D4.4.2 Rating Concept Evaluation

D4.4.2

Dissemination level	Public
Contractual date of delivery	Month 42 | 31.03.2017
Actual date of delivery	10.05.2017
Work package	WP4 | Corporate Social Responsibility Rating
Deliverable number	D4.4.2 | Rating Concept Evaluation
Type	Report
Approval status	Approved
Version	1.1
Number of pages	35
File name	D442_170510_V1_1_WikiRate_Rating_Concept_Ev.docx

Abstract

This deliverable covers the steps taken by the project consortium to test, refine and update WikiRate's rating functionality. It highlights processes, developments and futures recommendations.

SEVENTH FRAMEWORK PROGRAMME

Co-funded by the European Union

History

Version	Date	Reason	Revised by
0.1	4/26/17	Initial framing	Ethan McCutchen
0.2	4/27/17	Explain concept, structure report	Ethan McCutchen
0.3	04.05.2017	Draft content	Hala Khalaf
0.4	04.05.2017	Content contribution	Alex Henderson Theresa Heithaus Laureen van Breen
0.5	05.05.2017	Draft consolidation	Hala Khalaf
1.0	05.05.2017	For review	Hala Khalaf
1.1	10.05.2017	Review and final additions	Richard Mills

Author list

Organization	Name	**Contact information**
Decko Commons	Ethan McCutchen	ethan@decko.org
The WikiRate Project	Hala Khalaf	hala@wikirate.org
The WikiRate Project	Theresa Heithaus	theresa@wikirate.org
The WikiRate Project	Alex Henderson	alexandra@wikirate.org
The WikiRate Project	Laureen van Breen	laureen@wikirate.org
Cambridge University	Richard Mills	rm747@cam.ac.uk

Executive Summary

The following report evaluates the WikiRate metrics and ratings system that handles the wealth of data available or researchable on corporate social and environmental impacts. D4.4.1 discussed the implementation of the first rating features in M18 of the project. This report (D4.4.2) goes on to show the applications of the metric – company – topic approach, whilst criticizing and discussing the breadth of applications of the rating concept, alongside possible improvements, refinements and further directions for the overall system. This report covers the practical applications of the rating system – particularly the ongoing use of research metrics by NGOs and Academics in practical research contexts. With research ongoing into corporate performance on the Sustainable Development Goals (SDGs) alongside WikiRate's partnership with United Nations Global Compact's Principles for Responsible Management Education Initiative (PRME) and participating Universities, and the increasing number of NGO partnerships – there is much use of the research metrics which underpin the rating system at the heart of WikiRate. This report goes on to explore future directions for metrics and ratings, alongside future directions for research and use of metrics and ratings on WikiRate and further afield too.

Table of Contents

List of Figures

1 Introduction

WikiRate is working to make corporate sustainability data useful to more people. The rating concept, as originally proposed, was designed to do so by offering approachable high-level ratings of companies, using indicators gleaned from corporate social responsibility (CSR) data as building blocks. But efforts to produce clean ratings from disorganised CSR data soon made it clear that the rating concept needed to be integrated with a research concept.

International efforts to standardise corporate reporting have admirably defined a rich array of measures of corporate performance, but they've focused little energy in standardising the format in which these measures were presented. The Global Reporting Initiative collaborated with WikiRate to translate a select set of G4 indicators to 168 WikiRate *research* metrics. Yet, the underlying issue lies in the fact that most CSR reporting take the form of glossy PDF presentations rather than, say, XBRL documents. What's more, sustainability concerns and scholarly responses to them advance so rapidly that it's challenging for metric design to keep pace.

By offering WikiRate's model to interested user groups and communities, researching static reports becomes less strenuous and more coherent. Each metric on WikiRate asks a question, offers context behind that question and provides researchers with a methodology to follow specific to the metric.

The WikiRate model also attempts to unify data collection from the plethora of standards concerning corporate ESG behaviour. Say, two different standards may offer similar indicators for example on, Scope 3 Emissions, WikiRate's functionality for compare and contrast provides researches with the necessary tools to do just that, all in one place.

The Metrics Framework as defined in D4.4.1 remains valid in in concept but includes redefinition of technical aspects pursuant to the revision of objectives (D7.7.2-resubmission) and in line with the critical assessment that ran throughout the 18-month period following the first WP4 deliverable. The assessment -and subsequent evaluation- was mainly conducted by WikiRate partners (particularly the WikiRate e.V., Cambridge University and Decko Commons) and evaluated against criteria as outlined in this report.

Metrics Framework: ***Ratings****: A framework within which its members can gather existing metrics, create new ones, and bring the best among them all to the attention of a wider audience. The "best" (or "most important") metrics are based upon up-down voting.* ***Addendum****: Now, both Company and Topic pages have been streamlined, de-emphasising or hiding older qualitative functionality (Articles now Reviews, and claims now Notes, etc.) and emphasising on their respective relationships with Metrics. In concept, when we, as a community, collaborate to build better metrics, we are providing essential tools for making companies better, together.*

WikiRate's metric research approach is versatile, in that it can bring together an array of data from different reporting standards like the GRI and the Poverty Footprint, and align, or map these to developing frameworks like the SDG Compass. With the private sector now included as a stakeholder and contributor to the UN Sustainable Development Goals, it is critical to understand where companies are leading and lagging in order to inform change and improvement. *Topics* are one type of shell for capturing sets of *Metrics* on WikiRate that can be combined within higher level topics to create mappings of Metrics to multiple frameworks at once.

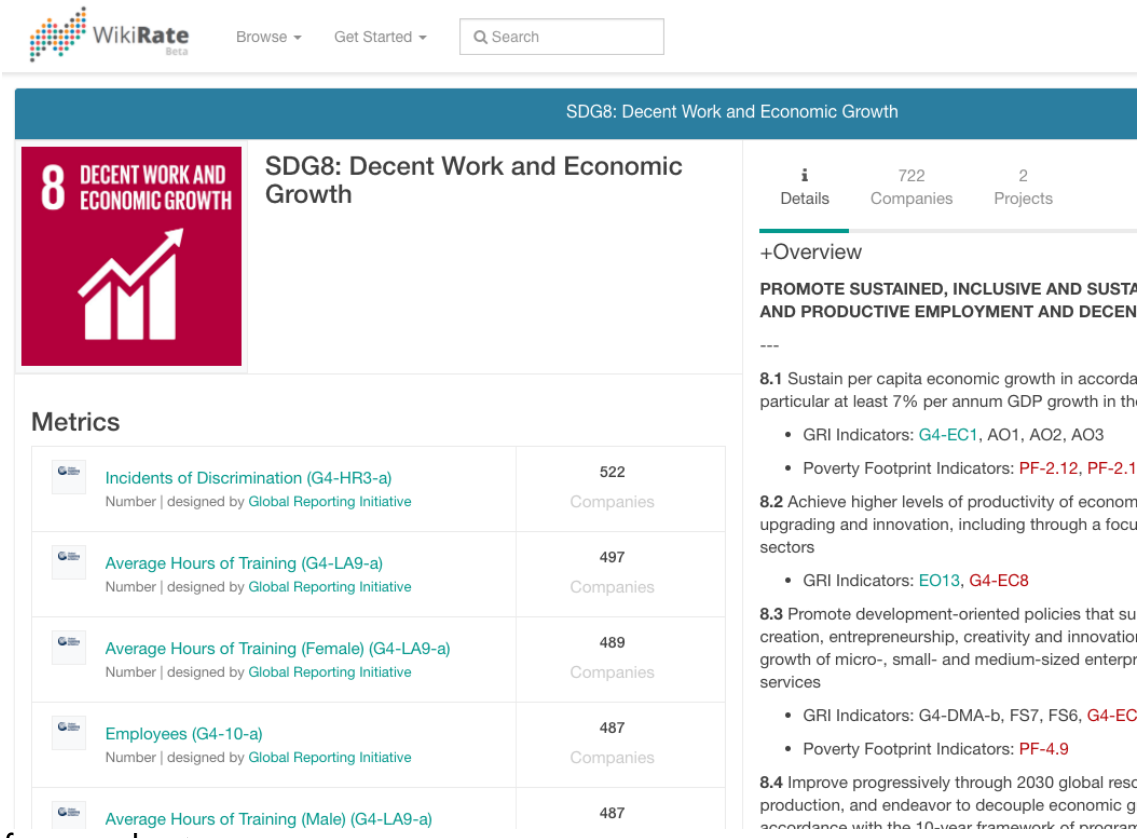


Figure 1 WikiRate Topics Screenshot

Applying multiple frameworks to data enables agile research, in this case around companies' contributions towards the SDGs. With goals as complex as the SDGs, many interpretations will be necessary to begin tracking and testing progress – the more diverse the individuals and groups working with the data, the better the wider public, along with policy makers, and companies themselves, will be able to connect to and advance the issues.

Academics, NGOs, employees, policy makers, app developers, investors, and companies themselves all stand to benefit from abundant, reliable, structured, clean data.

2 The Evolved Rating Concept

The Rating Concept has developed significantly since its initial presentation in the grant proposal and (to a lesser extent) since its description in D4.4.1 *Implementation of and Report on Rating Concept*. We will not review the entire evolution of every feature, but will note the progression of principles before going through the current features of the implemented system.

Principles

Many of the core principles framing what WikiRate.orgs' rating system should achieve and what WikiRate e.V's role should be in it have not changed since the original proposal.

1. **WikiRate e.V. will remain neutral with regard to all issues**. WikiRate e.V. will not be responsible for defining what constitutes good behaviour (and results in a high rating) in relation to particular topics or issues. It is important that the WikiRate organisation is not perceived as inherently for or against companies.
2. **…except corporate transparency**. Without sufficient transparency, any rating system falls apart. WikiRate therefore must seek to reward transparency.
3. **WikiRate's rating system must be based on sourced data points**. All researched metric answers (as opposed to those calculated automatically) must cite Sources.
4. **Companies should be welcomed as active contributors.** We want companies to be active on WikiRate in providing data and responding to questions, and we also want their most informed critics to play the role of assessing their performance.
5. **WikiRate's rating system should dynamically generate high-level company ratings**. High-level ratings make the data more compelling and support clearer narratives. The manner in which these ratings are generated must be entirely transparent. These ratings should be calculated dynamically. When data points are added or updated, the ratings should be re-calculated.

However, and as previously mentioned, some additional principles were agreed upon during the grant period:

1. Rather than yet-another-ratings -framework, **WikiRate.org *is* an arena in which many frameworks can co-exist and integrate**. This was a significant conceptual shift for the project, and lay the foundation for WikiRate's emergence as a tool for diverse actors in the CSR domain.
2. **Indicators, scores, ratings, etc. are all Metrics**. Unifying all of these concepts, so that they can all have shared properties and functionality and all can be combined to form new metrics, is central to our model.

3. Unlike the organization WikiRate e.V., **WikiRate.org community members are not necessarily neutral.** Different systems of measuring a company's performance may have different levels of inherent bias. Community members are welcome to advance different perspectives through metric design.
4. **Anyone can create a metric, but the best metrics should rise to the top**. Without over-specifying "best", the principle is that metrics that the community deems valuable should receive the most prominence on the site. The idea is that WikiRate e.V. should not be a gatekeeper, but the WikiRate.org community should be given power to convey and aggregate determinations of metric quality.

These design principles are also presented as more general knowledge statements that go beyond metrics and ratings in D3.3.4.

The fundamental goal underlying these principles is to make WikiRate a platform that produces useful knowledge about companies' behaviour and presents this in an accessible way.

Metrics Basics

The basic building block of the Rating Concept is the Metric - a standardised way of measuring some aspect of a company's performance.

A Metric can be very high-level, like a "company's respect for right to privacy" or very low-level, like "Annual Revenue". A Metric is like a question that can be asked about any company, and in fact each metric has a question field (card) in which the metric should be posed in the form of a question. Some sample Metrics:

Metric Designer	Metric Title	No. Companies
Researched Metrics		
Global Reporting Initiative (GRI)	Direct greenhouse gas (GHG) emissions (Scope 1) (G4-EN15-a)	499
Amnesty International	Conflict Minerals Report	661
Calculated Metrics		
Combined Scope 1 and 2 Greenhouse Gas emissions	Richard Mills	475
HESA: Total Renewable Energy Consumed (kWh)	John Callewaert	1

Figure 2 Metric Types

The above examples have been selected to illustrate the diversity of types of information that metrics will display.

Each metric asks a question and can be populated for a given company with a Metric Answer. Metric Answers can be numerical, categorical, or free text; the only constraint is that it must be possible to apply the metrics to companies in a standard way and provide at least one source to support the answers entered.

Each metric answer is connected to a year or range of years. We have chosen the year as the primary organizing time unit on WikiRate because of its dominant use in CSR reporting, government reporting, and conventional ratings systems, each of which provides significant source material for WikiRate. Year-based values will allow companies' performance to be tracked over time and will support archival handling of companies that are disestablished or reorganized.

Metric Examples and Interface

While any given example or interface is not, of course, germane to the Rating Concept, it is helpful to demonstrate the concept in practice.

Many users will first encounter metrics on WikiRate.org on a Company page. (When, in the future, WikiRate shifts its focus towards driving higher traffic, it is expected that a central strategy will be SEO of company search terms, with the ambition that WikiRate's company page follow immediate after the company itself.) Each Company page lists the Metrics for which values are available.

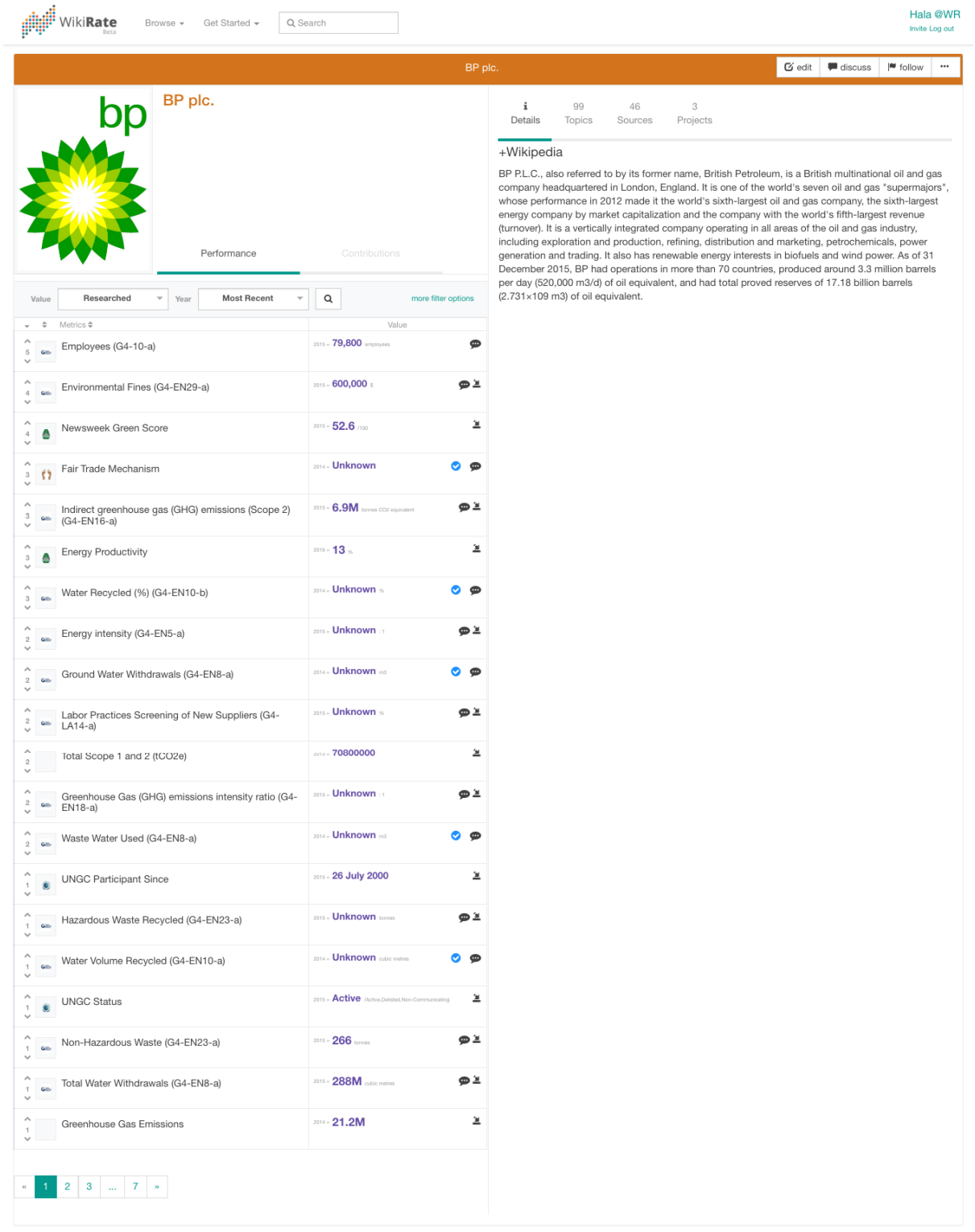


Figure 3 Company Page Screenshot

It is worth noting that, in addition to being the *subject* of a metric answer, a Company can also be a metric designer. The contributions tab links to page showing all contributions made to WikiRate by a formal company account (designing a metric, organizing a research group, etc). This approach is, of course, in keeping with our stated principle that *Companies should be welcomed as active contributors*.

Beneath the company you will see a list of metrics, or more precisely a list of *metric records.* A *Metric Record* is a group of all the Answers associated with a given metric and company (regardless of year).

By default, all metrics with answers are sorted by metric *importance*, currently implemented as a simple Reddit-style up-down vote with interface visible to the left of each metric designer logo. Other sorting options are available on column headers, the most common filters are available just above the headers, and *more filtering options* may be reached by the eponymous link.

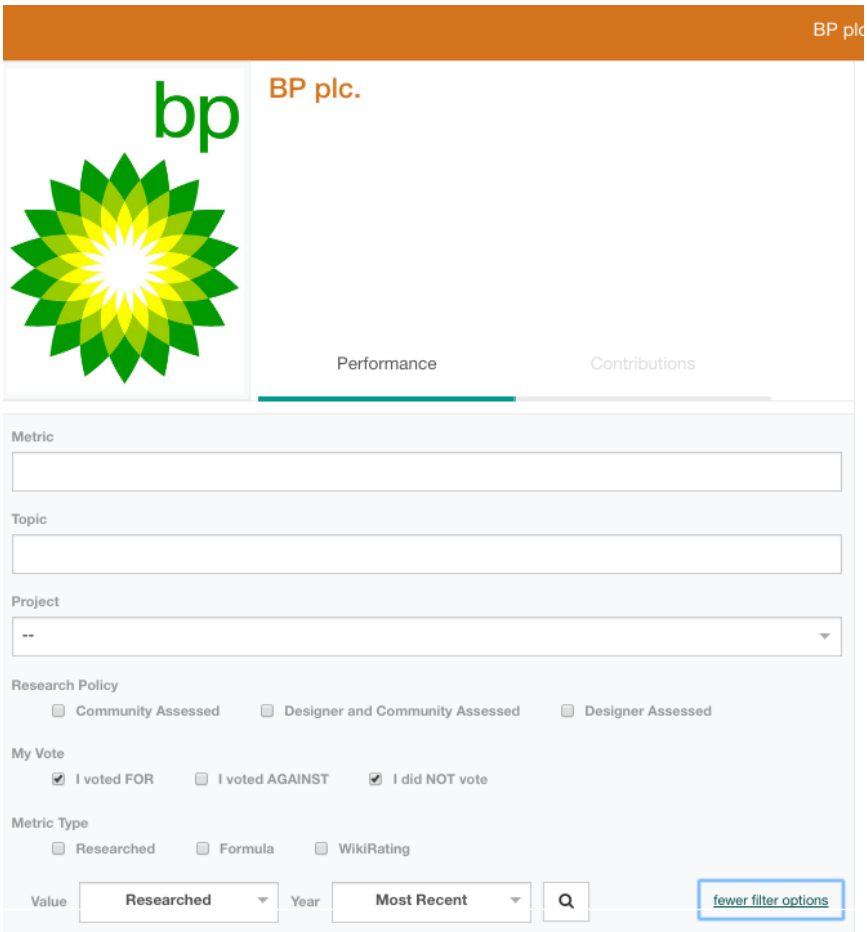


Figure 4 Filter Interface from a Company's Page

Clicking on a given record provides you a view of a *metric record detail;* a list of all the metric answers available for this metric and company, along with a visualization showing this company's place in the answer's distribution among all companies on WikiRate.org.

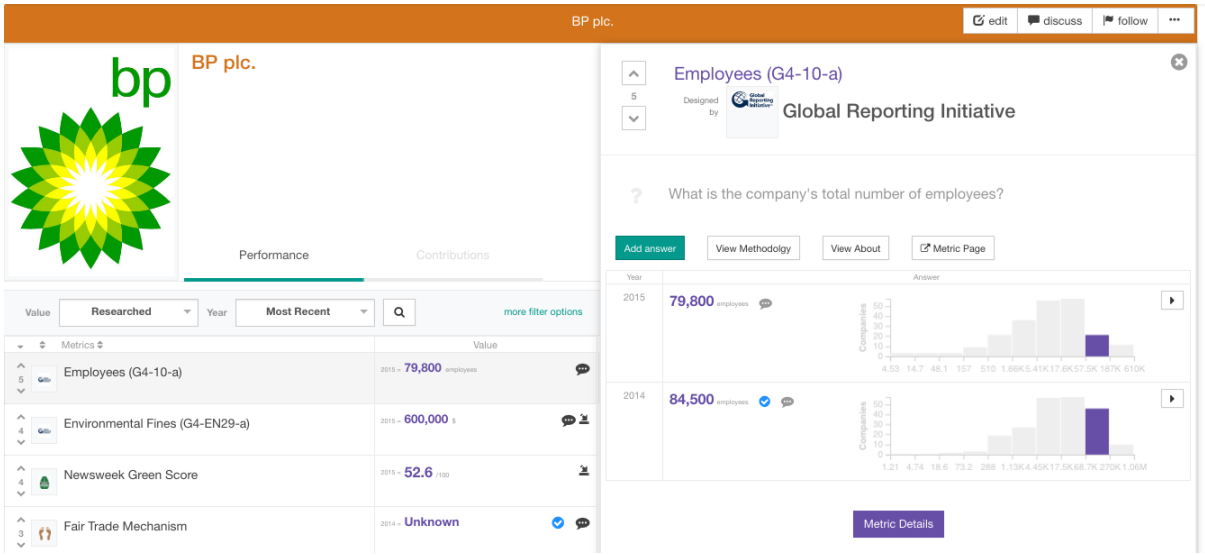


Figure 5 Expanded Record Listing

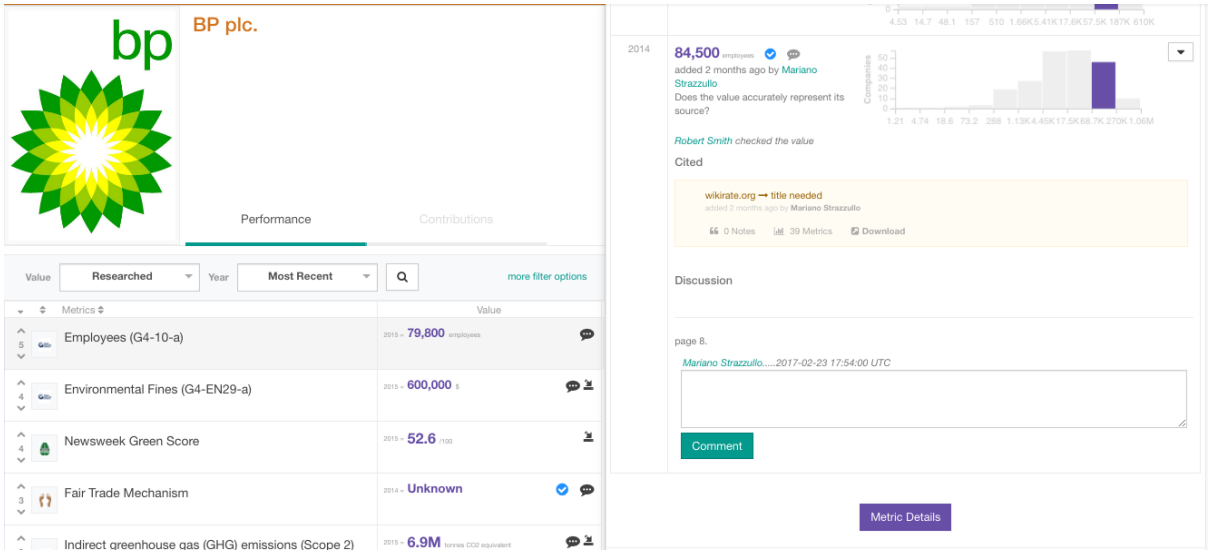


Figure 6 Expanded Metric Answer on Same Record

You can expand any given answer to see more detail about the answer's editors, verification (or double checking), source(s), and discussion.

From a company page you can navigate to any given metric to learn more about the metric and to compare company answers. Here is an example of a Metric page in its current form on WikiRate.org:

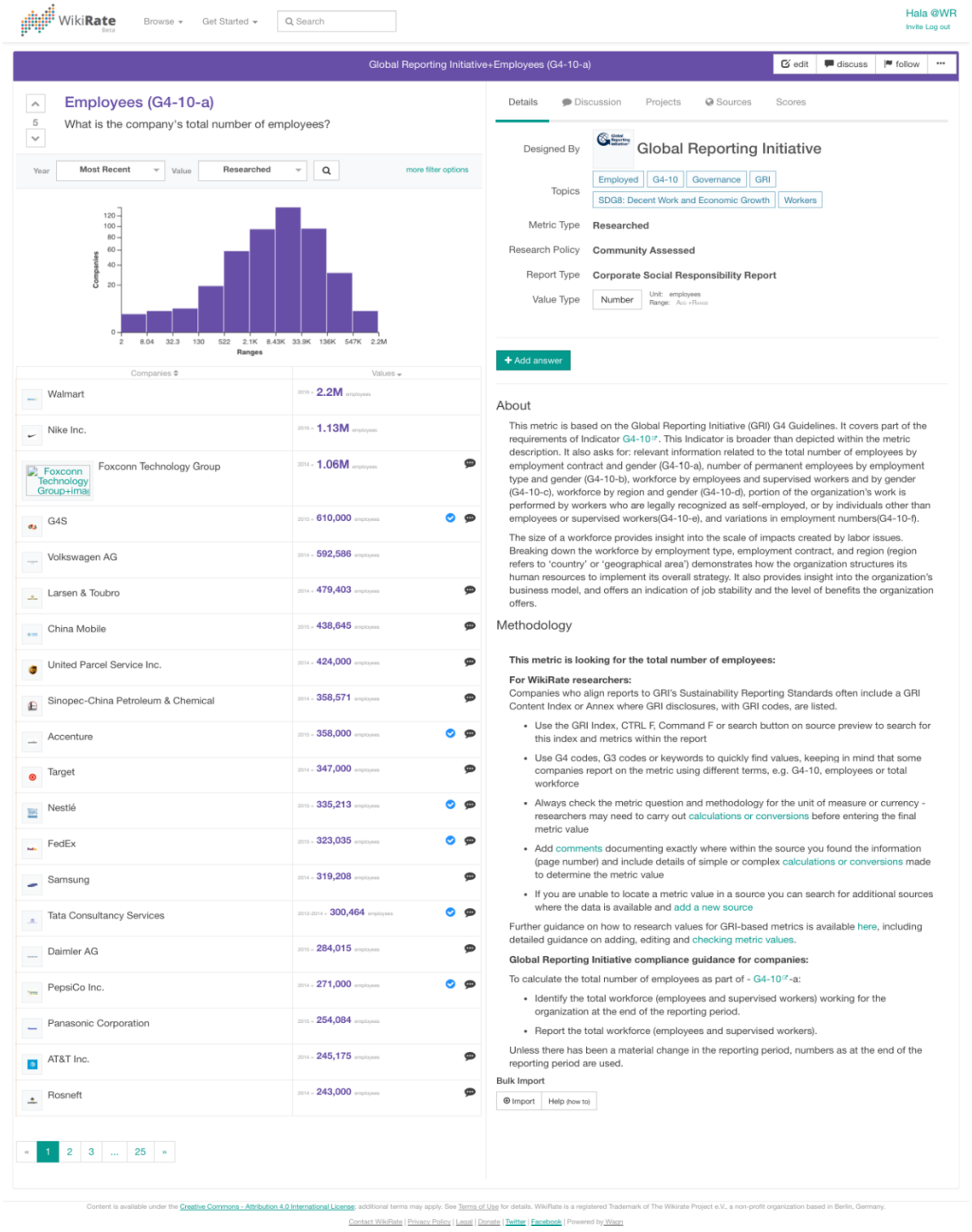


Figure 7 Metric View from Company Page

The metric page layout is largely the reverse of the company page. Here, of course, the company values are compatible, making them straightforward to visualize. Clicking a column on the prominent distribution graphic will immediately filter the company records below with applicable answers in that group.

Several fields in the details section warrant a brief explanation:

- **Research Policy** is a socially enforced rule governing who can edit a metric. Some metrics are designed for open community research (*the same metric above: Global Reporting Initiative+Employees (G4-10-a) is open for community research*). Others are intended to be edited and have data from only by the metric's designer (they have a *designer assessed* policy).
- **Report Type** specifies the type of source report – e.g. CSR Report, Modern Slavery Statement, etc. – in which the metric's answers are typically found. Associating a metric with a given report type facilitates automated source suggestions. WikiRate.org currently has 8 report type cards and users can create cards for the type of source they are citing.
- **Value Type** specifies the data type constraints of the metric's answer values.

Another field, Metric Types, warrants an entire subsection, which follows.

Metric Types

The metrics used in examples above are all known as *Researched* metrics (previously named simple), meaning that their values are directly entered into the database and not altered unless directly edited. This distinguishes them from *Calculated Metrics,* which are dynamically determined from other metric answers and are automatically updated when values change.

Researched and Calculated Metrics are not distinguished by the nature of their content – either can be low-level indicators or high level ratings – but by the method that determines their value.

There are three types of Calculated Metrics:

- **Formulas** generate new values from existing data using free-form mathematical formulas and make use of Wolfram Language integration
- **Scores** normalise answer values on a 1-10 scale
- **WikiRatings** combine scores with weighted averages

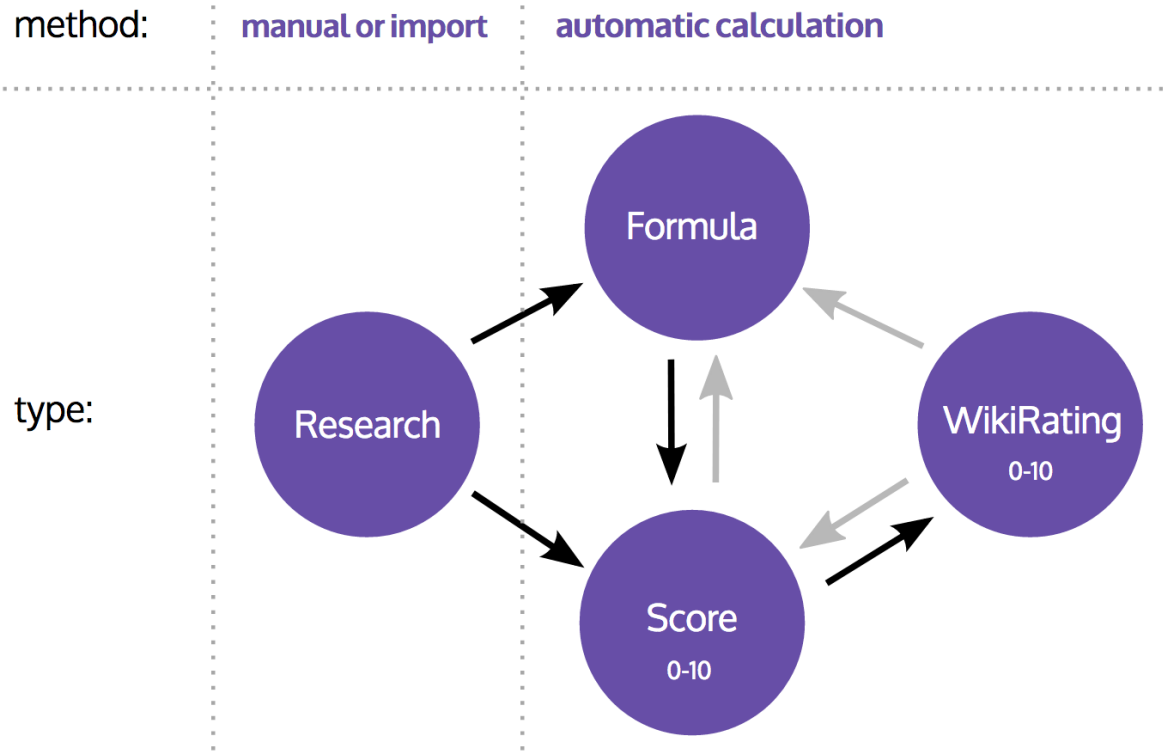


Figure 8 Metric Connections

As Figure 8 shows, the metric types are deeply related and reusable. It is key to understand that all calculated metrics ultimately derive their values from research metrics (even if, for example, a WikiRatings is based on Scores of Formulas, the Formulas must eventually connect to researched metrics.

As will be explained below (see *Application of Calculated Metrics*), calculated metrics are not yet in heavy use, but a brief perusal may help with context.

Formulas are very powerful but not yet very discoverable, as there is as of yet no substantial support for writing the syntax that makes it possible to refer to other answers. They are intended for mathematically inclined users or users who want to perform a specific mathematical/logical operation on metric data-sets.

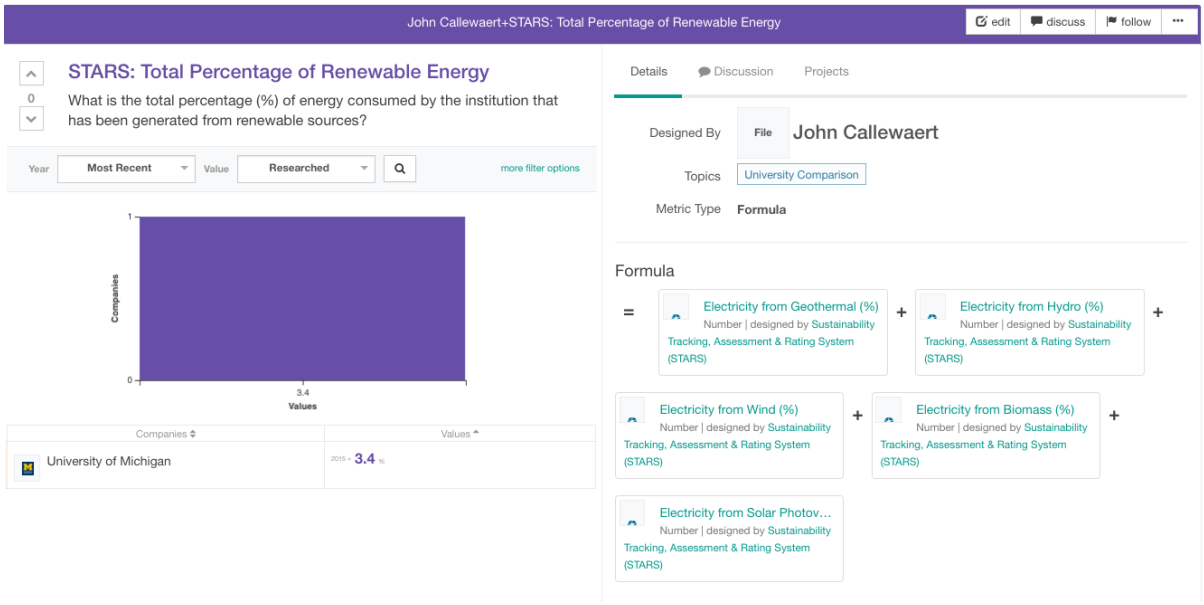


Figure 9 Formula Metric

Scores are far simpler than Formulas and can be used to translate numerical and categorical metrics into numbers on a 0-10 scale. This 0-10 scale reflects judgments about performance (low numbers represent poor performance, high numbers good performance). To Score a categorical metric, a user decides what score each of its response options should represent. To score a numerical metric, a user can specify how ranges map onto 0-10 scores – or, they could use a formula metric that converts the numerical data of interest onto a 0-10 scale algorithmically. 0-10 scores are the glue that holds the Metrics system together and ensures that any type of metric can be included in a WikiRating – because any kind of metric can be Scored, and it is these Score versions that feed into WikiRatings.

Score metrics allow for the expression of value judgments in a way which doesn't affect the underlying researched data. Any metric can be scored differently by multiple users, these scores do not change the data but instead offer different interpretations of how that data should be used in Ratings. Scores are intended to give an outlet for users' value judgments, in the hope that handling these in a structured way will prevent them from distorting researched data.

It is important for the quality of WikiRate's data that the community embraces the importance of neutral and dispassionate metric research, Scores and WikiRatings give a way to reflect on the relative performance of companies but for these to work well the underlying data must be reliable.

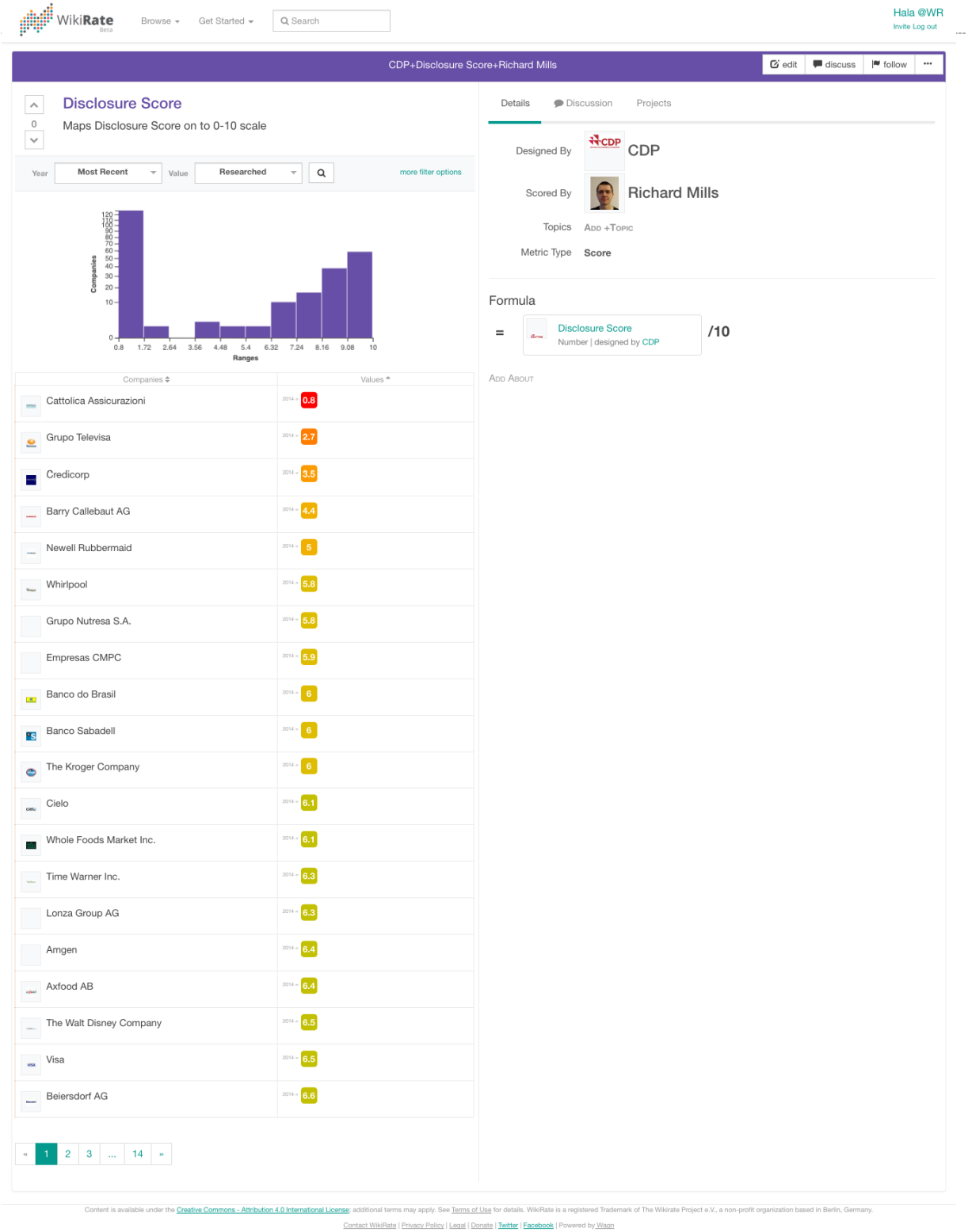
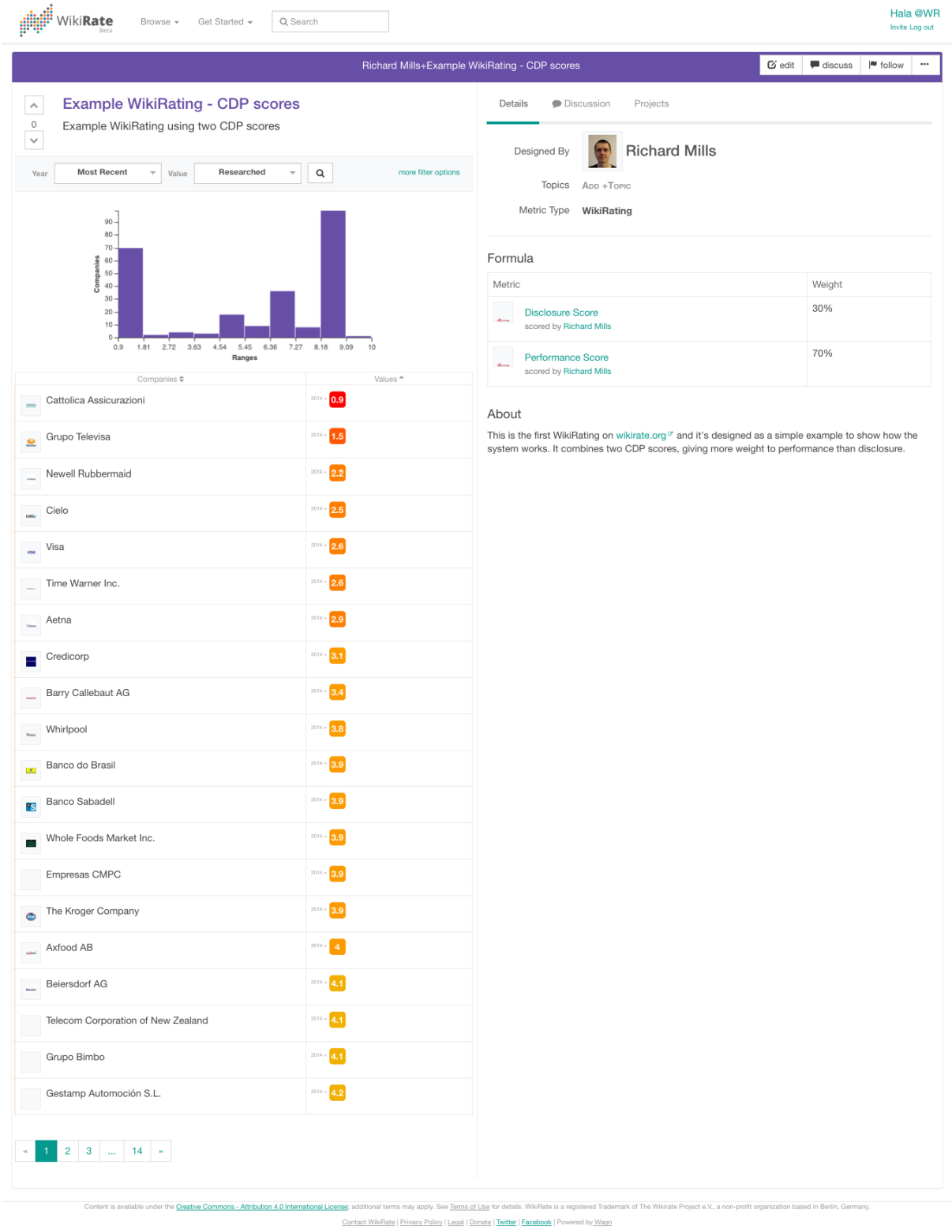


Figure 10 Score Metric Screenshot



WikiRatings are conceptually little more than weighted averages with a friendly interface.

Figure 11 WikiRating Metric Screeshot

This simplicity is intended to make them easy to create and understand.

These examples of calculated metrics can be used for more complex analyses: For example, the Centre for Sustainable Organizations have begun developing WikiRate metrics that measure progress on climate action (SDG 13). The CSO context-based carbon metric has been designed to connect live input data so that as new data is added, the metric is automatically applied to assess additional companies. This metric considers a company's carbon emissions, assigning each company a "fair share" of global carbon emissions based on how much they contributed to the global economy (using gross margins data), and then assessing whether their carbon emissions have been greater or less than their fair share with a single number (<1 means emitted less than fair share, >1 means more than fair share emitted).

This rationale can be adapted and expanded to analyse relationships between let's say, the World's Resource Institute's science-based targets and CSO's carbon context rating on WikiRate.

Researching Metrics

As explained above, Researched metrics are distinguished from Calculated Metrics by the method that determines their value – namely, values are directly entered into the database and not altered unless directly edited. Data in Researched Metrics serves as input for Calculated Metrics, so quantity, for certain types of data, and quality are key.

Researchers extracting answers to Metric questions, will use Researched Metrics to do so, and *Projects* on WikiRate.org serve to facilitate this research by providing a bespoke frame for conducting research. Each Project includes a set of Researched Metrics and a set of Companies to researched. The interface allows individuals to enter into a page wherein they can conduct research on a company according to those selected metrics in the page.

To increase quality and quantity of data on companies we scrape data from public sources, and engage individual researchers, students, and volunteers in reading public reports and extracting key data to answer Metric questions on WikiRate. There are a number of benefits to engaging individuals: (1) there is a certain learning that comes from direct research into sustainability reporting – students, particularly business students who may go on to work for companies and determine their CSR strategies, learn about theoretical approaches to CSR, but rarely engage with practical exercises like researching company reports to find useful data; (2) Researching Metrics often requires researching non-standardized documents for answers to new Metric questions. Refining metrics is iterative, and requires input and discussion from multiple researchers; and (3) increasing quality through verification.

Partnering with the Principles for Responsible Management Education, a pilot assignment was created to engage students with company sustainability performance as it relates to specific SDGs. Utilising the SDG Compass, which has conducted an initial mapping of different standards organisations' indicators to each of the 17 SDGs, students researched sustainability reports and UN Global Compact Communication on Progress reports, to pull out data associated with the SDG(s) related to their courses topic. The framing – the set of companies a class researches, alongside the specific SDG-related metrics – is structured by the Project pages on WikiRate, tailored to the needs of each classroom or group of researchers.

The first pilot group consisted of 13 courses at 9 different universities, with over 1,000 students involved.

3 Metric Application

Application of Research Metrics

In order to engage individuals in *researching* Research Metrics, we need individuals or groups to *design* metrics. The designer of a metric is the person or organisation that framed its question and established the methodology for answering it – this allows for different entities to design metrics with the same title and general purpose but to each apply their own methodology, keeping the associated data-sets separate and preventing a "land grab" for popular titles like "Climate Change rating" or "Scope 1 Carbon Emissions".

Once a Metric methodology is created by an organisation or individual and translated onto WikiRate.org alongside a Metric, the designer decides who is allowed to edit/augment it through adding a Research Policy. As mentioned above, there are two types of Research Policies: (1) a "Designer Assessed" policy that doesn't allow participants who are not the designer to add or edit its data-points. This allows data researched through the designers "official" methodology to be protected, and still allow data consumers to access the exitsting data; (2) a "Community Assessed" policy opens up the metric to allow additional unaffiliated researchers to add data for further companies. These two research policies are just a starting point, and we anticipate many metrics having a more nuanced policy that falls somewhere between these extremes. If a metric's creator doesn't find the available policies appropriate for how they want their metric to be used, they can create a bespoke policy (which will then be available for other metric designers).

As it stands, there are currently 1,000 Metrics, with 222,496 associated values on WikiRate. That means on average, each Metric has 222 values or answers, and each company has an average of 25 data points associated with it.

The bulk of Metrics on WikiRate are Community Assessed, which allows for the kinds of student and volunteer engagement cited above. The majority are also designed by standards organisations like the Global Reporting Initiative (GRI), which is the most commonly used sustainability reporting framework for companies. Over 150 metrics on WikiRate are designed by GRI. Universities (considered Companies on WikiRate) also report according to standards like the Higher Education Statistics Agency (HESA), who are the designer of 229 metrics on WikiRate. New import functionality is allowing bulk metrics and data to be imported onto the platform, which drastically speeds the process of adding data, but only applies in cases where data is reported in an open, machine readable format, as it often is with HESA compliant universities.

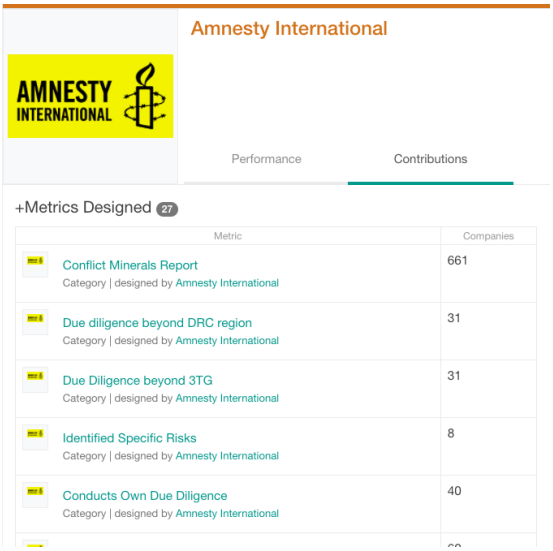


Figure 12 Amnesty Int'l Page on WikiRate.org

Beyond standards organisations, bespoke Metrics are designed by NGOs, researchers, and working groups, as well as individuals. We see uptake in bespoke metric development particularly with advocacy organisations, and with non-financial reporting regulations, where companies are required to report according to an issue, but little is done to analyse and compare those reports.

Below are two examples of such research campaigns.

Amnesty International

In the U.S., Section 1502 of the Dodd-Frank Act introduced a legal requirement for companies that file with the U.S. Securities and Exchange Commission (SEC) to report on their due diligence to ensure that their sourcing of certain minerals does not fund armed groups in the Democratic Republic of the Congo (DRC). Amnesty International collaborated with WikiRate to develop metrics that could be used to increase the transparency of these reports through introducing structured, public comparability and analysis.

Amnesty defined a set of questions and created metrics to record the answers to those questions, with support from the WikiRate team. Amnesty ran two "data sprints" within the last year to engage volunteers in researching reports, and adding their findings onto WikiRate. One university, the University of Western

Australia, has recently integrated a research assignment in their Business Ethics course, engaging students in the process of reading Conflict Minerals reports and adding answers to the WikiRate database.

The Walk Free Foundation

The Walk Free Foundation was similarly interested in developing research related to a government non-financial reporting requirement, in this case, the Modern Slavery Act (MSA). The MSA became law in the UK on 26th March 2015, and requires a commercial organisation that has a turnover of over £36 million with operations in the UK to publish a slavery and human trafficking statement each year, which sets out the steps it has taken to ensure there is no slavery or trafficking in its supply chains or its own business, or states that it has taken no such steps.

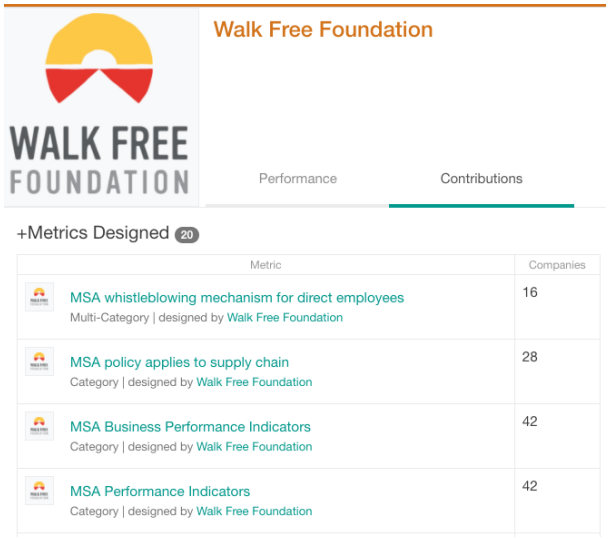


Figure 13 The Walk Free Foundation Page on WikiRate.org

The Walk Free Foundation wanted researchers, volunteers and consumers to be able to evaluate the quality of Modern Slavery Statements that are produced by required companies in accordance with the Modern Slavery Act. In order for this to take place, they first had to formulate a set of metrics based on the guidelines issued under the Modern Slavery Act by the Home Office, add these to WikiRate, and conduct initial tests and refinement of the metrics.

The first pilot group which engaged in researching and testing these metrics was a group of students at Columbia University, studying Business and Human Rights. The refined metrics have been utilised in a course at at Johns Hopkins University, and are set to being included for research in courses at the University of Melbourne, and Nottingham University.

Application of Calculated Metrics

Calculated metrics (*Scores*, *Formula* and *WikiRatings*) offer the WikiRate community further opportunities to analyze, examine and understand corporate performance and impact. These metrics act as a valuable application and extension for the basic *research metrics* which capture raw data points gleaned from public sources. The 222,496 data points currently on WikiRate.org (and increasing daily) are rich in exciting research possibilities and ripe for further exploration/exploitation. Although fully functional, *calculated metrics* are not yet

central to outreach and are currently considered to be in beta. The Consortium agreed the functionality should first be proven (and improved upon) with strong initial use cases. In parallel, the strategic team focused development efforts on major updates to WikiRate.org, including the homepage, projects, profiles, sorting/filtering and new functionality, such as research groups and badges.

Calculated Metrics have been live on WikiRate.org for one year, however at first glance they might not be so obvious. They are not being promoted or prominently linked to, nor is there abundant help text. However, we have been working collaboratively with partners to test the functionality, usability and performance of the calculated metrics. When we are satisfied the metrics are preforming at optimum, the outreach and engagement team will then develop relevant content, guidance and user pathways to promote the use of calculated metrics, once again demonstrating WikiRate.org as a research tool, as well as a data hub. We are consulting with a range of research partners at various stages of the process, including design and implementation. Following the pilots the development team will update the interface to support discoverability and usability, which will ultimately allow us to engage a wider audience.

Current applications of calculated metrics include: simple and complex Formula calculations using existing data on WikiRate.org (e.g. Center for Sustainable Organizations and University of Michigan) or a combination of Scores and WikiRatings to determine an overall rating on company performance (e.g. Richard Mills, Cambridge University). Simply put, calculated metrics are computed by combining other values. Figure 14, illustrates a simple *formula metric* which was designed in collaboration with the University of Michigan as part of a larger project researching university ESG data. Although a slight shift from corporate data, universities report at great length and are excited to share their data on WikiRate and demonstrate its potential. Following the import of over 40 metric answers we created some simple formula metric questions - below shows the formula created to answer this question:

What is the total (%) of energy consumed by the institution that has

been generated from renewable sources?

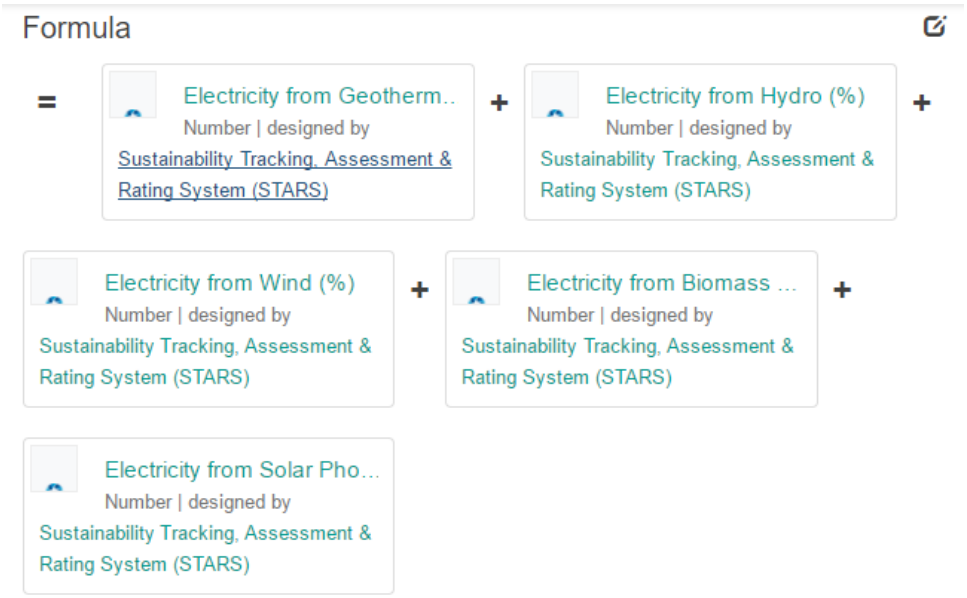


Figure 14 Simple Formula Metric Sceenshot

The formula metric allows us to combine the total (through simple addition) of multiple research metrics. However, Formula metrics are not limited to addition and can support complex mathematical equations and enable users to combine a range of research metric values. In this example (Figure 15), the Center for Sustainable Organizations created the following formula to answer the this metric question:

What are the total CO2 emissions relative to gross revenue?

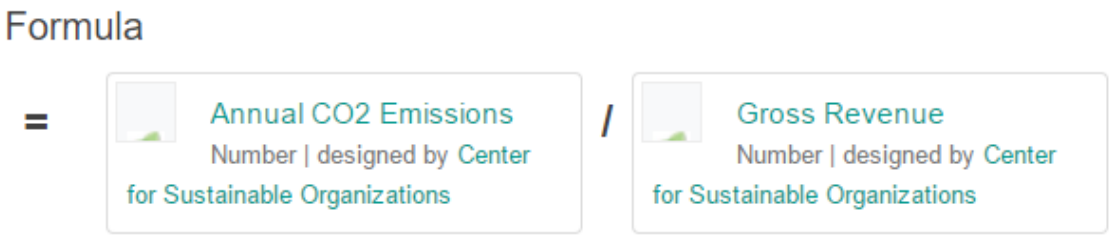


Figure 15 Sample Formula Metric: CSO

These formulae provide a distinct interface for mathematically reusing metric answers, but importantly always lead back to the source of the raw data, ensuring the greatest transparency.

A Score metric normalises another metric's value on a 0-10 scale and offer lots of mathematical value, specifically in preparing Research and Formula metric values for use in WikiRatings. Figure 16, shows the A-E metric answers on the left be converted into weighted numerical values.

Formula

Metric	Weight
A	10
A-	9
B	7
C	5
D	3
E	1
Unknown	0

Figure 16 Weighted Average: Calculated Metrics

This score was then used to create the CDP Scores metric which rates 270 companies against two scores, disclosure and performance (see above). Figure 17 shows this metric with a value expanded. A key requirement for the design of WikiRatings is to present them in an accessible way which allows a reader to easily understand which metrics have been used to produce the WikiRating and how the data for these metrics is used to determine a particular company's score. WikiRate wants to present ratings in an accessible way, but also to use these as an entry-point to the world of ESG performance data. When one expands the WikiRating score for a company one can see which metrics have been used, what the company's answer was for each of those metrics, how those answers have been scored (by the Score metrics the WikiRating uses) and the number of points each metric contributed to the WikiRating once its weighting has been taken into account.

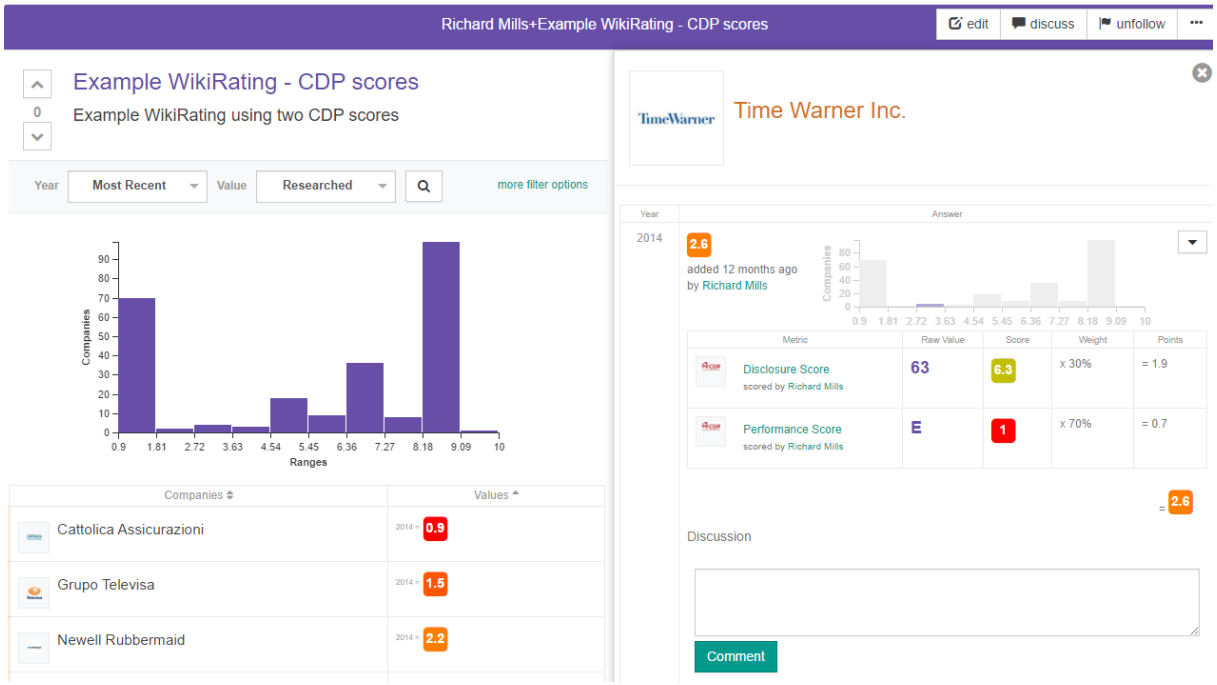


Figure 17 Example WikiRating: Expanded view

WikiRatings are designed to be accessible, easy to understand and even easy to create. The new metric interface is almost identical to the other calculated metrics, including the traffic-light coloured values (red, yellow, green) to indicate the lowest to highest ratings. However, one limitation is that first the community must populate Research, Formula and Score metrics to enable ratings to be created on a larger scale. *The more Research metrics are populated, the closer WikiRate is to being able to start generating meaningful WikiRatings.*

Over the coming months WikiRate will continue to work closely with partners to develop and improve the calculated metrics, but already we are starting to see the applications and benefits. Calculated metrics enable users to create one question and one calculation and get answer for all companies that have raw data associated with the relevant metrics, essentially allowing you to carry out hundreds of calculations through one metric. By creating the metric on WikiRate.org it will now be available for others to utilise for research, discussion and further exploitation through scores and ratings.

4 Recommendations

The WikiRate Project's ultimate goal is to "crowdsource better companies". It is important to reiterate that the organisation, WikiRate e.V. shall remain neutral at all times and in all

research results and implications. It is important that the WikiRate organisation is not perceived as inherently for or against companies. What WikiRate does is provide the structured tools to crowdsource better companies, and while doing so, provide a research environment that is structured, multi-faceted, and open.

As WikiRate progresses, it becomes clearer that additional features are as important as evaluating current ones. More organisations running corporate behaviour research are also looking at platforms to conduct their research through, and then link the results openly for deliberation, inquiry, verification and publication. The arena for calculated metrics opens grounds for testing and experimenting, and this raises the question of privacy of certain aspects along the research methodology, which also necessities development and integration of new models for collaborative research.

Applicability

Some metrics are best applied to only certain companies or indeed industries. When looking at metrics on particular niche topics, or for example: provided by industry initiatives such as the sustainable apparel coalition (SAC), some metrics are likely most applicable for a particular range of companies. A question around textile sourcing processes for example would not be as applicable or even sensible to apply to a corporation that only deals in financial services. For this reason applicability of metrics is an especially useful dimension to consider. As it stands the applicability of metrics is considered and defined by the materiality of indicators for a particular corporation. Materiality describes what indicators an investor or stakeholder within a company considers important and relevant within sustainability reporting and due diligence. Such considerations are becoming increasingly important with respect to considerations by investors and also in concert with e.g. Corporate performance on the Sustainable Development Goals (SDGs) where a leading accountancy firm (PWC)[1] and leading CSR member network (Business for Social Responsibility) are in dialogue with a number of corporations around coming to materiality judgements with corporations and within an industry. This builds on the work of the SDG Compass[2] and links with ongoing work that WikiRate is pursuing with PRME around researching corporate performance on the SDGs and as part of the Multiadvisory Stakeholder Council (MAC) to the Corporate Action Group (CAG)[3] towards reporting on the SDGs. Interestingly whilst a metric might initially only appear applicable to certain industries initially – it may become apparent over time that metrics are more broadly applicable than initially thought, or than particular companies may have previously judged to be material. This is where WikiRate's voting and reputation system around metrics and also the ability for NGOs and Academics (as well as

[1] https://www.pwc.com/gx/en/sustainability/publications/PwC-sdg-guide.pdf
[2] https://www.pwc.com/gx/en/sustainability/publications/PwC-sdg-guide.pdf
[3] https://www.globalreporting.org/resourcelibrary/GRI%20UNGC%20Corporate%20Action%20Group%20(002).pdf

standard setters) to create their own metrics on WikiRate can play into the judgement of what companies should consider important, relevant, material and ultimately applicable to them as a corporation and actor within a particular industry.

The mechanisms for applicability of WikiRate metrics have been designed but not yet developed. Applicability is a decision that will be taken at the level of the metric (by the metric designer, or the community in the case of collectively owned metrics). A metric will specify applicability by specifying the values a company will have on other metrics if the present one is applicable to them. Applicability can be specified through metrics like Industry, Headquarters Country, company type, or size (measured through revenue or employees).

This is a departure from how materiality is approached in CSR reporting at present. Each company makes decisions about what they perceive as material and reports on that basis. The details of these decisions are usually not included, and when seemingly relevant data is absent one is left wondering whether the company did not report it because the viewed it as immaterial, or excluded it because it would reflect negatively on the company. WikiRate wants materiality decisions to be visible and publicly discussed. For WikiRate, materiality of metrics will be handled at the metric level by the designers/community – in WikiRate' view it is for the community to decide whether a metric is relevant to a company with certain characteristics.

This system works in one direction – if a company views a metric as material but the designer does not, the company's data can still be added for that metric. However, if a company views a metric as immaterial but the designer says it is relevant, then that company's missing data will be interpreted as a lack of disclosure of relevant sought information. These mechanisms are important in creating the WikiRate Transparency Score, which will use the availability/missingness of data for *relevant* (applicable) metrics to score companies on their transparency.

Transparency Score

A growing strategy for WikiRate is to engage companies through ratings. By introducing the *inquiry* model to support direct communication between company representatives and community members; structured so as to reward corporate transparency and discourage opacity. For the community it's engaging through inquiries, for WikiRate is applying a transparency score to invite high-performance companies to do better, and to encourage performance-laggards to put in more effort.

The WRIT (WikiRate Index of Transparency) score, which has been conceptually designed but not yet developed, will perform calculations by considering whether the company has data for all of the metrics which are relevant to it. Where a company has a value for a metric this means that its status is known with regard to that metric - this implies that the company has disclosed the information which is necessary to populate that value on WikiRate. During this period of assessment, a decision was reached that the WRIT is still relevant to WikiRate's whole Metric Framework with its current functionality and setup.

WikiRate's WRIT aims to promote the idea that: 'unknown' is the worst possible value that a company can have on any given metric. One of the key functions of the Rating Concept as a whole is to identify where the gaps in our knowledge about a company are by breaking measures of their performance down into standardised metrics - the WRIT score serves to focus attention on the missing values, the size of the gaps in our knowledge.

For the purposes of calculating a company's WRIT score, researched metrics will be weighted by their importance - through a combination of the importance score for the individual metric and the importance scores for calculated metrics which use it. Through 'importance voting' in the metrics sphere, WikiRate users will identify what they regard as the most important metrics. The weighting of metrics by importance will allow (and encourage) them to prioritise the disclosure of the information which is most sought after. Rather than being presented with an unordered list of 100 missing metric values, the company will be presented with an ordered list showing the degree to which their WRIT score will be improved by providing each value.

Descendant Metrics

In concept, WikiRate has identified a new range of metrics called "descendant metrics". Descendant metrics are metrics that allow wider communities to build on the data collected by expert groups, while preserving the distinction between data from the organisation and data from the crowd. The data generated from these descendant metrics are free and open source, while the data generated from the original metric may have 'ownership' or provenance.

The collaboration between WikiRate and Oxfam India is one such case. Oxfam India runs an annual index on India's top 100 companies listed on the Bombay Stock Exchange. By running its index, Oxfam India publishes its results, but limits access to them (in agreement with the companies). Oxfam's indicators are very interesting and relevant to other markets, users may want to make use of these metrics in other research. This descendant functionality is a data importing feature to facilitate bringing existing data into the users' research

environment where it can be presented, discussed and analysed alongside data from other sources.

All this plays in the type of metrics defined in the metrics type section. Another reason for developing descendant metrics is to remove barriers to community extension and refinement of existing data-sets. A descendant metric pulls in data from existing metric(s), with this version of the data being "owned" by the creator of the descendant metric (with clear articulation of where the data came from originally). As the descendant metric is dissociated from the designer of the metrics it inherits from, its creator has the freedom to give it a different research policy, like "community assessed", and allow community members who are not affiliated with the designer to build upon the designer's data. In addition to the Oxfam India use case, WikiRate has another use case for Ranking Digital Rights: Ranking Digital Rights (RDR) have been approached by a number of CSOs that want to apply their methodology to companies not covered by RDR – usually companies that operate in the locality of the CSO. This is potentially a win-win situation, RDR want their methodology to be used in assessing additional companies, and the CSOs see value in conducting an assessment in a way which is consistent with RDR's approach. But the tools that RDR currently use are difficult to copy. Descendant metrics offer a solution to overcome these barriers. Having solid use cases as the ones mentioned-above provide compelling premise for extended application to include for example, metrics designed by GRI or the UNGC.

Relationship Metrics

Underway as part of a separate grant - called ChainReact - is the development of Relationship Metrics. This new breed of WikiRate metrics is designed to capture the connections that exist between different corporate entities. Whilst advancing transparency on the social and environmental performance of companies is a great step in the right direction, being able to position this performance data in its broader context of related business activities is a much needed next step.

In other words, these Relationship Metrics are the fundamental building block for a larger effort to map the social and environmental performance of companies across their corporate networks. More specifically, the metrics will facilitate three different types of corporate mapping; 1) supply chains, 2) ownership structures, and 3) investor relations.

The aim of connecting these "dots" is to make corporate networks transparent, understandable, and responsive - so that companies and their stakeholders can see, react to, and ultimately transform corporate network impacts. Bringing companies out of their performance silos and fostering a culture of corporate responsibility that does not just mean

improving your own practices but that also means holding those you are associated with accountable for their actions and leveraging your relationship to help them improve.

Whilst the technical proof of concept has already been delivered, the first Alpha version of these Relationship Metrics will be available on the WikiRate.org platform in the Summer of 2017, at which stage the first consultation partners will begin to test their functionality.

Geographic Specificity

To provide another relevant contextual frame that helps people understand and interpret a company's social and environmental performance data, WikiRate is also looking to capture the geographic specificity of companies and data points. The kinds of analysis that become possible through this geographic lens would include, but are not limited to:

- *Environmental* contextualization of corporate performance:

  For example, assessing a company's water consumption in relation to data on water availability. A company using what at first glance may seem like a moderate amount of water, could be having a much worse impact on its surroundings if it is in a region suffering from droughts, than a company using double the amount of water but which is located in an area where water is abundant.

- *Legal* contextualization of corporate performance:

  Companies' reporting is often much influence by the legal requirements outlined in national or regional policies. Apart from assessing compliance with the laws that apply to corporations operating in those localities, it would also help distinguish which companies produce, for example, a Modern Slavery Act statement because they are mandated and which do so voluntarily. It would as such help pinpoint and reward the companies that show leadership.

- *National* contextualization of corporate performance:

  Similar as to companies, countries often set performance goals, monitor their progress towards these goals, and report on them. Take for instance, the CO2 emissions targets formulated in the Paris Agreement or the UN Sustainable Development Goals. Contributing to these national targets, are both the performance of national governments as well as the companies registered within the countries. It would as such be of tremendous value for governments to be able to identify which companies and corresponding data-sets should feed into these targets.

Moreover, national contextualization of corporate performance would also make it possible to compare the operations of a multinational within a certain country, to the performance of smaller (nationally confined) businesses operating in the same country. For example, comparing Unilever's chocolate supply chain operations in Indonesia to those of Kakoa Chocolate. In other words, levelling the playing field in terms of performance assessments, making it possible to rate SMEs and multinational conglomerates alongside each other.

Such a localisation feature on WikiRate.org is thus likely to open the platform for broader areas of research and a wider network of collaborators. It would as such aid WikiRate's recommendation to accommodate a larger and more diverse audience.

Like providing support for geography, WikiRate also foresees importance of providing multilingual support to expand on the bonging conversation on a global level.

Multilingual Support

"Globalising" the platform remains limited unless integration tools are executed in other locations and using other languages. WikiRate's multilingual support is designed but remains a dormant functionality for the time being until research and calculated metrics are at higher functionality and utilisation. We have been monitoring user experience and feedback, and factored that feedback into our implementation timeline, we can report that there is still a lot that can be captured in one language by contributors from different parts of the world, by WikiRate will champion a multilingual set-up to ultimately engage local communities with other languages. We currently have one PRME university project from Columbia and another run by the South Korean Women's University EHWA. These two projects are researching data from reports published in their respective languages and are inputting their research in WikiRate's supported language, English. All that said, we still believe that having a multilingual platform is concurrent and relevant to the expansion of the use of WikiRate.org and the growth of its community and metrics.

Prominence of Voting

That through importance voting on metrics, the WikiRate community will be collectively preparing a prioritised list of questions which are being asked of each company. This prominence of voting will increase in visibility alongside the growing community on WikiRate.org and the evolution of research and methodology on the platform. The emerging

stakeholder mix and their voting prominence will ultimately become a factor that cannot be waned down or ignored. Although WikiRate produces the WRIT score, the power remains with the voting community. Hence the democratisation of sustainability becomes more relevant with each voting 'up/down' click.

Permissions

A proposal is on the table to develop WikiRate's permissions capabilities. For multi-stakeholder groups, conducting collaborative research with trust-worthy results requires the ability to restrict views and editing abilities of certain contributors at different stages throughout a research process. Different NGO partners for instance, request the ability to restrict different researchers from seeing the data points entered by other researchers to avoid errors arising from bias. This is particularly relevant to research projects that assess corporate statements with a lot of political lingo that require pro-active interpretation from the researchers.

Moreover, the permissions functionality enables project organisers to invite "external" stakeholders to verify the data that has been collected. For example, having company representatives come into the research tool, allowing them to view only the data on their company, comment on this data and provide additional sources in case they want to object to a certain finding. With nuanced permissions, they would not be able to see the data on other companies nor be able to directly edit the data without having a researcher review their suggested changes.

Nuanced permissions create new opportunity for verification around sensitive information like human rights violations, names of individual whistle-blowers, or information containing disturbing images. Organisations of investigative journalists and data journalists are interested in these functionalities to protect sources as well protect journalists themselves – so that those who work on a sensitive/disturbing task do so with appropriate training and within a supportive environment.

Permissions already exist on WikiRate to the extent that every card can be assigned specific rules. To expand this functionality to serve the use cases outlined above, user-accounts will need to be segmented and assigned roles that correspond to specific abilities. This initial step can be developed, tested, and refined for a pilot with Ranking Digital Rights, who want to advance their research and evaluation process for the 2020 Corporate Accountability Index through WikiRate.

Years

Years are a key component of every metric answer on WikiRate.org, every metric value must be associated with a year, but multiple years of data can be added to one metric question. WikiRate.org allows users to gain a glance at both the company's historical performance for and market context for the answer on a year-by-year basis (see Figure 18). When viewing a Company page, only one answer per metric is shown at a time on the left side. However, by clicking on the Answer, one can see the full record for that metric.

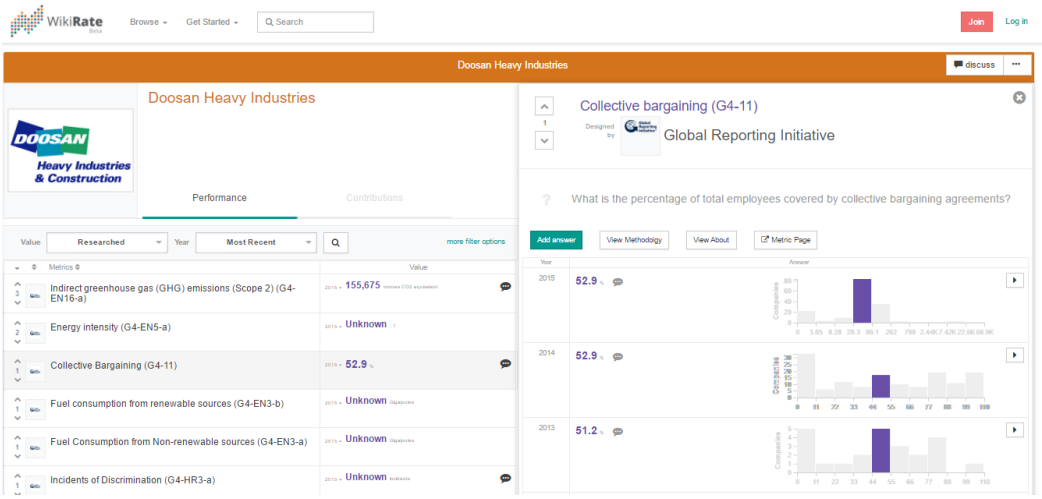


Figure 18 Year View

Presently, users are limited to selecting one year per answer but this has proven to have its limitations. In an ideal world every corporation would report between 1 January – 31 December of any given year, however the reality is that organisation report at different times, using different reporting cycles and ultimately this has been initially challenging to represent on WikiRate.org. For now WikiRate recommends users report the most recent year displayed on the report, e.g. 2014 – 2015 Sustainability Report would be reported as 2015 in the metric answer. In the future we intend for the platform to handle a larger variety of year ranges to account for organisations that report over multiple years, for example the academic year or financial year.

5 Conclusion

Achieving WikiRate's goal of "crowdsourcing better companies" relies first and foremost on community uptake by offering a working platform with relevant functionalities. Adopting a

strategy of reaching out and engaging advocacy groups and academics has been fruitful on many levels, notably that of testing calculated metrics. Metrics that have been designed by a reputable organisation can be considered more prominent or relevant than a metric with lots of votes. Insofar that researchers have been mainly engaged in researched metrics, their expertise is growing to include a more complex research criteria. Similarly, designing calculated metrics based on well-known research metrics might have more prominence than those which use metrics from less reputable designers. Such metric designer/user relationships are likely to be influenced by a user's value profile.

The relationship between the importance of voting and ratings is also relevant to the WRIT methodology for calculation. Where users on WikiRate are judging, which metrics are most important, either through the introduced voting mechanism, or through other criteria such as reliability of reputation of a metric designer.