Opinion piece

Uncovering hidden signals for sustainable investing using Big Data: Artificial intelligence, machine learning and natural language processing

Received (in revised form): 21st February, 2020

Madelyn Antoncic
Managing Partner, Global AI Corporation, USA

Madelyn Antoncic, PhD, an Economist, is Managing Partner at Global AI Corporation, a Big Data company providing artificial intelligence-driven environment, social and governance (ESG) and sustainable development goal (SDG) data for institutional clients, governments and corporations. She also serves as a member of the Board of Directors of S&P Global Ratings and FinTech Acquisition Corp III and serves on the Board of Overseers of Weill Cornell Medicine. Formerly, she has held a variety of leadership roles with global financial institutions in both the private and public sectors, including as Vice President and Treasurer of the World Bank, Goldman Sachs and Lehman Brothers, and most recently she was CEO of SASB.

205 Hudson Street Floor 7, New York, New York 10013, USA
Tel: 646-807-4759; E-mail: Mantoncic@globalai.co

Abstract Risk managers and investors have increasingly been seeking high-quality environment, social and governance (ESG) data in order to assess nonfinancial risks as well as allocate capital towards companies that manage themselves in a 'socially responsible' way and adhere to their contract with society. The problem is that due to the lack of agreed-upon standards for companies to use for reporting on sustainability issues, there is a paucity of high-quality firm-level data to serve as key inputs in assessing a company's risks and adherence to ESG criteria. Big Data, developed through cutting-edge statistical models, artificial intelligence (AI) and natural language processing (NLP) covering dozens of languages, provides the solution for ESG rankings and ratings and can help combat self-reported bias and 'greenwashing' and provide high-quality data. The 'next generation' measures of firms 'doing good' are the UN *sustainable development goals* (SDGs), which are this decade's benchmarks against which millennials and many investors are beginning to assess companies. The SDGs go beyond the more narrowly focused set of sustainability issues embedded in ESGs, and quality data to measure performance against the SDGs are even more sparse. Using Big Data, Global AI Corporation uncovers data measuring companies' and counties' performance on all 17 SDGs, which can enable the integration of SDG factors into investment, risk management and national policy decision-making processes. Big Data is providing statistical indicators and performance metrics data to national governments and the United Nations to benchmark progress towards achieving the SDGs. It is also producing the SDG footprint of the private sector at the regional and global levels for policy purposes as shown in the United Nations Conference on Trade and Development's (UNCTAD) SDG Pulse publication. Using Big Data, Global AI Corporation eliminates self-reporting biases and uncovers hidden data, which results in negative as well as positive ESG/SDG scores, while the self-reporting data only produces positive scores.

Keywords: *sustainable investing, risk management, Big Data, artificial, intelligence, ESG, SDGs*

INTRODUCTION

It was not that long ago that the top risks on the minds of investors and business leaders were economic in nature. Over the past decade, the profile of those risks has gradually shifted from traditional economic considerations towards an emphasis on environmental, social and governance (ESG) issues.[1]

Today, market participants, policy-makers and governments' highest priority risks — both in terms of probability and magnitude — include extreme weather events, natural disasters, water-related crises, and the failure of climate change mitigation and adaptation. In addition, investors are increasingly incorporating sustainability considerations into their investment decisions and using these indicators to screen potential investments in order to allocate capital to the most efficient users of that capital.

ESG data embed a set of criteria that one can use to evaluate a company's adherence to its contract with society. More latterly, however, millennials and many investors looking for firms 'doing good' have moved beyond the more narrowly focused set of 'material' sustainability issues embedded in ESGs and have embraced the UN sustainable development goals (SDGs) as this decade's true measure of 'doing good'. The UN SDGs,[2] otherwise known as the global goals, are a universal call to action established in 2015 by 193 countries to end poverty, protect the planet and ensure that all people enjoy peace and prosperity.

This has all led to a significant shift from short termism to focusing on the long-term sustainability of companies and 'fiduciary capitalism'. While the UN secretary general has noted[3] that the 'exclusive pursuit of short-term profit is incompatible with the promise to realize sustainable and inclusive development for all people, everywhere, on a healthy planet', business leaders are increasingly aware that the exclusive focus of even pursuing long-term corporate sustainability and profit is incompatible with their contract with society as one can infer from the recently published business roundtable statement, which moves away from shareholder primacy to a commitment to all stakeholders.[4] In other words, in order to ensure long-term corporate success and value creation, companies have to focus on what is material to all stakeholders to keep their contract with society and their 'license to do business'.

One could argue that this shift in focus and awakening was partly due to the lessons learned from the great financial crisis and partly a result of another awakening in 2007, when the United Nations Intergovernmental Panel for Climate Change published a report linking human action to global warming.[5] Whatever the catalyst, it is clear that from these events was born sustainable investing, which is one of the most important trends observed in the asset management industry over the past decade.

Meaningful progress against social and environmental goals — such as achieving the 17 SDGs by 2030 or the targets under the Paris Agreement — will require mobilising capital on a large scale to where it is needed most. This will require having material sustainability information in hand so investors can direct their financial capital to those companies — globally — that are being managed most effectively for the long term. Investors need sustainability information that is both relevant and reliable, if they are to make informed decisions concerning whether or not to invest in a company or whether to overweight or underweight an investment compared to a benchmark.

The problem is, regardless of all the effort and good intentions, there is a lack of high-quality firm-level data to serve as key inputs in assessing the sustainability risks and opportunities a company faces. While sustainability reporting has become near ubiquitous in recent years, the practice has been widely criticised for lacking the rigor of traditional financial reporting; thus, the data quality lags data volume.

Moreover, sustainability reporting typically covers policies and targets and not performance, and the data are not always comparable across companies, thus making the data of limited value. In addition, the sustainability information companies share is not always focused on the unique needs of investors. For example, it often lacks clear links to a company's financial condition, operating performance or market valuation, which drive investors' decisions. Perhaps, most importantly, more often than not, the data are often reported without the rigorous oversight typically applied to traditional financial reporting, nor are there consistent reporting requirements.

The reason for the lack of quality data and the inconsistent reporting is there are no generally accepted agreed-upon standards for companies to use for reporting on sustainability issues. Instead, there is an 'alphabet soup' of acronyms representing different measurement frameworks created by various competing non-governmental organisations (NGOs). There have been attempts at harmonising these standards, the most recent of which is laid out in the World Economic Forum's (WEF) International Business Council's discussion paper.[6] This proposal consists of mapping elements of the existing smorgasbord of standards into four basic 'pillars', which the authors created to represent four broad categories that are intended to encapsulate the SDGs. One can, however, reasonably question how useful this 'harmonisation' can be to investors as it lumps together 'competing standards' that are not constructed consistently but instead use different definitions of materiality, express risks over different time horizons and are not empirically tested as to their relevance to the SDGs. Moreover, these various ESG standards are not consistent with the broadest set of issues for sustainable economic development under the SDG framework, which is 'leaving no one behind'.

Not surprisingly, while corporations now largely self-report some sustainability data, given the lack of standards and the lack of reporting requirements, there is a significant amount of 'greenwashing' and data biases, further adding to the lack of usefulness of reported sustainability data. Moreover, companies that do provide sustainability information typically report these data only on an annual basis, rendering the information stale for any meaningful investment and risk management purposes.

ESG raters and rankers attempt to improve on the lack of data quality and consistency; the information they provide, however, just adds to the confusion not only for investors but also for corporations who wonder how they can be ranked differently by the different raters. A recent, high-profile example includes MSCI, an index provider, which has ranked Tesla as 'one of the top car manufacturers when it comes to assessing it on its ESG credentials. This means indices and fund managers following an "ethical" mandate and using MSCI's ratings are likely to put more of investors' money into the company. However, FTSE — an MSCI rival — took the opposite view and ranks Tesla as one of the least ethical car brands'.[7]

Comparing a company's ratings from the different raters and rankers in the following scatter diagrams shows (see Figure 1) that there is a high degree of non-correlation of a company's ESG rating and ranking across the data providers. This has led to significant noise and a lack of useful sustainability data for investment purposes.

BIG DATA AND ARTIFICIAL INTELLIGENCE PROVIDE A SOLUTION

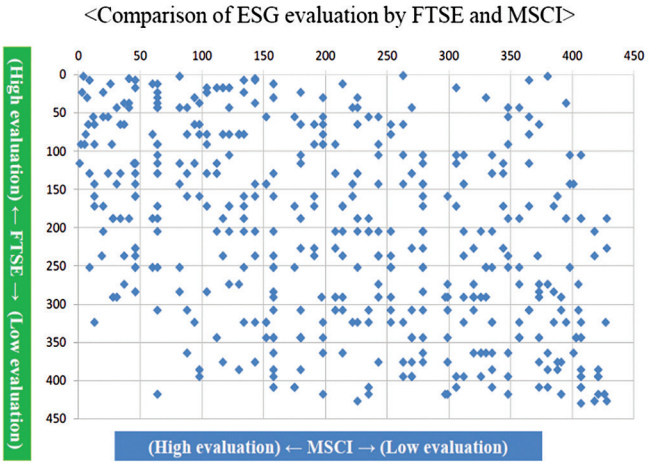
While the private sector has done a commendable job of trying to come up with standards, the world needs an agreed set of standards if we are to ever make progress mobilising capital to delivering on our SDG goals by 2030, and to transitioning to a low-carbon economy consistent with the Paris Agreement. The UN has developed such a system of worldwide reporting standards regarding SDGs, which provide metrics and standards for firm, entity-level reporting on their contributions to the SDGs. UNCTAD's Intergovernmental Working Group of Experts on International Standards of Accounting and Reporting (UNCTAD–ISAR) created the Global Core Indicators[8] (GCIs) through a multiyear, multi-stakeholder effort by governments, regulators, standard-setting agencies and investors. These standards have already received widespread support from over 100 countries, which are encouraging their home companies to report using these metrics. The International Integrated Reporting Council (IIRC) and the World Business Council for Sustainable Development (WBCSD) fully endorsed the core indicators as a baseline approach to facilitate 'harmonisation' and comparability on sustainability and SDG reporting.

The GCIs provide a common set of standards that fulfil all of the quantitative requirements — consistent, comparable, material, universal etc — and also facilitate convergence of financial and nonfinancial reporting, one of the other challenges of the other reporting frameworks, thus making them suitable for consolidated reporting and legal entity reporting. Moreover, the standards can link firm-level reporting to national statistics and

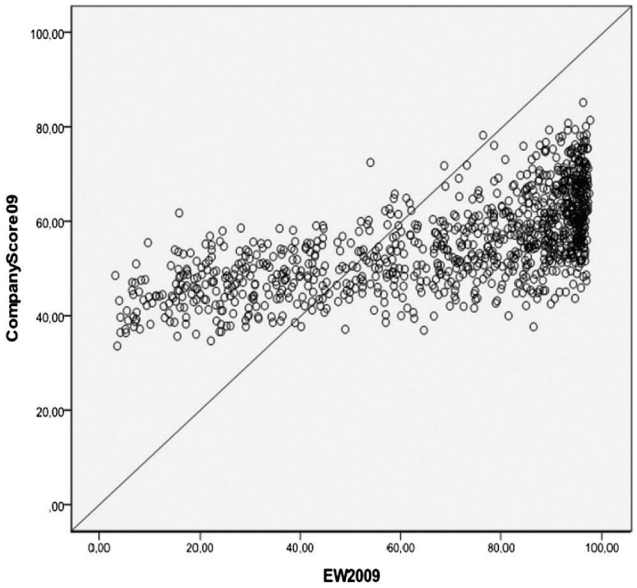
ESG Scores are Different Across Providers
(Cross Sectional Correlation for Constituents of the MSCI World Index, June 30, 2017)

	Sustainalytics	MSCI	RobecoSAM	Bloomberg ESG
Sustainalytics	1	0.53	0.76	0.66
MSCI		1	0.48	0.47
RobecoSAM			1	0.68
Bloomberg ESG				1

The ESG Data Challenge, SSGA, March 2019



Results of ESG Index Selection, GPIF, July 2017



How Robust Are CSR Benchmarks? Comparing ASSET4 with Sustainalytics, Tilburg University, October 2018

Figure 1: Comparison of ESG scores from rankers and raters
Notes: CSR, corporate social responsibility; ESG, environment, social and governance; SDGs, sustainable development goal.

reporting. The main problem with the UNCTAD–ISAR framework, however, as with any reporting by corporations, is that the frequency of reporting is typically only on an annual basis. A useful complement to what Big Data calls 'slow-moving' quantitative indicators ('slow moving' meaning the infrequency of reporting) is to leverage Big Data and AI technologies to extract, process and analyse large-scale 'structured' (firm reported) and 'unstructured' (uncovered through advanced statistical techniques) data on SDG-related factors, which can then enable the integration of SDG factors into the decision-making of global investors.

Using state-of-the-art statistical techniques, Big Data can replace 'dated' sustainability and materiality analyses with 'alternative data' using AI, machine learning and natural language processing (NLP) to cull through tens of thousands of news items, social media and reports in dozens of languages, providing up-to-date information going beyond what is present in unaudited, self-reported annual firm reports or firms' marketing efforts. An AI-driven approach can help uncover hidden material risks, substantially reduce positive biases and uncover negative scores resulting from an adverse SDG footprint. Figure 2 shows the self-reporting bias where all scores are positive. Using Big Data, Global AI Co. uncovers hidden data expressed in the 'Cloud Charts', which result in negative scores.

Exposing what companies are truly doing, not just what they are reporting creates incentives for corporations to quantify and increase their net

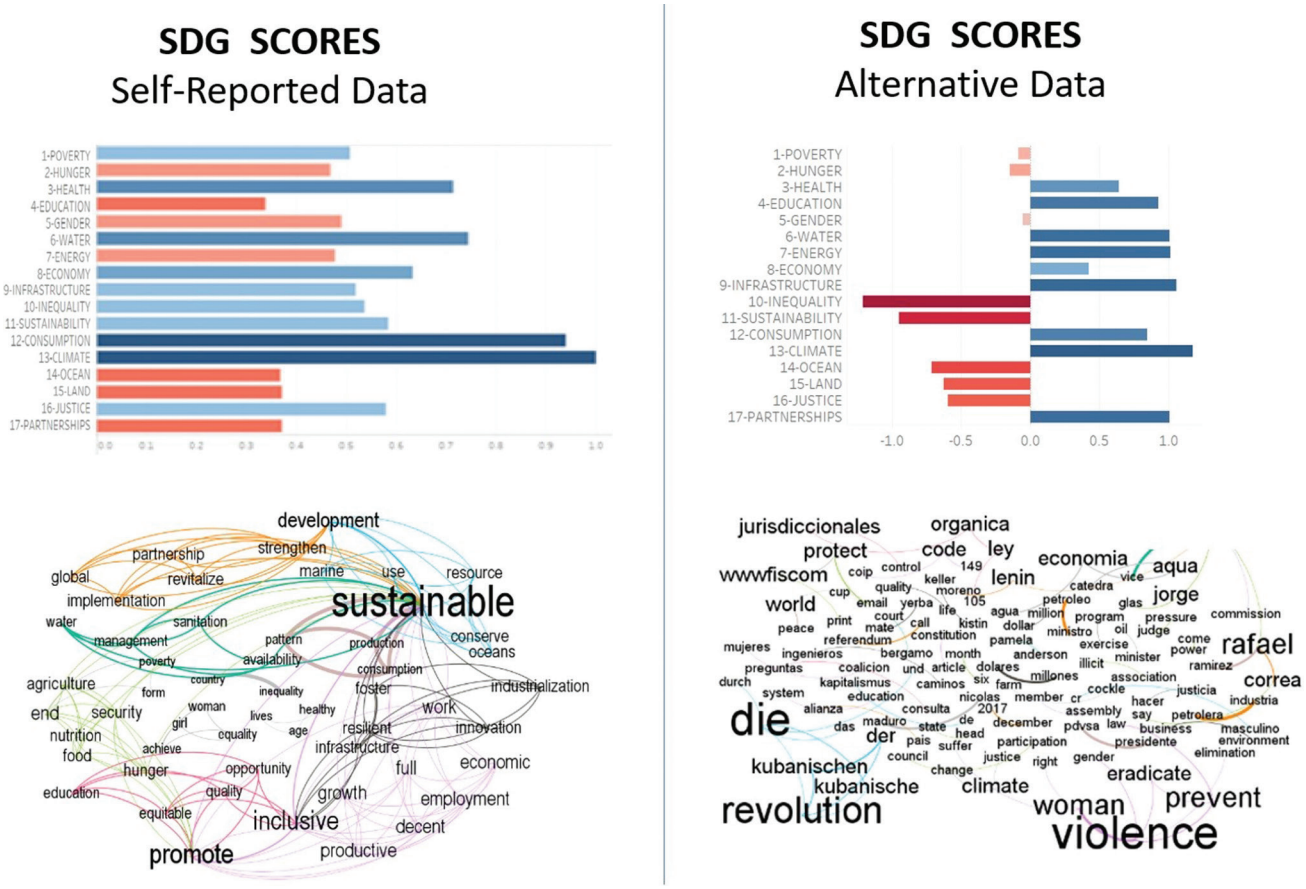


Figure 2: Single stock: Unstructured data
Source: Global AI Co.
Notes: ESG/SDG scores for stock screening, due diligence and integrity analysis.
ESG, environment, social and governance data; SDG, sustainable development goal.

SDGs contributions and SDG ratings in order to become more attractive for investors concerned with sustainable investments and impact investing. This can also provide increased transparency for investor engagement strategies and improve the investment process and enhance asset owner's engagement strategies by helping investors identify negative issues that might not have been reported by the company in a transparent manner. Moreover, Big Data can make this information available on a daily basis for investors, governments and all stakeholders — not just annually when a firm reports an unaudited sustainability report. Thus, a Big Data approach eliminates self-reporting bias and 'greenwashing' and can show which firms are effectively having a positive SDG footprint. Of course, not all news and public information is reliable. Preventive measures using advanced statistical techniques as well as performing extensive manual verification of data, if necessary, control for relying on publicly available information, such as newspaper articles, which may be fake news, or commemorate negative events from the past, which could, if left unanalysed, potentially lead to biased scores.

BIG DATA'S BROADER APPLICATION: IMPACT INVESTING AND PROVIDING THE NEEDED LARGE-SCALE SOLUTIONS FOR LARGE-SCALE PROBLEMS

In order to close the financing gap to achieve the UN SDGs by 2030, we need to mobilise capital

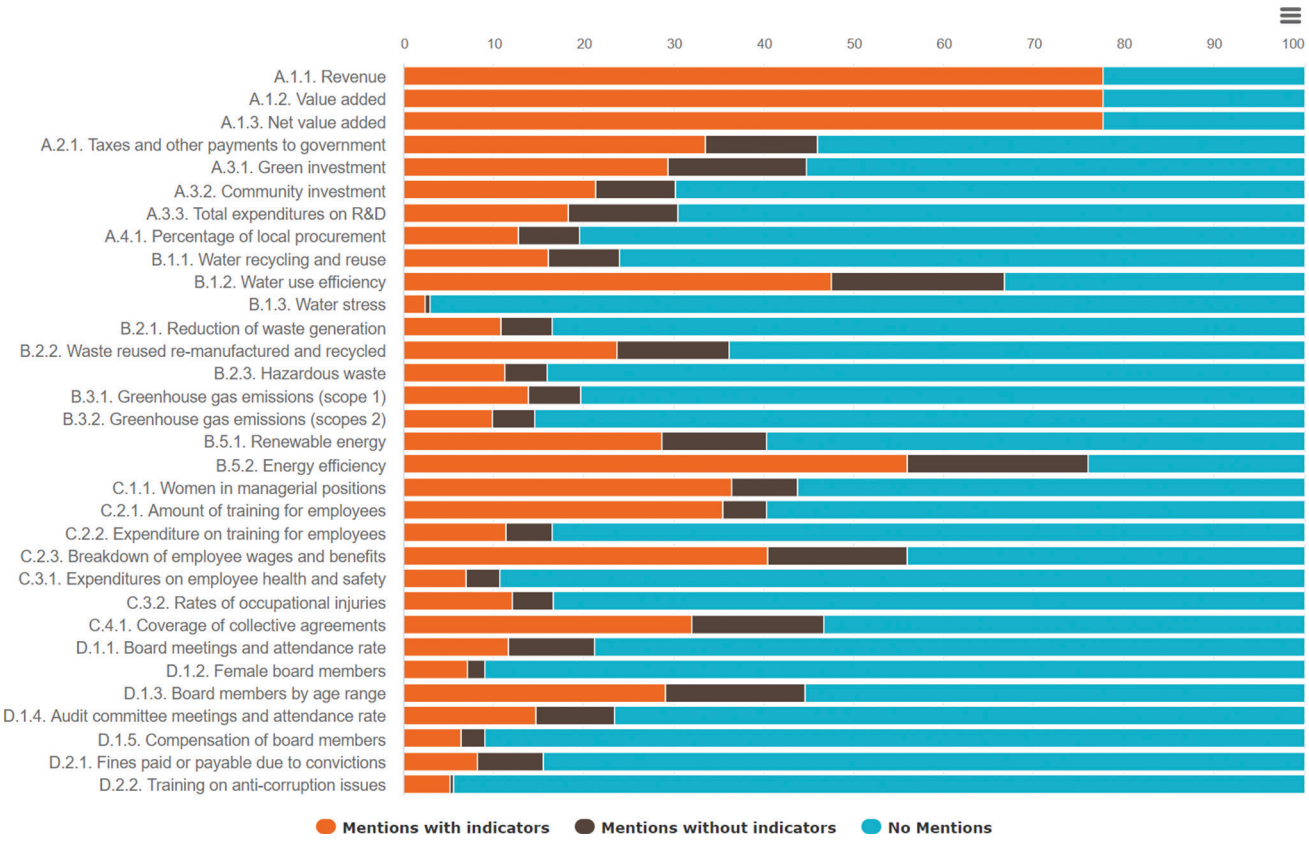
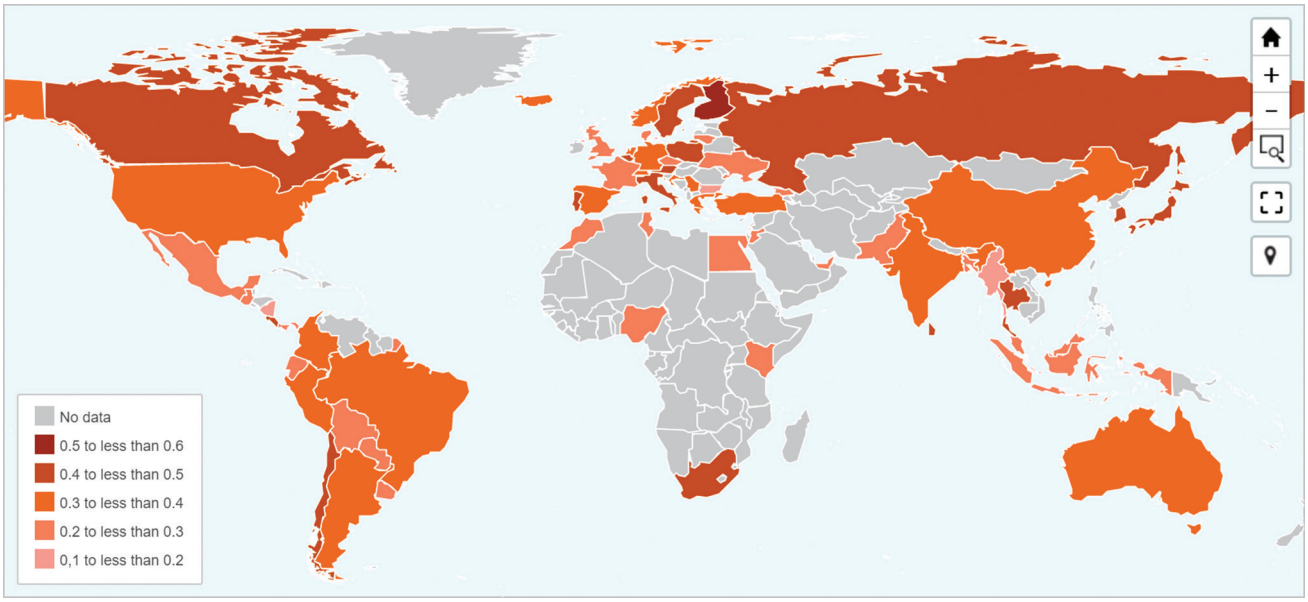


Figure 3: Compliance with sustainability reporting by UNCTAD core indicators (percentage)
Source: Global AI Co. (see SDG Pulse, 'Signs of a greening economy')[10]

on a large scale to where it is needed most. Key to the ability to achieve this shared goal is having consistent, quality data so that the private sector can play its role in supplying that capital. In order to track progress towards achieving the 169 targets of the SDGs, one needs to be able to measure performance metrics — the 231 unique statistical indicators — which are to be produced by every country in the world to benchmark progress towards the SDGs.

That is a daunting task especially considering that many developing countries, those which need the private sector capital the most in order to eliminate poverty and achieve sustainable economic growth, do not have the capacity to measure progress towards these goals. As the chief statistician of UNCTAD has put it, 'To Keep Track of the SDGs, We Need a Data Revolution[9] otherwise where will we get the funding in an "environment of faltering multilateralism" and persistent underfunding of Official Development Aid?' While estimates vary, at least US$1bn per year will be needed just to measure the progress towards achieving the SDGs, which far exceeds the existing funding. Therefore, the chief statistician puts forward a question as to 'why not introduce a mechanism to certify unofficial statistical indicators as official?' More importantly, he has astutely suggested that 'it is time to harness the work and intellectual creativity of those outside the official statistics tent'. This is where Big Data is playing an important role towards helping to achieve the SDGs.

Global AI Co. is using Big Data to provide statistical indicators and performance metrics data to the United Nations to benchmark progress towards achieving the SDGs and is producing the SDG footprint of the private sector at the regional and global level for policy purposes as shown in UNCTAD's SDG Pulse publication.[10] This is the first time AI will be used at the national policy level.



Map 1: Compliance with sustainability reporting, country averages, March 2019 (percentage)
Source: Global AI Co. (see SDG Pulse, 'Signs of a greening economy')[10]

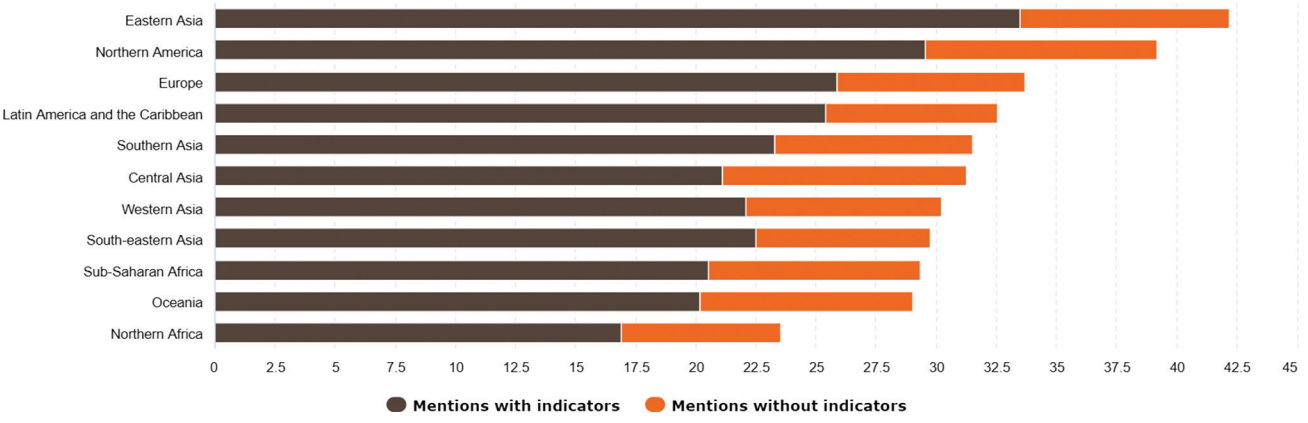


Figure 4: Compliance with sustainability reporting, regional averages (percentage)
Source: Global AI Co. (see SDG Pulse, 'Signs of a greening economy')[10]

Finally, for investors and companies alike, such measured SDG footprints can help quantify how investing in SDGs contributes to long-term investment performance. Building on an institutional investment framework, which incorporates and measures the net SDG impact of public and private entities and prices their long-term effects as externalities can then incentivise public corporations and investors to mobilise capital towards the SDGs at the scale needed, and ultimately contribute to long-term economic growth.

CONCLUSION

Big Data can help overcome data challenges by not restricting data only to self-reported documents from companies but by enhancing those 'structured data' with 'unstructured data', thus extracting

information from hundreds of thousands of sources from around the world. Publicly available sources include news, social media, government reports, blogs, twitter, industry-specific publications, NGOs, among others. Big Data developed through cutting-edge statistical models, AI and NLP covering dozens of languages, provides the solution for ESG/SDG reporting, rankings and ratings, and can help combat self-reported bias and 'greenwashing'. Investors can better understand the underlying risks in corporate behaviour, and ultimately, make more sustainable investment decisions. Also, Big Data can help policy-makers see where and how countries are progressing towards achieving their national policy goals. Armed with this information, they can determine where capital may need to be mobilised, helping to inform global policy decisions so as to achieve the UN SDGs by 2030. Ultimately, Big Data can enhance investors' decisions about allocating their capital to where it is needed most to make the greatest impact.

References and Notes

1 Antoncic, M. (2019) 'Why sustainability? Because risk evolves and risk management should too', *Journal of Risk Management in Financial Institutions*, Vol. 12, No. 3, pp. 206–216(11), available at: https://www.ingentaconnect.com/content/hsp/jrmfi/2019/00000012/00000003/art00002 (accessed 3rd March, 2020).

2 Sustainable Development Goals, United Nations Knowledge Platform, 'Sustainable development goals', available at: https://sustainabledevelopment.un.org/?menu=1300 (accessed 3rd March, 2020).

3 United Nations 'Global Investors for Sustainable Development Alliance' (p. 5), available at: https://www.un.org/esa/ffd/wp-content/uploads/2019/10/GISD-1015.pdf (accessed 3rd March, 2020).

4 Business Roundtable (2019) 'Business roundtable redefines the purpose of a corporation to promote "An economy that serves all Americans"', 19th August, available at: https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans (accessed 3rd March, 2020).

5 ScienceDaily 'PCC report on climate change – 2007', available at: https://www.sciencedaily.com/terms/ipcc_fourth_assessment_report.htm (accessed 3rd March, 2020).

6 World Economic Forum (2020), 'Toward Common Metrics and Consistent Reporting of Sustainable Value Creation', White Paper, 22nd January, available at: https://www.weforum.org/whitepapers/toward-common-metrics-and-consistent-reporting-of-sustainable-value-creation (accessed 3rd March, 2020).

7 The Telegraph (2019) 'Is Tesla an ethical investment? Even experts can't agree', 7th November, available at: https://www.telegraph.co.uk/investing/shares/tesla-ethical-investment-even-experts-cant-agree/ (accessed 3rd March, 2020).

8 United Nations Conference on Trade and Development, 'Guidance on core indicators for entity reporting on contribution towards implementation of the Sustainable Development Goals', available at: https://unctad.org/en/pages/PublicationWebflyer.aspx?publicationid=2469 (accessed 3rd March, 2020).

9 World Economic Forum, 'To keep track of the SDGs, we need a data revolution', available at: https://www.weforum.org/agenda/2019/01/its-time-for-a-data-revolution/ (accessed 3rd March, 2020).

10 SDG Pulse, 'Signs of a greening economy?, Businesses striving to close large gaps in sustainability reporting, Figures 5 and 6 and Map 2', available at: https://sdgpulse.unctad.org/ sustainability/ (accessed 3rd March, 2020).