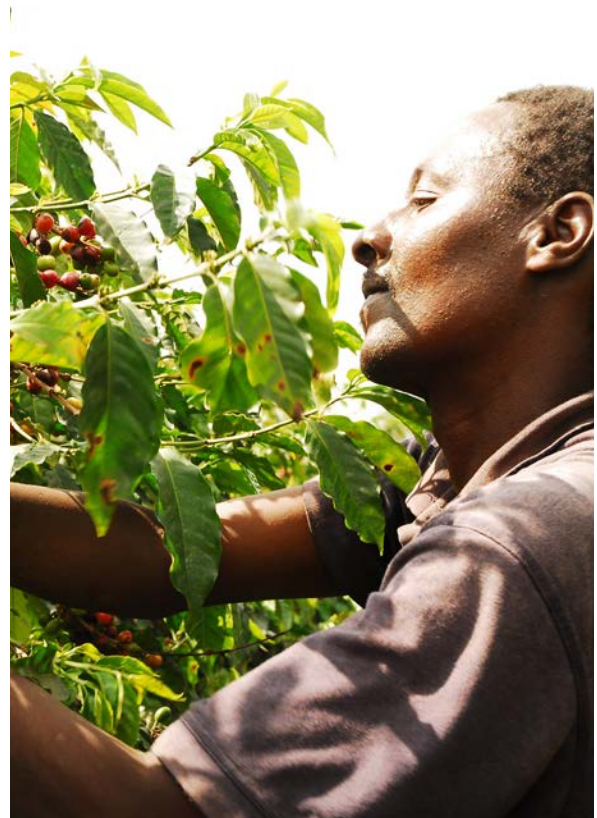


## 3. Audience

The primary audiences for this guidance are those organisations and initiatives involved in setting up or carrying out benchmarking exercises or programmes.

The guidance can be used as a reference tool for these initiatives.

The guidance is also applicable for users of benchmarks or the entities that are benchmarked, to better understand what information they should be looking for or questions they should be asking from the convenor of a benchmarking programme. Annex 2, a checklist of benchmarking good practices, is particularly helpful in this context.







## 4. Definitions

### Benchmark

- A benchmark is the reference point against which something is evaluated (noun)
- To benchmark is the act of determining (or judging) alignment with the fixed reference point (verb)

### Benchmarking exercise

- A 'one-off' benchmark, which is not part of a long-term programme, policy or strategy. For example, a scoping report delivered by a consultant to inform a policy process might include a benchmarking exercise but does not constitute a benchmarking programme.

### Benchmarking programme

- A structured and systematic way of carrying out evaluations against benchmarks, often coupled to specific organisational or policy goals.

### Entity

- In this context, the subject of a sustainability benchmark. This can be a sustainability standard, company, NGO or other stakeholder.

### Convenor

- The organisation that leads development of a benchmarking exercise or programme and makes key decisions about its purpose, structure and process. The convenor can implement the benchmarking programme or outsource this to external experts or consultancies.



## 5. Principles

Not all benchmarking exercises or programmes will look the same. Different programmes will have different goals or objectives that inform their structure. Regardless of individual approaches, a number of core principles are applicable across all benchmarking exercises and programmes and can be used as a point of reference for decisions about how to develop and implement the programme. These principles are derived from the [\*ISEAL Credibility Principles\*](#).

### Transparency

***Relevant information is made freely available in an accessible manner.***

In the context of benchmarking, this means that interested stakeholders have access to information about the mission of the benchmarking programme, the criteria, how the benchmark was set and how it is being implemented, the results of the programme, and what those results mean. Transparency also relates to the extent of publicly available information about the performance of the entities being benchmarked.

### Rigour

***Benchmarking exercises and programmes are structured and implemented in ways that are sufficient to produce quality outcomes.***

This means that the benchmark content is sufficiently detailed and clear, while the benchmarking process is robust<sup>1</sup> and implemented consistently by individuals who are competent for their roles.

---

1. A robust benchmarking process is one which is consistent with these guidelines

## Stakeholder engagement

***Interested stakeholders have appropriate opportunities to participate in and provide input to the process.***

This means that programmes have identified interested stakeholders and should determine whether and how they can provide input to the development of the benchmark or to its subsequent implementation, and to make those opportunities<sup>2</sup> available to them.

## Impartiality

***Benchmarking exercises and programmes identify and mitigate conflicts of interest throughout their operations.***

This means that organisations involved in the benchmarking programme are not engaged in activities that would compromise the integrity of the results. Where one organisation carries out an internal benchmark of other organisations against itself, this principle would imply that the results only be used for internal purposes.

## Efficiency

***Benchmarking exercises and programmes are structured as simply as possible and avoid redundancies.***

This means that a new benchmarking programme should only be developed if its objectives can't be fulfilled by existing initiatives, that the programme is not unduly complex, that the number of benchmarking criteria are fit for purpose and not overly prescriptive, and that the criteria are aligned as much as possible with the most relevant existing benchmarking initiatives

## Improvement

***Benchmarking exercises and programmes are structured to incentivise better practices in the entities that they cover.***

Ideally, this means benchmarking programmes stimulate a 'race to the top' rather than recognizing practices that meet the lowest common denominator. Benchmarking programmes support mechanisms such as gap analyses that show where benchmarked entities fall short of the benchmark and that encourage and reward improvements. Convenors of benchmarking programmes also revise and improve the benchmark itself, based on learnings from its implementation.

## Accessibility

***Benchmarking exercises and programmes avoid structures that create unnecessary barriers to participation and seek to minimise the reporting and engagement burden for entities being benchmarked.***

This means that the benchmark and accompanying procedures are appropriate, easy to understand, and broadly applicable, that any associated fees or other requirements do not create significant burdens that would prevent or inhibit participation, and that requests for information or engagement by benchmarked entities are limited and clear.

---

2. Appropriate opportunities could include providing input to the setting of the benchmark, providing feedback on the evaluation of entities against the benchmark, or having a complaints process available to stakeholders.





## 6. How Benchmarking is Applied

### Benchmarking is an important tool for a number of reasons:

- It provides the user with evidence to recognise, use or support a given sustainability initiative or entity in order to achieve defined sustainability objectives;
- It promotes consistency of performance and alignment between the benchmarked entities;
- It improves transparency about the operations and performance of the benchmarked entities;
- As a result, it helps to build awareness and uptake of these entities;
- At its best, it has the potential to improve the rigour and effectiveness of the benchmarked entities through a 'race to the top'; and
- It creates a guidepost for stakeholders and other users to know what acceptable practice looks like.







### Take Note

While there may be good reasons to develop a new benchmarking programme, the default approach should be to first determine whether it is possible to reference existing benchmarking initiatives, either in whole or in part. For example, industry association benchmarks can decrease the need for individual benchmarks carried out by individual companies. New benchmarking initiatives should only be developed where an unmet need has been identified through a formal needs assessment. Descriptions of a number of existing benchmarking programmes are available as fact sheets through the [\*UN International Trade Centre \(ITC\) Sustainability Map\*](#) initiative.



Benchmarking is used by multiple actors for multiple purposes. Defining the purpose of the benchmark is critical (see 8.1) and differences in organisational or policy objectives explain why a variety of benchmarks are developed. The table below illustrates some generic applications of benchmarks by different stakeholders.

**TABLE 1: Stakeholder applications of benchmarks**

	 Governments	 Buying companies	 NGOs	 Finance institutions	 Producing enterprises	 Sustainability standards
Implementation mechanism to fulfil regulations						
Sustainable procurement policies						
Transparency for users of sustainability initiatives or entities						
Risk filter for investment						
Reduce audit burden of multiple audits						
Alignment or equivalency tool for unilateral or mutual recognition						
Set expectations of benchmarked entities – drive consistency and improvement						



**Take Note**

It is important to note that benchmarking programmes do not always achieve positive results and may have unintended negative consequences. For example, where a benchmarking programme omits relevant criteria in its evaluation, the benchmark may create a mechanism to legitimise poor performing entities and thus encourage a race to the bottom. Where possible, these unintended consequences should be avoided.



## 7. Benchmarking Models

Benchmarks are used for variety of purposes and can apply to a wide range of entities. This guidance identifies four basic models that are commonly applied in sustainability benchmarks. This section provides information about each model, which can be useful for determining the structure of a benchmarking initiative.

A benchmarking programme or exercise can be structured in different ways, the most common of which are<sup>3</sup>:

### **Threshold**

focused on determining whether entities meet a performance level defined by the benchmark, possibly resulting in recognition.

### **Ranking**

focused on ordering the performance of similar entities against the benchmark, from highest to lowest performing.

### **Peer comparison**

often an internal exercise focused on assessing organizational or programmatic performance in comparison with similar entities

### **Improvement**

focused on gap analysis with an aspirational performance benchmark, to guide further development or strengthening of the benchmarked entities

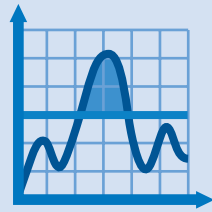
It is also quite common to develop benchmarks that are a combination of these models, such as when a threshold benchmark also includes mechanisms to stimulate improvement.

---

*3. See also Annex 1 for a brief comparison of these benchmarking models*



## Benchmarking Models



Threshold



Ranking

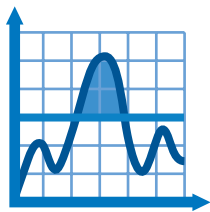


Peer comparison



Improvement

FIGURE 1: Common benchmarking models



### Threshold

This is the most common model of external benchmarking exercise or programme, enabling the convenor to establish a common baseline or threshold

for qualifying benchmarked entities. That qualification can have multiple applications and incentives attached to it, from governments recognising standards in their allocation of subsidies or procurement contracts, to finance sector investment risk screens, through to companies or business platforms determining which standards are acceptable for use in supply chain policies.

With recognition serving a variety of different purposes, it follows that the nature of the performance threshold will vary significantly between benchmarking exercises. The convenor of a benchmarking programme will need to choose whether the performance threshold should be 'broad and shallow', covering a range of issues superficially, e.g. for determining some broad concept of market acceptability, or whether it should have 'deep and narrow' thresholds that can be very specific and focus in-depth on one priority area like modern slavery or child labour.

An example of a threshold benchmark is the German government's *Siegelklarheit* portal which provides a simple framework for consumers to know which sustainability certification labels meet the minimum requirements set by the German government. The *Kompass Nachhaltigkeit* (Sustainability Compass) for sustainable public procurement is based on the same underlying architecture but provides users with the additional option to determine which practices are important for their own procurement specifications. Both platforms enable users to drill down into the criteria to understand more about the sustainability practices and operations required in each standard.



### Take Note

Implicit in choosing a threshold model benchmark is that the benchmark itself must reflect the minimum set of performance and operational requirements that are acceptable to the convenor. The benchmarking process will require sufficient rigour to determine whether the benchmarked entities meet that performance threshold (see also section 8.6).





## Ranking

Rather than setting a performance threshold, this approach focuses on measuring the performance of entities relative to each other. The convenor still has to decide which criteria to use as the basis for the benchmark and how to weight or score those criteria, but rather than stating what are sufficient practices, the entities are evaluated and ranked in order of performance. This model is particularly prevalent in corporate benchmarks but can be applied equally to sustainability standards and other entities. There is some overlap with improvement models to the extent that ranking of entities creates an incentive for lower ranking entities to improve their performance.

Examples of ranking benchmarks include the *Corporate Human Rights Benchmark (CHRB)*, which assesses and ranks 101 of the largest publicly traded companies in the world on a set of human rights indicators, and Oxfam's *Behind the Brands* benchmark, which assesses the agricultural sourcing policies of the world's 10 largest food and beverage companies. These benchmarks incentivise companies to improve their business practices and empower stakeholders (such as investors) with the information needed to make informed decisions.



## Peer Comparison

Many companies, sustainability standards, and other organisations carry out benchmarking exercises of their peers and competitors to understand how they are performing

in relation to those peers and, potentially, where there are overlaps and possible collaboration opportunities. In this model, the benchmark against which other entities are evaluated is the current practices or performance of the organisation carrying out the evaluation. These benchmarking exercises tend to be less formal, with the results not intended for broader distribution. They also share common purposes with both ranking and improvement benchmarks. As these benchmarks are most often carried out internally, examples are not readily shared.



## Improvement

Less common but no less important are benchmarking programmes that aim to create drivers for improvement in the benchmarked entities. These benchmarks are often structured

as aspirational performance bars that may go beyond current practices but that provide a roadmap for expectations about future performance. As practices improve, these benchmarks can be revised upwards to reflect future aspirations on sustainability performance and to instil an ethic of continual improvement. In practice, these have been applied especially by NGOs and by governments to incentivise sustainability standards and influence the definition of acceptable practices.



### Take Note

The implication for improvement model benchmarks is that they should be structured as aspirational criteria, sometimes going beyond good practice to look at best in class. They reflect the convenor's views of what best practice should look like. These benchmarks are regularly reviewed and revised as best practice evolves, and generally require a higher degree of engagement with the entities being benchmarked.





## 8. Benchmarking Development Process

The process to develop a benchmarking exercise or programme requires consideration of various design elements that will determine how the programme is structured and implemented. This section provides practical guidance for each of these elements, highlighting decisions that need to be made and the implications of different choices.

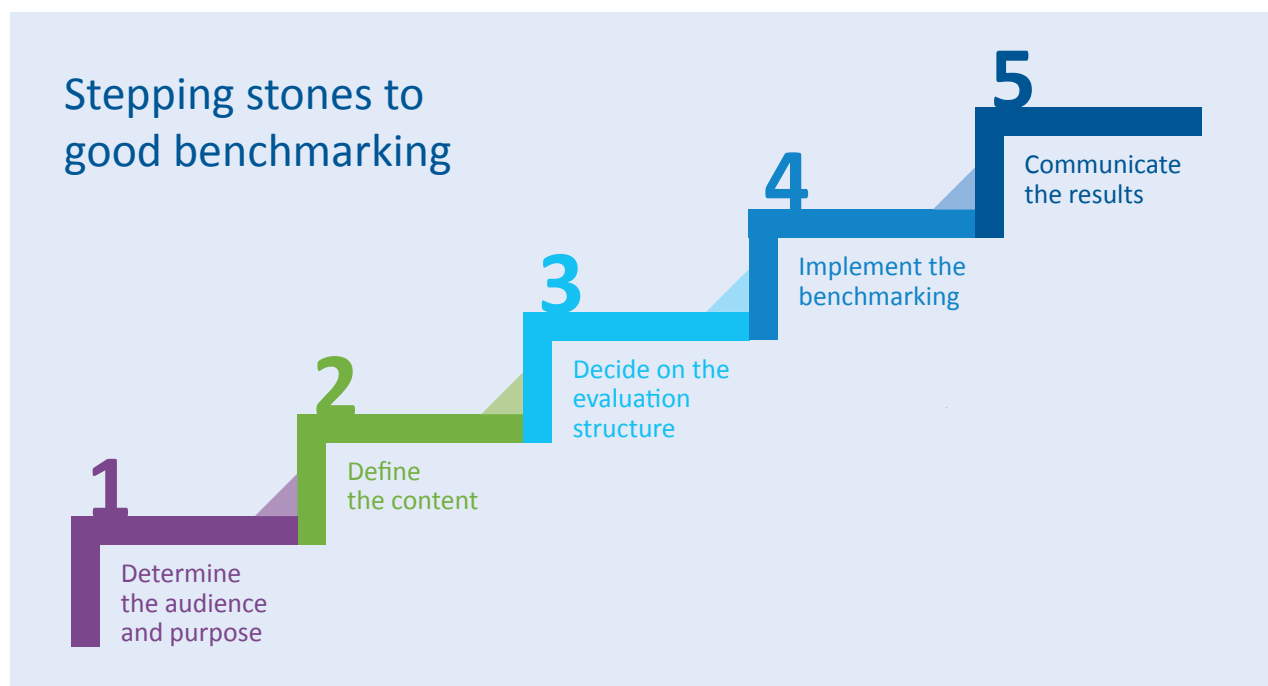


FIGURE 2: Stepping stones in the benchmarking development process

# 8.1 Determine the purpose and audience of the benchmark

The first step in thinking about a benchmarking programme is to identify the primary and secondary audiences and to understand how they would use the results of a benchmark. For example, government agencies interested in setting sustainability guidelines for procurement may have significantly different needs and expectations than NGOs wanting to strengthen performance requirements in sustainability standards. A useful exercise is to speak with potential users directly to better understand their expectations and desired uses for a benchmark.



A critical step in the development of a benchmarking exercise or programme is to then employ what has been learned about users' needs to define the purpose of the benchmarking, essentially what

you are trying to achieve. For example, the purpose may be to identify and reward leading companies in the mining sector, or to identify which sustainability standards need to strengthen the way they address gender issues. The purpose will help in determining the structure of the benchmarking programme, choosing what benchmarking model is used, setting the performance level of the benchmark itself, and communicating about the benchmark.

## ! Take Note

The purpose of any benchmarking programme should be made explicit and included in all communication about the programme. This will also help to inform the types of claims that benchmarked entities and users of the benchmarking programme can make about the results.






## ! Take Note

Convenors should also assess whether a benchmarking programme is the most effective strategy to achieve their purpose. Part of that assessment should include whether existing benchmarks already meet all or part of the articulated goal.





## To what will the benchmark apply?

 Sector or commodity	 Geography	 Type of entity	 Supply chain scope*	 Market segmentation
Agriculture	Global	Companies	Products or services	Only market leaders
Mining	Region x	Sustainability standards	Supply chain activities	All entities within scope
Etc.	Country y	Other	Production or extraction	

\*Applies primarily to sustainability standards

FIGURE 3: Determining the scope of a sustainability benchmark

## 8.2 Determine who or what is being benchmarked

Within the scope of sustainability benchmarks, there can be a range of entities being benchmarked, from companies and corporate performance to the sustainability standards and related tools that companies use to convey their sustainability credentials. A decision about who or what is being benchmarked will greatly influence how the benchmark content is set. The following diagram shows a series of decisions that need to be made to arrive at a definition of the scope:

### 8.2.1 Type of entity

Once the sector and geography have been determined, the principle decision a convenor needs to make is whether the benchmark will focus on evaluating sustainability standards and initiatives or on corporate action towards sustainability, or some combination of these. This decision is driven by the defined purpose or goal of the programme, as the following examples highlight:

- To give clear guidance on how companies can operationalize the OECD's Due Diligence Guidance for Responsible Business Conduct, the *OECD is carrying*

*out 'alignment assessments' or benchmarks for different sectors.* In 2018 the OECD's benchmarking programme on mining and minerals concluded its first alignment assessment of six related industry initiatives and sustainability standards, assigning them into one of three categories (not aligned, partially, or fully aligned) based on its benchmarking score. The benchmarking also resulted in a gap assessment that was used to strengthen alignment of these standards with the OECD Guidance.

- The *World Benchmarking Alliance (WBA)* develops benchmarks to incentivise the private sector to contribute to seven system transformations that are fundamental to achieving the SDGs. These benchmarks are developed through an extensive multi-stakeholder process and build on best available science, existing principles and normative standards, corporate reporting frameworks, and sector-, product- and issue-specific initiatives. In implementing the benchmark, WBA collects data from public sources such as company reports and websites, directly from companies through questionnaires, and from third-party sources wherever relevant. Based on the collected data, companies are then assessed, scored and ranked, with the resulting benchmarks made freely available to serve as sector roadmaps.

- The German government's Green Button initiative is an unusual hybrid in which the benchmarking programme seeks to evaluate both a company's performance and the sustainability credentials of its products, based on their compliance with responsible production standards. This approach can be achieved by combining what is essentially two benchmarking exercises.

### 8.2.2 Supply chain scope

Another decision to be made, appropriate primarily to benchmarks of sustainability standards, is the supply chain scope to be covered by the benchmark. Benchmarking programmes can focus on standards that apply only at the production phase of the supply chain (e.g. deforestation legality) or can look at performance across all stages of the supply chain (e.g. production and factory labour standards).

### 8.2.3 Market segmentation

The final decision focuses on which entities to include in the benchmark. For a corporate benchmark,

this means determining which types of companies to include. Is the benchmark focused only on companies of a certain size, sector, market presence, capitalisation, etc.? A similar exercise is necessary where the benchmark focuses on sustainability standards and similar tools. Decisions need to be made about the types of standards, sectors to which they apply, and scope.

#### Take Note

Where benchmarks aim to recognise products that derive from standards-compliant production, the benchmark also needs to consider the chain of custody systems that are in place to ensure the integrity of the product through the supply chain.

## 8.3 Determine who will manage the benchmarking process

Benchmarking programmes are set up by many different types of institutions, from governments to companies and NGOs to the finance sector. In some cases, the institution that is seeking to use the results of the benchmarking programme is the one to develop and manage it. This is most often the case for companies seeking to understand which standards are most relevant for their supply chains. This model is also used where an entity seeks to compare its performance against that of its peers.

The advantage of this model is that the owner-operator has complete control over the content of the benchmark, enabling them to tailor the process to their specific needs. However, the potential conflicts inherent in this model mean that it is not an appropriate basis for any public claims or communications. A common alternative is where the convenor carries out the benchmark itself and has no formal affiliation to the entities being benchmarked. This would include examples such as the [\*WWF Certification Assessment Tool\*](#) and the World Benchmarking Alliance.

It is not always the case that the end-user of a benchmarking programme is the one to manage it. Benchmarking programmes can also be managed by a

related organisation or institution on behalf of the user, such as in the case of trade associations ([\*SAI Platform Farm Sustainability Assessment\*](#) or [\*Global Sustainable Seafood Initiative \(GSSI\)\*](#)), or by a consultant, such as the contracting of Ecofys (now Navigant) to help the European Commission carry out their [\*Renewable Energy Directive\*](#) benchmarking.

#### Take Note

It should be noted that even though a convening organisation may lack a direct interest in the outcomes, they still bring vested interests to the table. It is useful for any benchmarking organisation to declare its interests and any biases or intentional positions it brings to the benchmarking exercise. As one means to manage its interests, the benchmarking organisation should explicitly state the goal or purpose of the programme in all communications.



## 8.4 Determine the content of the benchmark

The benchmark itself is the reference against which the benchmarked entities are assessed. Given the diversity of potential entities to be benchmarked, it is not surprising that the content of benchmarks differs significantly. For example, the types of criteria used to evaluate the sustainability performance of a retailer or manufacturer will look very different from the criteria used to evaluate the credibility of a sustainability standard.

This document does not prescribe where performance levels should be set for the content of a benchmark as these should be developed through a robust benchmark development process. Instead, we first describe the issues all convenors will need to explore as they build out the content of their benchmark. In addition to this, we provide guidance on the elements of a standards system that should be included in a benchmark and why these elements are important.

### 8.4.1 Things to consider when developing the benchmark

#### 8.4.1.1 Stakeholder engagement

Convenors of benchmarking programmes or exercises will need to decide how the content of their benchmark will be determined and the extent to which this should involve interested stakeholders. The value of engaging stakeholders in this process comes both from the legitimacy and acceptance of the benchmarking process and from improving the relevance and feasibility of the benchmark itself.

The goal or purpose of the benchmarking programme is likely to inform the extent of stakeholder engagement. For example, where a government agency is setting a benchmark to assess fulfilment of a specific policy objective or where a company is determining which standards are consistent with its supply chain sourcing policies, there may be less scope for having stakeholders influence the content. On the other hand, when an industry association is seeking to establish minimum acceptable practices for sustainability standards operating in their sector, it will want to involve a broad cross-section of stakeholders.

When a benchmarking initiative aims to mobilize and align corporate practices and performance on global agenda's such as the SDGs or human rights – as with the WBA and CHRB respectively – open dialogue and broad consultation can be considered as part of the purpose of the benchmark itself.



**Where stakeholder consultation or engagement is relevant, the following practices are useful to consider.**

**Further information on good practices for stakeholder engagement are found in the *ISEAL Standard-Setting Code of Good Practice*:**

- Identify the types or categories of stakeholders who may be interested in the benchmark and the key stakeholders within each of those categories;
- Provide to stakeholders a clear and concise synopsis of the benchmark development process, how long the process will take, how stakeholders can participate, and what happens to their input;
- Provide multiple different mechanisms through which stakeholders can share their opinions and feedback;
- Be transparent about how input is being taken into account and how decisions are made, potentially making available any comments received and responses;
- Test the applicability of the benchmark in the draft stage to ensure its feasibility and relevance;
- Make draft and final versions of the benchmark and benchmarking process freely available and easily accessible.

#### 8.4.1.2 Measurable and relevant criteria

The benchmark consists of a number of criteria against which performance of the benchmarked entities is measured. A goal in defining these criteria is that they are clear and incisive, enabling a consistent assessment that results in a robust picture of potential performance of the benchmarked entity. Characteristics of clear and incisive criteria include that:

- They are concise;
- They employ clear and unambiguous language;
- They are accompanied by explanatory guidance where necessary to ensure consistency of interpretation;
- They are based on a mapping of current practices of entities that can be benchmarked;
- Evidence to assess their fulfilment is available;
- They are relevant to the purpose of the benchmark; and
- They allow for differentiation between entities on issues that matter.

#### 8.4.1.3 Technical experts

While stakeholders can provide meaningful input to help shape a benchmark, it is also important to have technical experts in lead roles preparing the benchmark criteria, testing its applicability in the field, and considering stakeholder input. A particular challenge for benchmarking initiatives that warrants the use of technical experts is in determining the evidence required to meet the benchmarking criteria, particularly where the benchmarked entities may not have the same operational structures or procedures. Pilot assessments by technical experts of different entities could be a useful step to strengthen the applicability of the benchmark.

The choice that convenors need to make is whether to contract consultants to carry out this work or to form a technical committee of experts. There are merits to either option and in both cases it is most important to clearly define the roles these experts have in the decision-making process on the content of the benchmark. For example, do they play an advisory role, deferring to the convenor on decision-making, or are they the final arbiters of the benchmark's content?

#### 8.4.1.4 Conflicts of interest

For the legitimacy and acceptance of the benchmark, convenors need to manage for potential conflicts of interest. While the entities that will be benchmarked are keenly interested in the benchmark's development, and can provide needed expertise, it is necessary to ensure adequate mechanisms are in place to avoid these vested interests having undue influence over the benchmark. One approach is to seek input from potential benchmarked entities, as from other stakeholders, while not involving them directly in the decision-making on setting the benchmark.

### ! Take Note

While not strictly a conflict of interest, it is also important to remember that there are value judgements in setting the benchmark and in assessing alignment with the benchmark. Whoever makes the decision about what to include in a benchmark is in an influential position and this should be recognised and potentially mitigated by having multiple people or organisations agree on the benchmark. Similarly, for assessors, having a second pair of eyes to review any decisions on benchmarking evaluations will help to allay any potential conflicts of interest.

#### 8.4.1.5 Definitions

An often-overlooked part of a benchmark is the definitions of specific terms. Much like any standard, evaluators who are assessing an entity's alignment with a benchmark will have to interpret that benchmark and it is important they do so consistently. One way to support more consistency is to build definitions of key terms into the benchmark. Any terms that could potentially have multiple meanings or interpretations should be included in the list of definitions. Definitions can be complemented by additional guidance on interpretation of the benchmark requirements, further strengthening consistency of application.





#### 8.4.1.6 Context

Where the benchmarking programme has a broad geographic scope, e.g. for global application, it may be necessary to look at how contextual factors might influence the content of the benchmark. For example, if a ranking model benchmark looks at company performance across multiple countries, it is useful to consider whether the ecological or social contexts in which those companies operate are sufficiently different that chosen benchmark requirements could be less relevant in one country than in another, thus biasing the benchmark for certain countries or regions. Ideally, the benchmark requirements are chosen so that they are equally applicable across all of the contexts in which the benchmark is applied.

### 8.4.2 Core elements for benchmarking sustainability standards

*This section applies specifically to benchmarking of sustainability standards systems and not to benchmarking of other entities*

Many benchmarking exercises or programmes for sustainability standards focus primarily on the content of the standard – the sustainability requirements that need to be met by certifying enterprises. This is clearly an important component of benchmarking, but not sufficient. Two standards that look identical on paper can support very different sustainability outcomes depending on how they are implemented in practice. It is therefore necessary to look at both the performance requirements and the operational systems that support their uptake. This guidance does not propose content requirements as these will be unique to the specific context of a given benchmark.

#### 8.4.2.1 Performance requirements

Two sources of performance requirements should inform any benchmark of sustainability standards, while a third is important to keep in mind for future integration:

- **The standard** - sustainability performance levels delineated in the standard. This may include environmental, social and/or economic sustainability, covering issues such as environmental management, biodiversity, natural resources, social well-being, labour rights, and economic livelihoods. Standards can set requirements for these issues in a number of different ways: defining acceptable practices; performance metrics; processes that should be followed; or desired outcomes.

- **Assessment guidance** – additional information about how the requirements in a standard should be interpreted. This is often included either as a separate guidance document or embedded in assurance requirements as interpretation guidance that auditors can refer to during assessments. It provides necessary context and consistency of interpretation for the standard.
- **Outcomes or impacts information** – data about the actual results that the sustainability standards system has achieved. While it may not be realistic at this stage to have comparable information available about the sustainability outcomes of a standards system, it is important to recognise that other requirements listed here are proxies for impacts and that we should aim to move towards measurement and evaluation of actual results as the data becomes more reliable and available.



#### Take Note

It is possible that benchmarks can focus just on evaluating the content of sustainability standards and not the assurance mechanisms or operational systems, but this severely limits the types of conclusions that can be drawn from the evaluations and is not recommended. For example, benchmarks focused only on standards content would preclude users from making any claims of comparability of the systems being benchmarked. It also prevents the benchmark being used as a way to recognize and incentivize companies or entities that are compliant with the benchmarked standards.

The one instance where a limited scope for the benchmark may be warranted is in improvement model benchmarks that are seeking to improve standards' requirements around specific sustainability issues but, even here, insight about how the standard is assessed in practice would be valuable.

#### 8.4.2.2 Operational requirements

The following systems support implementation of a sustainability standard and, because they can have significant impacts on the sustainability results achieved by the standards system, should be included in the benchmark:

- **Assurance** - the means of assessing a company's compliance with the requirements in the standard. This is critical because assurance models vary significantly so that even when two standards look similar, verification practices may lead to different compliance results.

Assurance elements most commonly incorporated include the certification or verification procedures, competencies required of auditors or evaluators, and accreditation or oversight of the assurance providers. The level of independence of the assurance process is also a consideration, distinguishing between self-assessments (first party), interested parties (second party), and independent entities (third party). Often, assurance models are constituted by a combination of these approaches. Examples of what credible assurance looks like are contained in the [\*ISEAL Assurance Code of Good Practice\*](#).

- **Scheme governance** – how the standards system is structured and implemented. The legitimacy of a sustainability standard derives in part from its operating structure. This includes the extent to which stakeholders participate in the development of the standard and the governance of the scheme. It also encompasses the existence of checks and balances like the inclusion of robust grievance or dispute resolution mechanisms, and the extent to which the scheme makes information about its procedures and assessment results transparent and accessible.
- **Traceability (chain-of-custody)** – the extent to which products from a certified production unit or facility can be traced through the supply chain to the final buyer. There are a number of different traceability models of varying rigour (e.g. from identity preservation to book and claim) and each is appropriate for different purposes and enables different claims to be made about the end products. Benchmarks may stipulate the type of traceability required or could focus on increasing transparency around the variation in traceability solutions.
- **Sustainability claims** – the communication of performance by a certified enterprise. While not often included in the scope of a benchmark, there is a strong case to be made for the critical importance of assessing the appropriateness of the sustainability claims allowed within the standards system and by the participating enterprises. The legitimacy of a sustainability standard is based in large part on a meaningful correlation between the level of ambition of the sustainability practices or outcomes required and the commensurate sustainability claims allowed.
- **Scope and uptake of the standard** – the breadth and scale of influence of a sustainability standard. Where a government, company or finance institution is planning to use the results of the benchmark to determine which standards to recognise or work with, it is useful to include an assessment of the geographical scope of the standard, the range of strategies or services offered to stimulate uptake of the standard, and the extent of that uptake. For example, a rigorous standard with little

market uptake will not be of value in sourcing decisions. In exploring potential uptake of a sustainability standard, it is also useful to look at how accessible the standards system is, particularly for small-scale or disadvantaged producers or enterprises.

A core function of ISEAL's work is to build consensus around good practices for the operation of sustainability standards. Our [\*Codes of Good Practice\*](#) serve as a useful reference point when developing benchmarking criteria. For an introductory overview of the issues and criteria that are useful to include in benchmarking of sustainability standards systems, please see Annex 3 and the [\*ISEAL Introduction to Comparing and Benchmarking Sustainability Standards Systems\*](#).

### 8.4.3 Alignment with international norms and guides

In determining which criteria to include in a benchmark, it is often the case that the benchmarking programme will build off the convenor's existing organisational goals or more specific policy objectives. For example, a government might set criteria that seek to translate an existing regulation into performance criteria, or a company may use its sourcing policy as the basis for the benchmark. Even in these cases, it is important to consider referencing external frameworks when developing the content of a benchmark, as this improves both the quality of the benchmark and the potential consistency between benchmarks.

For the sustainability performance criteria, be it for standards or companies, it is useful to refer to the products of inter-governmental initiatives as these have inherent legitimacy and global relevance. Among the most relevant and recognised frameworks to consider are the following:

- [\*UN Sustainable Development Goals \(SDGs\)\*](#)
- [\*UN Guiding Principles on Business and Human Rights\*](#)
- [\*OECD Due Diligence Guidance for Responsible Business Conduct\*](#) (general and sector specific)
- [\*FAO Code of Conduct for Responsible Fisheries, and Guidance for Responsible Agriculture Investing\*](#)
- [\*ILO Core Labour Conventions and Fundamental Principles and Rights at Work\*](#)

While these international frameworks are useful to structure and delineate benchmarking content, they still need to be translated into criteria that are clearly measurable and relevant to the scope and purpose of the benchmark. For this reason, it is also useful to review and potentially build upon existing benchmarks, helping to promote consistency in benchmarking.



For sustainability standards specifically, there are additional common reference points for good practice in how these systems are implemented, including:

- *ISO Conformity Assessment standards (17000 series of standards)* – for core competencies in certification and accreditation
- *ISEAL Codes of Good Practice* – for standard-setting and governance, sustainability assurance going beyond ISO, and measuring impacts of sustainability standards.

## 8.4.4 Accommodating diverse approaches in a benchmark

One of the most significant challenges inherent in benchmarking is how to recognise a diversity of models or approaches in the benchmark, while at the same time being prescriptive enough to ensure relevance and rigour. The most direct (and quite obvious) implication of this dilemma is that entities can be benchmarked only for those sustainability characteristics that they share in common.

By way of example, we can use the analogy of a fruit bowl that is composed of a variety of fruit such as oranges, lemons and limes. If we were to “benchmark” these fruits, we could use a set of criteria related to nutrition facts (vitamins, sugars, acidity), as well as production criteria (amount of water used in production, type of tree and soil characteristics, temperature for optimal growth, etc.). If the fruit bowl also contains pears, apples, melons and grapes, some of the benchmarking criteria would have to be dropped or revised to ensure continued relevance of the benchmark: for example, the optimal level of acidity for citrus fruits will be different than for grapes or melons.

Broadening the variety of fruits in the bowl does not mean that benchmarking becomes less relevant or useful, or that it is only possible to benchmark apples with apples, but it does call for a careful review and possible adaptation of the benchmarking criteria. As the diversity of fruit in the bowl increases, there will be fewer characteristics in common that the benchmark can cover. If we then add bananas, kiwis, pineapple and mangoes to our fruit bowl, these are still fruits and an assessment of vitamins and sugar levels would still be feasible but the results may not be so relevant if, for example, our intended purpose is to inform growers which fruits are best suited for production on their farms. For the benchmarking of sustainability initiatives, this analogy illustrates the need to recognize that the wider the spectrum of initiatives to be benchmarked, the less characteristics they will have in common, and the more restricted the benchmarking claim and result will need to be at the end of the process.



Where a benchmarking programme includes entities with different characteristics, it is good practice to identify at an early stage the characteristics they have in common, as a starting point for then determining which of these common elements is relevant to include in the benchmark. Not all common elements need to be included in the benchmark if they don't support achievement of the programme's purpose.

### 8.4.4.1 Variations in sustainability standards and related tools

*This clause applies specifically to benchmarking of sustainability standards systems and not to benchmarking of other entities*

In the case of sustainability standards, there is increasing variation in how they and similar tools are structured and how they seek to incentivise sustainability improvements. This presents a fundamental challenge for benchmarking since benchmarks are, by nature, establishing common reference points.

The way a standard is structured will influence how the benchmarking performance criteria are set. Standards themselves can take many forms, the most common of which are:

- those that delineate desired practices;
- those that define processes or systems that should be in place;
- those focused on desired performance outcomes that should be achieved; or
- a combination of the above.

Similarly, variations in the types of assurance models that are used to assess compliance with the standard will influence the choice of benchmark criteria. Among the most significant variations are differences in how compliance is measured or scored. For example, improvement model standards can recognise enterprises as compliant once they have started on the improvement journey, compared to standards that require enterprises to achieve a high performance bar before they are considered compliant.



### Take Note

A sustainability standards system can set the compliance bar for their programme at any level in relation to the requirements in the standard, e.g. compliance can require meeting all requirements in a standard or only the core requirements plus 60% of the other requirements. Given the potential variability, understanding where the compliance bar is set is as important to consider in the benchmark as the content requirements themselves.

An additional challenge is that sustainability standards are now experimenting with different strategies for incentivising uptake of more sustainable practices. For example, they are delivering capacity building or are convening stakeholders for collaborative action. Benchmarks of sustainability standards do not often recognise or compare these additional strategies and yet these strategies are becoming increasingly important to the understanding of how sustainability standards deliver systemic impacts.

Finally, it is also worth considering the variety of sustainability tools that are emerging and that play similar roles to certification, employing slightly different strategies to assess and communicate sustainability performance of enterprises, such as geospatial mapping of deforestation or communication of improvements through performance dashboards.



**If a benchmarking exercise or programme chooses to incorporate sustainability standards and tools that employ quite varied strategies for uptake, then there are two approaches the programme can take:**

- 1.** The first is to focus the benchmark primarily on the content of the standard or performance requirements, since all initiatives incorporate these to some extent. Being less prescriptive on criteria related to the implementation strategies will mean that a broader suite of tools can be assessed. However, as noted earlier, initiatives with similar standards can look very different when implemented in practice so this strategy should be approached with caution.
- 2.** The second approach is to consider whether there are benchmark criteria specific to each general type of uptake strategy. Rather than have one set of criteria to evaluate all standards and similar tools, it may make sense to have multiple sets of criteria that are applied to different initiatives as appropriate. For example, those standards or tools that employ certification as a strategy would be evaluated against the certification benchmarking criteria; those that deliver training would be evaluated against training-related benchmarking criteria, etc. Initiatives could be evaluated against more than one set of criteria if they employ multiple strategies.



### Take Note

It should be noted that capturing multiple strategies in benchmarking programmes is an emerging field, and little pre-existing experience is available to define what good practice looks like. Therefore, some amount of trial and error should be expected.



#### 8.4.4.2 Capturing improvement models

*This clause applies specifically to benchmarking of sustainability standards systems and not to benchmarking of other entities*

A growing trend in sustainability standards and similar tools, representing a challenge for benchmarking, is the shift from a pass/fail certification approach to a continual improvement model. This trend recognises the value of working with enterprises at all levels of sustainability performance and creating the levers necessary to incentivise their improved performance over time.

The challenge is that while the standard becomes the aspirational goal for enterprises to meet, each of those enterprises is at a different stage in their improvement journey. It is not possible to say that the programme delivers a certain performance level across all participating entities, thereby making it challenging to know at what level to assess performance against the benchmark.



Traditionally, benchmarking programmes have tended to penalise improvement standards by evaluating only those minimum criteria that they are certain are being met by all participating enterprises (the core criteria included in most improvement standards). This is a logical approach if the goal is to convey what we know for certain about performance levels. However, a more nuanced approach is to carry out two assessments, of both the minimum performance level and the level required to be met over time, and then include whether and how quickly that upper performance level must be reached. For example, many improvement standards require that participating enterprises make regular progress each year, moving to full compliance within a set period of time. That would be considered more rigorous than having a baseline set of practices but not requiring any improvement towards aspirational goals.

At minimum, benchmarks should distinguish which standards include improvement approaches so that these can be compared with other like programmes. Special attention should also be paid to the claims that can be made when different approaches are being benchmarked against the same criteria.

## 8.5 Determine the evaluation structure of the benchmark

An important component of the benchmarking exercise or programme is a decision on the evaluation structure. This decision is also linked to the purpose of the programme (see 8.1). The simplest evaluation structure for a benchmark is that all criteria are equally weighted and mandatory. However, there are innumerable derivations of this approach to suit different circumstances, created by making choices related to the following variables:

- **Mandatory vs aspirational** – a distinction between mandatory or core criteria and those that are aspirational or advanced. In a threshold model benchmark, a clear delineation needs to be made between those criteria that are required to be met to qualify and those that enable differentiation above and beyond the core requirements.
- **Scoring** – even where all criteria are mandatory, there is an option to implement a scoring protocol for qualifying the benchmarked entities. There are many variations in potential scoring models, the simplest of which would require that x% of the criteria are met in order to qualify. Alternatively, the scoring model could require that all core requirements are met as well as a percentage of the additional requirements.

- **Progress models** – as with improvement standards, it is possible for benchmarks to incentivize progress over time. From the perspective of the evaluation structure, this means differentiating between mandatory and aspirational criteria and increasing the number of requirements that are mandatory over time. Alternatively, it could mean carrying out regular revisions of the benchmark to increase the mandatory criteria as good practice evolves.



### Take Note

As with other aspects of the benchmarking process, no matter which evaluation structure is applied to the benchmark, a critical element is to be transparent about what this model is and how it works by publishing the benchmarking methodology in full.



## 8.6 Determine the benchmarking methodology

### 8.6.1 Steps in implementing the benchmarking process

Once the content of the benchmark has been decided, the second core component is to determine how

the benchmarking exercise or programme will be implemented. The steps in the benchmarking process are relatively straightforward and tend to be similar across most benchmarking initiatives:

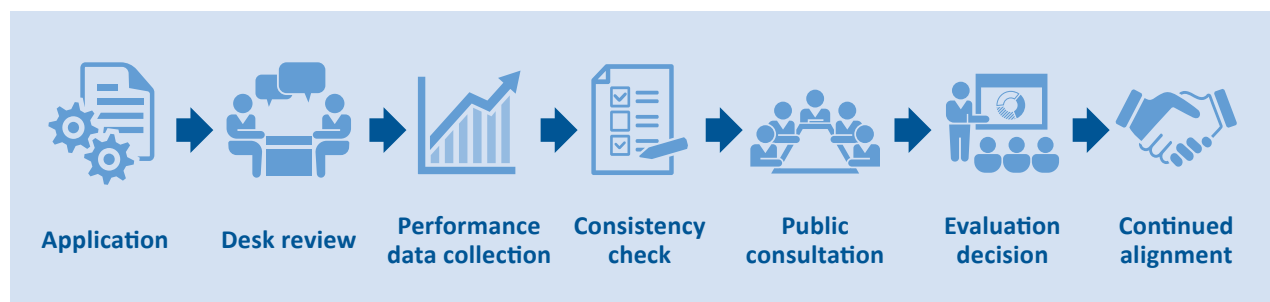


FIGURE 4: Steps in the benchmarking process



A consistent evaluation procedure (the benchmarking protocol or methodology) and reporting format should be developed to guide the review process. This procedure should include all the steps in the evaluation and should be made publicly available. This will help ensure consistency in application of the benchmark and how results are reported, providing stakeholders with a better understanding of the process. The methodology should include an indication of how long the benchmarking process takes and the points at which stakeholders and the benchmarked entities can engage.

#### Application

The first step in the process is for the convenor to decide whether to intentionally choose which entities to evaluate or to create an application process by which any qualifying entity can apply for evaluation.



- The process could begin with a self-evaluation, giving entities the opportunity to assess how aligned they are with the benchmark before formally applying to be evaluated.
- Where the entities are chosen by the benchmarking programme, it is important to reach out to them at an early stage to establish good lines of communication and build trust in the process, making it more likely that benchmarked entities will share performance information beyond what is publicly available.
- The desired outcome from the application stage is that the benchmarking programme has all information about the entity necessary to conduct an evaluation. Checklists can be developed that list the types of documents or information that should be made available.
- If there is a fee payable by the entity to participate in the benchmarking, this is the stage in the process at which to collect it.



### Desk review

The main step in the benchmarking process is the evaluation of how the entity is aligned with the benchmark. This consists most commonly of a desk review of detailed documentation about the entity's procedures and practices.



- The first decision is to determine which information will inform the evaluation. Evaluations can be based initially on publicly available information, then supplemented by information shared by the entity being benchmarked.

This additional input will be helpful in gaining a more complete understanding of how the entity operates.

- There are additional sources of information beyond what can be provided by the benchmarked entity. These include databases of information about the practices and requirements of different entities, such as ITC's Standards Map, which serves as the primary data source for several benchmarks of sustainability standards. The benefit of these sources of information is that the data about each entity is standardized and comparable.
- Convenors will need to decide who will carry out the reviews and ensure that these individuals or organisations are competent for the task. More recommendations on this element are included under 8.8 Further Considerations.



### Take Note

Engaging the benchmarked entity in the benchmarking process will strengthen the quality of the benchmarking results. Often, it is not possible to get a complete or accurate picture of an entity's work from publicly available information or procedures. Engaging in dialogue helps to ensure accurate interpretations, develop trust, and gain access to supporting materials. Benchmarked entities can be involved at the information gathering stage and can also provide feedback on the benchmarking decision, before it is finalized.

While it is important to develop an open relationship with benchmarked entities, the benchmarking programme should also retain a sufficient level of independence in arriving at the benchmarking results.

### Performance data collection

While information about the operational practices and performance of the benchmarked entity is important, this can be supplemented by performance data gathered in the field. This is an optional step, included in some benchmarking processes to give greater insight about how the entity operates in practice. The general premise is that looking at how the entity operates in practice will provide a better picture of potential performance. Some of the approaches to performance data collection include office visits, often used in the case of sustainability standards; interviews with relevant stakeholders, such as NGOs or community organisations impacted by a company's operations; research on sustainability outcomes; and stakeholder comment platforms. These additional activities will help to identify and bring to light entities that may have good procedures in place but make little effort to implement those procedures or vice versa.



- Site visits are optional because of the additional cost involved. In general, the costs of benchmarking processes are a major constraint and inhibit accessibility, so office

visits are only advised where the benchmarking programme is covering the costs or where a deeper level of trust-building is required, such as where benchmarking is a step towards mutual recognition between entities.

- Witness audits are an additional level of scrutiny that can be included, in addition to or instead of an office visit. A witness audit involves the evaluators witnessing the evaluation of an entity's performance. For example, they may witness a certification audit to understand better how compliance with the standard is evaluated in practice.
- Much like the desktop review, an office visit, witness audit or other performance data collection activities should be conducted according to consistent operating procedures that are made publicly available.



### Consistency check

In order to bring a level of consistency to interpretation of the criteria by different evaluators, it is valuable to employ a consistency check of some kind, whereby the results of an evaluation are checked by qualified experts. In a formal process, this can be a small benchmarking committee (2 to 3 people), but the structure can also be less formal (e.g. qualified individuals). The most important factors include: consistency in who participates, and the qualifications of the reviewers, their independence, and their knowledge about the benchmark and the sector.



- The first responsibility of a reviewer is to assess the work of the evaluators and ensure the evaluations are ready for public scrutiny.
- Where a committee or other independent party is not put in place, evaluators could also function as a peer group to review each other's work, aiming for a similar level of consistency in interpretation.

### Public consultation

An optional public consultation consists of making the draft evaluation publicly available for a period of time (e.g. 30 days) so that interested stakeholders can provide feedback, both on the evaluation and on insights they have about the benchmarked entity.



- While this step is also optional, it is highly recommended in most cases as a means to gather additional insight about the entity, providing a useful complement to the desk review, particularly where an office visit or witness audit is not included.
- If the benchmarking programme has built a strong and open relationship with the benchmarked entity, then the entity could help by raising awareness with interested stakeholders about the consultation.
- Consultation input should be considered as one source of information to be weighed alongside other information, like the evaluation results. It is important to assess the potential for bias or conflicts of interest in feedback received through consultation.
- The convenor can decide whether to make input received publicly available, along with a summary of how that input was considered. Good practice favours more transparency than less.

### Evaluation decision

In most benchmarking exercises or programmes, and threshold model benchmarks in particular, the process results in a decision about the extent of alignment with the benchmark.



- The convenor will need to determine who makes the decision. Is it the evaluator, a committee, or the benchmarking programme staff? As with the evaluation, key considerations are consistency and competence in who makes the decision, and transparency about the decision-making process and results, particularly to those that are being evaluated.
- Decisions do not always need to be about alignment. In ranking models, the decision could be about finalizing the assessment so that information about the ranking can be made publicly available.
- In coming to a decision, different sources of information will be used, some of which may conflict. It will be important to consider the source, robustness and integrity of each type of data or information to assess how much weight to assign to it in the decision-making process.
- Notifying the benchmarked entity in advance of making any decisions public is good practice, providing the opportunity to discuss any findings and clarify any misconceptions. If the notification is combined with providing the entities with insights about their performance, this increases the value proposition of the benchmark for these entities.
- It is also useful to have a dispute resolution mechanism in place, should there be differing opinions between the benchmarking programme and the entity that can't be reconciled.

### Monitoring of continued alignment

The operations and practices of both the benchmarked entity and the benchmarking programme itself can evolve over time, requiring a regular re-evaluation of alignment between the two.



- The benchmarking process can be arduous and expensive so it is important to balance the need for up-to-date evaluations with the extent of effort required. It makes the most sense to re-evaluate entities after the benchmark has been updated and for that to happen on a regular basis.
- If benchmarked entities incur time and costs to participate, that re-evaluation should take place no more frequently than every three years. An exception may be improvement model benchmarks where more frequent evaluations are needed to reflect progress made by benchmarked entities.
- A re-evaluation should be more streamlined than the original evaluation and structured so that the original evaluation acts as a starting point. Benchmarking entities can then provide information related to changes in their procedures and practices and any information relevant to new or revised requirements within the benchmark.



## 8.7 Determine how the results will be communicated

The final area to consider in the development of a benchmarking programme relates to the claims and communications regarding the programme and the results of benchmarking evaluations. This is another factor that is critical to the credibility and legitimacy of the benchmarking programme. Claims need to be grounded in, and consistent with, the actual results of the evaluation, and should not misrepresent the conclusions that can be drawn from the evaluation.

### Take Note

It is important to consider from the outset of the development process the types of claims that the benchmarking programme and the benchmarked entities can make, as this will inform how the benchmark and benchmarking process are structured. For example, if the desired claim is about the entity meeting a benchmark, then the convenor would likely need to construct a threshold model benchmark. Conversely, if the convenor wants to avoid compliance claims then it may be more appropriate for them to consider setting up a ranking model benchmark or a peer comparison model.



All benchmarking programmes should have a defined claims policy that is publicly available and that covers both claims that the programme can make and claims made by entities about their participation in the programme. It should also cover management protocols and ramifications in the case of misuse or miscommunication of the benchmarking results. The policy should be explicit about who can say what about benchmarking results at what point in time. It should also define the lifespan of applicable claims, recognising that benchmarking results and the claims that depend on them can become out of date after a period of time. Additional guidance is available through [\*ISEAL's Sustainability Claims Good Practice Guide\*](#).

Communication of benchmarking results should strive for simplicity. This is about making it easy for stakeholders to understand how the conclusions were arrived at and what those conclusions mean. For example, if a benchmark includes the evaluation of both mandatory and aspirational criteria, the programme should communicate clearly what this means and how benchmarked entities can be differentiated. The proposed language should be tangible and concrete, clearly describing what the results mean and ideally tying them to the expressed goals of the benchmarking programme. Additional guidance on core principles for sustainability claims can be found in the [\*UNEP Guidelines for Providing Product Sustainability Information\*](#).



## 8.8 Further considerations

### 8.8.1 Competence

The legitimacy and value of the benchmarking process depends on the competence of all those involved to implement the process professionally and consistently. This is particularly the case for evaluators and decision-makers.



- Define a set of skills and competencies required for each position in the benchmarking process and ensure that personnel or contractors meet those competencies. This can be achieved through consistent training and

ongoing calibration of personnel. To build trust with the entities being benchmarked, the qualifications and credentials of evaluators can be provided to these entities in advance of the evaluation.

- It is usually more effective to have a small group of well-trained evaluators than a large group, because it is easier both to maintain consistency of interpretation between a small group and to support evaluators maintaining a deep understanding of the benchmark and its intent over time.
- Where evaluators or decision-makers need to make decisions based on personal judgement, these should be noted and ideally incorporated into a systemic approach that seeks to minimize the subjectivity. One approach is to build out a body of case evidence of any interpretations so they can be consistently applied in the future.



#### Take Note

Evaluation results will almost always include some level of subjectivity, given the potential for differing interpretations of the benchmark. It is important both to ensure there is clarity of intent behind each element in the benchmark to inform how the evaluators should interpret it, and to document how different criteria are interpreted to build a body of knowledge for subsequent evaluations. However, benchmarking programmes should also recognise that a level of subjectivity will always exist.

### 8.8.2 Cost and complexity

As noted previously, benchmarking processes can be quite burdensome for the entities being evaluated as well as for the benchmarking programme. Convenors need to find a balance of costs and complexity that achieves meaningful results while still being manageable.



- High profile, international standards and multinational corporations, in particular, are the subject of numerous benchmarking processes. The most effective way to reduce cost and complexity is for

a new benchmarking initiative to use or adapt existing benchmarks rather than create a new one.

- Where a new benchmark is required, the costs to the benchmarked entities can be minimised if those entities are transparent and make relevant information about their systems easily available, such as by including data about their systems in publicly available, verified sources like *ITC's Standards Map*. Their time commitment to engage with the benchmarking programme is inversely proportional to the level of relevant information they make publicly available.
- Fostering collaboration between benchmarking initiatives is also valuable to increase consistency on what is being evaluated. This will help to reduce already existing duplication. Not all benchmarking requirements can be harmonized since different benchmarking programmes have different goals and objectives. However, catalysing conversations fosters a better understanding and alignment between benchmarks.
- It is good practice for benchmarking programmes to cover the costs of carrying out the evaluations. This should be budgeted for in the development of the programme, as should ongoing costs to maintain and update the benchmark and conduct re-evaluations. An exception to this is where successful alignment with a benchmark provides an entity with significant new market access, in which case there is more justification asking that the participating entity pays the cost of benchmarking.

### 8.8.3 Improving the benchmark and process

Through implementation of the benchmarking process, it is likely that the convenor will gain insight into what works well and how to improve both the benchmark and the process. A benchmarking programme is dynamic and should be revised regularly to reflect these insights.



- It is easier to capture insights and learning if a good data management system is in place. Some forethought should be given as to how the results of evaluations and other data will be stored and managed for easy analysis. Being able to analyse where entities are most aligned or not with the benchmark is an important foundation for a data-driven revision. Similarly, capturing data like time and cost allocations at different stages of the benchmarking process may highlight areas where increased efficiencies are possible.
- As noted above, the benchmark should be updated at regular intervals but not so frequently that it creates additional cost burdens. Every three to five years is recommended except where insights about potential improvements to the benchmark are sufficient to warrant more frequent updates.

### 8.8.4 Government use of benchmarks

Technical, political and/or economic obstacles might affect how public bodies implement certain benchmarking practices and how they integrate benchmarking results into policies that support sustainability.



- Where the results of a benchmark are tied to incentives for companies, such as access to subsidies, public procurement contracts, or exemption from government inspections, there is more pressure on the benchmarking initiative to make its process and evaluation decisions fully transparent in order to mitigate for potential conflicts of interest.
- In certain legal contexts, such as EU public procurement, governments need to accept products or services that are deemed 'equivalent'. For benchmarking, this means that a government body will need to consider any company's claim to meet their requirements and cannot choose to recognise only one or a few specific entities.

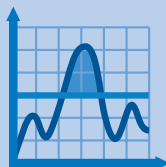



- Government policies that incorporate sustainability benchmarking might need to consider the supply of products that would meet the benchmark. If a policy affects imports, the government could be accused by exporting countries of setting up illegitimate barriers to trade. Some experience suggests that overly ambitious demand-side policies (e.g. banning all palm oil imports unless they are certified against acceptable standards) are politically unfeasible. One option to address this is to use broader benchmarking and recognition criteria, and then gradually introduce more stringent elements over time so that export markets have time to adjust.





# Annex 1: Summary of Sustainability Benchmarking Models

The following table provides a quick reference overview of the four main sustainability benchmarking models. Characteristics for each model are provided, along with an example. More information about these example benchmarking initiatives as well as others are provided in a series of Fact Sheets that have been produced by the UN International Trade Centre (ITC) in the context of their *Sustainability Map* initiative.

Benchmarking Model	Purpose	Distinguishing Features	Pros	Cons	Example
<b>Threshold</b> 	To qualify entities that meet or exceed a threshold. Often used for recognition	Performance bar set at level of acceptable practice	Simplifies message about which entities are acceptable to use	If poorly designed, can recognise poor performers without differentiating better performers	<i><a href="#">CGF Sustainable Supply Chain Initiative (SSCI)</a>; <a href="#">Global Sustainable Seafood Initiative (GSSI)</a></i>
<b>Ranking</b> 	To compare performance of similar entities through a ranked evaluation	Entities are scored against performance topics and compared	Public communication of results creates incentive for entities to improve	Potential subjectivity in how entities are ranked, particularly if based only on public information	<i><a href="#">World Benchmarking Alliance (WBA)</a></i>
<b>Peer comparison</b> 	To conduct an internal comparison of an entity's own performance against its peers	The reference benchmark is the practices of the benchmarking entity itself	Effective for understanding strengths and challenges compared to peers	Starting with own performance may neglect key issues addressed by others	<i><a href="#">CSR Hub is one example enabling companies to compare CSR performance with peers</a></i>
<b>Improvement</b> 	To encourage improved practices by showing progress toward good practice	Aspirational performance bar set beyond current practice to provide direction and incentive	Encourages improved performance over time	Can have narrow focus on benchmarking programme's topics of interest	<i><a href="#">WWF Certification Assessment Tool (CAT)</a></i>



# Annex 2: Sustainability Benchmarking Good Practice Checklist

This checklist distils key steps in the guidance for setting up and implementing a sustainability benchmarking exercise or programme. It is intended to be used as a quick reference both for those organisations and initiatives that are setting up new programmes and for stakeholders who want to assess the credibility of existing programmes and to hold them to account. The checklist should be used in conjunction with the full guidance.

## Develop the framework

- 1. Audience:** identify who is the intended audience and what are their needs; speak with potential users to understand more deeply their expectations (8.1)
- 2. Purpose:** define the goal or purpose of the benchmarking programme and ensure this is explicit and included in all communication about the benchmarking programme (8.1)
- 3. Strategy:** assess whether a benchmarking programme is the most effective strategy to achieve your purpose and whether existing benchmarks already meet all or part of the articulated goal (8.1)
- 4. Scope:** make a decision about who or what is being benchmarked, including the sector or commodity, geography, type of entity, and supply chain scope (8.2)
- 5. Openness:** decide whether to target specific companies or initiatives with the benchmark and which ones, or to

allow any qualifying entity to be evaluated (8.2.3)

- 6. Management:** determine who will manage the benchmarking programme (8.3)
- 7. Claims:** develop a publicly available claims policy that ensures claims are grounded in, and consistent with the actual results and do not misrepresent the conclusions that can be drawn from the evaluation (8.7)

## Determine the benchmark

- 8. Process:** set the process for determining the content of the benchmark, ensuring that contextual factors are taken into account where relevant (8.4.1)
- 9. Stakeholders:** determine whether and how to engage stakeholders in the content development and, where appropriate, follow good practices in the ISEAL Standard-Setting Code for how to engage stakeholders (8.4.1.1)

- 10. Definitions:** include definitions of key terms in the benchmark to support consistent interpretation (8.4.1.5)
- 11. Content:** for sustainability standards, include in the benchmark consideration of both performance requirements and operational requirements (8.4.2 and Annex 3)
- 12. Commonalities:** identify all the characteristics that entities to be benchmarked have in common, as a starting point for determining what is relevant to include in the benchmark (8.4.4)
- 13. Relevant criteria:** Ensure that benchmark criteria are clear and incisive, supporting a consistent evaluation (8.4.1.2)
- 14. Technical experts:** define the role for technical experts in content development, including their role in deciding on the content of the benchmark (8.4.1.3)
- 15. References:** use international reference documents to inform content and encourage consistency (8.4.3)
- 16. Alternative models:** for sustainability standards and related tools, consider how to accommodate different standards models, including those with different scoring models, different assurance models, and different strategies for incentivising uptake of more sustainable practices (8.4.4)
- 17. Evaluation:** determine the evaluation structure of the benchmark (8.5)
- 24. Public consultation:** consider a public consultation on draft evaluations and put in place the steps to do so where relevant (8.6.1)
- 25. Decision-making:** determine how decisions on alignment of benchmarked entities will be made (8.6.1)
- 26. Dispute resolution:** put in place a dispute resolution mechanism (8.6.1)
- 27. Alignment:** establish a process for monitoring continued alignment between the benchmark and the benchmarked entity over time (8.6.1)

## Additional considerations

- 28. Transparency:** make information about the benchmarking programme, including how it works, its governance, policies, decision-making and results publicly available and accessible (various clauses)
- 29. Impartiality:** manage for potential conflicts of interest in setting of the benchmark and implementation of the benchmarking programme (8.3, 8.4.1.4, 8.6.1, 8.8.4)
- 30. Improvement:** capture insights and learning from implementation of the benchmarking programme to inform its regular revision and improvement (8.8.3)

## Develop the benchmarking process

- 18. Effectiveness:** find a balance in the benchmarking process that achieves credible results in an accessible and cost-effective way (8.8.2)
- 19. Application:** determine the application process, where benchmarking programmes are open to qualifying entities (8.6.1)
- 20. Desk review:** carry out a review of detailed documentation about the entity's procedures and practices, engaging the entity to ensure accuracy of interpretation (8.6.1)
- 21. Performance data collection:** consider whether and how to gather additional information, such as an office visit or witness audit, to get a better picture of performance (8.6.1)
- 22. Competence:** ensure that evaluators, decision-makers and others involved in the benchmarking process are competent for their work (8.8.1)
- 23. Consistency check:** determine whether to put in place a benchmarking committee or some other mechanism (e.g. evaluator peer review) to support consistency of interpretation (8.6.1)





# Annex 3: Benchmark Criteria for Evaluating Sustainability Standards

*This annex applies specifically to benchmarking of sustainability standards systems and not to benchmarking of other entities*

This list of criteria represents a summary of core operating practices that a sustainability standards system should have in place in order to be considered credible. These practices derive from [\*ISEAL Codes of Good Practice and credibility tools\*](#), which represent a broadly shared understanding of what good practice looks like. With this list, ISEAL offers a snapshot for stakeholders to better understand what to look for when evaluating the likely effectiveness of sustainability standards. This list can also be used as a basis or starting point for defining the systems implementation criteria to be included in a benchmark.

## Scheme Management

1. The scheme owner has a sustainability-oriented mission or vision
2. The scheme owner has defined and makes publicly available its desired long-term sustainability impacts and strategy for achieving those impacts
3. Stakeholders have an opportunity to provide input on the intended sustainability impacts and possible unintended effects of the standards system
4. On a regular basis, the scheme owner monitors and evaluates progress towards its sustainability impacts and accurately and publicly communicates the results
5. The scheme owner applies adaptive management by using the learning from monitoring and evaluation to improve its standard and supporting strategies
6. The scheme owner carries out internal or external audits of its management system and operations at least annually and incorporates the findings
7. The scheme owner makes information on the governance structure and income sources or financing structure of the scheme publicly available
8. Stakeholders have the possibility to participate in or provide formal input on the governance of the scheme

## Standard-Setting

9. Information is made publicly available on standards development and revision processes and on decision-making
10. Consultations on the development or revision of the standard are open to all stakeholders
11. Input received during consultations is documented and there is a public report back on how issues raised are addressed
12. Decision-making on the content of the standard includes a balance of stakeholders and aims for consensus
13. The standard and consultation drafts are made freely and publicly available
14. Criteria in the standard contribute to addressing the key sustainability hotspots for the scope of the standard
15. The structure of the standard or accompanying guidance ensures consistent interpretation (e.g. auditable indicators)
16. There are provisions or mechanisms to ensure that the standard is locally applicable in the regions where it is applied
17. The standard is reviewed and revised on a regular basis (not exceeding five years)

## Assurance

18. The overall assurance methodology and structure for the scheme are publicly available
19. Assurance bodies are required to implement a management system that supports consistency, competence and impartiality (e.g. ISO 17065, 17021 or equivalent)
20. Full audits of at least a sample of clients are carried out regularly (from every year to every 5 years depending on sector)
21. Full audits include office visits and on-site assessments of at least a sample of operations
22. Stakeholders have an opportunity to provide input to the audit
23. There is a publicly available methodology for how compliance with a standard is determined (e.g. the scoring methodology)
24. Decision-making on compliance is impartial
25. There are procedures for how clients are required to address non-compliances, including when a certificate is suspended or revoked
26. There is a publicly accessible complaints and appeals process for certification decisions

27. Summaries of certification assessment reports are made publicly available
28. The certificate or license defines the scope of certification and duration of validity
29. A list of all certified enterprises is made publicly available
30. The scheme owner carries out regular reviews of its assurance programme and notifies assurance bodies and clients of any changes in requirements

## Group Certification (where applicable)

31. Groups are required to operate an internal management system that includes procedures for inducting, evaluating and removing group members
32. There is a representative sampling methodology for assessing group members during the external audit, and defined repercussions when a sampled member is found to be non-compliant

## Personnel Competence

33. Specific qualifications and competencies are defined for auditors and assurance body personnel
34. Auditors and assurance body personnel are required to have an in-depth understanding of the standard and its interpretation
35. New auditors have a probationary period during which their competence in an audit is assessed or supervised
36. Auditors and assurance body personnel are required to participate in regular training and professional development
37. The competence of auditors and assurance body personnel is demonstrated through regular evaluation
38. There are repercussions such as probation or suspension for misconduct or poor performance of auditors or assurance body personnel

## Oversight

39. There is an oversight mechanism that is independent of the assurance bodies being assessed
40. There are documented procedures for oversight and a management system that ensures consistent and competent application of these procedures
41. Individuals involved in oversight are competent to evaluate assurance bodies and possess knowledge of the standard and its intent

42. Oversight includes a review of the performance of assurance providers and auditors in the field

## **Chain of Custody (where applicable)**

43. Chain of custody verification is required if the scheme results in a communication of product origin from certified production
44. All enterprises that physically take products into storage are assessed (except where handling tamper-proof packaged products)
45. Enough information is documented in the chain of custody assessment to enable tracing of the product and to avoid fraud in the supply chain

## **Claims and Labels**

46. There are publicly available requirements for the use of claims and labels, including minimum levels of certified product content required for use of claims
47. A legal agreement is required for the use of claims and labels by enterprises in the supply chain
48. The types of claims allowed are appropriate considering the chain of custody models being applied
49. Allowable claims and labels contain enough information that their validity can be checked
50. The scheme employs surveillance strategies to monitor and rectify misuse of claims and labels



# About ISEAL

ISEAL is the global membership association for credible sustainability standards.

Our members are sustainability standards that meet our Codes of Good Practice and promote measurable change through open, rigorous and accessible certification systems. They are supported by international accreditation bodies, which are required to meet accepted international best practice.

By providing tools, training, events, resources and a community for standards systems, we help to shape an

effective standards movement. We support cooperation between our members and others to strengthen the effectiveness of that movement.

We also work with governments, businesses, NGOs and others in this field to support the use of credible sustainability standards as effective tools to achieve our collective sustainability goals for products and services worldwide.

## Photography credits

**Cover:** Fair Trade Certified Tea, Maud Tea Estate © Fairtrade for ISEAL Alliance

**Page 2:** Collecting baobab © David Brazier for the Union for Ethical BioTrade

**Page 3:** Coffee Nursery, Uganda © UTZ Certified for ISEAL Alliance

**Page 3:** Kenya © 4C Association, Claire Hogg for ISEAL Alliance

**Page 4:** Photo © Nishan Gunasekera for Sustainable Agriculture Network

**Page 5:** Vanilla traders, Madagascar © Sylvain Ly, Fairtrade International for ISEAL Alliance

**Page 7:** Fair Trade Certified Tea, Tonganagaon Tea Estate © Fairtrade USA for ISEAL Alliance

**Page 9:** Coffee berries in Costa Rica © SAN for ISEAL Alliance

**Page 10:** aluminium © RUSAL for ISEAL Alliance

**Page 12:** Cotton field, USA © BCI for ISEAL Alliance

**Page 13:** Coffee, Nicaragua © Steve Russel for GCP

**Page 17:** Goodweave textile worker © Robin Romano, GoodWeave

**Page 21:** Skipjack tuna © Fair Trade USA

**Page 23:** Bananas cultivation at Platanera Río Sixaola © Adrián Coto for Sustainable Agriculture Network

**Page 24:** Frogmary Green farm in Somerset © LEAF (Linking Environment And Farming)

**Page 26:** Textile manufacturing © Goodweave

**Page 28:** Palm oil harvest © RSPO

**Page 29:** early morning rib boat trip © ASC

**Page 30:** Catching yellow fin tuna © Fair Trade USA

**Page 32:** Bananas, Colombia © Marcel Koppen Fairtrade International

With financial support from:



Federal Ministry  
for Economic Cooperation  
and Development

**giz** Deutsche Gesellschaft  
für Internationale  
Zusammenarbeit (GIZ) GmbH



**ISEAL Alliance**

The Green House  
244-254 Cambridge Heath Road  
London  
E2 9DA

+44 (0)20 3246 0066  
[info@isealalliance.org](mailto:info@isealalliance.org)  
[twitter.com/isealalliance](https://twitter.com/isealalliance)  
[www.iseal.org](http://www.iseal.org)

Please address comments and queries to  
[info@isealalliance.org](mailto:info@isealalliance.org).