



Arabesque S-Ray

Methodology (v.2.6)

March 2020

Table of Contents

- Introduction** 3
- Step 1: Input Layer** 5
 - Gathering data** 5
 - Report-based metrics 5
 - News-based controversies 5
 - NGO-based activity 5
 - Cleaning data** 5
 - Organizing data** 5
- Step 2: Feature Layer** 6
 - Long-term trend** 8
 - Short-term correction** 8
 - Final feature scores** 9
- Step 3: Score Layer** 9
 - GC Score** 9
 - Category scores 10
 - Total score 10
 - ESG Score** 11
 - Total score 11
 - Preferences Filter** 12
- Appendix A: Examples of report-based inputs** 14

Introduction

S-Ray® is Arabesque’s unique tool that allows anyone to monitor the sustainability performance of around 7,000 of the world’s largest corporations. Smart and simple to use, Arabesque S-Ray looks beneath a company’s surface by assessing its extra financial performance, helping us to better understand its value to society.

Leveraging big data through the power of machine learning, Arabesque S-Ray systematically combines over 250 environmental, social and governance (ESG) metrics with news signals from over 30,000 sources published in over 170 countries.

S-Ray provides information on corporate sustainability in three complementary ways:

GC Score	A normative assessment of each company based on the core principles of the United Nations Global Compact. This provides a deeper understanding of reputational risk facing a company.
ESG Score	A sector specific analysis of each company’s performance on financially material environmental, social, and governance (ESG) issues. This can be used to identify companies that are more likely to outperform over the long run.
Preferences Filter	A search tool that allows anyone to check the business involvements of companies. This helps users check company activities against their personal values.

In this document, we describe the methodology used to construct the S-Ray scores, following a three-layer process (Figure 1).

1. **Input layer:** Collect and clean a wide variety of sustainability-related data
2. **Feature layer:** Aggregate inputs along well-defined topics
3. **Score layer:** Combine features into easy-to-use S-Ray scores

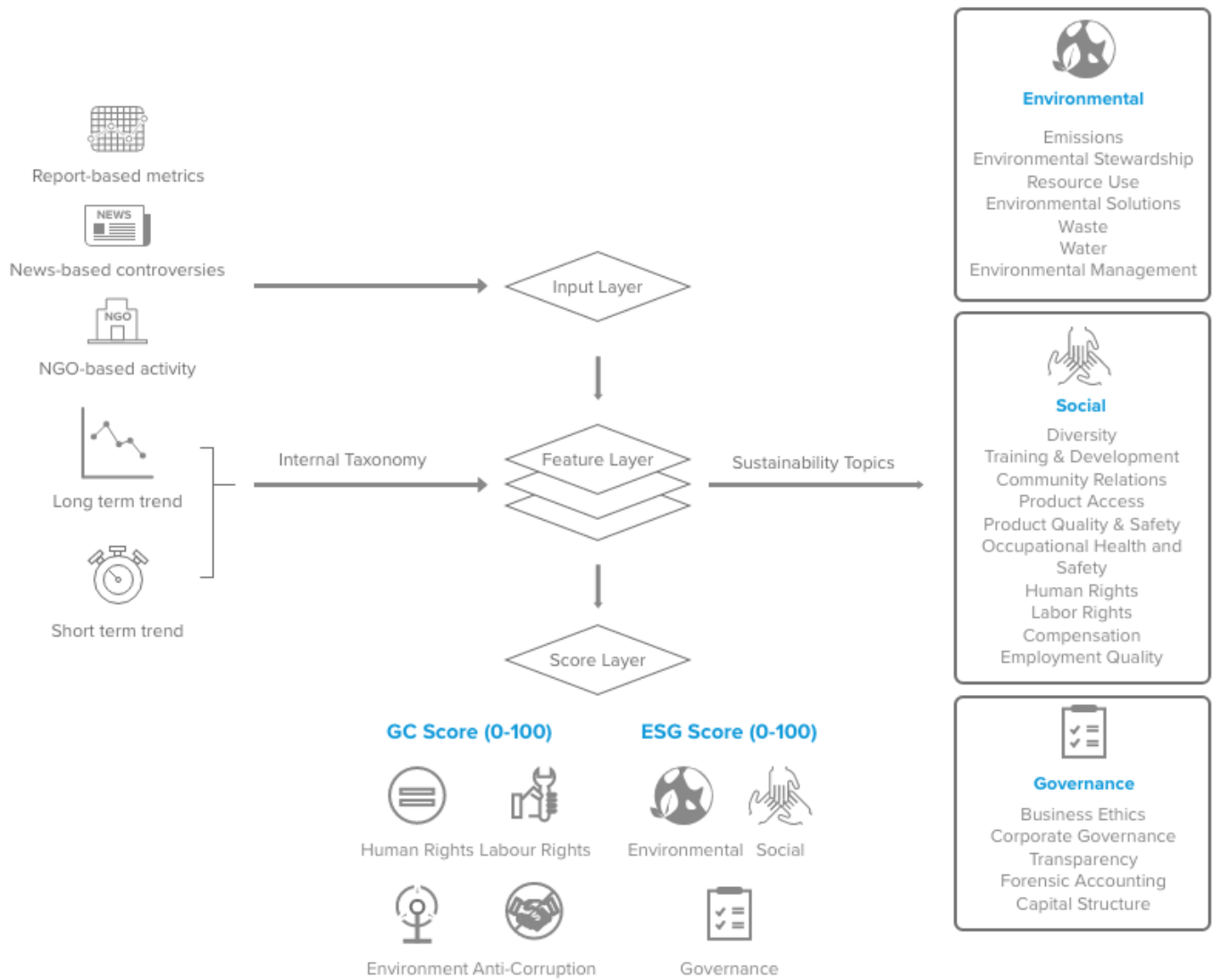


Figure 1. S-Ray ESG and GC methodology diagram

Step 1: Input Layer

Gathering data

During the first step, S-Ray collects a wide range of data from three types of sources.

Report-based metrics

To obtain a first understanding of a company's sustainability performance, S-Ray collects over 250 reported metrics from non-financial disclosures (e.g. sustainability or integrated reports).

News-based controversies

Following typical reporting cycles, report-based metrics are only available on an annual basis. To get a more up-to-date analysis of sustainability performance, S-Ray uses natural language processing (NLP) to scan over 30,000 public news sources published in over 170 countries daily for sustainability related controversies.¹

NGO-based activity

Alongside the news-based controversies, S-Ray also tracks NGO campaign activity across over 400 sustainability issues. NGO campaigns can be both positive or negative in nature.

Cleaning data

The data are gathered daily and cleaned to allow for further operations. For example, all inputs are subjected to a set of data quality checks (e.g. false outlier detection), and poor-quality data gets discarded. In addition, inputs are scaled and/or normalized to allow for comparison and aggregation of data across various sources. Sparse or infrequent time series are imputed and resampled to accommodate daily calculations. Additional checks also take place to ensure that data is up-to-date.

Organizing data

The cleaned inputs are further organized and labeled according to an internal taxonomy, which prepares the data for further calculations in the feature and score layers. Two questions are asked for every available input:

- What is the focus of the input?
 - Preparation (e.g. having a human rights policy)
 - Outcome (e.g. number of times major incidents have halted work on the factory floors)
 - Business involvement (e.g. fraction of sales derived from tobacco-related products)
 - News (e.g. controversies regarding the redistribution of private customer information)
 - NGO campaign (e.g. campaigns focused on plastic pollution of the marine environment)
- What is the topic of the input (e.g. employee diversity, energy use, board independence)?
 - We have defined a list of 22 sustainability-related topics (see Table 1) and 11 business involvements (see Table 3) from the set of inputs, which directly correspond to the features in the feature layer.

A comprehensive list of report-based input examples is provided in Appendix A.

¹ Note that S-Ray currently only considers negative news (i.e. controversies).

Step 2: Feature Layer

A major challenge with data from the input layer is that significant correlation and overlap can be found between inputs.² The feature layer is introduced to further structure the input data along 22 well-defined sustainability topics (Table 1) using (semi-supervised)³ dimensionality reduction techniques. Measures have also been taken to ensure that there is no single or dominant reliance on any one data provider. For every topic, we first construct two types of feature sub-scores reflecting the frequency of data input. These sub-scores are calculated separately for the annual report-based data and for the higher frequency data based on news controversies and NGO campaigns. These two scores are then aggregated into a final feature score. In addition to these core sustainability topics, S-Ray also flags business involvement across 11 areas to allow for further filtering based on personal preferences (Table 3).

Feature	Description
Environmental	
Emissions	The contribution of business activities to the emission of greenhouse gases and other air pollutants. Inputs into this feature include emissions data as well as reduction initiatives, objectives, policies and monitoring.
Environmental Stewardship	The impact of business activities on biodiversity and animal welfare. Inputs into this feature include the use and regulation of animal products/testing alongside biodiversity impact initiatives and targets such as Forest Stewardship etc.
Resource Use	The efficient use of energy and other natural resources including land and materials. Inputs into this feature include energy use/efficiency and land use reduction initiatives, recycled raw materials, toxic chemicals reduction and resource efficiency policies as well as targets and monitoring.
Environmental Solutions	The environmental impact of products and services and the contribution towards sustainable consumerism. This feature includes inputs including, but not limited to, the development of hybrid vehicles, smart water solutions and sustainable building products as well as offering environmental and renewable/clean energy products. Responsible asset management, the total supply of renewable energy and environmental products targets are also included.
Waste	The generation of waste and other hazardous output as part of business activities. This includes inputs covering hazardous and general waste generation and reduction policies, recycling practices and oil spill disclosure.
Water	The efficient and responsible use of water throughout company operations. Covering water pollution, recycling initiatives and water withdrawal.
Environmental Management	The mechanisms and policies employed to manage the overall environmental performance of the business. Inputs to this feature relate to environmental lawsuits, investments into resource efficiency and environmental impact reduction as well as the environmental management system of the company.

² For example, S-Ray ingests raw data on board membership from three different independent sources.

³ Typical dimensionality reduction techniques (e.g. principal component analysis or PCA) are labeled as unsupervised machine learning techniques. Rather than applying these techniques unconditionally, the S-Ray algorithm requires human oversight to avoid spurious data aggregations.

Social	
Diversity	The representation of and equal opportunity for women and minorities in the workforce and on the board. To calculate this board and employee diversity, discrimination lawsuits and commitment to supplier diversity among other inputs are used.
Occupational Health and Safety	The workplace-related health and safety performance. This includes inputs such as the disclosure of accident rates and workplace injuries, employee health and safety training and certification, policies/procedures and targets.
Training and Development	The opportunities and programmes in place to enable and support learning across employees and the supply chain. Considering employee training hours, costs, and policies as well as the monitoring and targeting of these.
Product Access	Providing access to products and/or services for disadvantaged communities. Product access is composed of the access to finance, medicine, education, food and affordable housing programmes.
Community Relations	The level of community involvement and public trust. Inputs include customer satisfaction and community relations policies, targets and the monitoring of these as well as a disclosure of employee volunteering hours.
Product Quality and Safety	The quality and safety of products and services and level of customer satisfaction. This feature is composed of lawsuits, policies, targets and monitoring relating to the product quality management as well as consumer satisfaction and others.
Human Rights	Adherence to and promotion of human rights throughout all business activities, including the supply chain. Inputs include suppliers' human rights consideration, employee human rights training and policy as well as targets and monitoring of these.
Labour Rights	Compliance with internationally recognized labour standards, both in-house and across the supply chain. This includes the Freedom of Association, Supplier Code Audit and child labour, forced labour and labour rights policies.
Compensation	The fair and equal compensation of staff and board members. Judged by average salaries and benefits, board member compensation, pension funding among many more.
Employment Quality	The working conditions and employee satisfaction. Accounting for employee turnover, work-life balance policy and worktime flexibility as well as several other inputs.
Governance	
Business Ethics	Fair business practices as it relates to issues like corruption, political contributions and anti-trust. Typical inputs include lawsuits relating to these issues and policies in place to monitor and reduce their impact.
Corporate Governance	The procedures and mechanisms in place that ensure proper long-term control and management of the corporation. The feature inputs relate to the board and committee as well as policies and targets surrounding governance issues like shareholder rights, insider trading and many other common problems.

Transparency	The level of transparency and disclosure of critical information about the business. This accounts for disclosures relating to director compensation and articles of association as well reporting standards and external verification.
Forensic Accounting	The overall earnings quality or the degree to which reported earnings properly represent a company's financial health.
Capital Structure	The relative level of leverage and how it might take away from a long-term focus in decision making. Inputs will typically include the debt to equity ratio and targets for this.

Table 1. List of 22 sustainability feature topics

Long-term trend

To get an aggregate view on a company's long-term performance trend per topic, we pull together all available report-based metrics from the input layer. The different metrics are then aggregated based on several considerations, including focus (see Appendix A), dimensionality⁴, and expert input.⁵

The report-based feature score also takes into account possible asymmetries between outcome-and preparation-based performance. For example, when a company does well in terms of preparation, but is lagging when looking at the actual outcomes, the outcome-based score will be dominant when constructing the total score.

Example: Occupational Health and Safety

We find 15 different report-based metrics related to occupational health and safety in the input layer. Separately, the 6 outcome-based and 9 preparation-based inputs get aggregated, also considering guidelines from the International Labour Organization (ILO) and dimensionality or correlation.

Note that not every feature follows the same formula for aggregating input metrics. Firstly, the aggregation relies on human inputs like the mapping and focus established in the taxonomy (Appendix A), and the expertise gathered from NGO's and academics. Next, the S-Ray model leverages unsupervised learning techniques to further improve aggregation.

Short-term correction

Because the report-based long-term trend is relatively slow to change – companies typically report on an annual basis – S-Ray also constructs a more frequent short-term signal based on news controversies and NGO campaigns. Looking back one year on a daily basis for each topic, S-Ray first aggregates news-based controversies using a proprietary present news value (PNV), which is a function of an article's controversy

⁴ For example, two highly similar and correlated inputs may be combined into one principal component.

⁵ When appropriate, we consider the expertise of NGO's and academics for specific topics (e.g. the work of Transparency International for the business ethics topic).

level, how long ago it occurred, and the impact of the source. Additionally, S-Ray takes into account relevant NGO campaign activities⁶ from the past year, which get added on top of the news-based signal.

Example: Community relations

In the past year, 3 different (and unique¹) news stories for a company related to community relations are found, in addition to 1 (negative) NGO campaign. To aggregate the three stories, S-Ray combines their respective present news values, which depend on its level of controversy, how long ago it was published and the influential nature of the source. The resulting community relations PNV, which is updated daily, is transformed into a news-based controversy correction (%) for every company. Additionally, a further correction is added based on the negative NGO campaign.

Final feature scores

To find the final feature scores, each of the 22 long-term trend scores (0-100) are multiplied with the matching short-term correction (in %) which is a combination of the news-based controversies and NGO campaign activity.

Example: Business ethics

In the first step, the 18 report-based inputs are aggregated into the long-term trend score (e.g. 65). Next, all business ethics-related news controversies and NGO campaigns over the past year are combined into a short-term correction (e.g. – 10%). The final business ethics feature score for the day is found by multiplying the long-term trend score with the controversy correction (i.e. $65 \times (100-10)\% = 58.5$).

For every company, the total feature scores represent a good approximation of sustainability performance across a range of 22 complementary sustainability topics, drawing from a variety of data types and inputs. The feature scores can now be used as the starting point to further calculate a variety of scores, which highlight different aspects of corporate sustainability performance.

Step 3: Score Layer

As sustainability is a multi-faceted concept, S-Ray offers a variety of scores which are available to subscribers through the score layer. Building from the 22 topics and 11 business involvements tracked in the feature layer, we currently offer three complementary S-Ray scores.

GC Score

The GC Score provides a normative assessment of companies based on the four core principles of the United Nations Global Compact (GC): human rights, labour rights, the environment and anti-corruption. With more than 9,000 corporate signatories from over 160 countries, the Global Compact is the world's largest corporate

⁶ Note that NGO campaign activity can be both negative (critique) or positive (appraisal) in nature.

sustainability initiative. It calls on companies and stakeholders to conduct business responsibly, and to pursue opportunities that advance sustainable development goals. With Arabesque S-Ray these principles are quantified for the first time, with the potential to inspire more companies to take shared responsibility and join the Global Compact in its commitment to achieve a sustainable and inclusive global economy.

Category scores

The first step of computing the GC Score consists of mapping the relevant⁷ features into each of the four GC categories (Table). A distinction is made between features that focus more on negative aspects and features that are more positive in nature, with the former taking precedence in case of poor performance to avoid compensation. For example, if evidence is found that a company is actively violating human rights but at the same donating a lot of money through its foundation, the S-Ray algorithm will almost completely discard the positive features and focus on the more negative ones. The result of the feature aggregations are four GC category scores (0-100) that reflect a company’s current performance across human rights, labour rights, the environment and anti-corruption.

Human Rights	Human Rights, Labour Rights, Occupational Health and Safety, Employment Quality, Diversity; Product Quality and Safety, Product Access, Community Relations
Labour Rights	Labour Rights, Occupational Health and Safety, Diversity, Compensation, Training and Development, Employment Quality
Environment	Emissions, Waste, Environmental Stewardship, Environmental Management, Resource Use, Water, Environmental Solutions
Anti-Corruption	Business Ethics, Corporate Governance, Transparency

Table 2. GC score feature mapping

Total score

In addition to the four GC category scores, S-Ray also computes an aggregate GC score using a non-compensatory aggregation approach that reflects the nature of the GC principles. Every GC category starts with a weight of 25%, but gets more weight allocated as the score starts dropping below 50 (i.e. the neutral center). As a result, it is not possible for a company to compensate poor performance in one category with great performance in another. As performance deteriorates for any of the GC categories, more weight gets shifted to that category, which will be driving the overall GC Score.

Both the four GC category scores and the total GC Score are scaled between 0 and 100, with higher scores indicating better performance. We believe the GC Score can be used to approximate reputational risk facing companies with lower scores correlating with a higher likelihood of a company losing its license to operate in the future (e.g. due to a major scandal or controversy). As such, the GC Score also provides a helpful tool to investors to help limit downside risk.

⁷ Note that features which are considered out of scope for the GC principles are not used.

ESG Score

While the GC Score builds from a normative framework to approximate reputational risk, the ESG Score is all about identifying companies that are better positioned to outperform over the long term. To understand the potential for long-term performance, the S-Ray methodology considers the principle of financial materiality.⁸ That is, when computing the ESG Score of a company the algorithm will only use information that significantly helps explain future risk-adjusted performance. Materiality is applied by more heavily weighting features with higher materiality, and weights are rebalanced on a rolling quarterly basis.

Each quarter for every company, materiality is assessed on a sector- and industry-level, using both equal- and marketcap-weighted monthly index returns over a period of the past 1, 3 and 5 years. This results in 12 different portfolio index returns for each company. For each portfolio, the materiality is then calculated with the following steps:

Step 1: Static materiality

During the first step, the S-Ray model assigns a baseline materiality to each feature in each portfolio. The baseline values assigned to each feature are based on third-party research and industry reports that have looked at the materiality question. These theoretical frameworks are a valuable input to obtain a first understanding of which categories are material in understanding a company's ability to outperform in the long run.

Step 2: Data-based materiality adjustments

In the second step, the S-Ray model then considers how much of the variation in returns can be explained by each of the features. This is achieved by first applying a multi-factor asset-pricing model to obtain residuals (i.e. the unexplained part of variation in returns), followed by a recursive feature elimination procedure with cross-validation to identify those features that can explain a significant part of the variation in residuals from the multi-factor regression. Features which are found to be material during this process are assigned more weight.

The static materiality score from Step 1 is then added to the dynamic materiality score from Step 2 and normalized to obtain a total feature weight.

Total score

The total ESG score is calculated as a weighted sum of the feature scores using materiality-based weights. The E, S and G pillar sub-scores are calculated by considering only the features within each of these themes.

Like the GC Score, the three pillar scores and the total ESG Score are scaled between 0 and 100, with higher scores indicating better performance. While the GC Score takes a normative approach to sustainability and approximates reputational risk, the ESG Score is calibrated using the principle of financial materiality and can be used to help compare companies on their ability to outperform on a risk-adjusted basis over the long run. For investors, the ESG Score can be instrumental to help identify investment upside, as a complement to the GC Score in protecting downside.

⁸ We refer to the materiality definition by the U.S. Supreme Court, which states that information is material when "there is a substantial likelihood that the disclosure of the omitted fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available." Source: TSC Indus. v. Northway, Inc., 426 U.S. 438, 449 (1976).

Preferences Filter

As sustainability is also a personal matter, which depends on individual preferences, S-Ray provides a Preferences Filter covering a set of 11 business involvements (Table 3). To determine a company's business involvements, S-Ray collects revenue-based inputs, which are transformed into 11 business involvement feature scores. These feature scores approximate the level of involvement for each business activity and company. In the last step, the 11 feature scores are transformed into 11 corresponding flags, which provide an answer to the below questions for each involvement.

Adult Entertainment	Does the company derive significant revenues from adult entertainment products?
Alcohol	Does the company derive significant revenues from the production and/or sale of alcohol?
Defense	Does the company derive significant revenues from defense contracting?
Fossil Fuel	Does the company significantly exploit fossil fuel-based energy sources?
Gambling	Does the company derive significant revenues from gambling?
GMO	Does the company significantly engage in research and/or production of genetically modified organisms (GMO) based products?
Nuclear	Does the company significantly rely on nuclear power and/or the non-military use of uranium?
Pork	Does the company derive significant revenues from the sale and or/production of pork-based products?
Stem Cells	Does the company derive significant revenues from stem cell (research)?
Tobacco	Does the company derive significant revenues from the sale and/or production of tobacco?
Weapons	Does the company significantly engage in the sale and/or production of weapons?

Table 3. List of 11 business involvements

The preference filters flags are binary indicators which alert users on the involvement of companies across a range of 11 business activities. S-Ray is programmed to be unbiased across all the flagged activities; the flags are not used to adjust GC or ESG scores. Rather, we see the flags as a tool for users to check companies against their personal values.

arabesque s-ray

Disclaimer

Arabesque S-Ray® is a service (the “**Service**”) provided by Arabesque S-Ray GmbH (“**Arabesque S-Ray**”) and its branch and subsidiary companies. Arabesque S-Ray is a limited liability company (*Gesellschaft mit beschränkter Haftung*) incorporated in Frankfurt am Main and organized under the laws of Germany with registered number HRB 113087 in the commercial register of the local court with its seat and business address at Zeppelinallee 15, 60325 Frankfurt am Main, Germany. The Service is unconnected to any of the asset management activities conducted within the wider group of Arabesque companies, and is not investment advice or a solicitation or an offer to buy any security or instrument or to participate in investment services. RELIANCE - Arabesque S-Ray makes no representation or warranty, express or implied, as to the accuracy or completeness of the information contained herein, and accepts no liability for any loss, of whatever kind, howsoever arising, in relation thereto, and nothing contained herein should be relied upon. ENQUIRIES – Any enquiries in respect of this document should be addressed to Arabesque S-Ray.

Appendix A: Examples of report-based inputs

Name	Topic/Feature	Focus
Adult Entertainment Revenue	Adult entertainment	Business Involvement
Alcohol Revenue	Alcohol	Business Involvement
Anti-Corruption Policy	Business ethics	Preparation
Employee Anti-Corruption Training	Business ethics	Preparation
Whistle-Blowing Procedures	Business ethics	Preparation
Anti-Corruption Audits	Business ethics	Preparation
Leverage	Capital structure	Outcome
Corporate Responsibility Awards	Community relations	Outcome
Crisis Management Systems	Community relations	Outcome
Employee Volunteering Disclosure	Community relations	Outcome
Community Involvement Policy	Community relations	Preparation
Compensation Committee	Compensation	Outcome
Long Term Objectives	Compensation	Outcome
Sustainability Compensation Incentives	Compensation	Outcome
Equity Compensation Plan	Compensation	Outcome
Senior Executives' Code of Ethics	Corporate governance	Preparation
Audit Committee Independence	Corporate governance	Outcome
CEO Chairman/Board Separation	Corporate governance	Outcome
Board Diversity Policy	Corporate governance	Preparation
Defense Revenue	Defense	Business Involvement
% Women Managers	Diversity	Outcome
Diversity and Opportunities Policy	Diversity	Preparation
Proportion of Women on the Board	Diversity	Outcome
Supplier Diversity Programs	Diversity	Preparation
Staff Transport Impact Reduction Initiatives	Emissions	Outcome
Greenhouse Gas Emissions Disclosure	Emissions	Outcome
Emissions Trading	Emissions	Outcome
Emissions Reduction Targets	Emissions	Preparation

Day Care Services	Employment quality	Outcome
Flexible Working Schemes	Employment quality	Outcome
Employee Satisfaction Measurement	Employment quality	Outcome
Work-Life Balance Policy	Employment quality	Preparation
% ISO 14000 or EMS Certified	Environmental management	Outcome
Environmental Partnerships	Environmental management	Outcome
Environmental Grievance Mechanisms	Environmental management	Preparation
Environmental Expenditures	Environmental management	Outcome
Hybrid Technology	Environmental solutions	Outcome
Eco-Design Products	Environmental solutions	Outcome
Clean Technology	Environmental solutions	Outcome
Sustainable Building Products	Environmental solutions	Outcome
% Labelled Wood	Environmental stewardship	Outcome
Biodiversity Impact Reduction	Environmental stewardship	Outcome
Animal Testing	Environmental stewardship	Outcome
Environmental Restoration Initiatives	Environmental stewardship	Outcome
Forensic Accounting - Overall Earnings Quality	Forensic accounting	Outcome
Fossil Fuel Revenue	Fossil fuel	Business Involvement
Gambling Revenue	Gambling	Business Involvement
GMO Revenue	GMO	Business Involvement
Human Rights Policy	Human rights	Preparation
Supports UDHR	Human rights	Preparation
Monitoring Suppliers' Human Rights	Human rights	Outcome
Employee Human Rights Training	Human rights	Preparation
Labour Rights Policy	Labour rights	Preparation
Trade Union Representation	Labour rights	Outcome
Abides by ILO Core Labour standards	Labour rights	Preparation
Supplier Code of Conduct	Labour rights	Preparation
Nuclear Revenue	Nuclear	Business Involvement
Lost Working Days	Occupational health and safety	Outcome
HIV-AIDS Program	Occupational health and safety	Preparation

Health and Safety Certifications	Occupational health and safety	Preparation
Safety Training	Occupational health and safety	Preparation
Pork Revenue	Pork	Business Involvement
Health Care Access Programs	Product access	Outcome
Low Price Product Access	Product access	Outcome
Affordable Housing Access Programs	Product access	Outcome
Food Access Programs	Product access	Outcome
FDA Warning Letters	Product quality and safety	Outcome
Customer Data Privacy Policy	Product quality and safety	Preparation
Product Responsibility Policy	Product quality and safety	Preparation
Customer Satisfaction Measurement	Product quality and safety	Outcome
Renewable Energy Targets	Resource use	Outcome
Green Buildings	Resource use	Outcome
Toxic Substances Reduction Initiatives	Resource use	Outcome
Environmental Management Team	Resource use	Preparation
Adult Stem Cells Revenue	Stem cells	Business Involvement
Embryonic Stem Cells Revenue	Stem cells	Business Involvement
Tobacco Revenue	Tobacco	Business Involvement
Internal Promotion	Training and development	Outcome
Supplier ESG Training	Training and development	Outcome
Career Development Policy	Training and development	Preparation
Average # Training Hours	Training and development	Outcome
Board Remuneration Disclosure	Transparency	Outcome
CSR Committee	Transparency	Preparation
Executive Incentive Plan Disclosure	Transparency	Outcome
Stakeholder Engagement	Transparency	Outcome
Oil Spill Disclosure	Waste	Outcome
Waste Recycling Ratio	Waste	Outcome
Waste Programs	Waste	Preparation
Numerical Waste-Related Data Disclosure	Waste	Outcome
Water Strategy	Water	Preparation

Water Recycling/Reuse Disclosure	Water	Outcome
Numerical Water-Related Data Disclosure	Water	Outcome
Total Water Use	Water	Outcome
Chemical and Nuclear Weapons Revenue	Weapons	Business Involvement
Cluster Bombs Revenue	Weapons	Business Involvement
Firearms Revenue	Weapons	Business Involvement
Landmines Revenue	Weapons	Business Involvement