

Rate the Raters Phase Three

Uncovering Best Practices

February 2011



Rate the Raters

Rate the Raters is a four-phase research program in which SustainAbility is working to better understand the universe of corporate sustainability ratings in order to influence and improve their quality and transparency.

First and foremost, we express our sincere thanks to the ratings organizations which participated in phase three of *Rate the Raters* (listed on page 5). This work would not have been possible without the considerable time they invested in responding to our questionnaire as well as providing context and sharing perspectives via multiple phone calls and emails. Through our phase three research, we deepened our appreciation for the important work they do.

We also thank the sponsors of *Rate the Raters* — Abbott Laboratories, Autodesk, Brown-Forman, Cisco, ExxonMobil, Ford Motor Company, GlaxoSmithKline, Ingersoll-Rand, Novartis, Sara Lee and Scotiabank — whose support has made this project possible. Our gratitude extends to our advisory panel members especially for their input on our methodology and approach: Seb Beloe (Henderson Global Investors), Suzanne Fallender (Intel), Allen White (Tellus Institute) and Mike Wallace (Global Reporting Initiative).

Finally, we thank our colleagues: Chris Guenther, Geoff Kendall, Mark Lee and Deirdre Lizio for their invaluable contributions to this paper.



Michael Sadowski

Director
SustainAbility Inc.
+1 718 210 3630
sadowski@sustainability.com



Kyle Whitaker

Manager
SustainAbility Inc.
+1 202 315 4173
whitaker@sustainability.com



Alicia Ayars

Analyst
SustainAbility Ltd.
+44 20 7269 6905
ayars@sustainability.com

Executive Summary

SustainAbility is pleased to present the findings of phase three of *Rate the Raters*. Over an intense four-month period, we conducted in-depth evaluations of 21 ratings to better understand how they work in practice. In getting to know these ratings, we gained a first-hand appreciation for their diligent efforts to assess corporate sustainability performance and the difficulties involved in this task. So we begin by giving credit to ratings organizations for their challenging and important work.

Having reviewed these 21 ratings in detail, we offer the following conclusions:

Transparency = Trust = Value

Nearly every rating in our selection values corporate transparency and seeks to encourage it. We in turn place the same value in ratings transparency, believing that when raters open up across a variety of facets — their methodologies, their results, the stakeholders they engage, how they manage potential conflicts of interest, etc. — they build trust and strengthen their brands. Indeed, we find that the stronger ratings in our selection are the most transparent. When raters provide strong disclosure of their methodologies and results, they give companies clear blueprints for improving performance in the future (which is of course what ratings want). This also gives clients, ratings users and other stakeholders a sound basis on which to decide whether to use rating information when making decisions (which drives ratings adoption or uptake). This thinking runs counter to the prevailing logic of the many raters who strive to maintain “black boxes” for commercial purposes. Instead, we see evidence that ratings acceptance and use increases as the lid opens wider. Perhaps the evolution of the software industry over the past two decades from proprietary to open-source models offers lessons. While greater transparency may result in less control, the potential upside — more rapid adoption and crowd-sourced enhancement of core intellectual property — is potentially enormous for those willing to embrace new ways of generating value.

Simple is Beautiful

In the race to build a better rating, ratings organizations have introduced considerable complexity in terms of their criteria, questions and scoring schemes. Yet we find some of the best ratings in our review to be the simplest. When ratings are straightforward, they are more likely to be understood, trusted and therefore used by consumers, investors and others. The same qualities make ratings more likely to be accepted and minded (i.e. responded to) by companies. As [Lucy Kellaway](#)¹ put it recently in the *Financial Times*, we often make extremely simple things seem fantastically complicated. In measuring “best places to work” as *Fortune* does, we could likely trim the multitude of questions and instead evaluate how long employees stay with a company.

“The problem with *Fortune*’s rankings — and with all similar ones — is not that the exercise is a daft one. Actually, it is quite worthwhile. It is helpful both to prospective employees and to managers. The problem is that it makes an extremely simple thing seem fantastically complicated.”

Lucy Kellaway

The Financial Times

Double Down on Quality

As an increasing number of investors, consumers, employees and other stakeholders follow and make greater use of sustainability ratings, it is imperative that raters improve their efforts on quality assurance and control. Currently, there is too little focus on quality control among ratings, with too many raters relying on others to ensure quality (e.g. trusting media or NGO reports or asking companies to review draft reports). We expect this approach won't fly in the future as users of ratings go mainstream. For example, as major asset managers increasingly consider sustainability factors in their investment processes, they will look to ratings as "seals of approval" on companies and will thus demand strong processes behind these seals. We do see promising approaches emerging on quality, for example the Corporate Sustainability and Responsibility Research Voluntary Quality Standard (CSRR-QS 2.1, see the section on Quality Management for more details).

Focus on the Future

Too many ratings focus on companies' past or current performance and not how they are positioned for the future. While such past / current ratings have merit, and can give some indication of future performance, we need more ratings to evaluate how prepared companies are for what lies ahead. Indeed, we believe the ratings most likely to emerge as winners from the expanding ratings universe will be those which take a distinct, forward-looking perspective on company performance and positioning. For example, instead of scoring companies on GHG emissions and diversity, they will measure revenue driven by climate-oriented products and services and market share of diverse consumer segments; and they will focus less on how a company fares in a Western-oriented world of seven billion than how this company will fare in a multi-polar world of nine billion.

Get Cozier with Companies

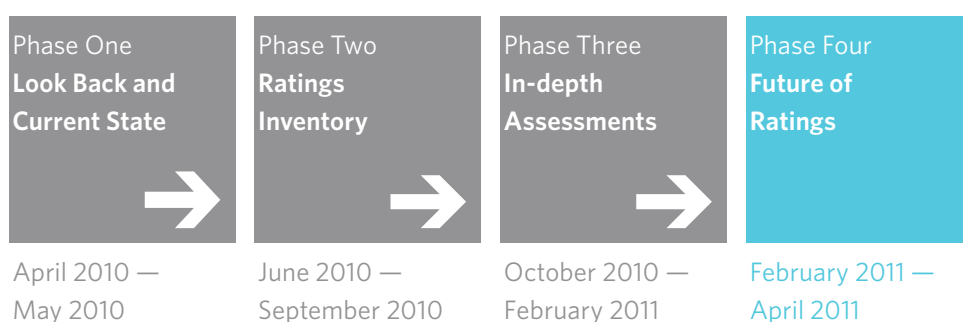
Lastly, raters won't be able to "Focus on the Future" unless they invest considerably more time directly engaging with the companies they rate. The majority of ratings today are based on arms-length assessments of performance. This is possible due to the nature (i.e. type of criteria) of today's ratings and often necessary given organizational limits and ratings' scope (i.e. small teams assessing large universes of companies). Going forward, we believe that every rater can and should spend more time with the companies they assess, and that there are myriad benefits to doing so (e.g. gain stronger understanding of their businesses, verify publicly-reported information, help companies understand ratings and improve sustainability performance). We say this with humility — as our own work on *Rate the Raters* demonstrates how difficult it is to develop a comprehensive understanding of a company (in our case a rating) without close engagement.

Looking ahead to phase four of *Rate the Raters*, we will explore — through conversations with stakeholders across the ratings universe — how ratings need to evolve to add greater value. In particular, we will explore more deeply a number of the most fundamental questions and dilemmas that we have touched upon in phases one through three (see Next Steps section).

Approach

Rate the Raters is a four-phase research program in which SustainAbility is working to better understand the universe of corporate sustainability ratings in order to influence and improve their quality and transparency (see figure 1). In phase one² we explored the evolution of the ratings agenda over the last decade and identified key trends and challenges. In phase two³ we inventoried 108 ratings and identified key themes and patterns across them. Now, in phase three, we dive deeper to better understand the inner workings of a representative subset of these 108 ratings.

Figure 1 **Project Phases**



Rating Selection

From the phase two universe of 108 ratings, SustainAbility selected 23 ratings for deeper evaluation in phase three. We contacted the organizations behind each rating individually, and 21 accepted our invitation to participate (see Figure 2). Participation was voluntary, and we did not evaluate any rating that did not agree to participate. In making our selections for phase three, we deliberately sought a mix of ratings in terms of audience, issue focus and geographic focus. Importantly, we do not submit that these are the 21 most important or prominent ratings extant. Rather, we believe the set collectively offers good lessons and insights for ratings, companies and other stakeholders.

Figure 2 **Participating Ratings**

- Access to Medicine Index
- ASSET4 (Thomson Reuters)
- Bloomberg ESG Disclosure Scores
- Carbon Disclosure Project
- Murky Waters: Corporate Reporting on Water Risk (Ceres)
- Climate Counts
- CR Magazine 100 Best Corporate Citizens
- CSRHub
- Dow Jones Sustainability Indexes
- EIRIS
- Ethisphere's World's Most Ethical Companies
- FTSE4Good Index Series
- The Global 100 Most Sustainable Corporations in the World (Global 100)
- GoodGuide
- GS SUSTAIN
- Maplecroft Climate Innovation Indexes (CIIs)
- Newsweek Green Rankings
- Oekom Corporate Ratings
- Sustainalytics
- Trucost Environmental Impact Assessment
- Vigeo

Questionnaire

Each participating rating completed the [questionnaire](#)⁴ SustainAbility developed specifically for *Rate the Raters*. We emailed questionnaires in mid-October 2010 and gave raters three weeks to respond (granting extensions in a number of instances). We held conference calls with each rater to further explain *Rate the Raters* and to address questions or concerns about the questionnaire and project more broadly.

Scoring Methodology

Our questionnaire is comprised of 13 criteria across four categories: Governance & Transparency, Quality of Inputs, Research Process and Outputs (see more detail in Figure 3 and in the full [scoring methodology](#)⁵). For each criterion, we developed a three-level scoring key — strong, moderate and weak — which we define in the methodology. We scored each rating against the criteria and documented areas of strength and opportunities for improvement. We then asked questions of clarification of each ratings organization by email and in conference calls.

Figure 3 **Evolution Criteria**

Governance & Transparency	Disclosure of Methodology	Conflict Management	Regular Review	Stakeholder Involvement
Quality of Inputs	Information Sources	Company Engagement	Input Verification	
Research Process	Experience and Capacity of Team	Quality Management	Sector Specificity	Basis for Rating
Outputs	Validation of Results	Accessibility		

Phase Three Outputs

For each rating, we developed an overview report describing its approach and performance against the criteria. We did not include scores in these reports as our focus is not on the actual scores but on identifying strong practices and areas for improvement across these 21 ratings. We shared the overview reports with the participating ratings organizations, asking for verification and suggested revisions. Based on the score sheets and report overviews, we summarized our findings and identified good practices for this *Rate the Raters* phase three report. Given the number of good practices for certain criteria, it is not possible for us to cite every example encountered. Thus, readers should not view cited cases as the only ones of merit.

Rating Objectives, Impact & Audience Findings

To better understand the context of each rating, we first asked raters to describe the primary objectives and motivations for their ratings, whether (and how) they evaluate the impact of their ratings, the intended audience for the ratings and how this audience uses their results. From their responses, we offer the following findings.

Ratings Objectives and Motivations

Nearly all ratings cite some variation of “improving corporate performance and / or transparency on sustainability issues” as their objective. There were a number of particularly lucid articulations of objectives and motivations, including those of Climate Counts (“to motivate deeper awareness among consumers” on climate change), Access to Medicine Index (“to stimulate pharmaceutical companies to improve their access to medicine practices”) and GS SUSTAIN (“to provide an objective measure of the effectiveness with which companies address the ESG issues facing their industry”). At least one rater — oekom research AG — has a [mission statement](#)⁶ which outlines its objectives towards different stakeholders including companies, investors and shareholders.

These examples aside, we find significant overlap and too much generality among the objectives of different ratings. Too many fail to explain clearly why their rating is distinct and how it creates value for companies and other stakeholders. We do not believe that the typical subject of ratings — large, multinational companies — need to be faced with multiple ratings to understand that improving performance and transparency is important. This may have been the case a decade ago, but today most companies appreciate this. Thus, ratings require greater purpose than the act of rating to increase transparency to justify themselves.

We also note inconsistency between some raters’ stated objectives of helping companies improve their performance and disclosure and their own lack of transparency and availability. Too often, raters do not sufficiently disclose their methodologies and results and offer limited time to explain results to companies. If raters truly want to help companies improve, they must provide clear blueprints for how to improve ratings performance (e.g. an “A” requires a company to do X, Y and Z). Some raters withhold information or limit engagement saying this might help companies “game the system.” Yet if their methodologies truly measure good performance or transparency, we should *want* companies to game the system and improve their results. Others cite confidentiality concerns regarding commercially-sensitive information, which we can understand — to a point. But even where a competitive issue exists, raters should offer significant disclosure at least to rated companies.

Evaluating the Impact of Ratings

When asked to describe the impacts of their ratings, only a few raters cite specific instances where they believe their influence changed company behavior (e.g. “We pushed a company on X and it changed course as a result”) or broader sustainability wins (e.g. “As a result of our rating, this industry supply chain changed for the better”). Among the few that do, Carbon Disclosure Project (CDP) mentions the use of its information by the UN Principles for Responsible Investment and Robeco to encourage better disclosure from low-scoring companies, and Access to Medicine Index describes how it helps drive improved performance and transparency from pharmaceutical companies specifically on the issue of access. Looking ahead, GoodGuide has the perhaps unique potential to demonstrate impact in a direct way: its interface allows it to track product page views, user recommendations for and against and user click-throughs to other websites (e.g. Amazon) for purchases.

Rather than explain the direct impact they have on companies, most raters cite things such as the number of visitors to their websites, the percentage of companies participating over time and media coverage. While these are important and might be indicative of outcomes, these data points don’t tell us much about influence on company performance or how stakeholders actually use ratings — more than anything, they reveal how popular these ratings are.

While we recognize that assessing direct impact is difficult, as ratings grow more numerous, companies and other stakeholders will pay greater attention to those ratings that bring the most value to them. Ratings that cannot demonstrate impact — by definition those that are not transparent in their approaches and results — will struggle for attention and adoption.

Ratings Audiences

We observe a clear connection between audience focus and the clarity of ratings’ objectives. Investor-focused (e.g. Maplecroft, GS SUSTAIN, Vigeo) and consumer-oriented (e.g. Climate Counts, GoodGuide) ratings typically have clear objectives and specific audiences, while ratings that try to meet the needs of multiple audiences — consumers, investors, companies, the general public, etc. — generally fail to meet any specific audience’s needs.

Rating Objectives, Impact & Audience Recommendations

- 1 Raters: clearly articulate the objectives and motivations behind your ratings and explain to companies, clients and other stakeholders how your rating is distinct and adds value. Otherwise, you risk being lost in the expanding universe of ratings.
- 2 As part of this effort, consider and describe the impact you have or strive to achieve — go beyond popularity / recognition measures (e.g. web page visits) to how you effect change with companies and your audience.
- 3 Narrow your audience — trying to meet the needs of multiple stakeholders generally ends up meeting the needs of none. With your audience identified, consider how your rating uniquely adds value to them.

“One of the key goals of the Access to Medicine Index is to encourage companies towards greater transparency and disclosure of their policies and practices in the developing world on access to medicine. If we expect transparency of companies, we should also apply the same principles to ourselves.”

Joop Muller

Executive Director, Access to Medicine Foundation

Governance & Transparency Findings

Sound governance and transparency underpin strong and credible sustainability ratings. Users and rated companies are more likely to pay attention to ratings when they understand how ratings work, conflicts that may exist and how they are managed, how ratings change over time and how external stakeholders are involved.

Definition

The degree to which a rater discloses its methodology to stakeholders (including the public and rated companies).

Strong Practice

A rater fully discloses its methodology to the public, including its selection process, information sources, criteria, areas of evaluation, scoring schemes, assumptions and rules. This information allows the user to fully understand and replicate how the rating is constructed.

Examples of Good Practice

Access to Medicine Index Provides a thorough description of its methodology in its [Methodology & Stakeholder Review](#)¹² and the [2010 Index Report](#)¹³ (both of which are publicly available). These include the key performance indicators, an overview of the methodology review process, detail on the scope of the index and the scoring guidelines.

Disclosure of Methodology

Looking across the 21 ratings, a majority provide what we consider to be strong disclosure of their methodology to at least one target audience (e.g. clients, companies, the general public). Several ratings, including [Access to Medicine Index](#)⁷ [CDP](#)⁸ [Climate Counts](#)⁹ and [FTSE4Good](#)¹⁰ provide strong public disclosure on their websites or in publicly-available reports (and we note that several of these are backed by non-profit organizations — a point we will further explore in this project). A greater number of ratings (e.g. ASSET4, EIRIS, GS SUSTAIN, Vigeo) provide strong disclosure to clients and / or rated companies, but not the general public. Several raters (e.g. [Ethisphere](#)¹¹) make public the questions they ask of companies which can help stakeholders better comprehend how ratings work.

Interestingly, we find the ratings that consider the general public to be their primary audience tend to disclose the least about their methodologies (to the public or to rated companies). While there may be instances where this makes sense (e.g. safeguarding commercially-sensitive information or minimizing complexity), we believe — and the strong practice examples demonstrate — that every rating can publicly disclose many aspects of their methodology and should provide more complete disclosure to evaluated companies. See the end of the Governance & Transparency section for specific recommendations on these topics.

CDP Makes public its methodology in a detailed [Excel file](#)¹⁴ which includes eligibility requirements, an explanation of how scoring works as well as all of the survey questions and their respective point allocations.

Climate Counts Website visitors can download the [scorecards](#)¹⁵ for every rated company. These include Climate Counts' 22 criteria and a detailed scoring key for each.

FTSE4Good Provides the full set of inclusion criteria in its publicly-available [Inclusion Criteria document](#)¹⁶ which allows companies to see exactly what they need to do to make the Index in the future.

Global 100 Provides a detailed [methodology](#)¹⁷ on its website that includes descriptions and weights for each key performance indicator, how each is calculated and the rationale for each.

GS SUSTAIN Provides full detail on its methodology (e.g. criteria, scoring key) in a published research report. This is intended for clients, but GS SUSTAIN also shares this with rated companies.

Definition

How well a rater avoids and manages potential conflicts of interest in the ratings process.

Strong Practice

Rater has formally articulated its approach to managing conflicts in a policy, guidelines or some other written document. The policy covers key aspects such as services and the independence of partners. In addition, the rater provides no services — related or unrelated to the rating — to rated companies.

Conflict Management

More than half of the analyzed ratings have a formal conflict of interest policy, though the format and nature of these vary considerably. Most of the raters with such policies disclose them to clients and / or evaluated companies and some do so to the public via their websites. The better policies explicitly address topics such as independence, the provision of services to rated companies, the use of company-provided information and confidentiality. The three ratings that have been certified to the CSRR-QS 2.1 research quality standard — [EIRIS](#)¹⁸, [oekom](#)¹⁹ and [Vigeo](#)²⁰ — have strong conflict of interest policies — which is not surprising given the principles and requirements of this [standard](#)²¹ (see the Quality Management section for additional perspectives on this standard).

Roughly a third of the analyzed ratings do not provide paid services or products to rated companies — this list includes Access to Medicine Index, EIRIS, FTSE4Good and the Global 100. The majority offer some type of service or product — typically detailed assessment or benchmark reports. Several ratings provide consulting services to rated companies, though state that these services are separate from their ratings business and thus present no conflict. Despite the firmness in our scoring key (“...the rater provides no services...to rated companies”), after reviewing the 21 ratings we respect that raters can consult or sell products to organizations they rate and still produce credible ratings. Still, the potential for bias exists anytime a rater is compensated by a rated company, and this potential is greater when the methodology is opaque and subjective. Yet, we see examples of well-regarded ratings that provide products or services — for example DJSI sells benchmark reports and oekom sells standardized industry reports and corporate ratings reports.

We believe full transparency is the best antidote for potential conflicts. Raters should: state publicly whether they are compensated by rated companies (as [EIRIS](#)²² does); describe the nature and cost of any products and services (as [CR Magazine](#)²³ does); and disclose the percentage of rating organization revenues that come from these activities (as [Ethisphere](#)²⁴ does). On this last point, Ethisphere produces a [Media Responsibility Report](#)²⁵ in which it discloses the number of listed companies with which Ethisphere has a material economic relationship. While we think this disclosure might go a step further by citing the nature and dollar value of these services, we applaud Ethisphere for their candor and encourage other raters to follow their example.

Examples of Good Practice

Access to Medicine Index

[Conflict of Interest Policy](#)²⁶ articulates risks associated with conflicts, identifies potential conflicts and describes how the Access to Medicine Foundation manages and corrects conflicts. All employees and contractors must sign this policy. The Access to Medicine Foundation does not have any financial relationships with rated companies.

[EIRIS](#) Publicly discloses [policies](#)²⁷ on independence, confidentiality and code of conduct and [states publicly](#)²⁸ that it does not charge companies for being rated nor does it provide any paid consulting.

[Ethisphere 2010 Media Responsibility Report](#)²⁹ includes a breakdown of the financial relationships Ethisphere and partner Corpedia have with listed companies and the names of sponsoring organizations.

[Oekom](#) Makes public its [policies](#)³⁰ on independence and conflicts. Does not provide any consulting services to rated companies, though it does sell standardized industry reports and benchmarks to interested parties.

Definition

How a rater manages the ongoing evolution and improvement of the rating. This may include, for example, updating a rating to reflect new context or understanding of issues.

Strong Practice

Rater takes a formal approach to reviewing, and updating as needed, its methodology to reflect new / improved information and context. The approach explicitly takes into account stakeholder feedback. The rater publicly discloses these changes and engages companies to explain the modifications. Changes are announced well in advance of when they take effect.

Regular Review

Generally, raters update their methodologies in the same frequency as they publish results (typically once per year) and most updates involve minor modifications rather than wholesale changes. For example, ASSET4 incorporates new indicators based on their relative importance and whether a significant number of companies (at least 10%) are reporting on said indicators. A handful of ratings explain and / or gather feedback on proposed methodological changes from stakeholders through in-person workshops, webinars or conference calls (e.g. CR Magazine, CDP, Newsweek), which we believe is good practice.

Among ratings organizations that make regular revisions, we find too little disclosure of changes. This is unfortunate as we believe that disclosure on such changes is critical to allowing rated companies and other stakeholders to understand the basis for the rating. Some exceptions (i.e. raters that do disclose for each review cycle) include Access to Medicine Index, which provides a fulsome description of the revisions it made from 2008 to 2010 in its [Methodology & Stakeholder Review](#),³¹ the DJSI, which outlines key changes in its [2010 Review](#),³² and Maplecroft, which summarizes key methodological changes in its 2010 cycle two report (available on its [website](#)³³).

Examples of Good Practice

CDP Based on stakeholder feedback on the previous year's process and staff review of random company responses, CDP develops and shares a draft questionnaire for consultation with a variety of stakeholders. CDP captures and discloses this feedback and proposed changes in a consultation document which it makes public.

Maplecroft CII's While the CII's have only been in existence since January 2010, Maplecroft reviewed its methodology prior to the launch of the cycle two report in September 2010 (see the summary of results in the cycle two final report). Maplecroft maintained roughly 95% of the questions from cycle one to two for consistency.

Sustainalytics Regularly consults clients on improving research processes and methodologies and will deploy a more formal annual client survey in 2011.

Definition

The degree and nature of external stakeholder input into the rating's construction and ongoing maintenance.

Strong Practice

The rater has convened an external advisory panel and / or systematically engages external stakeholders in the development and ongoing maintenance of the rating. The rater is transparent about the nature and outcomes of this convening and engagement.

Stakeholder Involvement

More than half of the analyzed ratings have some type of formal, external advisory panel — typically populated with prominent academics, NGO representatives and other experts (see [CR Magazine](#),³⁴ [Global 100](#),³⁵ [Newsweek](#)³⁶ and [Trucost](#)³⁷ for examples).^{*} While such panels can lend expertise and constructive challenge, few ratings organizations provide more than brief, general statements on the remit of these bodies. This is a missed opportunity for raters to build trust in their products — and also to squash cynicism that these “guru panels” provide more polish than substance. Notable good practices include FTSE4Good, which makes explicit the roles of its [Policy Committee](#)³⁸ in several locations online, Vigeo, which presents the responsibilities of its [Scientific Council](#),³⁹ and Access to Medicine Index, which outlines the remit of its [Advisory Committee](#)⁴⁰ in its Governance Charter.

Similarly, most raters cite some sort of stakeholder engagement in their methodology reviews. At least one — EIRIS — has a formal stakeholder policy that defines its philosophy on and approach to engagement (see [summary](#)⁴¹). But, generally raters do not publicly disclose the details and outcomes of such engagement, exceptions including CDP and Access to Medicine Index.

Disclosing the details and outcomes of engagement is important for several reasons. One, in order to hold raters accountable, companies and users of ratings should be able to see who raters engage and how the feedback is leveraged. Two, we know from experience that stakeholders face an increasing number of requests to engage (from ratings, companies, investors, etc.) and will be more apt to do so with those organizations that demonstrate that their feedback is being utilized. Lastly, we observe from our review of these 21 ratings that, in general, those with more robust engagement mechanisms are also stronger overall.

* John Elkington, SustainAbility's founder and Board member is a member of the Newsweek Green Rankings advisory panel. Sophia Tickell, a former SustainAbility Board member, serves on the Expert Review Committee for the Access to Medicine Index.

Examples of Good Practice

Access to Medicine Index

Methodology Report provides considerable detail (i.e. over 10 pages worth) about their engagement efforts, including identifying stakeholders by name and disclosing summary feedback.

CR Magazine 100 Best Corporate

Citizens The Methodology Committee convenes a series of meetings throughout the year to discuss methodological issues and potential changes. The Committee also holds an open comment period and public conference calls through which any stakeholder can comment.

Ethisphere When Ethisphere updated its methodology last – in 2010 – the changes were communicated via the FAQ section of the Ethisphere website. In addition, Ethisphere offered rated companies and external stakeholders a free public webinar and PowerPoint presentation on the updates that are scheduled to take place.

FTSE4Good The FTSE4Good [Policy Committee](#)⁴² — comprised of roughly ten external [experts](#)⁴³ — approves new or revised inclusion criteria and acts as an independent judge of the ability of companies to meet the criteria.

Vigeo The Scientific Council examines the principles and rules of the methodology, ensures that the methodology is applied well and resolves disputes between Vigeo and companies under review. Vigeo also works with experts to develop new criteria (e.g. WWF and biodiversity).

Governance & Transparency Recommendations

We offer the following Governance & Transparency recommendations to ratings:

- 1 The single most important step ratings can take to earn trust and to increase stakeholder use — and this mirrors our survey findings from phase two⁴⁴ of *Rate the Raters* — is to be more transparent about their methodologies. We don't mean simply providing high level overviews of approach and focus areas as many ratings currently do, but rather offering details on criteria, weightings and scoring schemes. While we understand sensitivities associated with disclosing commercially-sensitive information, ratings should — at a minimum — be fully transparent with rated companies and primary users.
- 2 All raters should: fully disclose whether they are compensated in any way by rated companies; describe the nature and cost of any products and services; and disclose the percentage of ratings organization revenues that come from these products and services. Only then can stakeholders form a sound opinion on potential conflicts of interest a rating may face.
- 3 Ratings should continue to use external advisory panels and engage stakeholders and tell us more about the nature, feedback and outcomes of this engagement (including how feedback has been incorporated). Such disclosure will allow users and rated companies to better hold ratings to account.
- 4 Lastly, ratings should disclose their ownership and funding models as this helps stakeholders make their own determination of raters' motives, biases and methods.

“The value of advisory panels can be at several levels. First, internal staff can be encouraged to think and try harder by the fact that they know they are going to have to open up at least parts of their black box methodologies to critical faculties. Second, external advisors can propose better ways of rating and ranking. Third, advisors can engage some of the companies being rated, helping get a better sense of the reality behind the questionnaires. Lastly, advisors can help communicate and test both the methodology and findings in public.”

John Elkington

Co-Founder & Executive Chairman, Volans Ventures Ltd.,
Co-Founder, SustainAbility Ltd.

Quality of Inputs

Findings

Inputs are the quantitative and qualitative data and information that raters use to evaluate companies. Raters draw these inputs from a variety of sources: company publications (e.g. annual reports, sustainability reports, websites), direct engagement with companies (e.g. questionnaires, interviews) and third-parties (e.g. media reports, government databases, NGO publications, stakeholder interviews, input from advisory panels). Put simply, the quality of a rating depends on the quality of its inputs — their breadth, diversity and credibility. Without high quality inputs, raters cannot develop an accurate picture of company strategy and performance. To assess the quality of inputs, we examined three factors: information sources, company engagement and input verification.

Definition

The breadth, diversity, consistency and credibility of information gathered by the rater to use in analyzing companies.

Strong Practice

Rater gathers information from sources which are current, consistent, credible and diverse. The rater goes beyond company-submitted and public information, for example obtaining information from third-party data providers or stakeholders.

Information Sources

The typical rating in our selection relies considerably (and sometimes solely) on public information to evaluate companies, then corroborates this information through desk research (e.g. comparing company-reported information and media reports) and / or by sharing draft profiles or scoring with companies for fact checking and verification. Generally, raters use similar sources (e.g. sustainability reports, Factiva) to gather information. Several raters leverage reputation measurement tools to ground their analysis, for example DJSI uses RepRisk for its Media and Stakeholder Analysis, while Newsweek uses a survey from CorporateRegister.com. A handful of raters distribute surveys to companies (e.g. DJSI, Ethisphere) and ask them to input all required information, but more raters develop draft profiles then share them with companies for review and completion. CSRHub — a slightly different animal in our universe of ratings — compiles its ratings from over 70 sources⁴⁵ (including several major ESG research providers) which allows it to assemble large data sets on companies and adjust for biases across its sources.

At face value, the trend towards raters using more public information is good for companies — this should lessen the burden they face in completing surveys. While companies still spend time reviewing draft reports from raters, stronger ESG disclosure should reduce time spent on reviewing such reports.

In terms of disclosure of sources, most raters outline source types (e.g. media scans, NGOs) as opposed to identifying specifics (e.g. *Wall Street Journal*, Sierra Club). At least one — Bloomberg — makes it easy for users to trace the information used in its scoring by linking each data point to the original source (which are generally corporate websites or sustainability reports). We expect to see greater disclosure from raters on their sources as this will allow users and companies to better judge the quality and credibility of ratings. Stakeholders should know, for example, if a rating is based solely on public information or if input from academics, NGOs, advisory panels and others influences the final outcome. And importantly, raters should explain how they corroborate conflicting information, for example when an NGO makes an accusation against a company. Too often we see these claims, seemingly unsupported or fact-checked, in the company profiles raters compile.

Examples of Good Practice

CSRHub Has five major sources of ESG data — ASSET4, GMI, IW Financial, Trucost and Vigeo — and pulls supplementary information from over 70 sources which are listed [here](#).⁴⁶

Global 100 First compiles research from ten providers (e.g. EIRIS, GS Sustain) to identify the top 10% of 3,000 global companies in terms of sustainability performance. It then assesses and ranks these roughly 300 companies based on public information (e.g. sustainability reports, proxy statements).

DJSI Draws on a combination of information sources including company responses to its proprietary questionnaire, publicly available information and reputation analysis provided by RepRisk.

Newsweek Green Rankings Sources information from Trucost (performance), MSCI (policies) and CorporateRegister.com (reputation). Trucost and MSCI rely primarily on publicly-available information, while CorporateRegister.com administers a reputation survey with responses from roughly 1,500 individuals

Definition

The degree and nature of engagement with companies throughout the ratings process.

Strong Practice

The rater takes a systematic approach to engaging companies in the ratings process, including spending equal time to gain an in-depth understanding of each company's performance and context. Rater regularly solicits feedback from all rated companies to improve the ratings process.

Company Engagement

Most ratings interact with companies at least twice during the rating process — at the beginning to announce the start of a review cycle, and again at or near the end to solicit feedback on draft reports. For most ratings organizations, the period in which they get feedback on draft reviews represents the bulk of the time spent engaging companies. A select group of raters go further, engaging companies more intensively throughout the process (e.g. Maplecroft), and a fair number hold webinars or in-person workshops to provide updates to and gather feedback from companies along with other stakeholders.

All told, we believe too few raters spend the requisite time to truly understand the future prospects of a company. Most criteria are backward-looking in nature (e.g. annual GHG emissions, number of fines, percentage of women in the workforce) — information which can be captured from public information. Sustainability ratings organizations generally do not do the sort of homework that mainstream equity analysts undertake when evaluating companies — convening meetings with senior management, participating in earnings calls, making site visits, etc. While both backward and forward looking ratings criteria have value, we find too few of the latter. Some certainly strive to get this future view via their questions — DJSI and Maplecroft ask important forward-looking questions on topics such as innovation — but few organizations couple such questions with adequately robust company engagement.

There are challenges related to more deeply engaging companies. Most raters are small organizations (employing less than 100 people) striving to build large portfolios of rated companies to meet client needs. This makes reliance on criteria that can be addressed from public sources necessary, but relying only on the most easily accessed information leaves the central sustainability performance question — present and future — inadequately answered.

Examples of Good Practice

FTSE4Good Through FTSE4Good's in-house engagement program, FTSE engages with companies in the Index on an ongoing basis and helps companies not included in the Index understand what they need to do to meet the criteria. See EIRIS (FTSE4Good's research provider) for more information.

Maplecroft Engages companies over a three to six month period via telephone conversations, email correspondence, in-person meetings and review of the questionnaires. Maplecroft achieved a company engagement level of 87% in its latest round of engagement.

Definition

The means by which a rater verifies the information it collects on companies.

Input Verification

While more than three-quarters of the ratings examined in phase three of *Rate the Raters* check for third-party verification when assessing the quality of information sources they use, not one reported requiring that all data is verified. Some raters give credit in specific areas for verification — typically operational environmental data such as GHG emissions (e.g. CDP and Climate Counts allot points for assurance of this type of data) — but the practice is not common to all raters and not applied to many data types.

Strong Practice

The rater should have a formal policy and robust process for verifying information used in its ratings.

A majority of the inputs to the rating are verified.

The primary means raters use to validate inputs is sharing draft reports or profiles with companies so they can suggest revisions. Other informal approaches include using multiple sources of information as a means to corroborate information, statistical analysis to identify outliers in data and gathering opinions from external stakeholders. Several raters use their advisory panels as both information sources and as corroborators of compiled data (e.g. FTSE4Good). As mentioned previously in this paper, CSRHub takes an innovative and intriguing approach in which it compiles data from a wide variety of sources, allowing it to identify inconsistencies and biases across data. Some raters accept at face value all company-reported information as accurate in the belief that companies would not publicly disclose erroneous information (a concern even if true, as it would not mean that information reported is complete). Several others cross check information against other ratings (e.g. GS SUSTAIN uses Bloomberg information to spot check some of its data).

Given the typical approach to input verification, it is likely that some ratings are built on erroneous or incomplete information. For example, if a company does not have the time to review draft reports — a likely scenario given the growing number of such requests — it is possible that a rater will publish final ratings based on incorrect data. Where third-party verification or assurance exists, such assurance should result in more robust research, which in turn should result in more accurate ratings. Yet, third-party verification of data in sustainability performance ratings is extremely limited in in scope and thus effectiveness, and is an issue likely to be more closely scrutinized as this field matures. We touch upon this topic more in later sections.

Examples of Good Practice

Bloomberg Within the Bloomberg system, subscribers can click on any figure within the ESG scores to see the original source for the figure. Bloomberg's methodology gives additional points to companies that have externally-verified data.

Global 100 Cross-checks data points against information in Bloomberg, ASSET4, Trucost and CDP. Gives credit to companies which have obtained third-party verification of their reports. Shares draft profiles with companies before publication.

Trucost Has calculated an average environmental impact for each of the 464 industry sectors and 700-plus indicators in its model. When Trucost receives ESG data on a specific company, it cross-references the data against its own expected (or modeled) data. If Trucost finds discrepancies between the reported data and expected data, it will enter into a dialogue with the company to resolve the issue.

Vigeo Has established guidelines that address information collection (e.g. policies on confidentiality, stakeholders, and quality control) and specifically considers the relevance and legitimacy of information sources.

Quality of Inputs Recommendations

We offer the following Quality of Inputs recommendations to ratings:

- 1 Publicly disclose the sources of your information (including your own questionnaires) and describe your approach to prioritizing and corroborating this information. Such disclosure will allow companies and users to better understand the basis for your ratings and in turn engender greater trust.
- 2 Spend more time engaging with companies during evaluation (to understand their businesses) and afterwards (to explain results). This is particularly important for raters that are trying to pick sustainability leaders of the future — you can't do this without spending significantly more time meeting with management, posing questions on earnings calls and diving deeper into company value chains. Such engagement will also allow you to better understand the degree of sustainability integration within a company.
- 3 Give more credit to companies that have invested in formal and informal assurance. This is a sign of company commitment and should result in more robust information.

Research Process Findings

We next asked raters about the processes and procedures they use to analyze data and ultimately generate sustainability ratings. We considered four factors — experience and capacity of research teams, quality management, sector specificity and the basis of ratings.

The research process is critical, as it ensures that reported data is treated consistently and fairly. Companies, in particular, are interested in the research process, as it can determine the credibility of a given rating organization and the level of effort required to participate in their rating process.

Definition

The experience and capacity of the individuals conducting the assessment of companies.

Strong Practice

At least 75% of a rater's analysts have 3 or more years of experience in the industry they are covering or on related topics. The rater has a formal approach to ongoing education. The rater's analysts cover no more than 20 companies each.

Experience and Capacity of Research Team

Raters exhibit considerable variation in the average tenure of their analysts — between two and ten-plus years — with a median tenure across the 21 ratings of four years. Only a handful of raters disclose the names (e.g. [oekom](#)⁴⁷) and experience (e.g. [Vigeo](#)⁴⁸, [EIRIS](#)⁴⁹) of their analysts, which we consider to be good practice.

In regards to training, nearly all raters cite informal efforts such as peer learning, conferences and other networking. A handful reference formal training programs, but none provide information on these programs (either to us or on their websites). Several ratings organizations are organized by sector which allows analysts to concentrate in particular sectors and learn from their peers.

The average number of companies covered by each analyst varies significantly as well. While the median number of companies covered per analyst is approximately 100, analysts for several ratings organizations cover several hundred companies each annually. To the points made elsewhere in this paper, this variation depends largely on each rater's approach, criteria and business model. For example, analysts for ratings driven solely by public information are able to cover more companies than analysts for ratings that spend more time engaging with companies.

Examples of Good Practice

Oekom Analysts have on average three years' experience conducting oekom Corporate Ratings; senior analysts have nearly six years' average experience and research directors seven years. Before an analyst takes charge of an industry, he must work under the supervision of an experienced staff member for a minimum of one year. On average, each analyst covers 50 companies or fewer per year.

Trucost Analysts have three to five years of experience working in or analyzing the industries they cover. Qualified accountants begin by segmenting each company's sales according to the 464 business activities in Trucost's model. This analysis drives the initial environmental profile which is used as the starting point for Trucost's environmental economics experts.

Sustainalytics Provides each analyst with an educational credit of up to \$1,000 per year to pursue educational opportunities related to his / her work. Sustainalytics also provides — in certain cases — additional leave time for educational purposes.

Definition

The policies, processes and systems established by a rater to ensure quality control throughout the ratings process.

Strong Practice

The rater must have a robust and well-documented approach to ensuring quality control throughout the ratings process. The rater's research process has been verified by a third party.

Quality Management

Nearly every rater cited some type of quality control process, with the most common practices being peer review (where a secondary analyst checks the work of a primary analyst) and statistical analyses (e.g. investigating data that appears inconsistent with past performance or the performance of industry peers). As mentioned above, advisory panels play a quality management role for some ratings in corroborating whether raters' assessments of company performance meets the panels' views of company practice "on the ground." Four raters obtain external validation of their research processes. EIRIS, oekom and Vigeo have been certified to the CSRR-QS 2.1 standard (the certifications and transparency matrixes for each can be found on the [AI CSRR website](#)⁵⁰). DJSI has received assurance on the application of its methodology from Deloitte, for which a letter is included on [DJSI's website](#).⁵¹ Notably, these four ratings are investor-focused, and their attainment of verification may speak to the higher expectations of their clients compared to users of other ratings.

External validation appears aligned with overall quality — the four ratings with such validation are among the strongest ratings in our selection. A read of the (quite demanding) principles and commitments of the [CSRR-QS 2.1 standard](#)⁵² indicates why certification may result in stronger ratings, but the value clients and other stakeholders assign such certification is unclear. Finally, even where certification or assurance exists, errors may still occur, and all ratings users need to exercise their own judgment as to the final accuracy and quality of ratings.

Examples of Good Practice

ASSET4 A primary analyst conducts company updates, and a senior analyst validates these updates. An independent auditing team conducts a second, full validation for roughly 50% of the companies. In addition, a team of seven senior quality analysts examines individual data points across all companies.

EIRIS Has developed a formal quality management system which has been [certified](#)⁵³ to CSRR-QS 2.1

DJSI Deloitte provided reasonable assurance on the application of SAM's corporate sustainability assessment methodology and limited assurance on the accuracy of participants' scores for inclusion in the DJSI (see [assurance report](#)⁵⁴ for more details on what these types of assurance entailed).

Oekom Has a formal Quality Management System (QMS), which is described on the [oekom website](#),⁵⁵ and a Quality Manager who is responsible for the ongoing development of the QMS. Oekom's Managing Board carries out spot checks on adherence to the QMS standards and external auditors regularly conduct quality checks of the system. Oekom's research approach has also been [certified](#)⁵⁶ to CSRR-QS 2.1.

Vigeo Has formally documented its quality control procedures which include controls within the rating process (e.g. expert assistance, quality checks and validation) and outside of the rating process (e.g. selective audits by Vigeo's Methodology Department). Vigeo's methodology has been [certified](#)⁵⁷ to CSRR-QS 2.1.

Definition

The degree to which raters consider industry and company specificity in their ratings.

Strong Practice

The rating must be based predominately on sector-specific criteria and weightings, and the rater incorporates company-specific issues and context in its analysis.

Sector Specificity

Of the 21 ratings organizations surveyed, nearly three-quarters incorporate sector-specific criteria and weightings into their analysis. The remaining quarter use the same criteria and weightings for all sectors and companies, citing the universal nature of the issues under evaluation (e.g. holding that all companies should be judged on GHG emissions). The approaches of most raters does not allow for them to consider factors distinct to particular companies. For example, raters typically judge companies equally on water risk regardless of whether the regions in which they operate are water-rich or water-poor.

It is difficult to gauge how well raters evaluate the most material issues in sectors as few raters disclose (publicly or to us) how their scoring schemes accommodate sector-specific criteria. However, from what we know about typical ratings processes and company engagement, it is unlikely that many ratings are able to judge companies closely on what we would consider to be their most important issues. For example, a rater can give credit to a bank for signing the Equator Principles, yet it takes considerably more time and research to understand the quality of said bank's environmental and social risk management processes (and thus the rater is not likely able to score this bank on its performance against this commitment). Similarly, in the IT sector, it is relatively easy to compare companies' operational GHG emissions, but more difficult (though arguably more important) to compare the green product strategies of such companies.

Examples of Good Practice

DJSI Questionnaires are sector specific (58 versions), following the [Industry Classification Benchmark](#)⁵⁸ sector classification. More than 50% of the criteria used are industry-specific. SAM discloses some of these industry-specific criteria in their [Annual Yearbook](#).⁵⁹

FTSE4Good Applies a risk matrix⁶⁰ that determines the criteria that companies must meet based on their sectors. Generally these criteria span a company's policies, performance and disclosure. Companies in high impact sectors (e.g. oil and gas) must meet more criteria than those in low impact sectors (e.g. media) and these additional criteria are more stringent.

GS SUSTAIN The GS SUSTAIN methodology is based on sector-specific indicators it discloses in its published research.

Definition

The underlying basis for / approach to rating company performance. For example, if a rater grades a company on diversity, is it making a relative comparison? Factoring in external norms or standards?

Strong Practice

The rating's scoring scheme is clearly defined and articulated and incorporates the broader sustainable development agenda (e.g. rewards companies that are taking action in line with what IPCC calls for on climate). The rater explicitly ties key external norms, standards or principles to its questions and scoring.

Basis for Rating

The original intent of this criterion was to evaluate the clarity of rating scoring schemes and how these schemes incorporate external norms, standards and expectations (e.g. GHG reduction targets and reporting standards such as the GRI). This proved a challenge as we received a wide variety of responses from which we reflect the following.

First, with a few notable exceptions (e.g. GS SUSTAIN, Access to Medicine Index, CDP, FTSE4Good), there is generally weak disclosure of rating scoring schemes. As discussed above under Disclosure of Methodology, such disclosure would help companies improve their performance and would boost the understanding and trust of users and other stakeholders.

Second, there is wide variation in the number of criteria or indicators raters use to evaluate companies. For example, GS SUSTAIN uses roughly 20 to 25 indicators per company while Trucost looks at over 700. This begs the question about the "right" number and the quality of criteria. Can GS SUSTAIN ably evaluate a company's ESG performance with just 20 to 25 indicators? Does Trucost need over 700 indicators to do the same? These questions have implications for both the basis for judgment of a company and how much time a company will spend responding to or engaging with raters, and there is a threshold above which performance indicators stop being "key" and we applaud efforts exploring this threshold from groups such as the Initiative for Responsible Investment.⁶¹ (see below)

From Transparency to Performance: Industry-Based Sustainability Reporting on Key Issues

In June 2010, the Initiative for Responsible Investment (IRI) and Arup released a report⁶² proposing a system for developing key performance indicators (KPIs) specific to particular industries. The framework, created by co-authors Steve Lydenberg, Jean Rogers and David Wood, is intriguing to anyone struggling to identify the "right" number of criteria for a given company or sector. The co-authors argue for KPIs based on three core principles — simplicity, materiality and transparency — and apply their method to six sample industries to demonstrate how such a system might be implemented within the United States.

For each of the sample industries, the co-authors arrive at fewer than 30 KPIs, which is aligned with GS SUSTAIN and other ratings in our selection which assess companies using relatively fewer criteria. The report ultimately contends that defining a select number of KPIs that relate to core business activities can help contribute to a corporate sustainability reporting regime that balances comprehensiveness and practicability. We think that a similar approach in the ratings world could do the same.

Third, there exists varying levels of complexity across ratings in terms of how scores are calculated. Some, like FTSE4Good, are relatively straightforward — if a company meets its inclusion criteria, and the Ratings Committee confers, it is on the Index. Other such as CSRHub and GoodGuide have more complicated calculation schemes. Again, this raises questions — are more complex scoring schemes better? Perhaps, but this may come at the expense of user and company comprehension and thus may limit rating usability.

Lastly, nearly all of the ratings cite that they incorporate external norms, standards or principles. Most commonly, ratings consider such norms in devising their criteria (e.g. questions on labor practices are aligned with ILO Conventions). Some raters give points or credit to companies based on things like GRI reporting (e.g. Global 100, Ethisphere) or endorsement of principles (e.g. ASSET4, CR Magazine).

Examples of Good Practice

Global 100 Assesses companies against a transparency indicator and ten equally-weighted KPIs. If a company does not report on a particular indicator, Global 100 removes that indicator from the calculation and assigns additional weight to the other indicators and the transparency score. Global 100 then normalizes the sum of the scores and then ranks companies on the basis of these scores.

GS SUSTAIN Using 20 to 25 indicators for each company, GS SUSTAIN scores companies on a relative scale, using normalized indicators in certain areas (e.g. indicators include Energy Consumption / Gross Cash Invested, GHG Emissions / GCI) and absolute indicators in others (e.g. separate CEO / Chair, reporting on sustainability).

Oekom Based on the methodology and 800 indicators of the Frankfurt-Hohenheim Guidelines (FHG), oekom developed a set of approximately 500 indicators for use in its Corporate Rating. Oekom uses roughly 100 indicators to evaluate each company depending on its sector and grades companies on each indicator on a scale of 1 (worst) to 4 (best) before computing an overall grade using industry-specific weightings.

Trucost Assigns economic costs to a company's operational and supply chain environmental impacts based on the company's specific business activities.

Research Process Recommendations

We offer the following Research Process recommendations to ratings:

- 1 Tell us more about the full team (including third-parties) behind the ratings: their identities, backgrounds and focus areas, how you train them, etc. Such transparency will engender greater trust in your process and improve engagement.
- 2 Firm up and better disclose your policies and processes around quality management. Sharing draft reports with companies — as many raters do — does not necessarily qualify as quality management as this depends on full — not to mention unbiased — company participation (which is quite rare). Look to models such as the CSRR-QS 2.1 standard for inspiration on the principles that might guide your quality management efforts.
- 3 Better explain (and disclose publicly) how you consider sector-specific issues in your ratings, in particular the more qualitative and future-looking areas (see above for examples). If using the same criteria and weightings for all sectors, better explain your rationale for this.
- 4 Focus and simplify your ratings. While we understand the inertia towards complexity given the subject matter and widening stakeholder expectations, we recommend that you pare down your evaluation criteria to focus on what's most important and make it easier for companies and others to understand your process. Complex does not necessarily equal robust or accurate.

CSRR-QS 2.1 Research Quality Standard

Launched in 2003, the Voluntary Standard for Corporate Sustainability & Responsibility Research Groups (CSRR-QS 2.1) is a quality standard for socially responsible investment (SRI) research firms. A number of independent SRI research firms developed the standard and contribute financially to the organization that owns and manages it (the Association for Independent Corporate Sustainability and Responsibility Research).

Signatories to the standard endorse its commitments and principles ([see standard⁶³](#)) and are certified to the standard by an independent auditor. Each signatory agrees to publicly disclose its [Certificate of Conformity⁶⁴](#) as well as its audited [Transparency Matrix⁶⁵](#) which serves as proof of conformity to the standard.

Outputs

Findings

Lastly, we asked raters how they ensured the veracity of their final ratings and the degree to which they disclosed the final results to companies, clients and other stakeholders. It goes without saying that accuracy is important for ratings — and the bar will continue to rise as sustainability considerations become further embedded into decision making (i.e. investors, consumers). It is also important for raters to share sufficient details on the results so that rated companies and users of ratings understand the basis for evaluation (and that companies can learn from them).

Definition

The means by which a rater checks the results (including scores and reports) prior to their finalization or publication.

Strong Practice

Rater has a formal, consistent and robust process for checking final results, including giving rated companies the opportunity to review the results prior to their finalization. Rater has a formal process for addressing challenges or disputes.

Validation of Results

As mentioned above, the most common method for validating results is sharing draft reports or profiles with companies for review and revision. While this can increase accuracy, raters do not get 100% response rates from companies. Other methods of validation include internal cross-checks (where one analyst reviews the work of another) as well as statistical techniques (e.g. comparing company scores against industry averages or past performance to identify aberrations). Some ratings use other ratings and data sets to verify the information in their reports. Several ratings receive external verification or assurance of their research process (e.g. DJSI, EIRIS, oekom, Vigeo), which should — all else being equal — result in fewer inaccuracies in the final product.

Several raters leverage advisory bodies for validation purposes. For example the FTSE4Good Policy Committee determines — based on research from EIRIS — whether companies are added to or deleted from the index. Newsweek's Advisory Panel provides a high level "reality check" of the scoring provided by its research partners and flags aberrations for further review.

Finally, roughly half of the ratings have a formal policy or process for addressing challenges or disputes, though none publicly disclose these policies.

Examples of Good Practice

CSRHub In compiling company information from its many data sources, CSRHub uses statistical analysis to remove outliers in the data.

FTSE4Good Includes several layers of validation including sharing information with companies for review and cross-checking information from EIRIS with findings from FTSE's own engagement with companies. Also, the Policy Committee validates the results and determines whether companies are added or deleted. This Committee also hears appeals from companies.

Trucost Process includes an inherent check on company reported information in that Trucost calculates the average or expected impact for companies given their industries. Trucost also gives companies four weeks to review compiled data and provide feedback and additional information.

Definition

The degree to which raters make the results accessible to companies and other stakeholders.

Accessibility

Generally, raters provide high-level reports or profiles (including scores, select strengths and areas for improvement) to rated companies and clients free of charge. Some sell more detailed reports or benchmarks. In either case, raters rarely share the underlying analysis (e.g. detailed scoring sheets, assumptions) used to determine results. Notable exceptions include EIRIS, which provides thorough company profiles to each company and clients. Only a few ratings make their results available free of charge to the general public — Climate Counts and Access to Medicine Index being among them.

Strong Practice

Rater fully discloses the details of its assessment and/or report to rated companies, and also provides a good level of detail to other stakeholders.

Examples of Good Practice

Access to Medicine Index Publicly discloses full report results including a profile for each rated company with scores, leading practices, changes from the previous Index and suggestions for improvement. The Access to Medicine Foundation holds a two hour conference call per rated company to explain the results.

Ceres Provides a detailed breakdown of each company's score in the Murky Waters report. Ceres also engages with interested companies, providing oral briefings on how to improve water disclosure.

Climate Counts Anyone can download the full scorecards for every company rated by Climate Counts — these include the 22 criteria and a detailed scoring key for each.

GoodGuide Product and company rating summaries and indicator scores are publicly available and free to all stakeholders on the GoodGuide website.

Outputs Recommendations

We offer the following Outputs recommendations to ratings:

- 1 As mentioned in the Research Process section, raters should reconsider how they validate their ratings before they are finalized. With growing interest in ratings from a broader set of constituencies, it's becoming ever more important to ensure that the end results are accurate.
- 2 As we recommended that raters disclose more on their methodologies, so too should they share more about their results. In particular, raters should provide more detail on results to rated companies — both to help these companies improve their performance and as a reward for taking part in the ratings process. We also recommend that raters provide greater disclosure to their primary clients — whether they be investors, consumers or the general public.

“I certainly appreciate the challenge many raters face in deciding how much “free” information to publish. However, when too little information is provided, it undermines the credibility and usefulness of the rating — there is no context to understand why companies are listed or not listed. The end result is that the rating is then not used within companies to drive performance improvements, which I see as the main value / contribution of these ratings.”

Suzanne Fallender

Director, CSR Strategy and Communications,
Intel Corporation

Next Steps

Exploring the Future of Ratings

With phase three complete, we now move into phase four in which we will explore — through a series of conversations with stakeholders from all parts of the ratings universe — how ratings need to evolve to add greater value. In particular, we will explore more deeply a number of the most fundamental questions / dilemmas that we have touched upon in phases one through three, for example:

What is the “right” funding model for sustainability ratings to ensure their objectivity and own sustainability? The predominant model for ratings today is “client pays” — that is clients such as investors (who buy ESG research), consumers (who buy magazines) and companies (who buy benchmark reports) generally fund the ratings. This model may encourage ratings to push the envelope on the products and services they provide — particularly to rated companies — at the potential risk of objectivity. We also know that many raters struggle to reap what we believe is the true value of their information to companies and clients. What other models might raters consider?

Can raters also be consultants or service / product providers? If so, under what conditions? Many of the raters in our selection both rate and consult or sell to companies, which at a minimum creates the perception of potential bias or conflict. Yet, raters need to earn income. How should raters approach this? What is a workable model?

How can raters provide greater details on their approach and results to companies and clients without giving away commercial secrets? Some of the ratings in our selection show potential paths forward on this, yet most disclose too little information on approaches and results. We take seriously concerns about giving away hard-earned intellectual property. But what are the ways in which raters can share more, maintain their commercial edge and create more value (for themselves, companies and others)?

Could the major raters agree on a core set of standard criteria and format (to reduce survey fatigue), then compete / differentiate themselves only on a smaller set of future-oriented criteria? Perhaps this is analogous to the common application process for universities in the United States, where aspiring students fill out one application that is used by multiple universities. These universities would form their own opinions of students based on their own judgment. Might raters develop a similar scheme?

What might a universal quality standard look like for ratings? As mentioned above, the CSRR-QS 2.1 quality standard is meant specifically for SRI research firms, yet its principles and commitments might be universally applied. How might this work? Would stakeholders value this sort of seal of approval?

We look forward to hearing perspectives on these questions and learning more over the next few months, then sharing our final reflections and findings. We welcome any feedback on the questions that will shape our path forward and any comments or questions about *Rate the Raters*.

Appendix A

Hyperlinks

- 1 <http://www.ft.com/cms/s/0/8691574c-2b0f-11e0-a65f-00144feab49a.html#axzz1CdnIbd1E>
- 2 <http://www.sustainability.com/library/rate-the-raters-phase-one>
- 3 <http://www.sustainability.com/library/rate-the-raters-phase-two>
- 4 <http://www.sustainability.com/library/rate-the-raters-phase-three>
- 5 <http://www.sustainability.com/library/rate-the-raters-phase-three>
- 6 http://www.oekom-research.com/homepage/english/oekom_MissionStatement_0409_en.pdf
- 7 <http://www.accesstomedicineindex.org/content/methodology-index-2010>
- 8 <https://www.cdproject.net/en-US/Results/Pages/leadership-index.aspx>
- 9 http://climatecounts.org/scorecard_overview.php
- 10 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Downloads/F4G_Criteria.pdf
- 11 <http://ethisphere.com/wp-content/uploads/2010/12/2011-WME-Questionnaire-Final.pdf>
- 12 <http://www.accesstomedicineindex.org/content/test-3>
- 13 <http://www.accesstomedicineindex.org/publication/access-medicine-index-2010-full-report>
- 14 https://www.cdproject.net/en-US/Respond/Documents/Rating_Methodology_2010.xls
- 15 http://climatecounts.org/scorecard_overview.php
- 16 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Downloads/F4G_Criteria.pdf
- 17 http://www.global100.org/images/stories/global100_website_pdfs/criteria_weights_final.pdf
- 18 <http://www.eiris.org/about-us/eiris-policies.html>
- 19 http://www.oekom-research.com/index_en.php?content=unsere-werte
- 20 http://www.csrr-qs.org/pdf/Vigeo_2009_CertifiedTM.pdf
- 21 http://www.csrr-qs.org/pdf/CSRR_QS_2_1_FINAL.pdf
- 22 <http://www.eiris.org/about-us/documents/independencepolicy-summary.pdf>
- 23 <http://www.croassociation.org/files/100-Best-Corporate-Citizen-Research-Report-Information-Sheet.pdf>
- 24 http://ethisphere.com/wp-content/uploads/2010/03/2010_media_responsibility_report.pdf
- 25 http://ethisphere.com/wp-content/uploads/2010/03/2010_media_responsibility_report.pdf
- 26 http://www.accesstomedicineindex.org/sites/www.accesstomedicineindex.org/files/publication/Conflict_of_interest_Policy_0.pdf
- 27 <http://www.eiris.org/about-us/eiris-policies.html>
- 28 <http://www.eiris.org/about-us/documents/independencepolicy-summary.pdf>
- 29 http://ethisphere.com/wp-content/uploads/2010/03/2010_media_responsibility_report.pdf
- 30 http://www.oekom-research.com/index_en.php?content=unsere-werte
- 31 http://www.accesstomedicineindex.org/files/general/100621_ATMMethodologyReport_with_correct_back_cover.pdf
- 32 http://www.sustainability-index.com/djsi_pdf/publications/Presentations/SAM_Presentation_100909_Review10.pdf
- 33 <http://www.maplecroft.com/cii/introduction/>
- 34 <http://www.croassociation.org/content/committees#corp100>
- 35 <http://www.global100.org/about-us/global-100-council-of-experts.html>
- 36 <http://www.newsweek.com/2010/10/18/green-rankings-2010-full-methodology.html>
- 37 <http://www.trucost.com/advisory-panel>
- 38 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Index_Rules.jsp
- 39 <http://www.vigeo.com/csr-rating-agency/en/qui-sommes-nous/equipes/conseil-scientifique.html>
- 40 http://www.accesstomedicineindex.org/sites/www.accesstomedicineindex.org/files/publication/GovernanceCharter_0.pdf
- 41 <http://www.eiris.org/about-us/documents/stakeholderpolicy-summary.pdf>
- 42 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Index_Rules.jsp
- 43 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Downloads/CommitteeMembers.pdf
- 44 <http://www.sustainability.com/library/rate-the-raters-phase-two>
- 45 http://www.csrhub.com/browse_data_sources_old
- 46 http://www.csrhub.com/browse_data_sources_old
- 47 http://www.oekom-research.com/index_en.php?content=team
- 48 http://www.vigeo.com/csr-rating-agency/index.php?option=com_content&task=blogcategory&id=106&Itemid=132&lang=en
- 49 <http://www.eiris.org/about-us/eiris-team.html#Research>
- 50 <http://www.csrr-qs.org/signatories.html>
- 51 http://www.sustainability-index.com/07_html/assessment/verification.html
- 52 http://www.csrr-qs.org/thestandard_02.html
- 53 http://www.csrr-qs.org/pdf/Certificate_EIRIS2009_2012.pdf
- 54 http://www.sustainability-indexes.com/djsi_pdf/AssuranceReport_2007.pdf
- 55 http://www.oekom-research.com/index_en.php?content=qualitaetsstandard
- 56 http://www.csrr-qs.org/pdf/Certificate_oekom2009_2012.pdf
- 57 http://www.csrr-qs.org/pdf/Certificate_Vigeo2009_2012.pdf
- 58 <http://www.icbenchmark.com/>
- 59 <http://www.sam-group.com/html/yearbook/?CFID=1586710&CFTOKEN=acbb8c97c2b47b7e-6794D107-B13A-CCCC-99AF9C3CB898520F>
- 60 http://www.ftse.com/Indices/FTSE4Good_Index_Series/Downloads/F4G_Criteria.pdf
- 61 http://hausercenter.org/iri/wp-content/uploads/2010/05/IRI_Transparency-to-Performance.pdf
- 62 http://hausercenter.org/iri/wp-content/uploads/2010/05/IRI_Transparency-to-Performance.pdf
- 63 http://www.csrr-qs.org/pdf/CSRR_QS_2_1_FINAL.pdf
- 64 http://www.csrr-qs.org/signatories_03.html
- 65 http://www.csrr-qs.org/signatories_04.html
- 66 <http://www.sustainability.com>



SustainAbility is a think tank and strategy consultancy working to inspire transformative business leadership on the sustainability agenda. Established in 1987, SustainAbility delivers illuminating foresight and actionable insight on sustainable development trends and issues. The company operates globally and has offices in Europe, North America and India. For more information, visit www.sustainability.com⁶⁶

Sponsors

