



THE DANISH
INSTITUTE FOR
HUMAN RIGHTS

AUGUST 2022

SUSTAINABILITY REPORTING AND HUMAN RIGHTS

WHAT CAN BIG DATA ANALYSIS TELL US ABOUT
CORPORATE RESPECT FOR HUMAN RIGHTS?

SUSTAINABILITY REPORTING AND HUMAN RIGHTS

WHAT CAN BIG DATA ANALYSIS TELL US ABOUT CORPORATE
RESPECT FOR HUMAN RIGHTS?

AUGUST 2022

Author: Gabrielle Holly, Sarah Kajander Holmer

e-ISBN: 978-87-7570-097-4

Cover photo: Omar Flores, unsplash.com

Layout: Michael Länger

© 2022 The Danish Institute for Human Rights
Denmark's National Human Rights Institution
Wilders Plads 8K, DK-1403 Copenhagen K
Phone +45 3269 8888
www.humanrights.dk

DIHR acknowledges the Global Reporting Initiative for making this project possible by providing access to company sustainability reports contained within the GRI Sustainability Disclosure Database for the period 2010 to 2020. The Database was decommissioned in March 2021.

This publication was possible thanks to the support from the Swedish International Development Cooperation Agency (Sida). Responsibility for the content rests entirely with the Danish Institute for Human Rights.

Provided such reproduction is for non-commercial use, this publication, or parts of it, may be reproduced if authors and source are quoted.

At the Danish Institute for Human Rights we aim to make our publications as accessible as possible. We use large font size, short (hyphen-free) lines, left-aligned text and strong contrast for maximum legibility. For further information about accessibility please click www.humanrights.dk/accessibility

CONTENTS

1	INTRODUCTION	4
2	METHODOLOGY AND LIMITATIONS	6
3	WHAT IS THE DATA TELLING US? OR IS IT TELLING US ANYTHING?	8
3.1	HYPOTHESIS 1: COMPANIES ARE INCREASINGLY REPORTING ON HUMAN RIGHTS SINCE THE ADOPTION OF THE UNGPS IN 2011	8
3.2	HYPOTHESIS 2: LEGAL AND REGULATORY DEVELOPMENTS HAVE DRIVEN IMPROVEMENTS IN REPORTING	9
3.2.1	Global initiatives	9
3.2.2	Regional initiatives	10
3.2.3	Particular indicators	11
3.3	HYPOTHESIS 3: LARGER COMPANIES ARE REPORTING MORE ON HUMAN RIGHTS THAN SMALLER COMPANIES	14
3.4	HYPOTHESIS 4: THERE ARE REGIONAL DIFFERENCES IN REPORTING ON HUMAN RIGHTS TOPICS	16
3.5	HYPOTHESIS 5: COMPANIES ARE INCREASINGLY REPORTING ON PROCESSES FOR IDENTIFYING AND ADDRESSING HUMAN RIGHTS IMPACTS	17
4	CONCLUSIONS: WHAT CAN AI DO FOR BHR	19
4.1	WHAT CAN AI DO FOR THE ANALYSIS OF COMPANY REPORTING ON HUMAN RIGHTS ISSUES?	19
4.2	CHALLENGES OF USING ALGORITHM SUPPORTED ANALYSIS TO REVIEW COMPANY DISCLOSURES	19
4.3	WHAT WOULD IT TAKE TO MAKE THIS WORK EFFECTIVELY?	21
4.3.1	Common standards	21
4.3.2	Central repository of reports in machine readable format	22
	ANNEX 1: DESCRIPTION OF THE DATASET	23
	ANNEX 2: HUMAN RIGHTS INDICATORS	25
	ANNEX 3: EXAMPLE TEXT BITS	26

1 INTRODUCTION

Since the adoption of the UN Guiding Principles on Business and Human Rights (UNGPs) in 2011, corporate respect for human rights has become an expected standard of conduct for businesses, discharged primarily through the process of human rights due diligence.

Despite the expectation set out in the UNGPs that companies “know and show” how they identify and address their human rights impacts, one of principal weak points of human rights management in companies continues to be a lack of transparency and communication of their efforts, progress, outcomes and challenges. This creates difficulties for stakeholders, including state entities, civil society, investors and consumers, to understand and assess whether and how companies are conducting human rights due diligence and facilitating access to effective remedy.

The capacity to easily access, analyse and compare company reporting is therefore crucial for a range of stakeholders including ESG investors and civil society groups focused on corporate accountability. However, corporate sustainability reports use a variety of different standards as a point of departure, including domestic legal requirements as well as third party standards such as those developed by the Global Reporting Initiative (GRI) which were revised in 2021 with a view to better align with human rights instruments such as the UNGPs and OECD Guidelines. Companies also use their own formats when preparing their sustainability reports, often adopting a narrative or visual style to communicate information. In addition to variation in standards used, there is considerable variation in the presentation of data on human rights issues in current company reporting which presents considerable difficulties when attempting to assess and compare company practice.

These matters may be addressed in part by certain regulatory developments in Europe, including the introduction of the Corporate Sustainability Reporting Directive (CSRD), which will replace the existing Non-Financial Reporting Directive (NFRD), which was enacted in 2014 with an obligation for large companies to report on a range of non-financial matters, including human rights. The CSRD foresees that companies will be required to report on a range of “sustainability matters” and provides for a new European Sustainability Reporting Standards (ESRS) that all reporting companies will be required to use.

Useful work has been done by the [Alliance for Corporate Transparency](#), the [Corporate Human Rights Benchmark](#) and others to analyse both what companies report in relation to human rights as well as assess the human rights performance of companies through manual analysis of company reporting.

Each of these projects has resulted in data on the state of company reporting on human rights by large companies capable of assisting a range of stakeholders identify trends and gaps in current company practice on human rights. However, efforts to undertake large scale qualitative analysis of company reports are limited by the

resource intensive nature of the review, requiring manual review of company reports which provide data in often quite different formats and without reference to common standards. This makes qualitative analysis challenging to scale up.

To address these challenges and supplement efforts to undertake qualitative analysis of company sustainability reporting on human rights, DIHR developed a project which aimed to use algorithm assisted analysis of a large number of company reports maintained in the GRI Sustainability Disclosure Database against a set of human rights indicators. This discussion paper sets out the methodology used, the challenges and limitations encountered and a summary of what we were able to see in the data in response to a series of hypotheses:

1. **Hypothesis 1:** Companies are increasingly reporting on human rights since the adoption of the UNGPs in 2011
2. **Hypothesis 2:** Legal and regulatory developments have driven improvements in reporting
3. **Hypothesis 3:** Larger companies are reporting more on human rights topics than smaller companies
4. **Hypothesis 4:** There are regional differences in reporting on human rights topics
5. **Hypothesis 5:** Companies are increasingly reporting on processes for identifying and addressing human rights impacts

Further, a range of actors, including Environmental, Social and Governance (ESG) ratings agencies and other market analysts have advertised services which rely in part on analysis assisted by artificial intelligence (AI) applied to large pools of data, including company sustainability reports, in order to make assessments of the ESG performance of companies analysed. Through our own use of similar technology through this project we also sought to test whether these ambitious claims could be substantiated.

2 METHODOLOGY AND LIMITATIONS

In collaboration with [Specialisterne](#), the DIHR developed a text mining algorithm capable of being applied to large datasets, which is a form of AI that classifies text to sustainable development and human rights topics. The text mining algorithm is the engine behind DIHRs [SDG-Human Rights Data Explorer](#) which currently classifies about 145,000 UN system recommendations to the 169 SDG targets. In this project we use the text mining algorithm to classify company sustainability reports published between 2010 and 2020 to a set of human rights indicators. The sustainability reports were accessed through the GRI Sustainability Disclosure Database (the Dataset), a database maintained by GRI until December 2020, which contained links to around 50,000 company reports in the period which have used the GRI reporting standards. The reports in the Dataset were tagged by GRI according to their size, sector and geography, enabling us to segment the data at the analysis stage, assisting us to identify trends among companies of a particular category.

However, the DIHR has faced a number of challenges in adapting the algorithm to the company reports in the Dataset. This has stemmed in part from the lack of standardisation of reporting and the machine readability of the company reports. From a review of the Dataset, potentially relevant information was contained in a range of formats, including tables or text boxes in picture format which were challenging to convert into a machine-readable format. Another challenge was the ability to download the reports from the database. DIHR was able to download and read only 21,759 (44 %) of the reports due to either broken links within the database, or links not leading directly to the reports. Out of the downloaded reports, 9,374 (43 %) of the reports are in English and are included in the analysis. This corresponds to only 18.78% of the original Dataset. Further, because of these limitations it became challenging to analyse data segmented by size, sector or geography as some categories contained only a small number of reports. For example, as will be discussed below, in some cases we were not able to identify reporting trends in the Dataset responsive to certain national laws because of a limited number of reports from the particular national context. A full description of the Dataset can be found at Annex 1.

To use the algorithm to classify the company reports, the reports were first converted from pdf-files into analyzable “text bits”, usually the length of a paragraph. The result was 3,391,615 text bits from the 9,374 sustainability reports. A set of human rights indicators were developed which the reports were to be classified to. The human rights indicators focus on a range of issues including: disclosure of the policies and processes used by companies to identify and address human rights impacts; how companies report on fundamental labour rights; how companies report on the needs of a range of stakeholders; how companies report on human rights impacts in the supply chain; and other human rights issues. A full list of the 18 indicators included in the analysis can be found at Annex 2.

The next step in the process of categorizing the text bits from the reports to the chosen human rights indicators was to train the algorithm. The algorithm was trained on a subset of text bits by tagging the text bits that were responsive to a particular indicator. This tagging was done by a human expert. From these text bits the algorithm analyses the patterns which it then uses to evaluate and categorize the full set of text bits. The output of this evaluation process is that each text bit was assigned a likelihood of being linked to a particular indicator. We use a threshold of 50% to determine if a text bit is linked to the indicator. If a text bit has a likelihood of being linked to an indicator of 50% or more, we say that the text bit is linked to the indicator. In order to improve the final model this process was repeated several times. A quality assessment was conducted by a human expert of a sampling of text bits that had been identified as linked to a particular indicator in order to assess whether the algorithm had correctly identified a particular text bit as relevant or irrelevant to the indicator. Our final training data consists of 4,371 text bits, which we have categorized based on an assessment of whether they contain text linked to the chosen indicators. Examples of text bits are included in Annex 3.

To analyse the data, a series of hypotheses were developed as follows:

1. **Hypothesis 1:** Companies are increasingly reporting on human rights since the adoption of the UNGPs in 2011
2. **Hypothesis 2:** Legal and regulatory developments have driven improvements in reporting
3. **Hypothesis 3:** Larger companies are reporting more on human rights topics than smaller companies
4. **Hypothesis 4:** There are regional differences in reporting on human rights topics
5. **Hypothesis 5:** Companies are increasingly reporting on processes for identifying and addressing human rights impacts

Our observations in relation to each of these hypotheses are outlined in the section below.

When we analyse the data, we use the following two different measures:

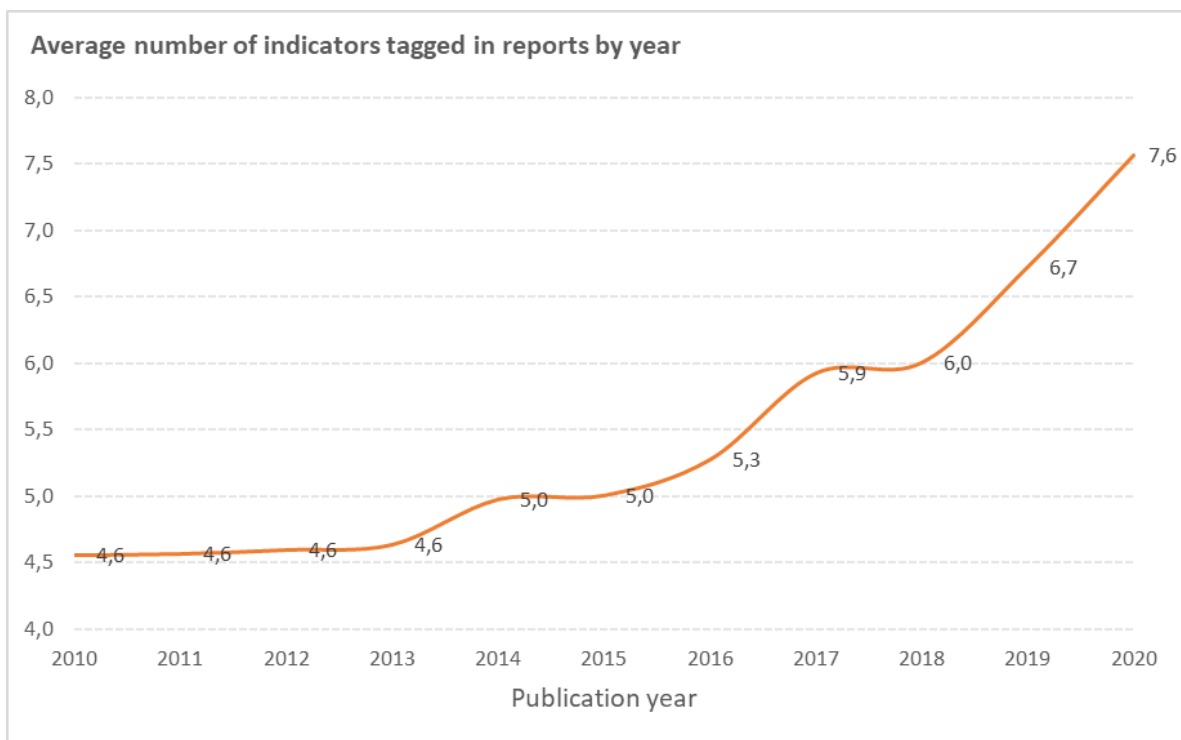
- The average number of human rights indicators tagged in company reports; and
- The percentage share of company reports containing text bits tagged to a particular indicator.

The graphs included in this discussion paper identify which of these measures has been used to analyse the data. We look at trends in terms of development over time and differences between groups in terms of company size and region. When we do this, we test if the observed development over time or difference between categories is due to statistical uncertainty since we are working with a sample, or if we can say that the observed development, or difference is statistically significant.¹

3 WHAT IS THE DATA TELLING US? OR IS IT TELLING US ANYTHING?

3.1 HYPOTHESIS 1: COMPANIES ARE INCREASINGLY REPORTING ON HUMAN RIGHTS SINCE THE ADOPTION OF THE UNGPS IN 2011

To begin, we wanted to test whether we could see a trend in reporting across all human rights indicators over time since the adoption of the UNGPs in 2011. To test this hypothesis, we identified the average number of human rights indicators tagged in reports in a particular publication year. This is represented in the visualisation below.



We observed a statistically significant increase in the average number of indicators per report over time for each year from 2010 to 2020, as is shown in the visualisation above. From 2016 onward, we see a more rapid increase. This suggests that companies are reporting on a wider range of human rights indicators over time, and that this trend has escalated in recent years.

3.2 HYPOTHESIS 2: LEGAL AND REGULATORY DEVELOPMENTS HAVE DRIVEN IMPROVEMENTS IN REPORTING

We next wanted to delve into the data to see whether we could observe any trends in reporting which could potentially be attributable to the introduction of legal and regulatory initiatives relevant to a number of the indicators. We did so by developing a non-exhaustive indicative list of potentially significant measures comprised of laws, reporting initiatives and guidance at the global, regional and national level as follows:

- 2011 UN Guiding Principles on Business and Human Rights
- 2011 OECD Guidelines for Multinational Enterprises
- 2014 EU Non-Financial Reporting Directive
- 2015 UK Modern Slavery Act
- 2015 UN Sustainable Development Goals
- 2016 GRI standards
- 2017 French Duty of Vigilance Law
- 2018 GRI standards
- 2018 OECD Due Diligence Guidance for Responsible Business Conduct
- 2018 EU General Data Protection Regulation

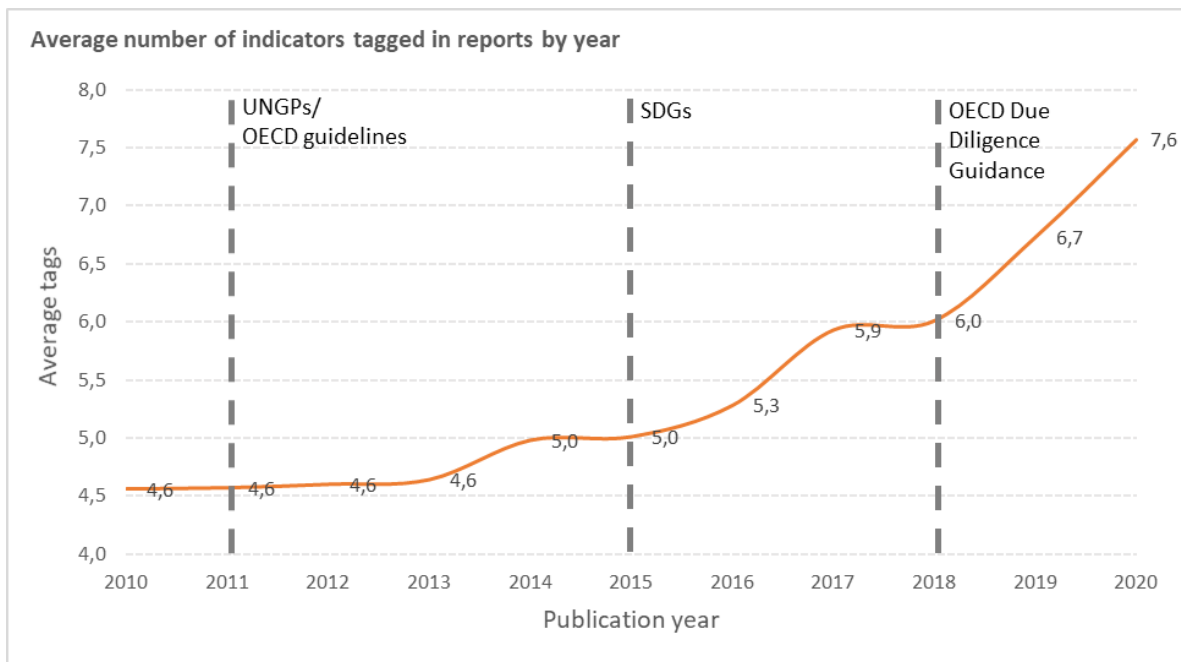
We were able to observe trends in relation to each of these measures as described in the section below, with the exception of the French Duty of Vigilance law, as there was an insufficient number of French companies included in the Dataset to generate statistically significant findings.

While we are able to observe trends in the data, we are not able to make conclusive findings on whether there is a casual relationship between the instruments and the developments that we see. Overall, we see upward trends in reporting from the years that the laws became operative, ie the first years where reporting obligations were imposed. However, in many cases, we can also observe upward trends from the date that the laws or measures were enacted or published. This could be attributable to an indirect effect of the measures as companies prepared for the upcoming legal obligations to disclose or took steps to implement guidance or adopt new reporting standards.

3.2.1 Global initiatives

We began with four major international soft law frameworks of global significance: the UNGPs which were adopted in 2011; the OECD Guidelines for Multinational Enterprises (OECD Guidelines) which were revised to align with the UNGPs in 2011; the UN Sustainable Development Goals (SDGs) which were introduced in 2015; and the OECD Due Diligence Guidance for Responsible Business Conduct which were launched in 2018.

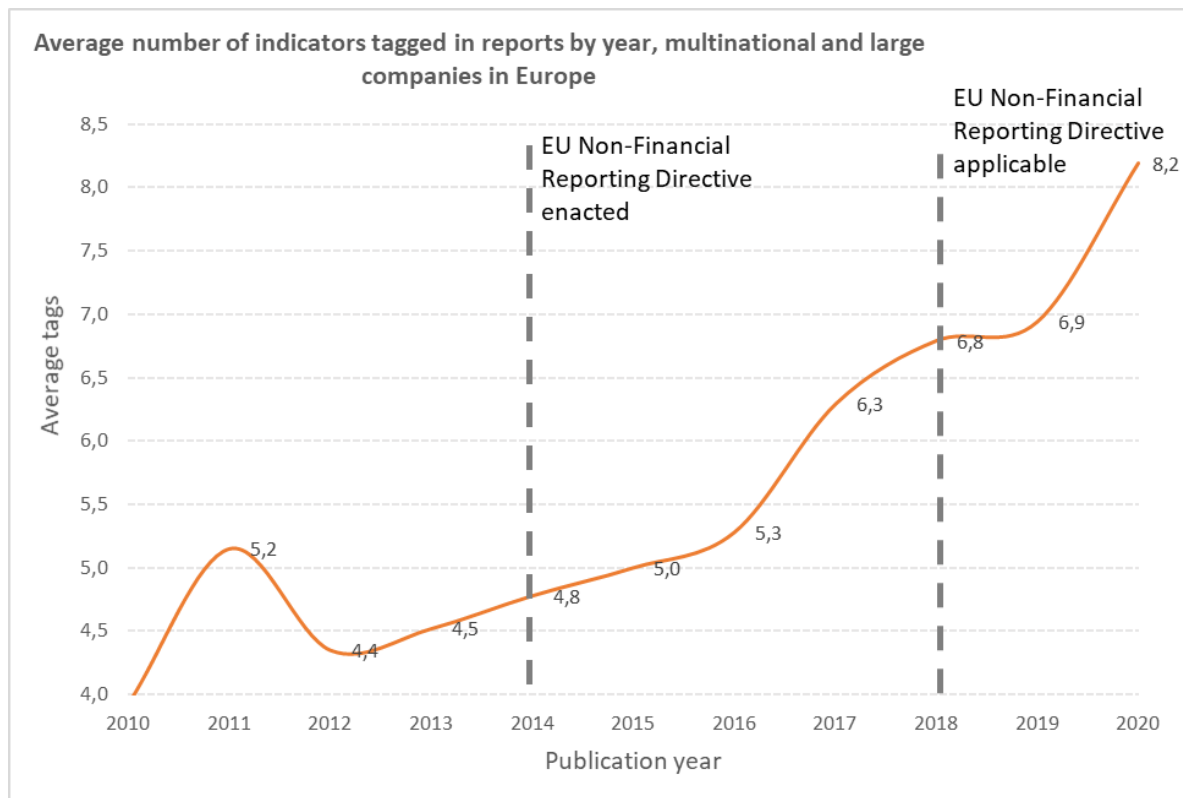
To do this we identified the average number of human rights indicators tagged in reports in a particular publication year, identifying the year that each of these instruments was adopted or relevantly revised. This is represented in the visualisation below.



From the visualisation above, we do not observe any statistically significant effect around the year 2011. However, we see a statistically significant increases between 2015 and 2017. We can observe an additional statistically significant increase from 2018 to 2019 and to 2020.

3.2.2 Regional initiatives

Next, we wanted to see whether there were any observable trends at the regional level. We isolated reports from multinational and large European companies from the dataset and again generated an average number of indicators tagged by reference to the publication year of that report. We then considered whether there was an observable trend in reporting which could be responsive to EU Non-Financial Reporting Directive (NFRD), which was enacted in 2014 with an obligation for large companies to report on a range of non-financial matters, including human rights, commencing in 2018. This is set out in the visualisation below.



The graph shows a statistically significant change from 2018 to 2020 when the reporting obligations became operative. However, we also see an upward trend from 2014 when the NFRD was enacted which could be attributable to preparations for the new reporting requirements.

3.2.3 Particular indicators

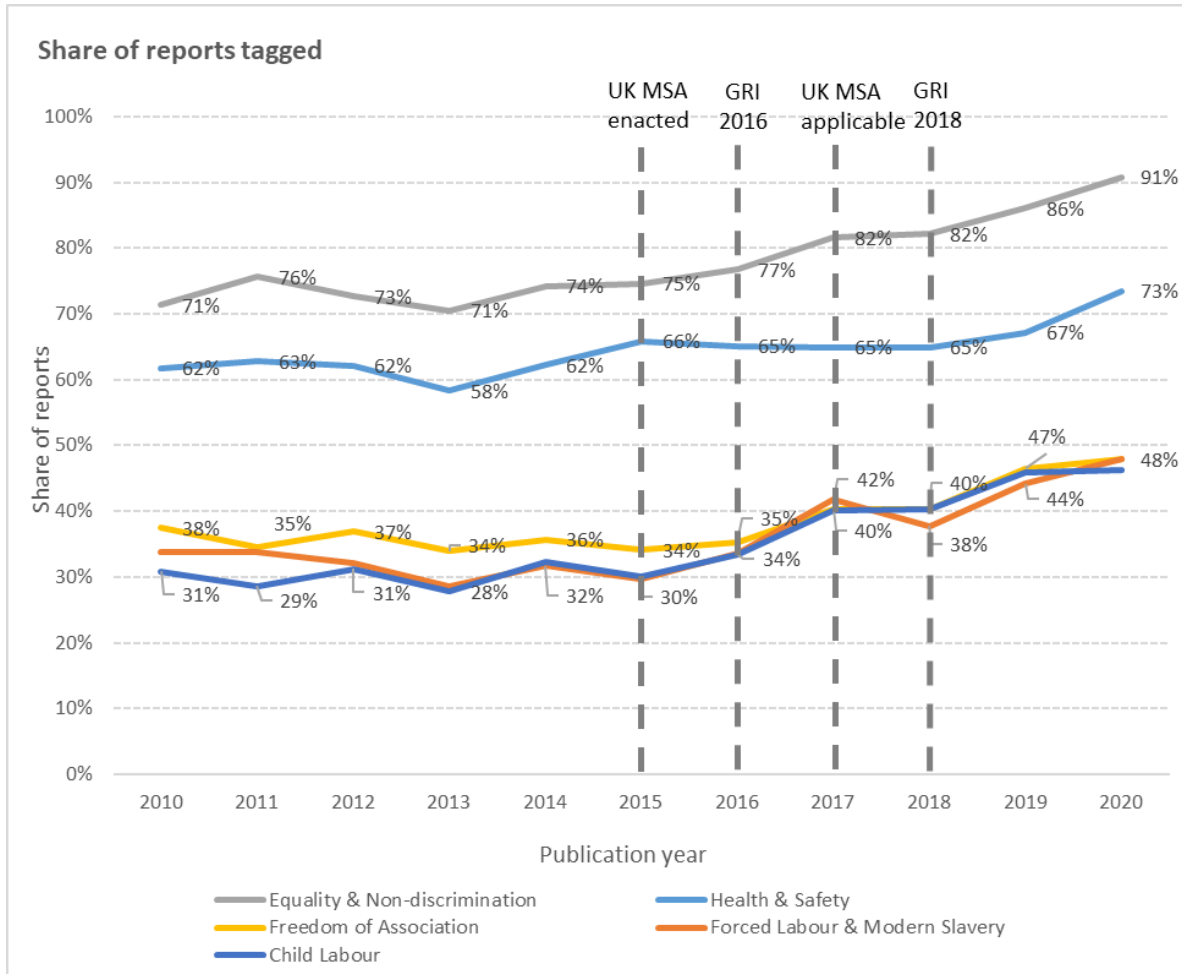
We then considered whether we could observe trends which could be attributable to regulatory initiatives and the introduction of standards focused on particular human rights issues when the data was segmented further.

Focus on labour indicators

We identified five fundamental labour rights indicators (child labour; equality and non-discrimination; forced labour and modern slavery; freedom of association; and health and safety) and considered whether there were observable trends which could be responsive to the following initiatives:

- UK Modern Slavery Act (UK MSA) which was enacted in 2015 and included obligations to report on action taken to address modern slavery in the supply chain from 2017 onwards;
- 2016 GRI standards which included voluntary reporting standards on: labour; diversity and equal opportunity; non-discrimination; freedom of association and collective bargaining; child labour; and forced labour; and
- 2018 GRI standards which included voluntary reporting standards on health and safety.

This is represented in the visualisation below.



For all the indicators we see a statistical significance development over time between 2010 to 2020. In relation to particular indicators:

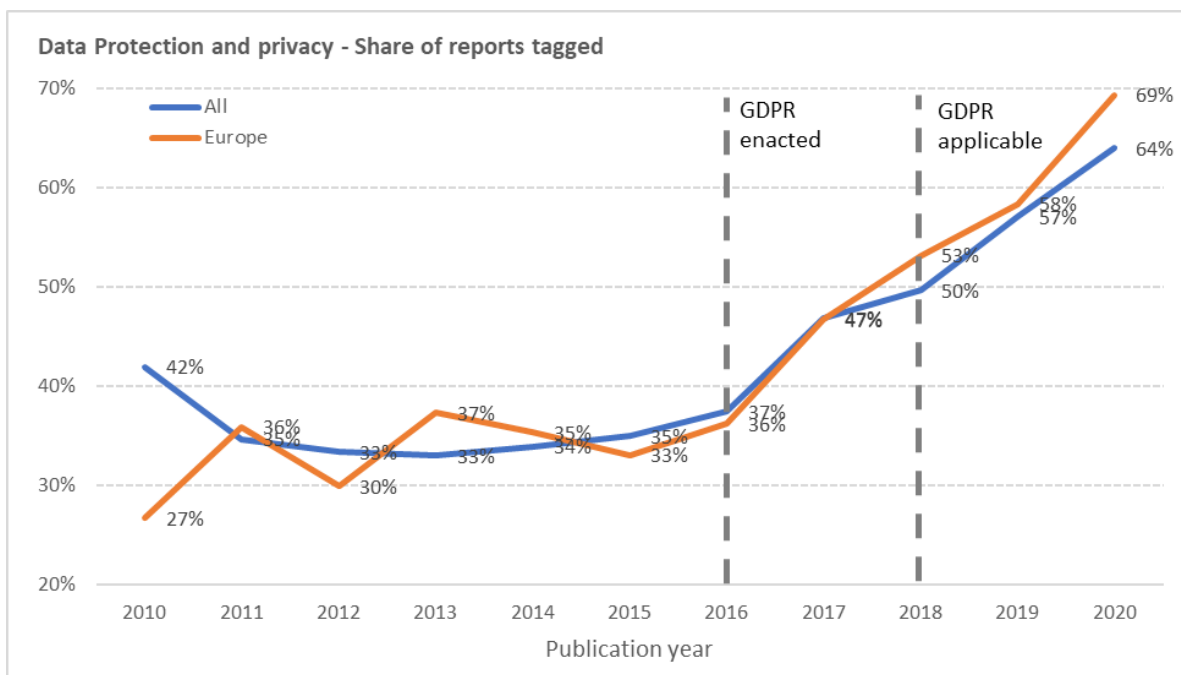
- Forced labour and child labour:** For both the indicators for forced labour and child labour we see a statistically significant development between 2010 to 2020 with a particular increase from 2015 onwards. However, we are not able to make conclusive findings attributing particular developments to the introduction of the UK MSA which was enacted in 2015 and required reporting on how companies address modern slavery in their supply chains from 2017, or the 2016 GRI standards which included reporting standards on child labour and forced labour. For example, from the period 2016 to 2017 we observe a statistically significant increase in reporting. For Child Labour we see a statistically significant development from 2017 to 2019, but not from 2017 to 2018. For Forced Labour, we see increases from 2015 when the UK MSA was enacted, but the development from 2017 to 2019, when the reporting obligation under the UK MSA was operative, is not statistically significant.
- Equality and non-discrimination & Freedom of association:** For both indicators we see a statistically significant development between 2010

to 2020 with a particular increase from 2016 onwards. For Equality and non-discrimination, we also see that the development from 2016 to 2017 and likewise from 2016 to 2020 is statistically significant. For Freedom of association we also see a statistically significant development from 2016 to 2017 and likewise from 2016 to 2020. This could potentially be attributable to the introduction of GRI reporting standards on these issues in 2016, but we are not able to make conclusive findings.

- **Health and safety:** Overall, we see a statistically significant development from 2010 to 2020. In 2018 the GRI standards, which included health and safety reporting standards, were published and we can observe some trends which may be responsive. From 2018 to 2020 the increase in reporting on health and safety is statistically significant, with a larger increase between 2019 and 2020.

Focus on data protection and privacy

We then isolated the data protection and privacy indicator to see whether we could observe any trends in reporting which could be attributable to the introduction of the General Data Protection Regulation (GDPR) which was enacted in 2016 and became operative in 2018.



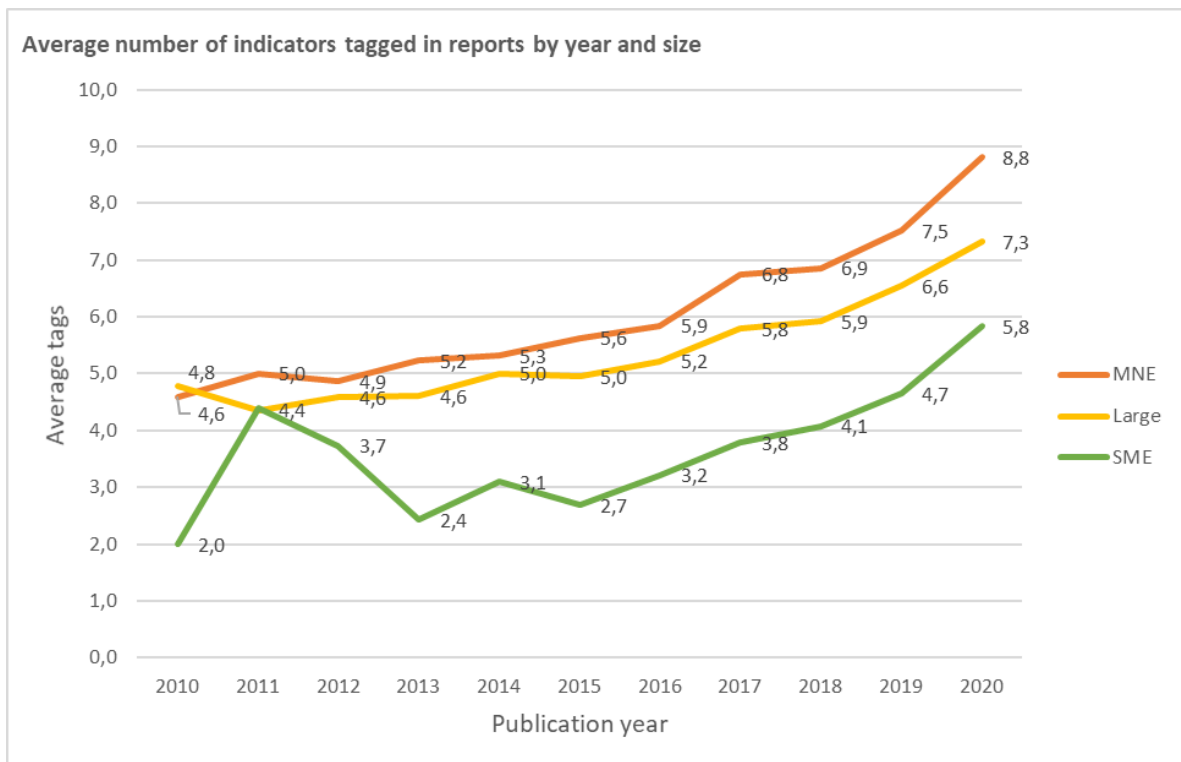
We see a statistically significant development from 2010 to 2020. Likewise, we also see significant increases in reporting between certain years following the GDPR becoming operative in 2018, where the developments from 2018 to 2019 and from 2018 to 2020 are statistically significant. Further, from 2016 to 2017 and the following year we see a more dramatic increase in reporting on data protection and privacy issues, which may indicate that companies were taking steps to prepare for the upcoming regulation, which is reflected in their disclosures.

Lastly, we see increases among European company reports, but in addition we see corresponding increases among company reports globally. This could indicate that the GDPR has had a broader impact globally, and not only among European companies who are directly captured by the regulation.

3.3 HYPOTHESIS 3: LARGER COMPANIES ARE REPORTING MORE ON HUMAN RIGHTS THAN SMALLER COMPANIES

We then wanted to segment the data in a different way to identify whether we could observe any trends in reporting among companies of different sizes. To do this we analysed the data in two different ways. Firstly, we identified the average number of human rights indicators tagged in reports in a particular publication year and by company size in accordance with how the reports were identified in the Dataset by GRI: Multinational Enterprises (MNE); Large Enterprises; and Small or Medium Sized Enterprises (SME).

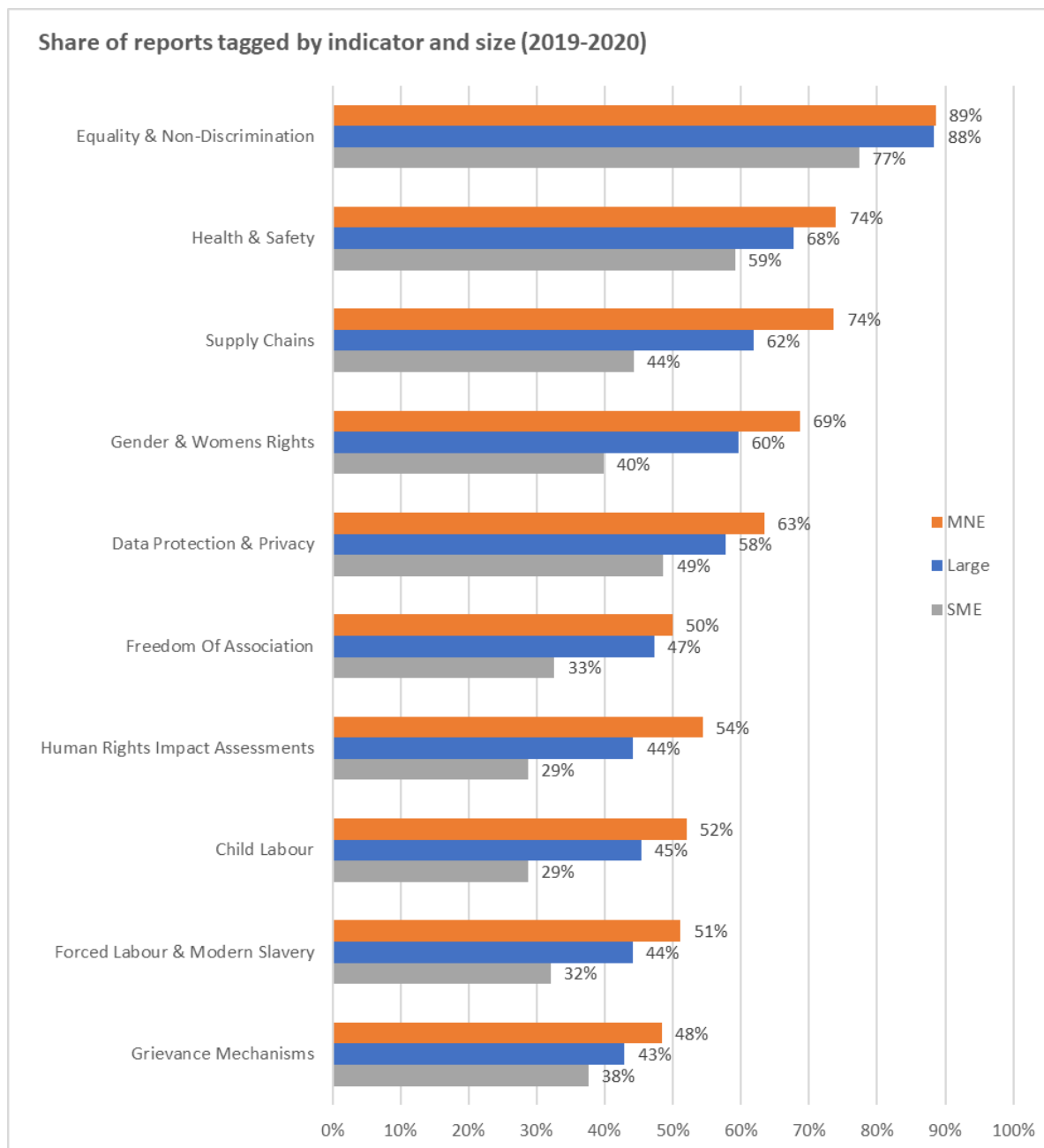
This is represented in the visualisation below.



As is shown in the visualisation above, MNEs report significantly more on human rights topics than companies in both the large and SME categories, with the exception of an outlier for SMEs between 2011 and 2012. Further, we can observe a statistically significant increase between MNEs and large companies in the period 2015-2020.

We then wanted to identify whether there were any observable trends in reporting when we looked at the kinds of human rights issues that the companies of the various size classes were reporting on.

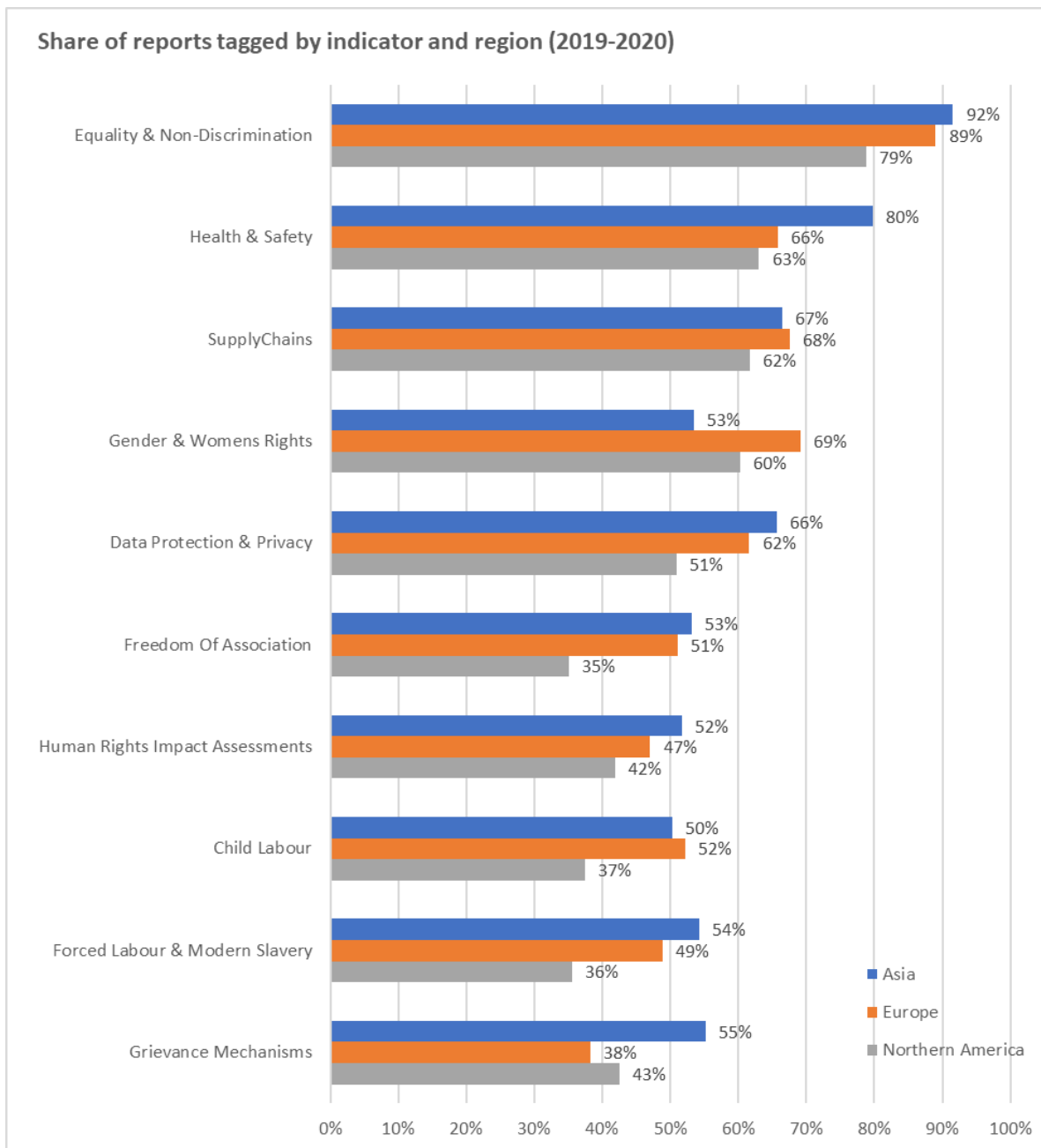
To do this, we isolated reports from 2019-2020 and identified the percentage of company reports in the different size segments that were reporting on the top ten indicators (ie, indicators that companies were reporting on the most). This is set out in the visualisation below.



We see that MNE companies are reporting more on almost all of the top ten indicators than both large and SME companies, and likewise large companies report more than SME for almost all ten indicators.²

3.4 HYPOTHESIS 4: THERE ARE REGIONAL DIFFERENCES IN REPORTING ON HUMAN RIGHTS TOPICS

We then wanted to test whether we could observe any regional differences in reporting. To do this, we analysed reporting from three regions: Northern America, Europe, and Asia, as these regions accounted for over 80% of the companies in the sample. We then isolated reports from 2019-2020 and identified the percentage of company reports in the three regions that were reporting on the top ten indicators (ie, indicators that companies were reporting on the most). This is set out in the visualisation below.



As an overall trend, for the period 2019-2020 it appears that companies from Northern America report less on human rights issues than Asian and European companies, whereas the difference between Asia and Europe is for most of the indicators not statistically significant. When we look into each of the top ten indicators, we can see some further trends in reporting:

- **Equality and non-discrimination:** Companies from Northern America (79%) report less on this issue than Asian (92%) and European (89%) companies, but is still the most reported indicator overall.
- **Health and safety:** Asian companies (80%) report more than companies from Northern America (63%) and Europe (66%).
- **Gender and women rights:** European companies (69%) report more on this than companies from Northern America (60%) and Asia (53%).
- **Data Protection and privacy:** Companies from Northern America (51%) report significantly less on this issue than Asian (66%) and European (62%) companies.
- **Child labour & Forced labour and modern slavery:** Companies from Northern America (Child labour 37%; Forced labour 36%) report significantly less on these issues than Asian (Child labour 50%; Forced labour 54%) and European companies (Child labour 52%; Forced labour 49%).
- **Freedom of association:** Companies from Northern America (35%) report significantly less on this issue than Asian (53%) and European companies. (51%)
- **Grievance Mechanism:** Asian companies (55%) report significantly more than companies from Northern America (43%) and Europe (38%).

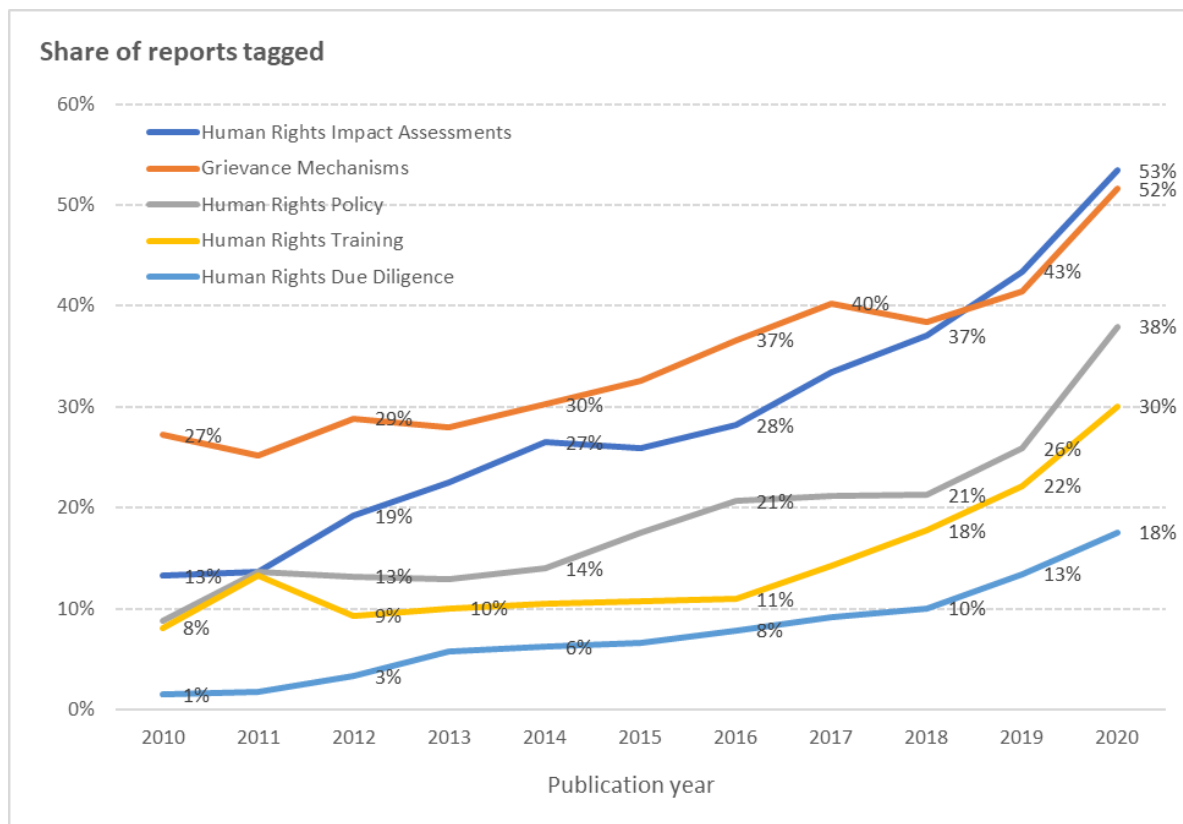
It may be the case that some of these variations can be attributable to differences in regional regulatory environments, for example, while larger European companies are required to report in accordance with the NFRD, there is no equivalent obligation in Northern America. However, this does not account for the generally high percentage of company reports which include the top ten indicators in the Asia region, which in many cases is similar to reporting found in European company reports and for some indicators has a significantly higher percentage of reports making a relevant disclosure (for example, in the case of health and safety disclosures).

3.5 HYPOTHESIS 5: COMPANIES ARE INCREASINGLY REPORTING ON PROCESSES FOR IDENTIFYING AND ADDRESSING HUMAN RIGHTS IMPACTS

Lastly, we wanted to test whether we could observe any trends in reporting on the process for identifying and addressing human rights impacts. To do this, we isolated a series of “process indicators” which identified what policies and processes for managing human rights were disclosed, as follows:

- Human Rights Policy
- Human Rights Due Diligence
- Human Rights Impact Assessments
- Human Rights Training
- Grievance mechanisms

This is set out in the visualisation below.



For all indicators we see a statistically significant increase over time from 2010 to 2020. For example, disclosures on human rights impact assessments increased from 13% in 2010 to 53% in 2020. Similarly, we see significant increases in reporting on human rights due diligence from 1% in 2010 to 18% in 2020, and human rights policies from 8% in 2010 to 38% in 2020.

4 CONCLUSIONS: WHAT CAN AI DO FOR BHR

4.1 WHAT CAN AI DO FOR THE ANALYSIS OF COMPANY REPORTING ON HUMAN RIGHTS ISSUES?

As is set out in the previous section, we were able to use AI technology to analyse the company reports in the Dataset and generate data which we reviewed against a series of hypotheses. From that review we were able to observe a number of high level trends set out in this report, which are indicative and by no means exhaustive. In some cases we were able to make observations which suggested that trends in company reporting may be responsive to regulatory developments. Further, we were able to segment the data, and explore whether there were observable trends by size, geography and by sector. The full range of data analysed by the algorithm could be of use to researchers and other actors interested in exploring company reporting on human rights issues, and potentially be a valuable supplement to quantitative analysis.

Further, as we have now trained the algorithm to identify text bits which are relevant to a range of human rights indicators, it is possible that it may be deployed in the future should company sustainability reporting mature in a manner which addresses some of the challenges outlined in the section below. This may indeed be possible as European companies begin to report against a common set of sustainability disclosure standards and place machine readable versions of their reports into an accessible database, as would be required by the forthcoming CSRD.

However, the challenges introduced in the methodology section above and elaborated on in the section below reveal the limitations of the use of this kind of technology for analysis of company sustainability reporting in its current form.

4.2 CHALLENGES OF USING ALGORITHM SUPPORTED ANALYSIS TO REVIEW COMPANY DISCLOSURES

Although we see promise in big data technology to scale up review of company disclosures on human rights, in addition to challenges associated with assessing the substance of reporting, there are challenges in relation to the accessibility of company reports and how reported information is prepared and presented. DIHR has successfully undertaken a similar project using big data analysis of 145,000 recommendations from the international human rights system, assisted by an algorithm resulting in the creation of the SDG-Human Rights Data Explorer. That project was assisted by the standard format used in Universal Periodic Review (UPR) and other similar reporting which simplified the process of extracting text for analysis by the algorithm.

As noted above in the methodology section, company sustainability reporting is significantly different from the standardised data sources such as UPR reporting to which DIHR has applied the algorithm to date with great success. Companies are

diverse, operate in different sectors, different geographic contexts, different cultures, and use different standards despite the common frameworks of the UNGPs and OECD Guidelines, and the introduction of disclosure standards such as those developed by GRI.

Accordingly, there have been challenges in adapting the algorithm to the Dataset, summarised as follows:

1. **Accessibility of reports:** while we had access to a database of around 50,000 sustainability reports maintained by GRI, as noted in the methodology section above, only 18.8% of those reports were able to be downloaded. Further, GRI's database is no longer being maintained, meaning that no reports post December 2020 are included. Although there have been initiatives such as GRI's Sustainability Reporting Database to generate repositories of reports uploaded by, in that case, users of the GRI reporting standards, there is currently no central repository of company sustainability reports. Each of these challenges makes it difficult to accurately track progress in company reporting over time.
2. **Format and presentation of data:** There is considerable variation in the presentation of data on human rights issues in current company reporting, which presents difficulties when attempting to assess and compare reporting practice. Potentially relevant information is contained in a range of formats, including tables or text boxes in picture format, which has been challenging to convert into a machine-readable format.
3. **Limits on sectoral analysis:** Further, because only 18.78% of reports in the Dataset were ultimately accessible and able to be included in the analysis our sectoral analysis was limited because of the way that sectors were tagged in the Dataset. In the GRI Sustainability Database, there was a large number of sectors included, which meant that ultimately there was often an insufficient number of reports for a particular sector by year to make statistically significant findings on trends in reporting over time. Further, the Dataset included an "other" category which included a large number of reports for which it was not possible to identify a sector. Out of the 9374 reports analysed, 831 reports are in the "other" category (representing 9% of all of the reports).
4. **No common set of standards:** reports in the database all used the voluntary GRI standards. However, even though the reports in the dataset all used the same standards, there is little standardisation in the language used in company reporting. Most reports adopt a narrative style which leads to considerable variation in the way that text bits relevant to the various human rights indicators were expressed. This feature of company sustainability reporting does not lend itself to comparability.
5. **Limitations of algorithm:** There are also limits on what this kind of algorithm assisted analysis can tell us about the data. The algorithm cannot assess quality of reporting, only whether a text bit meets a threshold of relevance. As we can see from the examples extracted in Annex 3 some text bits clearly disclose more substantive information than others. This places limitations on the conclusions that can be drawn from the overall trends observable in the data set out in the previous section. Further, more disclosure does not necessarily equate to stronger performance. For example, previous research has found that there is often a size bias when it comes to ESG ratings whereby larger

companies tend to receive higher ESG scores.³ A similar trend was observed in this project, whereby MNEs reported more than large companies or SMEs on human rights indicators. However, it is possible that it is the case that larger companies simply have more resources to devote to reporting. Caution must be exercised when making findings that company reporting is improving on human rights issues simply because the algorithm has identified text relevant to particular indicators with greater frequency.

As noted above, a number of actors have advertised services which use AI technology similar to that deployed by DIHR in this project. This includes ESG ratings agencies and market analysts, many of whom purport to be able to accurately assess the ESG performance of companies through analysis supported by machine assisted review of sustainability reports and other data sources. While AI assisted analysis could be a component part of a broader analysis, as we have found in this project there are limitations on the ability of an algorithm to assess the quality of reporting and therefore make a qualitative assessment of a company's performance on social and human rights matters.

4.3 WHAT WOULD IT TAKE TO MAKE THIS WORK EFFECTIVELY?

The challenges elaborated above could be addressed in part through greater standardisation and accessibility. Regulatory developments at the EU level could go some way to addressing these challenges, as described below:

4.3.1 Common standards

The capacity to easily access, analyse and compare reporting is crucial for a range of stakeholders including ESG investors and civil society groups focused on corporate accountability. However, as this project has already highlighted, corporate reports are prepared in different ways with differing formats. This lack of standardization and accessibility has a deleterious effect on the ability to assess and compare company reporting. Standardisation of reporting would greatly assist this kind of big data analysis, enabling efforts to analyse company reporting to be scaled up. Further, as noted above, only 18.78% of reports in the GRI database were ultimately accessible by the DIHR. This makes it difficult to track progress in company reporting over time. While there exist a range of voluntary reporting standards such as those developed and disseminated by GRI, there is a considerable degree of discretion, and therefore variation, concerning how companies utilise these standards, if at all.

As noted above, the development of the CSRD, which will replace the NFRD, foresees that companies will be required to report on a range of "sustainability matters" and provides for a new EU sustainability reporting standards that all reporting companies will be required to use.

Exposure drafts of these standards were published and at the time of writing this discussion paper were available for consultation. The current drafts of these standards require companies to report on a range of cross cutting matters, including: A description of the undertaking's business model and strategy, including plans to ensure

compliance with net zero targets, the effect of sustainability risks and opportunities, and how the strategy/plans reflect broader stakeholder interests; Sustainability targets and progress against them; Due diligence processes implemented for sustainability matters; Principal and potential adverse impacts throughout the value/supply chain; The role of management in sustainability matters; and Principal risks of the undertaking related to sustainability matters.

In addition, the exposure drafts include detailed reporting requirements on a range of social topics, which include human rights matters. These are organised by reference to affected stakeholders, requiring disclosures on impacts on the company's own workers, workers in the value chain, affected communities and end users or consumers. The development of a common set of compulsory European Sustainability Reporting Standards could have a significant impact on the comparability of company reports, improving the ability to analyse reporting and gain greater insight into how companies are identifying and addressing their human rights impacts, including through the introduction of numbered and specific disclosure requirements allowing for easy comparison.

4.3.2 Central repository of reports in machine readable format

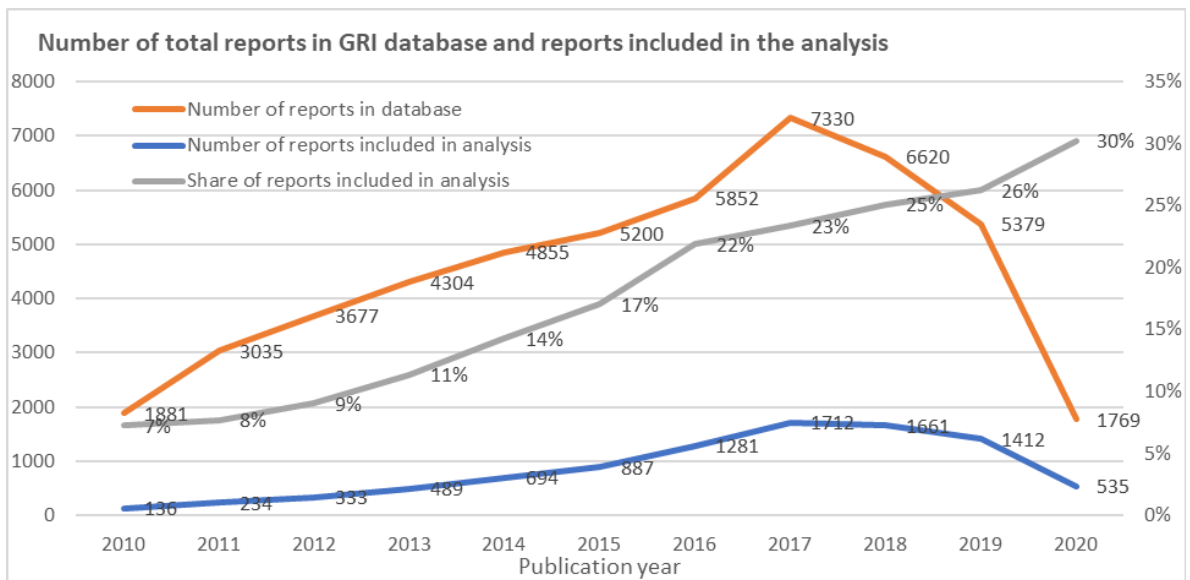
Ensuring that company reports are made accessible and digitised in a machine-readable format is essential not only for big data analysis projects, but for all stakeholders wishing to assess and compare the human rights performance of reporting companies. An official EU repository, where businesses are required to deposit up-to-date reports, is important for a range of stakeholders including ESG investors and civil society actors focused on corporate accountability.

The CSRD anticipates the creation of an EU Single Access Point for financial and non-financial information. Such a database could address a number of the challenges faced in this project concerning the accessibility of the reports in the Dataset, including challenges faced with respect to machine readability and processing of the reports into text-bits. The ability to access a comprehensive database of reports would also facilitate analysis of changes in company reporting over time. Lastly, it would be of great assistance if reports in the Single Access Point could be tagged by size, sector and geography to enable segmentation and analysis of the reports.

ANNEX 1: DESCRIPTION OF THE DATASET

The number of reports contained in the Dataset increased from 2010 to 2017 and decreased from 2017 to 2020. However, the period 2017-2020 represents some of the years with the highest number of reports. There were significantly fewer reports for 2020, which could be attributable to the impact of COVID 19. We have nonetheless included data for this year, as analysis on the company reports for this year indicated that it included a representative sample of companies across sizes, sectors and geographies comparable to previous years.

The share of reports included in the analysis increased during the whole period, from 7 % in 2010 to 30 % in 2020. This also means that we have more reports from the most recent years in our analysis. In the period from 2010-2015 we have 2,773 reports included in our analysis and in the period from 2016-2020 we have 6,601 reports.



Out of the 9,374 reports included in our analysis we have reports from all company sizes, from companies all over the world and in a wide range of sectors. Most of the reports are from large and multinational companies and from the regions Asia, Europe, and Northern America. When we compare the reports included in the analysis (sample) with the reports in the GRI database (population) on size, geography, and sectors we see that the distribution of our sample is not considerably different from that of the population (see the table below). In both periods we see that multinational companies are a bit overrepresented in the sample whereas the small and medium companies are a bit underrepresented. When looking at regions we see that companies from Africa and Northern America are a bit overrepresented in the sample and companies from Latin America & the Caribbean are underrepresented. Companies from Asia are a bit underrepresented in the latest period whereas this is the case for companies from Europe in the first period.

These features of the dataset are set out in the table below.

	2010-2015		2016-2020	
	Population -Number of reports in database	Sample -Number of reports in analysis	Population -Number of reports in database	Sample -Number of reports in analysis
Number of reports	22,952	2,773	26,950	6,601
Size				
Large	62%	61%	59%	57%
MNE	28%	32%	27%	33%
SME	10%	7%	14%	10%
Not defined	0%	0%	0%	0%
Region				
Africa	7%	13%	6%	10%
Asia	29%	30%	38%	33%
Europe	36%	31%	30%	31%
Latin America & the Caribbean	12%	2%	14%	2%
Northern America	12%	19%	10%	19%
Oceania	4%	6%	3%	5%
Not defined	0%	0%	0%	0%
Sector				
Construction	6%	4%	5%	5%
Energy	10%	9%	9%	9%
Financial Services	13%	13%	14%	15%
Health	4%	3%	4%	4%
ICT (Information Communications Technology)	7%	8%	7%	7%
Mining and metals	6%	7%	5%	6%
Other	43%	43%	45%	42%
Retail and consumer goods	5%	6%	4%	6%
Transport and logistics	7%	7%	7%	7%

ANNEX 2: HUMAN RIGHTS INDICATORS

Process indicators

- Human Rights Policy
- Human Rights Due Diligence
- Human Rights Impact Assessments
- Human Rights Training
- Salient Issue Analysis
- Grievance mechanisms

Fundamental labour rights indicators

- Child labour
- Equality and non-discrimination
- Forced labour and modern slavery
- Freedom of association
- Health and safety
- Rightsholder indicators
- Gender and women's rights
- Human rights defenders
- Indigenous people

Supply chain indicators

- Conflict minerals
- Supply chain

Other rights indicators

- Data protection and privacy
- Land rights

ANNEX 3: EXAMPLE TEXT BITS

A review of the text bits tagged by the algorithm as relevant to a particular indicator reveals that some text bits clearly disclose more substantive information than others. Extracted below are some examples of text bits responsive to the “human rights due diligence” and “human rights impact assessment” indicators. As can be seen, some text bits disclose meaningful information within the limits of the length of the text bit (see examples 1, 2, 7 and 8 below), whereas others are significantly less meaningful. In the examples below, we see less significant disclosures being made either because a text bit discloses generic information (see examples 3, 4, 9 and 10 below) or because the text bit extracts text from an index or similar, containing only a text reference to the relevant indicator while disclosing no information of substance (see examples 6 and 11 below).

HRDD: Stronger/more meaningful disclosure

Example 1: The goals of the due diligence include: the identification and mapping of Human Rights risks resulting from the Group’s operations; confirmation that each topic is governed by a specific internal regulatory framework (e.g., policy, procedure), has a management system that regularly monitors and traces the performances (if possible through appropriate indicators, such as those concerning health and safety), and that the relative responsibilities have been assigned; the definition of a gradual improvement path which, starting with simple respect for the local laws, guides the policies and processes of Human Rights towards sharing with the appropriate stakeholders through appropriate involvement initiatives; the highlighting of any gaps or inconsistencies between the company’s various departments and/or between the different companies of the Group; the possibility to discuss Human Rights with the other companies.

Example 2: During 2015, we rolled out our previously launched Human Rights Policy across the Group and used our human rights framework to guide our approach to identifying and addressing our salient human rights risks. Human rights requirements have also been integrated into the Social Way and SEAT tools. As a consequence, operation-level due-diligence processes have been conducted at the majority of sites either through our mandatory SEAT process or through a specific human rights risk assessment, to determine salient human rights issues for their local context. Our most salient human rights risks from a Group perspective and throughout all stages of the mining life cycle are shown in the graph below. Our progress on managing these is addressed throughout this report.

HRDD: Weaker/less meaningful disclosure

Example 3: We are maintaining our focus on understanding risks and integrating mitigating actions into our business processes, for example through implementing human rights due diligence across our business units. We are approaching this from a continuous learning and improvement perspective. As a company we face many challenges, and a rights respecting approach can therefore be demanding.

Example 4: The three main groups we impact are our own people, the people in our supply chain and our customers. We are committed to upholding the UN Guiding Principles on Business and Human Rights and the OECD Guidelines for Multinational Enterprises. Our approach is to carry out effective and robust due diligence to prevent, mitigate, remedy and account for adverse impacts of our activities.

Example 5: ISO 26000 Index Core Subjects and Issues Related CSR Report Section Page(s) Explanatory Notes Organizational governance Decision-making processes and structures 4. Corporate Governance 4.1 Governance Structure 4.2 Board of Directors 4.3 Audit Committee 4.4 Compensation Committee 17, 18 Human rights Due diligence Our new sites are in Science Park and compliant with Science Park's Environmental Impact Assessment commitments and legal requirements.

Example 6: Human Rights / Due diligence / Human rights risk situations / Avoidance of complicity / Discrimination and vulnerable groups / Fundamental principles and rights at work 54 HR8 Percentage of security personnel trained in the organization's policies or procedures concerning aspects of human rights that are relevant to operations.

HRIA: Stronger/more meaningful disclosure

Example 7: In 2013, we developed a joint project involving our business units and corporate areas to bring together the practical lessons learned and develop a corporate guide to assessing human rights impacts for the business units. The project consists of four human rights impact assessments: two in Peru and two in Bolivia. Three of these studies involved onshore exploration and production operations where indigenous communities live in the direct area of influence of our activities. We also carried out a study on a downstream facility in Peru. These four studies were carried out with the support of leading international human rights experts. We also involved a number of our operating units in the study so that we could analyze and assess the overall human rights situation. The assessments enabled us to define a specific methodology, consistent with the Guiding Principles, for identifying and assessing the potential human rights impacts of our activities. They also helped us improve the integration of human rights specifically into our impact assessment processes.

Example 8: In addition, human rights issues are integrated in the due diligence process we conduct before entering new markets (either as an operator ourselves or through partnerships with other operators). This process uses internationally recognised indexes and evaluations of particular issues such as corruption, political affiliations, respect for privacy, internet freedom, freedom of expression and workers' rights to assess and highlight the potential impacts or risks associated with entering new markets. In 2013/14, we further strengthened our human rights impact assessment process for potential new markets identified as high risk. The findings are considered in the decision-making process before entering a new market.

HRIA: Weaker/less meaningful disclosure

Example 9: Our Approach to Engagement on Human Rights As our business grows and reaches more people across the world, we must continue to identify and enhance the way we address our key human rights opportunities, risks, and challenges. To help achieve our goals, we regularly engage and collaborate with our stakeholders, including nonprofits, industry peers, investors, and CSR practitioners, to help us gauge expectations and understand the ongoing effectiveness of our work.

Example 10: Security and human rights We developed a new tool this year to help us implement our security and human rights programme, including our policy for security personnel on the use of force, arms and firearms. We also started a security and human rights risk assessment in Algeria using this framework. Based on the findings, an action plan has been developed for implementation in 2013.

Example 11: Indexes About this Report OTE Group Subsidiaries 218 Marketplace Employees Society Environment CR Strategy and Management GRI Summary Description Status Report Section / Notes Social: Human Rights Disclosures on Management Approach HR10 Human Rights Reviews or Risk Assessments F CR STRATEGY AND MANAGEMENT / Compliance Management System and Policies EMPLOYEES / Fair Employment Policy Framework MARKETPLACE / Supplier Evaluation HR11 Grievances Filed on Human Rights Issues F EMPLOYEES / Fair Employment Policy Framework MARKETPLACE / Supplier Evaluation Index Key: F = Fully P = Partially Not = Not Reported 8.3.4 Social Performance Indicators (Cont'd)

ENDNOTES

- 1 The test tells us the probability that the relationship we observe is due only to random chance. When we test the average number of indicators we use a t-test, whereas we use a χ^2 -test when we test the share of company reports tagged to an indicator. We use a significance level of 5% when we conclude on the test results.
- 2 We can observe a statistically significant difference in reporting between each of the company size groups for almost all indicators. The only indicators where we do not observe a statistically significant difference is Freedom of association and Equality and non-discrimination where the difference between Large and MNE companies are not significant. And likewise for the indicator Grievance mechanism we do not see a statistically significant difference between Large and SME companies.
- 3 Timothy M Doyle, ACCF, Ratings that Don't Rate: The Subjective world of ESG Ratings, July 2018, p9 available at https://accfcorpgov.org/wp-content/uploads/2018/07/ACCF_RatingsESGReport.pdf



THE DANISH
INSTITUTE FOR
HUMAN RIGHTS