

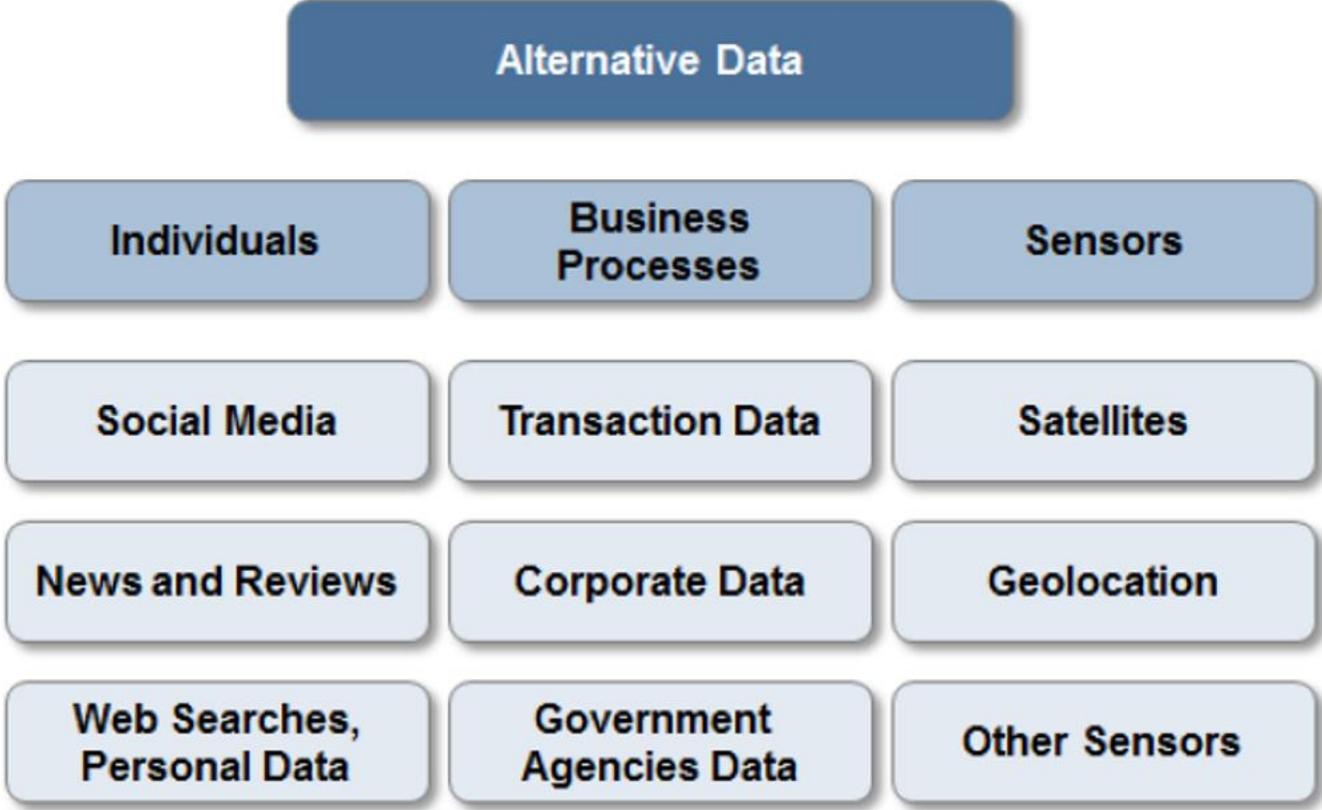


# **Big Data and AI-driven Approaches for SDG Nowcasting**

Richard V. Rothenberg

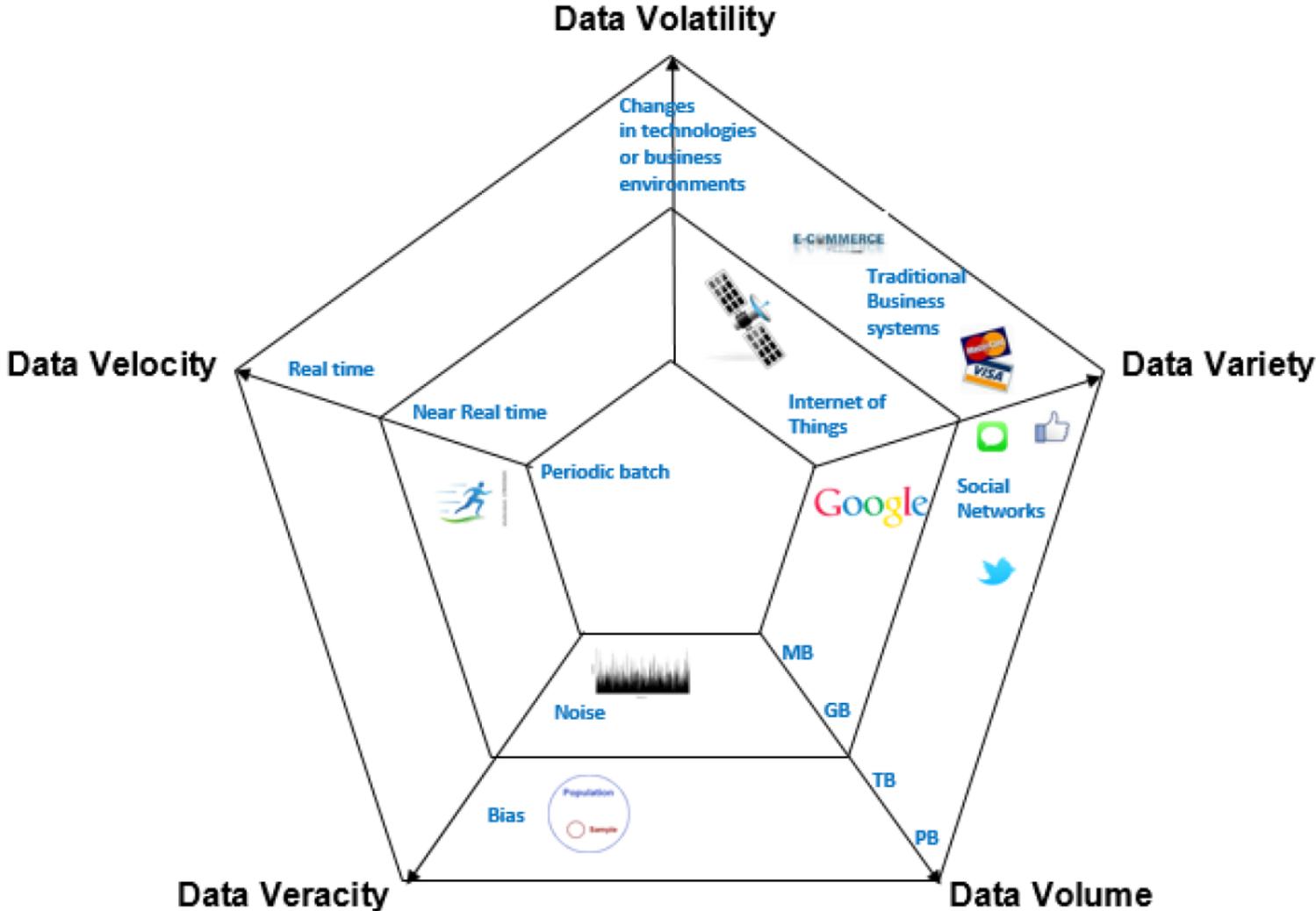
Executive Director, Global A.I. Corporation; Research  
Affiliate, Lawrence Berkeley National Laboratory

# Classification of Alternative Data Sources



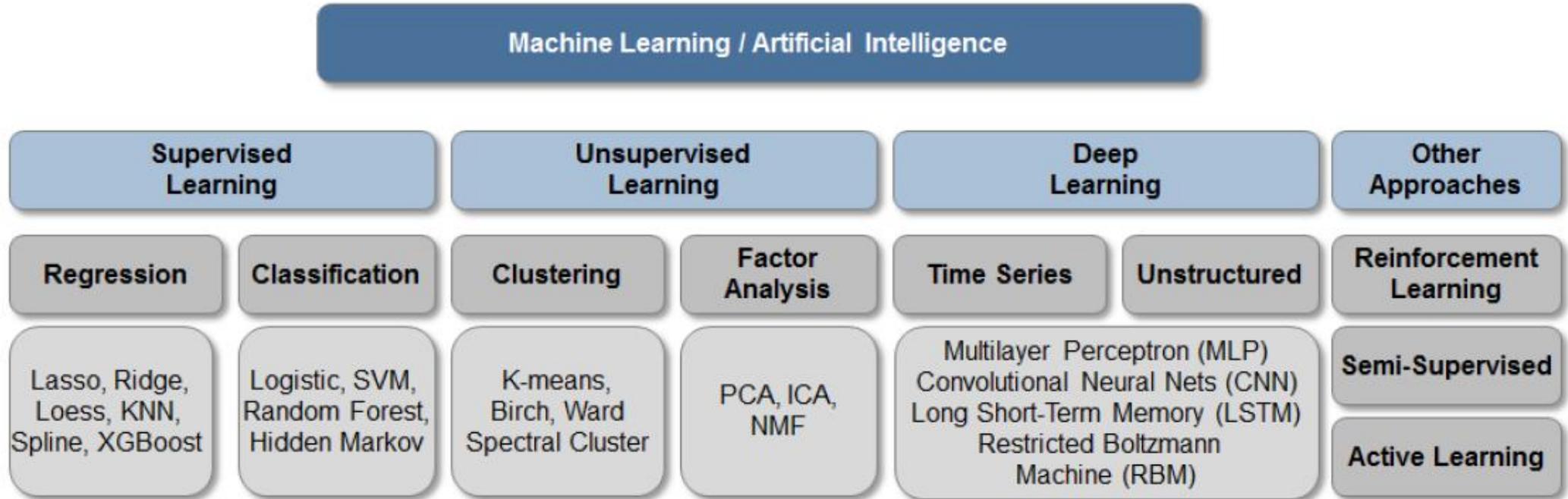
Source: J.P.Morgan Macro QDS.

# Alternative Data Features

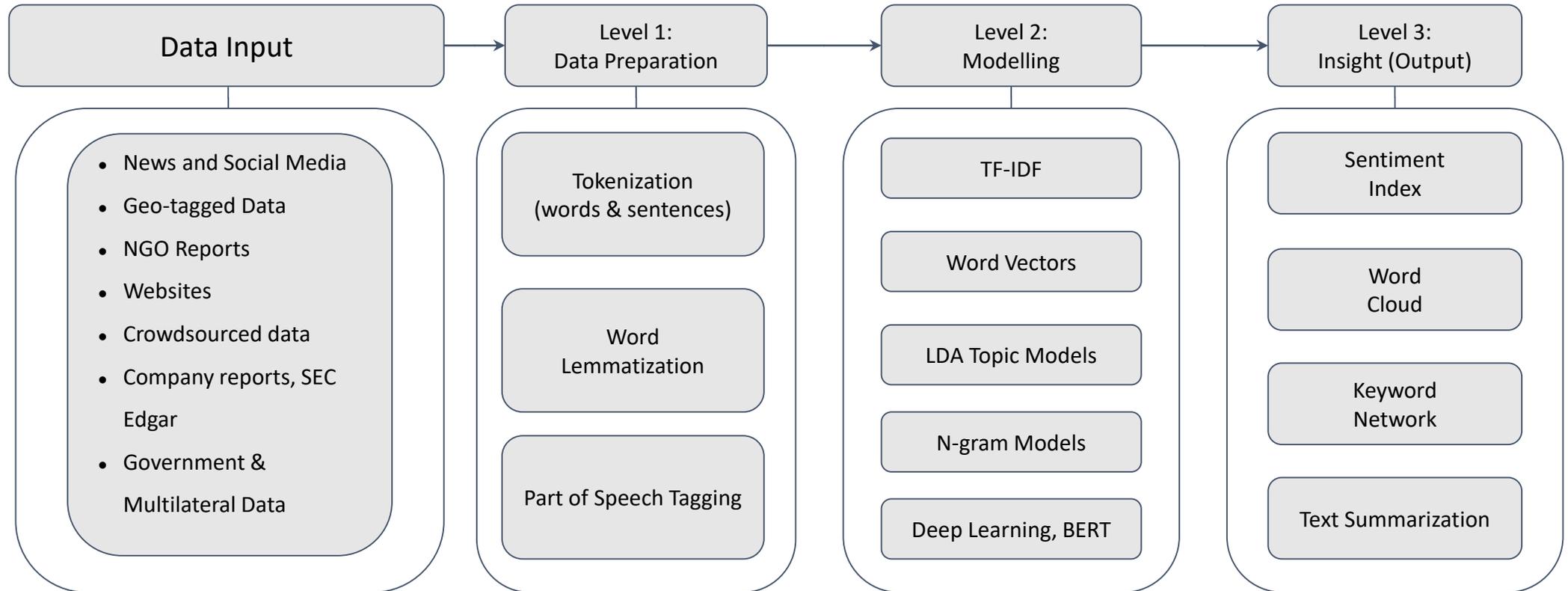
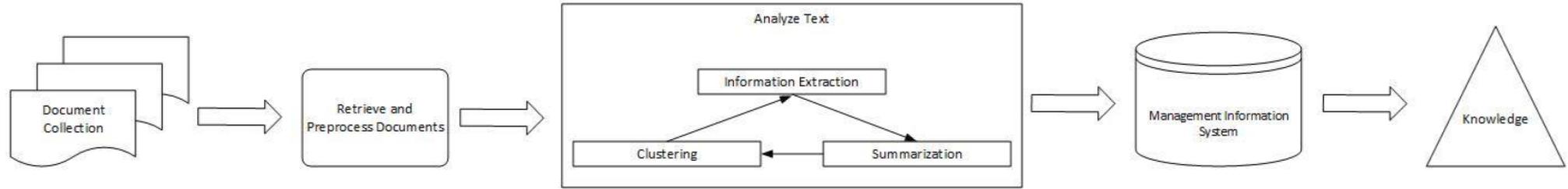


Based on Doug Laney, 2001

# Classification of Machine Learning Techniques



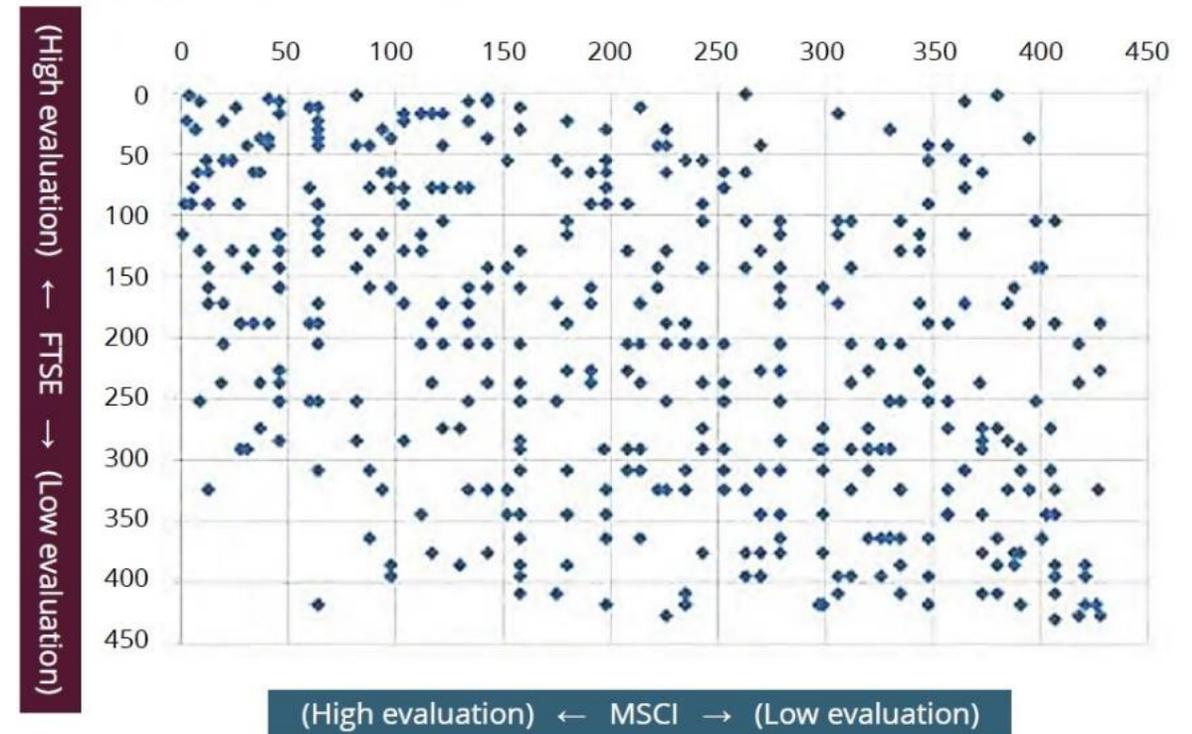
# Natural Language Processing Analysis



## About ESG

- Environmental, Social and Governance (ESG) is used to screen investments
- Based largely self-reported data
- the lack of ESG standards and metrics results in significant 'green-washing' and data biases.
- ESG metrics are updated infrequently, typically on an annual basis.
- Due to the lack of agreed ESG standards major discrepancies exist across company ESG ratings
- This has led to significant noise and a lack of useful ESG data for investment purposes.

Comparison of ESG Scores from FTSE vs MSCI



Source: CLSA, GPIF

## About the SDGs

The SDGs are emerging as the new standard to measure the sustainability footprint of Companies, Countries, and other investable assets.

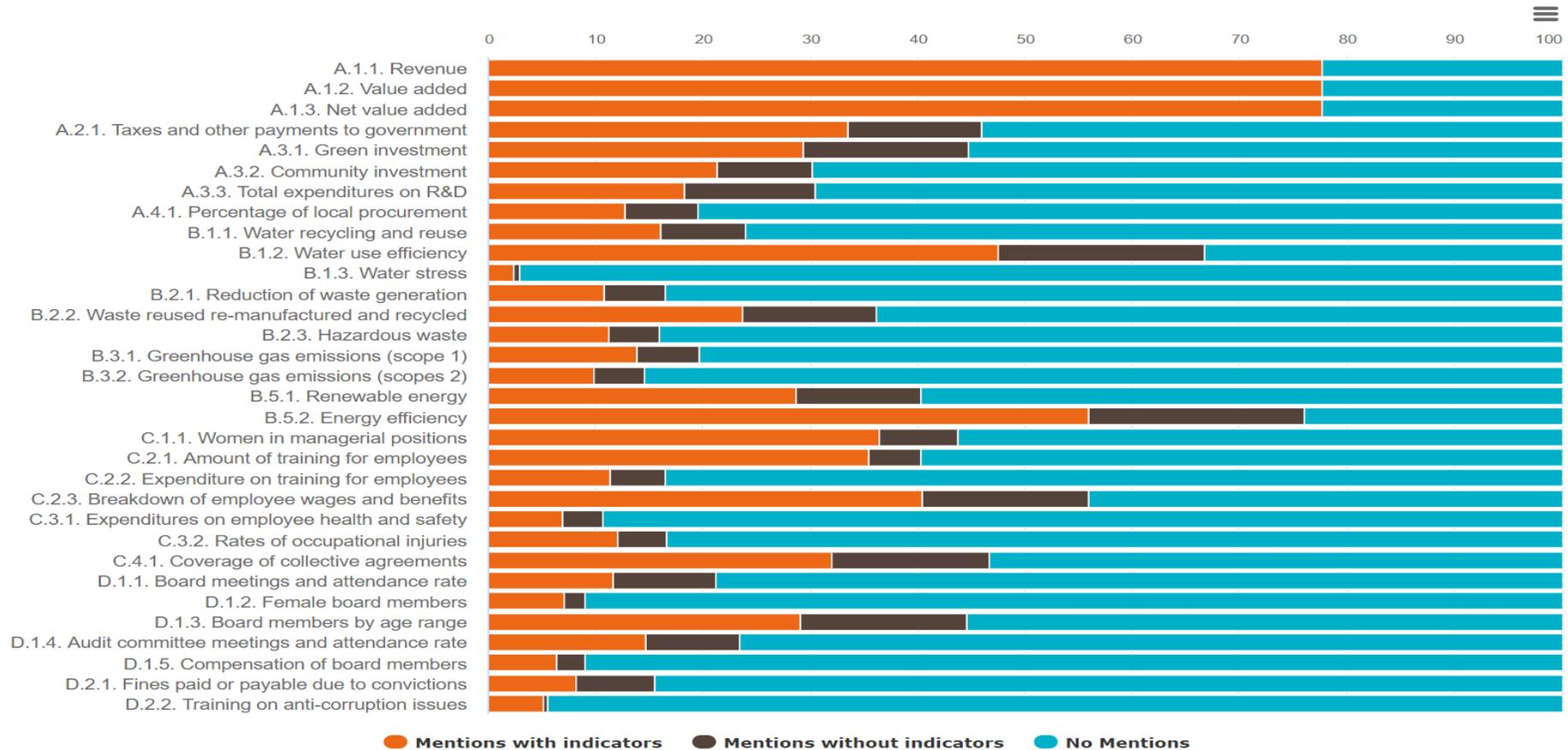
According to the UN Principles for Responsible Investment, the SDG's relevance to responsible investors can be summarized in 5 categories:

1. The SDGs are a critical part of investor's Fiduciary Duty
2. Macro Risks: the SDGs are an unavoidable consideration for "Universal Owners".
3. Macro Opportunities: the SDGs will drive Global Economic Growth.
4. Micro Risks: the SDGs as a Risk Framework
5. Micro Opportunities: the SDGs as a Capital Allocation Guide.



- The UNCTAD-ISAR is the Intergovernmental Working Group of Experts on International Standards of Accounting and Reporting (ISAR), the United Nations focal point on accounting and corporate governance matters, as well as sustainability standards for companies.
- Our UNCTAD Global Core Indicators (GCI) Ratings are based on the UNGC-ISAR set of indicators which have resulted from several years of multi-stakeholder discussions among Governments, leading Regulators and Standard-setting agencies such as GRI, FRSB, among others. These indicators incorporate the G20's FSB Task Force on Climate Disclosure (TFCD) indicators, among others.
- The GCI enable availability of comparable indicators at a company level on the rational use of resources such as water, energy, land; on emissions and waste reduction; good governance, human resource development and gender equality.
- Consistent with financial reporting requirements and in alignment with the SDG macro indicators on the use of financial, natural and human resources at a national level
- Core SDG indicators for companies are instrumental for measuring the only SDG target where the private business sector is mentioned: 12.6.1. "number of companies publishing sustainability reports", which UNCTAD is developing jointly with UN Environment as co-custodians of this indicator.

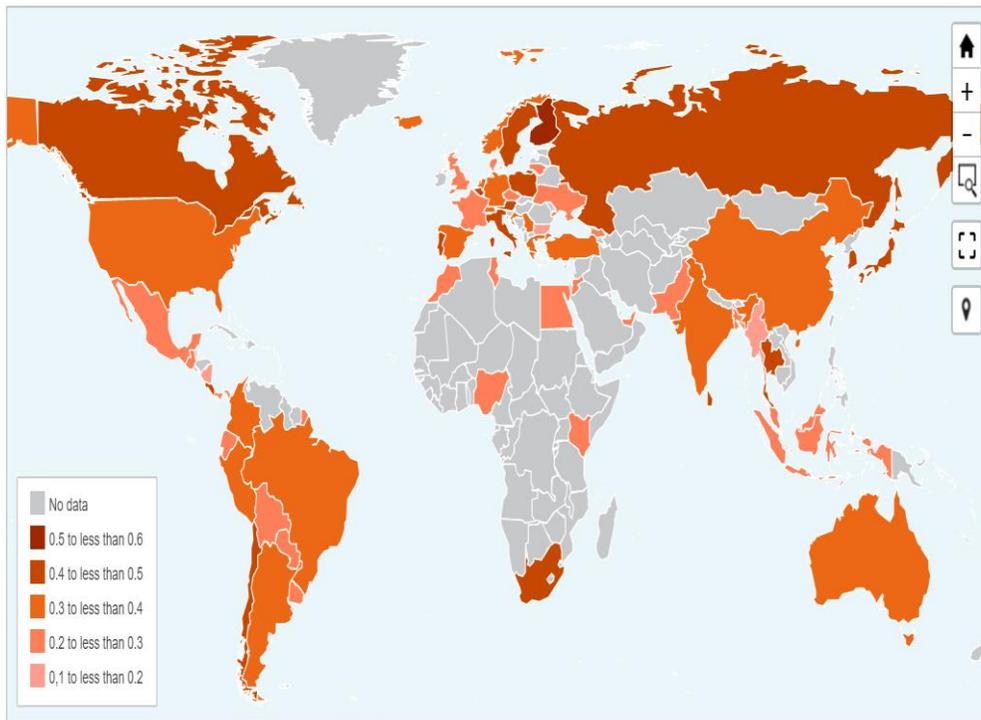
**Figure 5. Compliance with sustainability reporting by UNCTAD Core Indicators**  
(Percentage)



Source: Global AI Corporation with data from United Nations Global Compact (2019) and UNCTAD (2018).

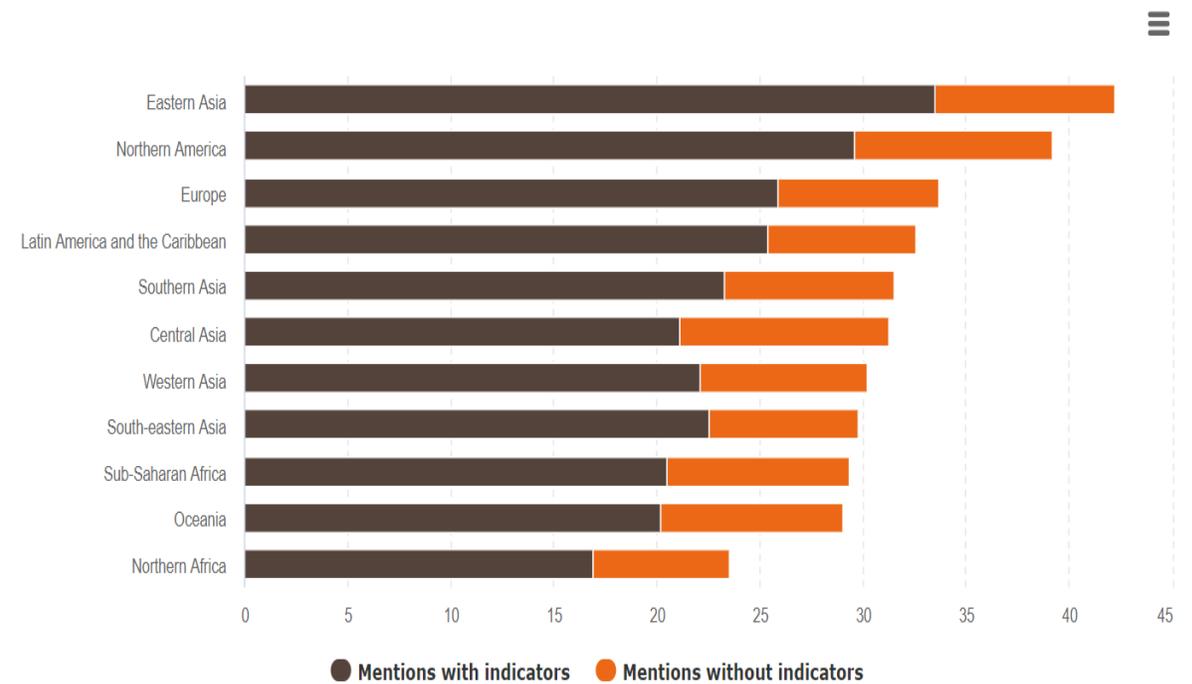
<https://sdgpulse.unctad.org/sustainability/>

**Map 2. Compliance with sustainability reporting, country averages, March 2019**  
(Percentage)



Source: Global AI Corporation with data from United Nations Global Compact (2019) and UNCTAD (2018).  
Notes: Countries with less than five reports available in the United Nations Global Compact database were excluded.

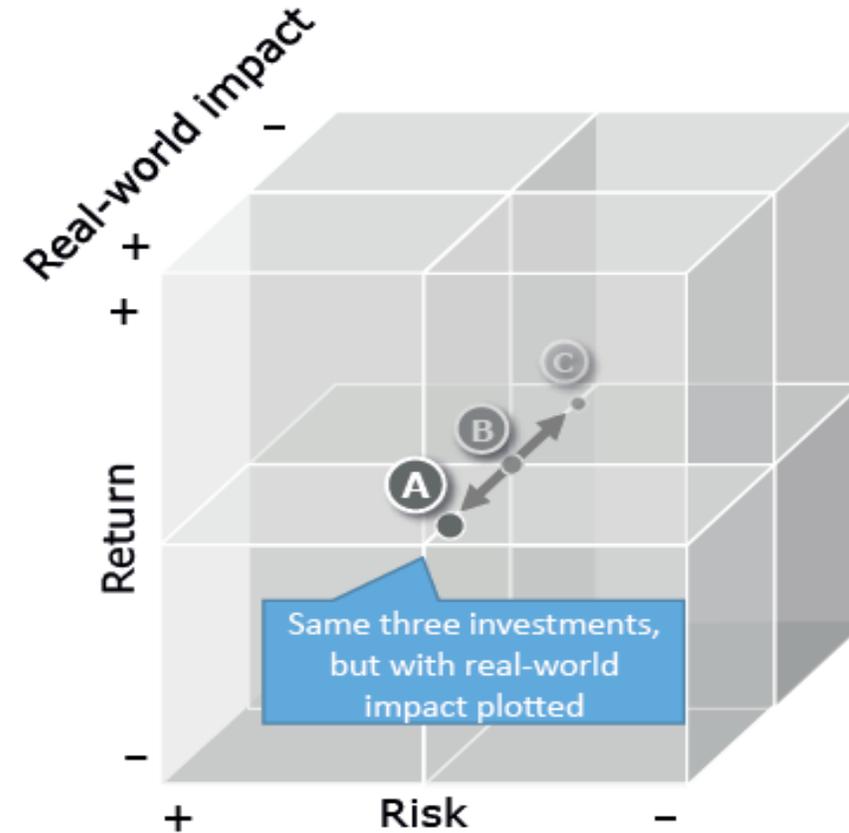
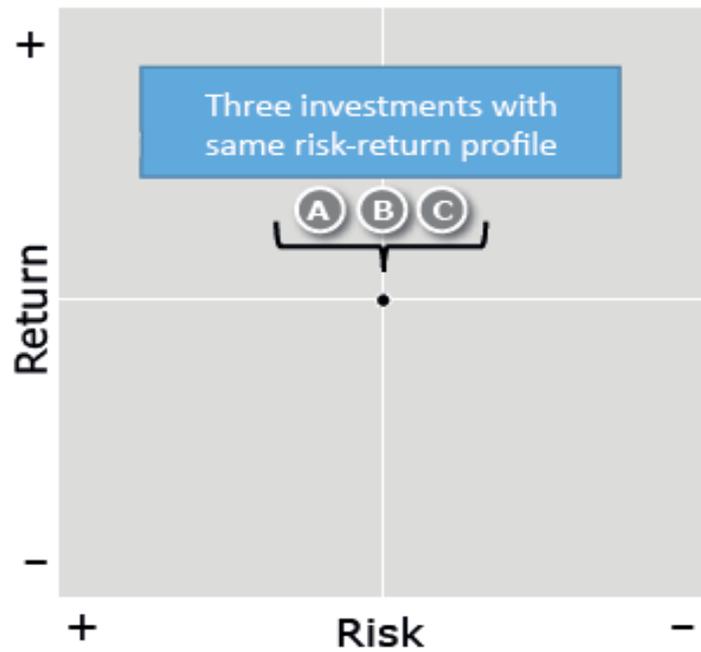
**Figure 6. Compliance with sustainability reporting, regional averages**  
(Percentage)



Source: Global AI Corporation with data from United Nations Global Compact (2019) and UNCTAD (2018).

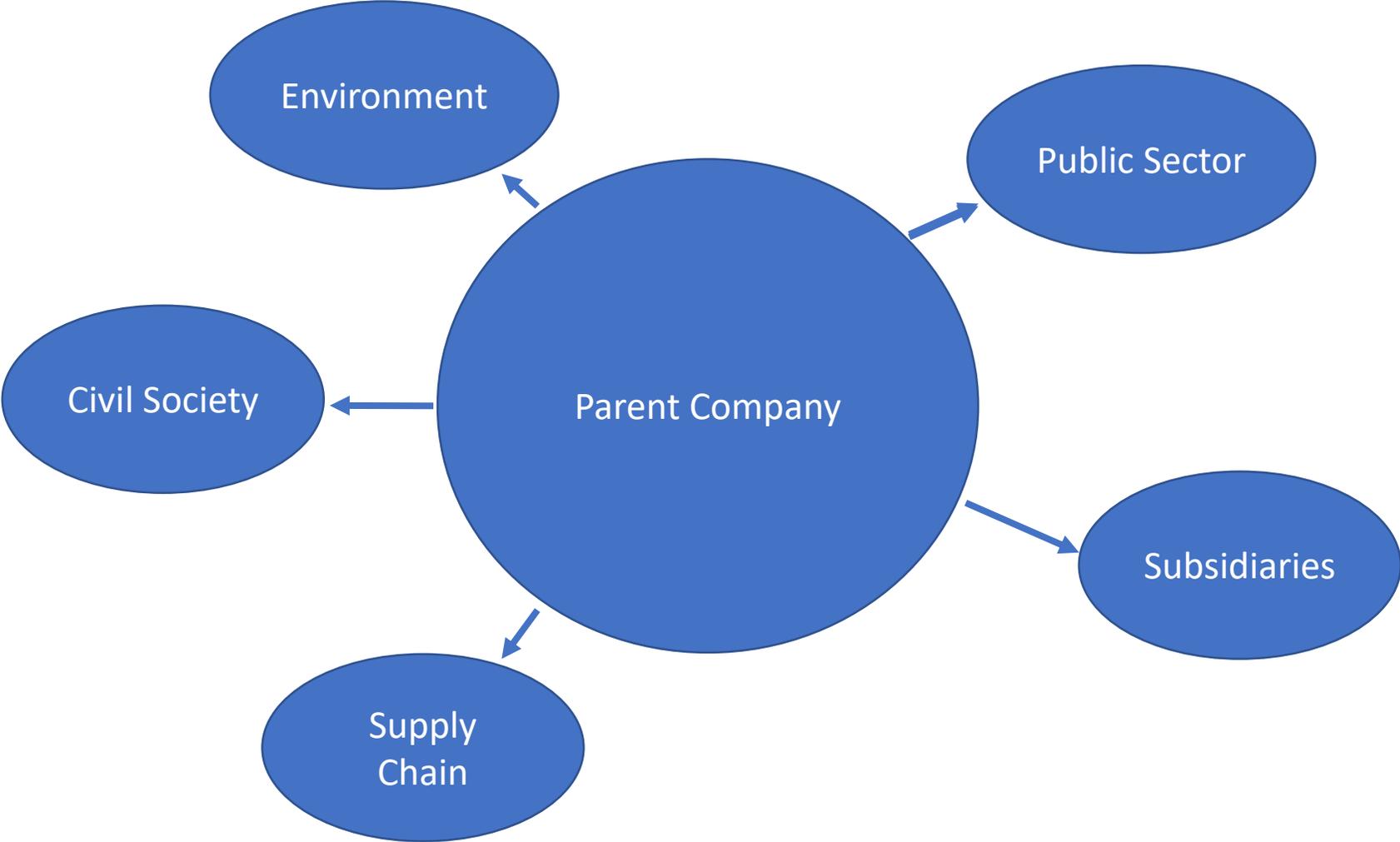
<https://sdgpulse.unctad.org/sustainability/>

# Risk-Return-Impact Framework



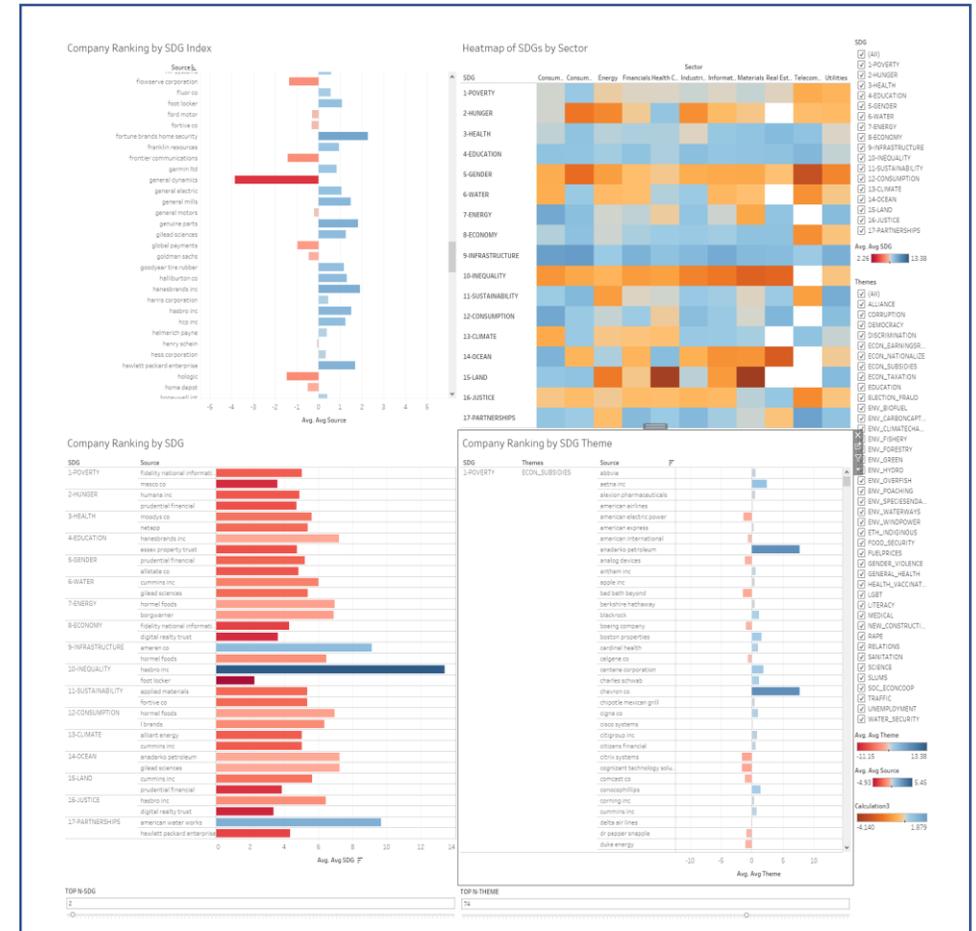
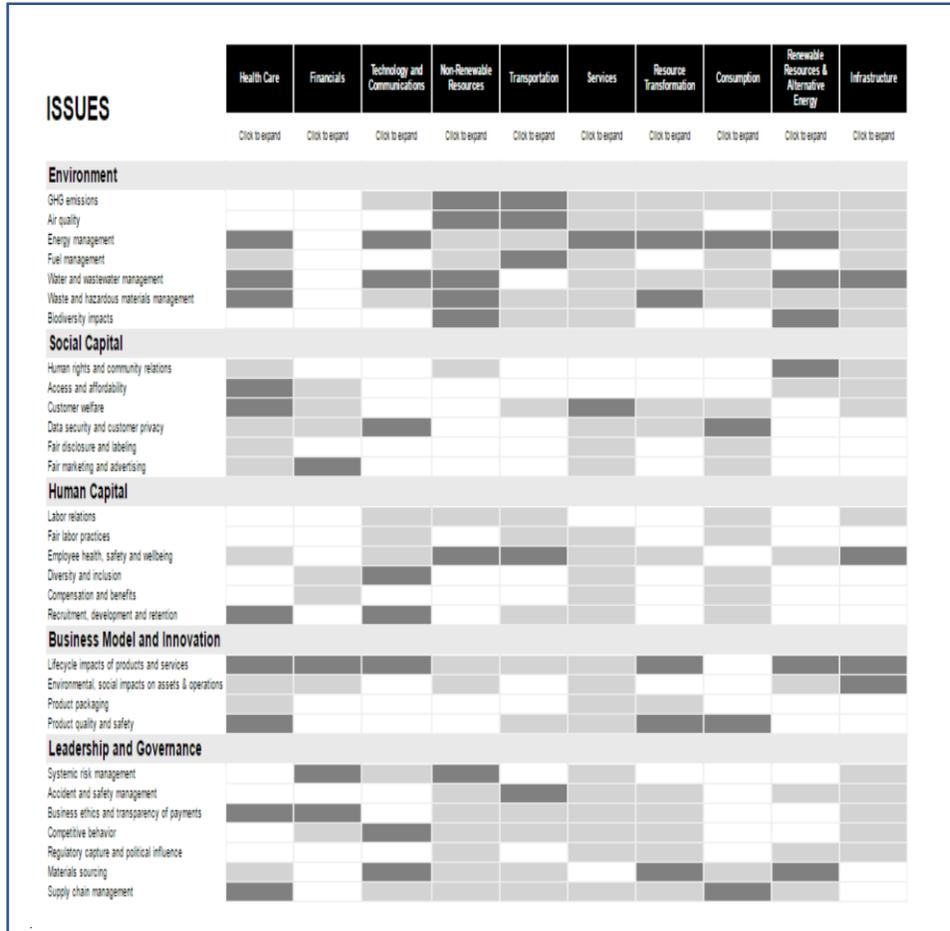
Source: UN PRI, SDG Investment Case

# SDG Footprint: Entity Network Mapping



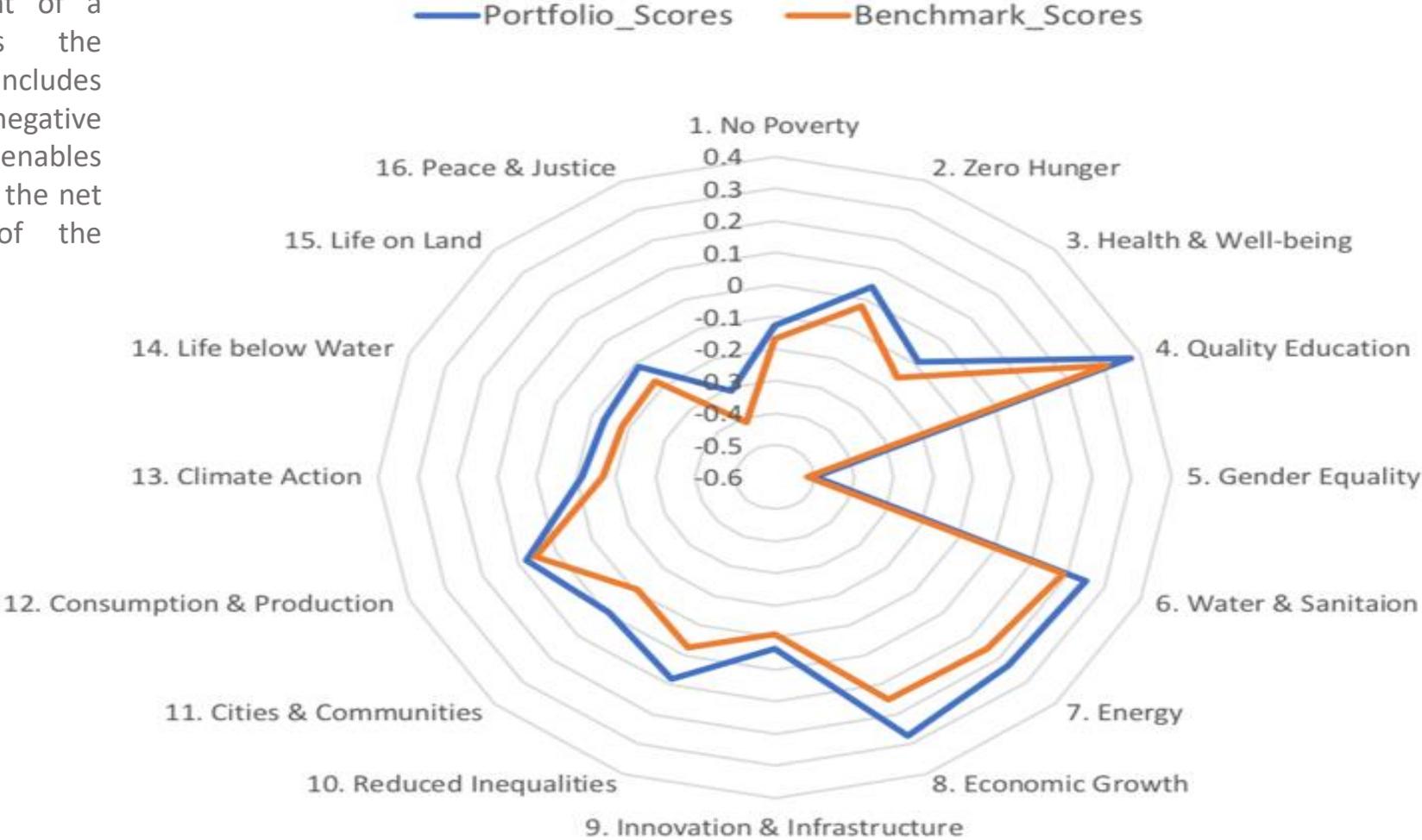


# SDG Footprint: Materiality Analysis



# SDG Footprint: Portfolio vs. Benchmark

The following graph shows the SDG footprint of a portfolio versus the benchmark. This includes both positive and negative SDG scores and enables the assessment of the net SDG footprint of the portfolio.

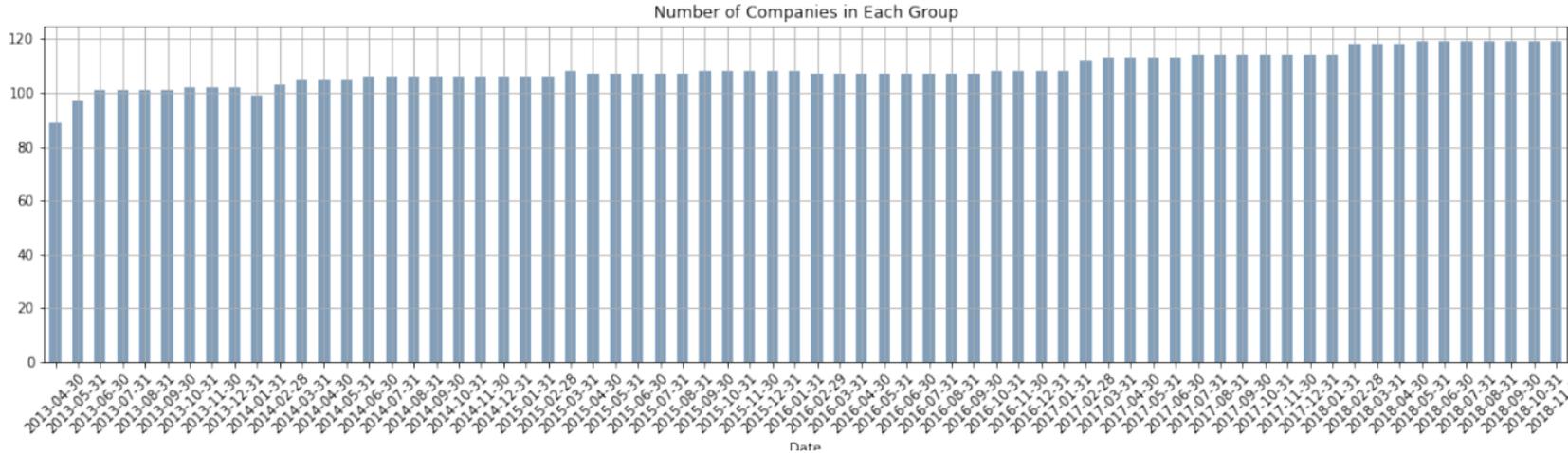
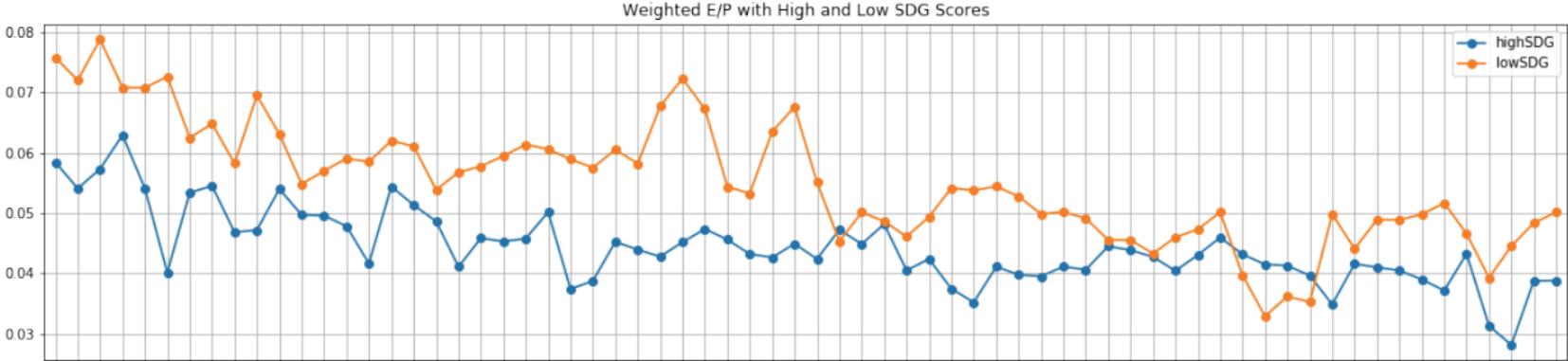


# SDG Footprint: SDG Scores & Company Valuation



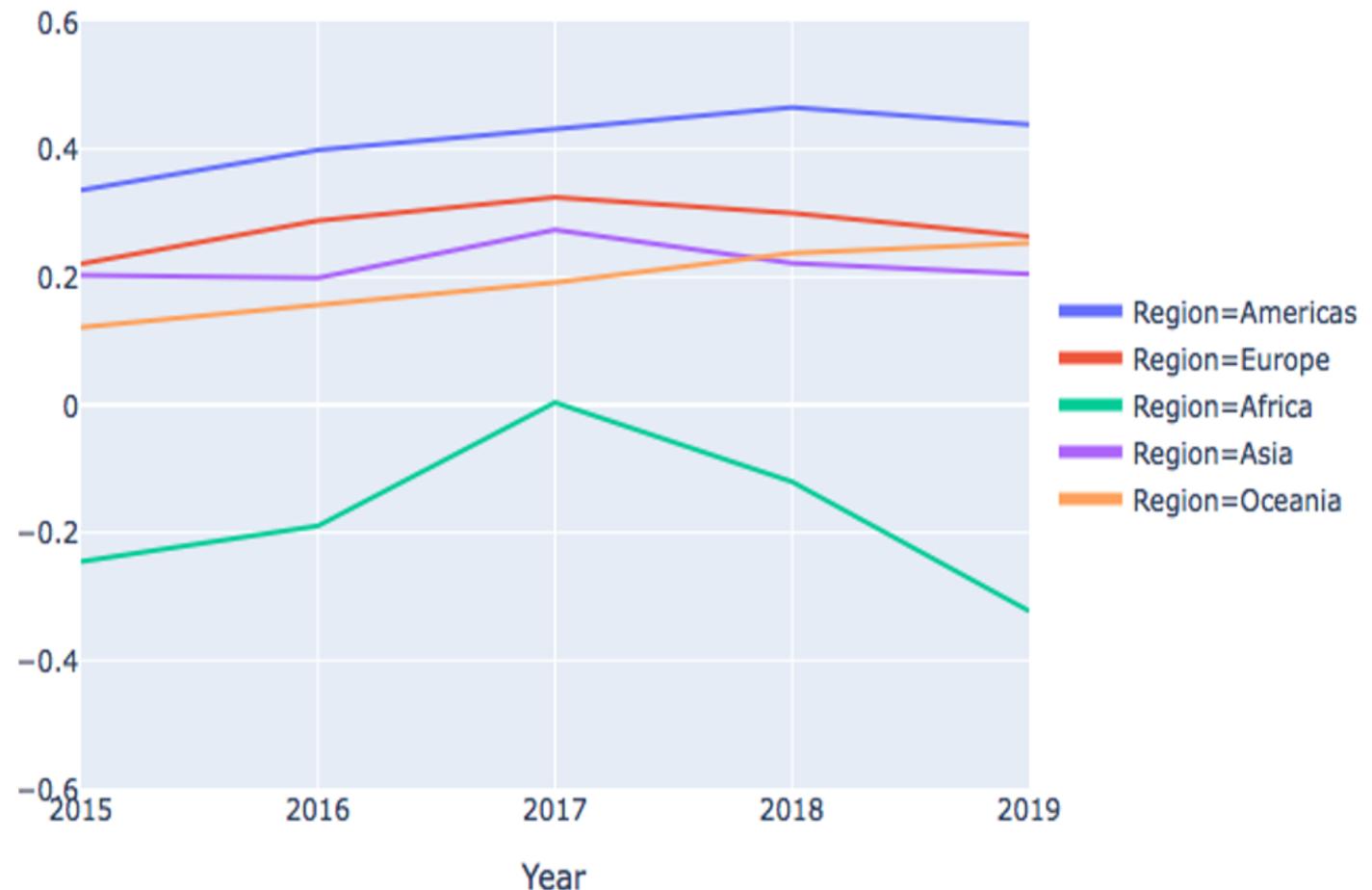
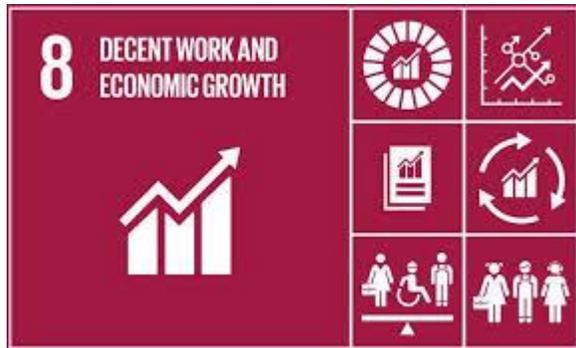
There is a statistically significant relationship between our SDG scores and companies valuation and fundamental ratios. For example, the following graph shows that industry sectors with high SDG scores across all sectors tend to have higher valuations and lower cost of capital.

Earnings Yield  
E/P:



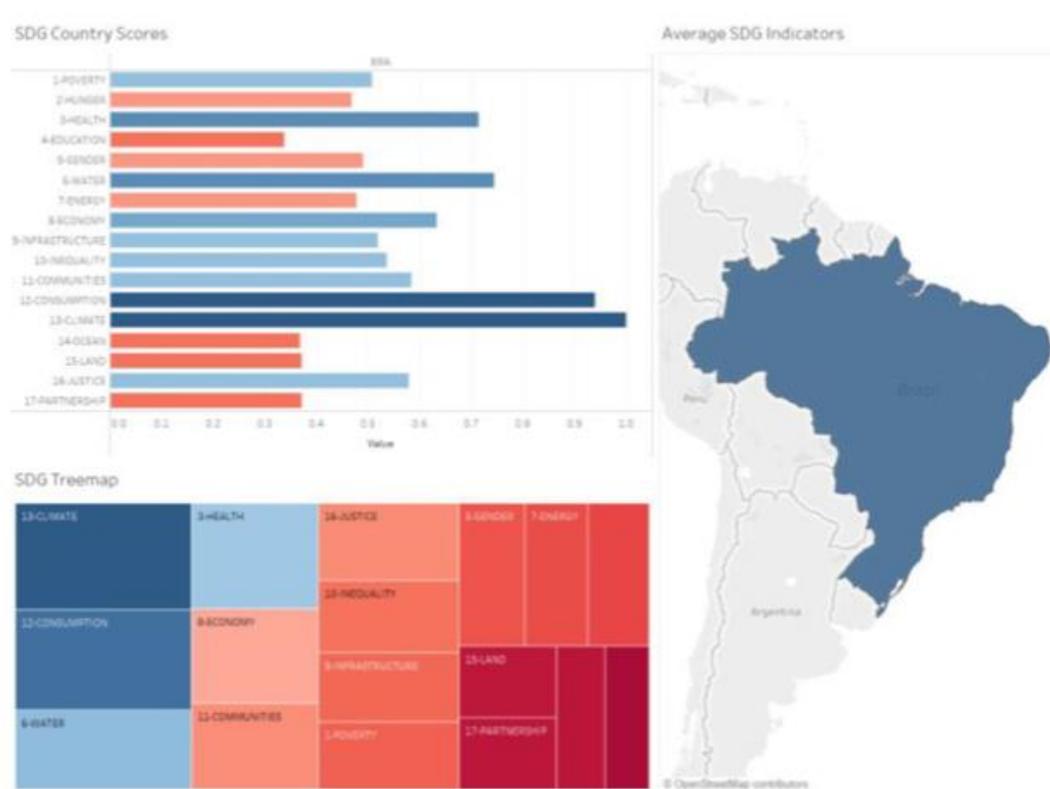
# Global SDG Private Sector Footprint: Regional Trends

This is the output of a global analysis of the SDG footprint of the private sector. The results incorporate data from 19,819 companies across Africa, Asia, Europe, Oceania and the Americas. The following illustrates the results for the SDG # 8- Decent Work and Economic Growth.



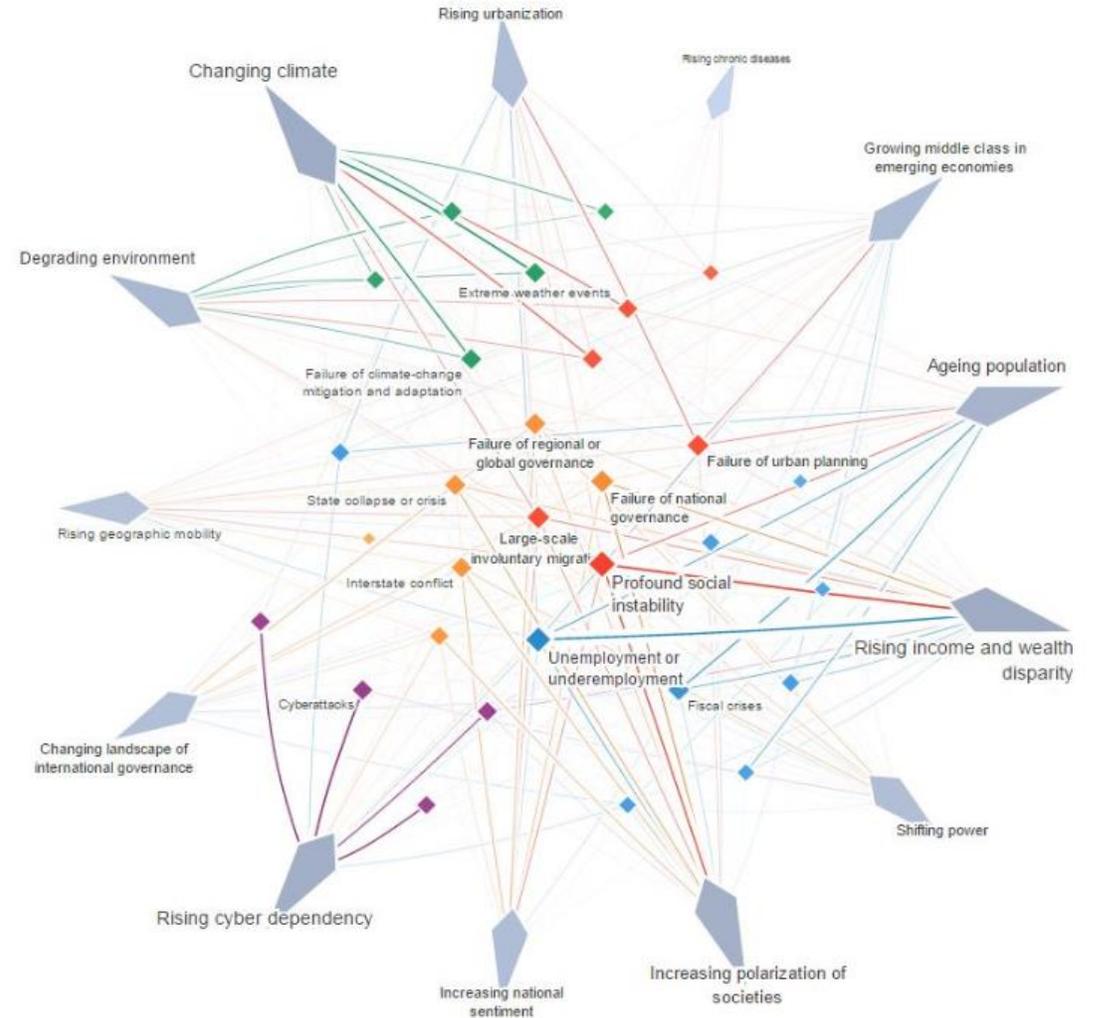
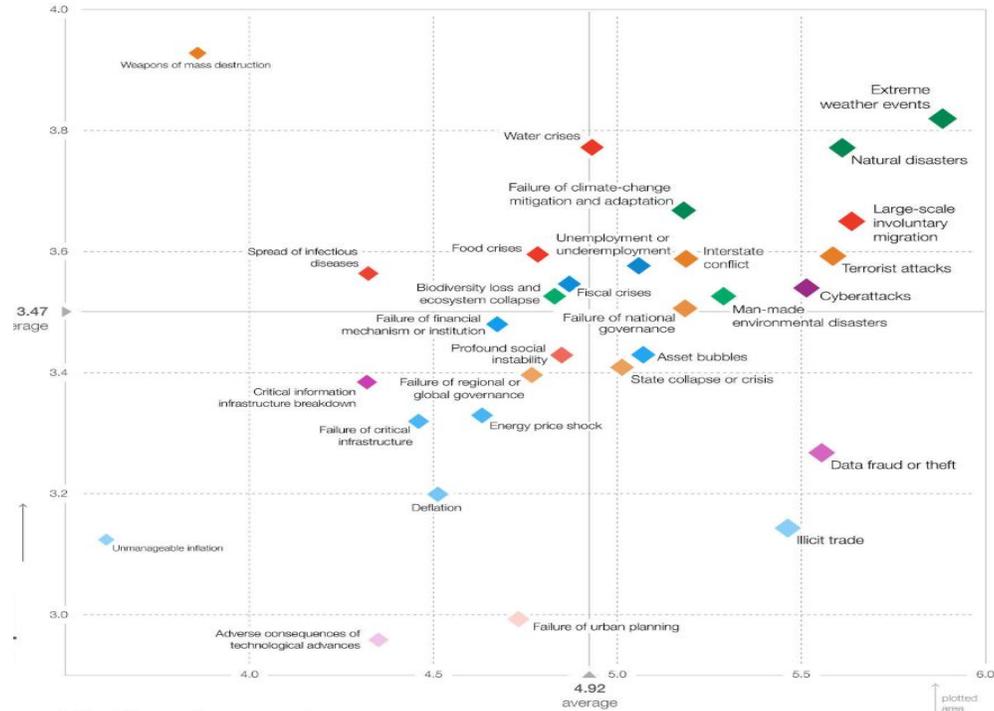
# Country SDG Footprint: Non-Financial Country Risk

Identify trends for each SDG at the Country level which serves as a proxy for non-financial Country Risk. This scores can be linked to asset price returns using liquid securities.



Symbol	ETF Name	Aspect				
		Economic Environ..	Geopolit..	Societal	Technolo..	
ADRE	Invesco BLDRS Emerging ..					
BBRC	Columbia Beyond BRICs E..					
BICK	First Trust BICK Index Fund					
CEY	VictoryShares Emerging ..					
CEZ	VictoryShares Emerging ..					
DBEM	Xtrackers MSCI Emerging ..					
DEM	WisdomTree Emerging M..					
DEMG	Xtrackers FTSE Emerging ..					
DGRE	WisdomTree Emerging M..					
DGS	WisdomTree Emerging M..					
DVYE	iShares Emerging Market..					
EDBI	Legg Mason Emerging Ma..					
EDIV	SPDR S&P Emerging Mark..					
EDOG	ALPS Emerging Sector Div..					
EEB	Invesco BRIC ETF					

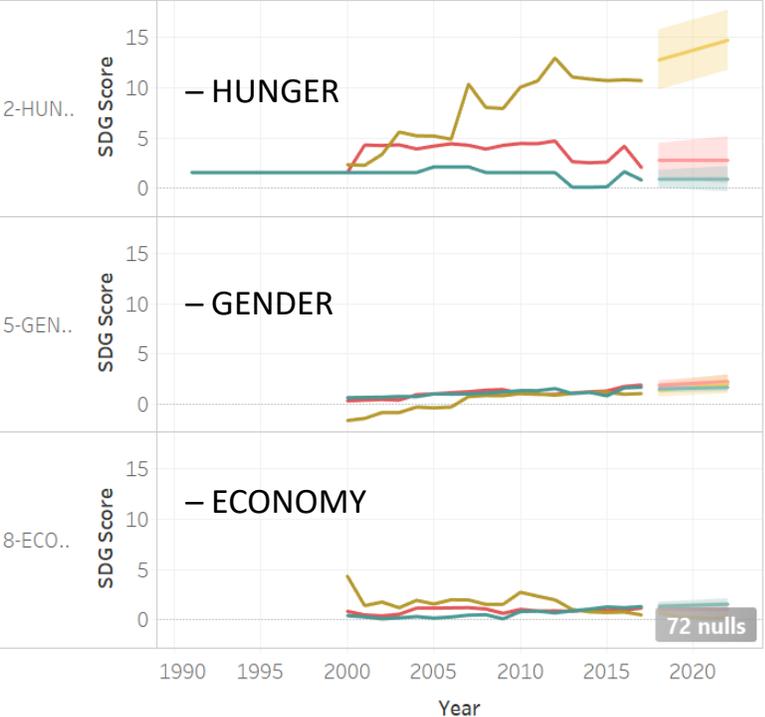
# WEF Global Risks & Inter-linkages



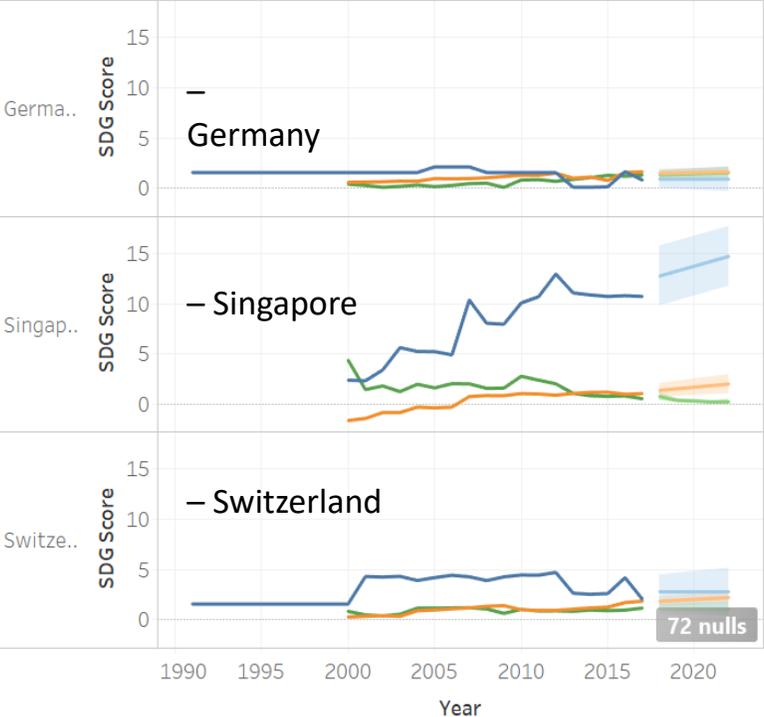
# SDG Country Scores: Time Series Analysis



Time Series Plot and Prediction of Germany, Singapore, Switzerland for Goal2-HUNGER, 5-GENDER, 8-ECONOMY



Time Series Plot and Prediction of Goal2-HUNGER, 5-GENDER, 8-ECONOMY for Germany, Singapore, Switzerland



Country or Area ...  
 (Multiple val... ▾

SDG Goal

- (All)
- 1-POVE...
- 2-HUNG...
- 3-HEAL...
- 4-EDUC...
- 5-GEND...
- 6-WATER
- 7-ENER...
- 8-ECON...
- 9-INFR...
- 10-INE...
- 11-CIT...

Country or Area ..

- Germany, Ac..
- Germany, Es..
- Singapore, A..
- Singapore, E..
- Switzerland,..
- Switzerland,..

SDG Goal, Forec..

- 2-HUNGER, ..
- 2-HUNGER, E..
- 5-GENDER, A..
- 5-GENDER, E..
- 8-ECONOMY,..
- 8-ECONOMY,..





Richard V. Rothenberg

Executive Director, Global A.I. Corporation; Research  
Affiliate, Lawrence Berkeley National Laboratory

[rrothenberg@globalai.co](mailto:rrothenberg@globalai.co)

+1(917)399-7840



# APPENDIX

# SDG Company Footprint: Introduction



## BACKGROUND

The lack of ESG standards and metrics based primarily on self-reported data by companies has led to significant 'green-washing' and data biases. A new alternative based on the intersection of Artificial Intelligence and the United Nations Sustainable Development Goals (SDGs) can help overcome existing shortcomings of ESG to measure the sustainability footprint of companies, including a more standardized taxonomy and the use of a large-scale unstructured data that can provide more comprehensive and timely insights.

There are significant practical challenges to quantify the SDG footprint of companies. One major issue is that in contrast to ESG approaches which only have 3 categories, the SDGs contain 17 categories with more than one hundred goals and indicators. Furthermore, the vast majority of data available on US and International companies is unstructured and highly fragmented. In addition, most of the data is not available in English but in local languages - particularly in Europe, Asia and Latin America - and this presents enormous challenges for institutional investors to extract, analyze and quantify complex, fragmented data associated with the SDGs.

Another issue with existing ESG metrics is that they are updated infrequently. This makes them largely irrelevant for investors who need to react quickly to emerging negative sustainability issues. The capacity of new technologies to quantify and track thousands of SDG factors and events globally in a more timely manner can contribute to make SDG indicators more relevant for investors and provide more up-to-date signals that can be used for both ESG and mainstream strategies, including tactical and strategic asset allocation, bottom-up equity selection based on SDG scores, long-short equity and other investment strategies.

For this purpose, it is important to leverage Big Data and Artificial Intelligence technologies to extract, process and analyze large-scale structured and unstructured data on SDG-related factors, which can then enable the integration of SDG factors into the decision-making of global investors. Typically, companies carry out voluntary reporting on their sustainability performance in order to assure their shareholders and investors of their compliance to regulations.

However, as more companies are wary of the adverse impact of negative sustainability performance on investor decisions, they may fail to disclose negative information. With regard to environmental issues, greenwashing, where companies use deceptive marketing to appear more eco-friendly, has been on the rise. Big Data enhances reported data with "alternative data" using artificial intelligence, 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.

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 significantly reduces self-reporting bias and 'greenwashing' and can show which firms are effectively having a positive SDG footprint. Of course, there are scenarios in which the technology can go wrong or provide imperfect information; relying on publicly available information such as newspaper articles, may lead to false or biased scores, for example. Other issues include fake news, articles that commemorate negative events from the past, major discrepancies between reported and third-party data, among others. For these reasons, it is necessary to perform extensive manual verification of data to evaluate if the analysis corresponds to reality and implement preventive measures.

The SDG footprint can show how companies can have an either positive or negative net impact on SDGs and potentially reveal hidden risks. This creates incentives for corporations to quantify and increase their net SDGs contributions and SDG ratings in order to become more attractive for investors concerned with sustainable investments, which control trillions in assets under management. It can also provide increased transparency for investor engagement strategies. 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 incentivize public corporations and investors to mobilize capital towards the SDGs at the scale needed, and ultimately contribute to long-term economic growth.

## ABOUT GLOBAL AI DATA

Global AI technology uses state-of-the-art Big Data and Artificial Intelligence techniques to access massive amounts of structured and unstructured data from more than 100,000 sources across more than 150 Countries and 60 languages to replace dated, slow and expensive manual processes used for sustainability and materiality analysis.

These new technologies can be used to mitigate 'SDG washing' by not restricting data to self-reported documents from companies and instead extracting data from tens of thousands of sources from around the world. Publicly available sources include news, social media, regulatory filings, government reports, blogs, twitter, industry-specific publications, sustainability reports, NGOs, among others.

For this report, we use Global AI's firm specific SDG scores and ratings. The company provides raw scores, a short-term and long-term rating. While we use the short-term ratings, the information is averaged over a year, so their measurement represents a relatively long-term measurement of SDG footprint. As a background to the Global AI scores, the company extracts, filters and cleans massive amounts both structured and unstructured data, including self-reported company data, news articles, blogs, NGO Reports, Social Media, etc. Specialized algorithms map the raw data to specific companies and associated entities such as subsidiaries, using different combinations of company names, abbreviations, tickers, ISINs and subsidiaries. Proprietary technology then ranks and filters content by relevance using domain-specific taxonomies based on the UN Sustainable Development Goals. Examples of SDG taxonomies include the Global Core Indicators (GCI), which resulted from extensive multi-stakeholder consultations led by the UNCTAD Intergovernmental Working Group of Experts on International Standards of Accounting and Reporting (ISAR); other examples include the United Nations Global Compact 10 Principles, which are divided in major categories such as human rights, labour rights, environment and anti-corruption.

The algorithms subsequently analyze the filtered content at a daily level, recording the number of relevant news items, providing a sentiment score per news item, which thus reflects both positive and negative SDG related issues, and also tracks volume and dispersion of sentiment across news items. This information is then aggregated into daily company specific scores, which are further aggregated in 7 day and 180 day ratings. The raw scores represent aggregate sentiment of the SDG data. The mapping from scores to ratings aggregates data from 7 days of information, uses statistics on the precision of the scores and the volume of the news sources, accommodates sparsity in the data and depends most heavily on recent information. Scores and ratings are available for each of the 17 SDGs and the system also provides an overall score, measuring the overall SDG footprint of a company. The ratings can be interpreted roughly as "z-scores", varying mostly between -1 and +1, and have a standard deviation of roughly 1.

The higher the score, the more positive the text is in relationship to each SDG, and vice versa. Thus, the sign represents positive or negative and the score indicates the degree of positivity and negativity. For example, for the SDG #5 (gender equality) the system would give a better score to a company that doubles the number of women on their board of directors from 20% to 40% than a large company that announces the hire of two female analysts. For the SDG #13 (climate action), a company would get a more negative score after a chemical spill that pollutes an entire ecosystem than a company that increases its carbon emissions by 5%. The scores are adjusted by Sector.

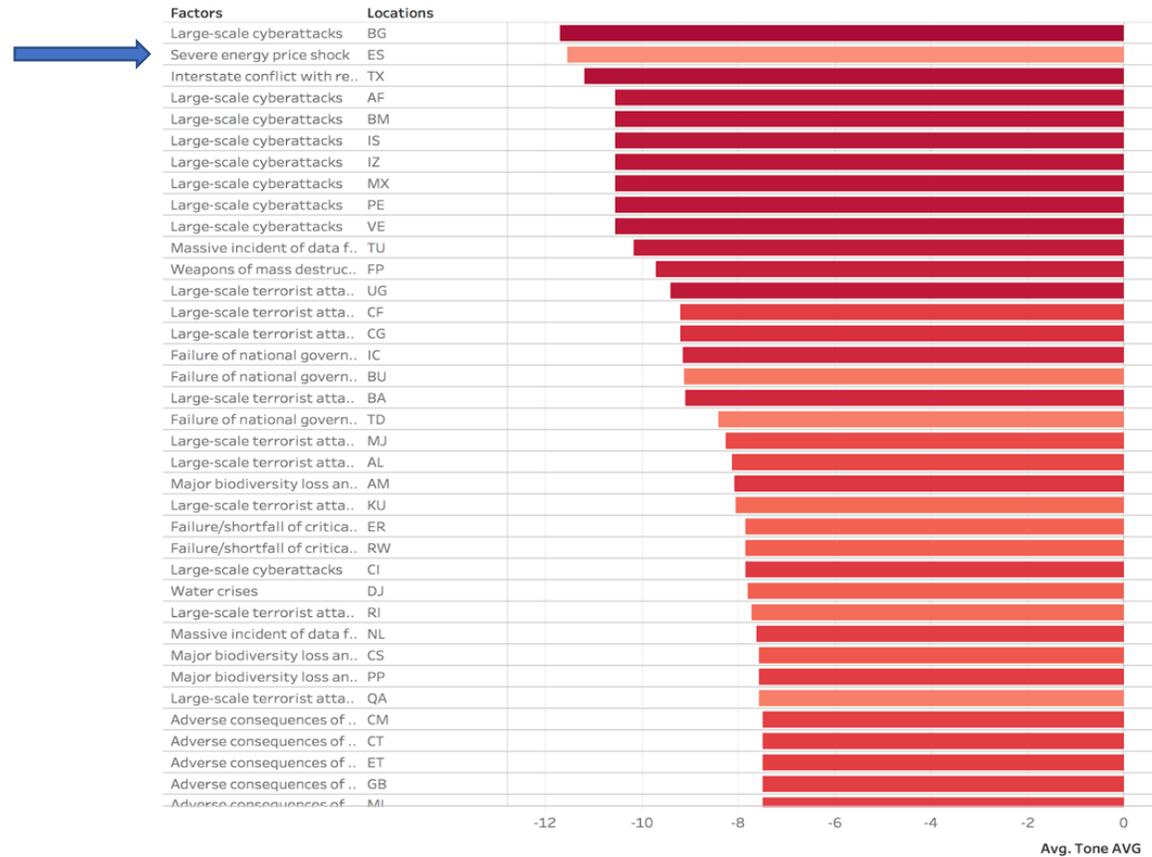
Furthermore, the combination of positive and negative SDG scores can be used to better assess non-financial risks and calculate a 'net' SDG footprint that account for the netting effect of positive and negative externalities at both long and short-term frequencies. This enables the algorithm to better identify both positive and negative trends in companies.

Thus, an AI-driven approach can help uncover hidden material risks, substantially reduce positive biases and uncover negative scores resulting from an adverse SDG footprint. This can 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.

- Financial markets are affected by sentiment, and bearish sentiment can make a down market worse and lessen the impact of positive news
- Firms which take advantage of sentiment information quickly can gain an edge
- Sentiment can be discovered in news articles, social media, blogs and other sources across multiple languages and regions
- Computers analyzing sentiment can work at millisecond speeds and process more information than human analysts
- NLP-driven approaches can be applied for both Companies and Countries
- The use of taxonomies and deep learning enable the decomposition of Sentiment analysis into multiple risk factors which can be tracked separately

# SDG Footprint: Relevance of Multi-language Sources

- The majority of the news available worldwide is not in English, particularly in Latin America, Europe, Africa and Asia.
- In many cases, there is a time lag between the time the news is reported in the local language and when it is published in mainstream English-based media
- The following shows a negative event in El Salvador which was not available in English on the day it was released.



**La energía eléctrica en El Salvador es la más cara de la región**  
 elsalvador.com - Aug 20, 2018  
 El Salvador es el país que más caro paga el precio de la **energía** .... contó que dada la **crisis** de insumos que hay en este momento se ven ...



**Combustible sigue elevando el precio de la energía en el país**  
 elsalvador.com - Aug 20, 2018  
 La **energía** en El Salvador es más cara porque, en promedio, .... que impondrá el Ejecutivo desde el lunes para intentar salir de la **crisis**.

# NLP for SDG / Event Monitoring



## Event:

**A US Federal judge dismisses some but not all criminal charges against Fedex Corp. in a case alleging it knowingly shipped illegal prescription drugs.**

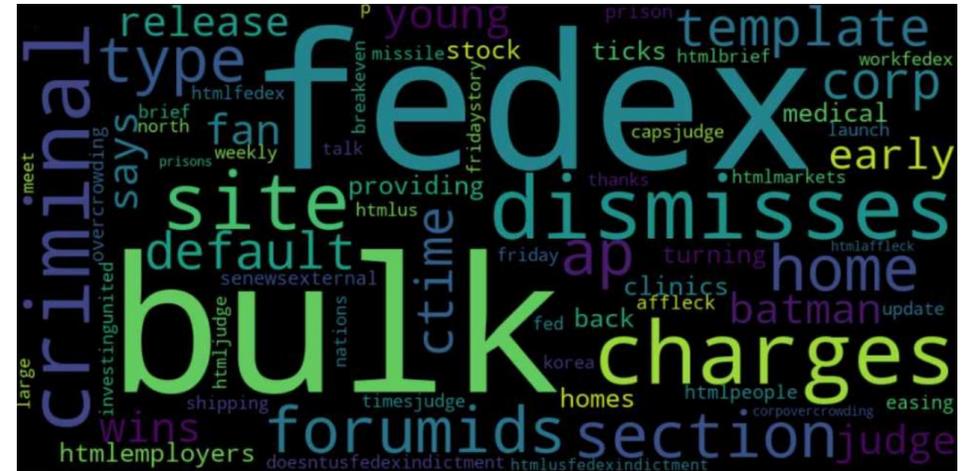
Sample News  
During Event Period

```
'http://www.4-traders.com/FEDEX-CORPORATION-12585/news/FedEx-stock-ticks-up-again-Friday-BRIEF-22040980/',  
'http://www.4-traders.com/FEDEX-CORPORATION-12585/news/FedEx-stock-ticks-up-again-Friday-BRIEF-22040980/',  
'http://forum.pafos.org/external.php?m=4d300d467b2a0212632b0d8ef473ddf0&type=RSS2&forumid=2',  
'http://seekingalpha.com/article/3959674-s-and-p-500-weekly-update-fed-talk-back-investing',  
'http://seekingalpha.com/article/3959674-s-and-p-500-weekly-update-fed-talk-back-investing',  
'https://mrtpstep.com/markets-back-to-breakeven-thanks-to-5-large-caps/',  
'https://mrtpstep.com/markets-back-to-breakeven-thanks-to-5-large-caps/',  
'https://www.equities.com/news/employers-turning-to-on-site-clinics-providing-medical-homes-at-work',  
'https://www.equities.com/news/employers-turning-to-on-site-clinics-providing-medical-homes-at-work',  
'http://cnnews.com/news/article/judge-dismisses-bulk-criminal-charges-against-fedex-corp',  
'http://cnnews.com/news/article/judge-dismisses-bulk-criminal-charges-against-fedex-corp',  
'http://www.pe.com/articles/charges-797552-judge-bulk.html',  
'http://www.pe.com/articles/charges-797552-judge-bulk.html',  
'http://www.abcfomontana.com/story/31518213/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://www.abcfomontana.com/story/31518213/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://bsoccomment.com/2016/03/19/united-nations-to-meet-on-north-korea-missile-launch.html',  
'http://kstar.com/story/971450/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://kstar.com/story/971450/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://hosted.ap.org/dynamic/stories/U/US_FEDEX_INDICTMENT?SITE=AP&SECTION=HOME&TEMPLATE=DEFAULT&CTIME=2016-03-19-15-53-15',  
'http://hosted.ap.org/dynamic/stories/U/US_FEDEX_INDICTMENT?SITE=AP&SECTION=HOME&TEMPLATE=DEFAULT&CTIME=2016-03-19-15-53-15',  
'http://kimt.com/ap/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp/',  
'http://kimt.com/ap/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp/',  
'https://www.equities.com/news/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'https://www.equities.com/news/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://www.hometownstations.com/story/31518213/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://www.hometownstations.com/story/31518213/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://kdhnews.com/business/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp/article_cd7b3263-ed46-5c77-1e87-8ddf7150d88.html',  
'http://kdhnews.com/business/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp/article_cd7b3263-ed46-5c77-1e87-8ddf7150d88.html',  
'http://www.commercialappeal.com/business/medical/employers-turning-to-on-site-clinics-providing-medical-homes-at-work-2e42a072-4575-51e3-e053-0100007-372623501.html',  
'http://www.caller.com/news/372738171.xhtml',  
'http://www.caller.com/news/372738171.xhtml',  
'http://www.brandonsun.com/business/breaking-news/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp-372728181.html',  
'http://www.brandonsun.com/business/breaking-news/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp-372728181.html',  
'http://www.newschannel10.com/story/31518214/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://www.newschannel10.com/story/31518214/judge-dismisses-bulk-of-criminal-charges-against-fedex-corp',  
'http://www.wpxi.com/ap/ap/business/judge-dismisses-bulk-of-criminal-charges-against-f/nqNk/',  
'http://www.wpxi.com/ap/ap/business/judge-dismisses-bulk-of-criminal-charges-against-f/nqNk/'
```

Sentiment Scores:  
Negative Outliers

COMPANY		
fedex corp	2016-03-01	0.552099
fedex corp	2016-03-02	0.601585
fedex corp	2016-03-03	0.453747
fedex corp	2016-03-04	-0.77439
fedex corp	2016-03-05	0.288858
fedex corp	2016-03-06	-0.13897
fedex corp	2016-03-07	0.848232
fedex corp	2016-03-08	0.737506
fedex corp	2016-03-09	1.46221
fedex corp	2016-03-10	0.606354
fedex corp	2016-03-11	0.127822
fedex corp	2016-03-12	-0.774874
fedex corp	2016-03-13	0.3852
fedex corp	2016-03-14	0.540515
fedex corp	2016-03-15	0.851443
fedex corp	2016-03-16	0.489121
fedex corp	2016-03-17	-0.229605
fedex corp	2016-03-18	-0.234236
fedex corp	2016-03-19	-6.02716
fedex corp	2016-03-20	-6.824
fedex corp	2016-03-21	-0.186601
fedex corp	2016-03-22	0.375192
fedex corp	2016-03-23	-0.728203
fedex corp	2016-03-24	-0.288793
fedex corp	2016-03-25	0.516625
fedex corp	2016-03-26	-0.0278542
fedex corp	2016-03-27	-0.242469

World Cloud  
Negative Keywords



# Country Risk Monitoring: WEF Taxonomy

- WEF-based Risk scores can be used as proxies to identify emerging risks and trends at the Country Level to better assess Country risks.
- WEF taxonomy is used to generate Country-specific NLP-based Risk Signals across 5 major categories and 30 sub-categories – including Economic, Environmental, Geopolitical, Technological, and Societal risks– based in geotagged data from 100,000+ sources in over than 60 languages.
- Risk scores for each category at the Country level.

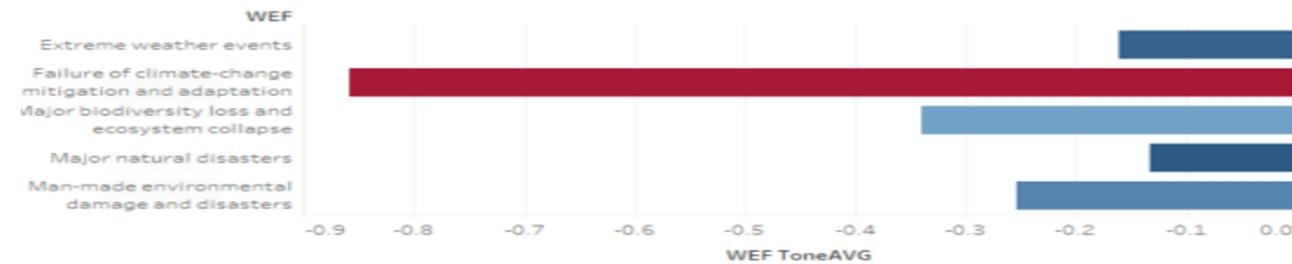
## WEF



### October 23, 2018: Latest Environmental WEF Tone



### October 23, 2018: Risk Breakdown of Latest WEF Tone



### Last 5 Days: WEF Risk Inter-linkages Outlier Map with breakdown of Economic Risk



#### Global Risk

Economic	Asset bubbles in a major economy
	Deflation in a major economy
	Failure of a major financial mechanism or institution
	Failure/shortfall of critical infrastructure
	Fiscal crises in key economies
	High structural unemployment or underemployment
	Illicit trade (e.g. illicit financial flows, tax evasion, human trafficking, organized crime, etc.)
	Severe energy price shock (increase or decrease)
	Unmanageable inflation
	Environmental
Failure of climate-change mitigation and adaptation	
Major biodiversity loss and ecosystem collapse (terrestrial or marine)	
Major natural disasters (e.g. earthquake, tsunami, volcanic eruption, geomagnetic storms)	
Man-made environmental damage and disasters (e.g. oil spills, radioactive contamination, etc.)	
Geopolitical	Failure of national governance (e.g. failure of rule of law, corruption, political deadlock, etc.)
	Failure of regional or global governance
	Interstate conflict with regional consequences
	Large-scale terrorist attacks
Societal	State collapse or crisis (e.g. civil)
	Failure of urban planning
	Food crises
	Large-scale involuntary migration
	Profound social instability
Technological	Rapid and massive spread of infectious diseases
	Water crises
	Adverse consequences of technological advances
	Breakdown of critical information infrastructure and networks
	Large-scale cyberattacks
	Massive incident of data fraud/theft