

# Supplementary Materials for Glitter or Gold? Deriving Structured Insights from Sustainability Reports via Large Language Models

Marco Bronzini\*, Carlo Nicolini, Bruno Lepri, Andrea Passerini and Jacopo Staiano

\*Corresponding author. Email: marco.bronzini-1@unitn.it

## 1 Semantic clustering

Cluster label	Cluster elements
ESG Commitments	ESG Metrics ESG Incentives ESG Promotion ...
Product Sustainability	Sustainability Sustainable Products Supplier Sustainability ...
Supplier Assessment	Supplier Audits Supplier Relationship Supplier Risk Assessment ...

**Table 1:** Three examples of the semantic clusters of the ESG categories (cat) grouped through thresholded cosine similarity on embedded categories.

Cluster label	Cluster elements
Assessment of	Monitoring of Assessment on Assessment and monitoring of ...
Partnership with	Collaboration of Working together with Partnering with others to ...
Opportunity to	Creation of opportunity Opportunities to work on Opportunity to contribute to ...

**Table 2:** Three examples of the semantic clusters of predicates (pred) grouped through thresholded cosine similarity on embedded predicates.

Attribute	Original	Modified text
cat pred obt	<b>Health and Safety</b> Designed to continually <b>improve</b> Workplace safety culture	<b>Employee Safety</b> <b>Improving</b> Workplace safety culture
cat pred obt	<b>Taxes</b> Increase <b>in</b> Tax and royalty payment	<b>Tax</b> Increase <b>of</b> Tax and royalty payment
cat pred obt	<b>Greenhouse Gas Emissions</b> Follows The WRI and WBCSD GHG Protocol	<b>Air Emissions</b> <b>Following</b> The WRI and WBCSD GHG Protocol
cat pred obt	<b>Resource Efficiency</b> Increasing <b>usage of</b> Renewable and recycled raw materials	<b>Energy Efficiency</b> Increase of Renewable and recycled raw materials

**Table 3:** Three examples of replacing the attributes of the triples with the respective cluster labels.

## 2 Full model instruction

Your goal is to extract structured information from the user's input that matches the form described below. When extracting information please make sure it matches the type of information exactly. Do not add any attributes that do not appear in the schema shown below.

```
““TypeScript
esg_actions: Array<{ //actions related to corporate's environmental, social or governance aspects
esg_category: "access.to.basic.services" | "access.to.healthcare" | "animal.welfare" | "anti.competitive.practices" |
"audit" | "biodiversity" | "board" | "board.diversity" | "business.ethics" | "chairperson.ceo.separation" | "child.labor" |
"climate.risk.management" | "clinical.trials" | "collective.bargaining" | "community.and.society" | "corporate.governance" |
"corruption" | "customer.relationship" | "diversity" | "esg.incentives" | "electromagnetic.fields" | "employee.development"
| "employee.turnover" | "energy" | "environmental.fines" | "environmental.management.system" | "environmental.policy"
| "environmental.reporting" | "financial.inclusion" | "forests" | "ghg.emissions" | "ghg.policies" | "gmso" |
"global.compact.membership" | "green.buildings" | "green.products" | "hiv.programs" | "hazardous.waste" | "health.and.safety"
| "human.rights" | "indigenous.rights" | "labor.practices" | "lobbying" | "non.ghg.air.emissions" | "ozone.depleting.gases" |
"packaging" | "philanthropy" | "privacy.and.it" | "product.safety" | "public.health" | "remuneration" | "reporting.quality" |
"resource.encyency" | "responsible.marketing" | "shareholders" | "site.closure" | "supply.chain" | "sustainable.finance" |
"systemic.risk" | "taxes" | "toxic.spills" | "unions" | "waste" | "water"
// an issue related to an ESG aspect
predicate: string //a nominalized verb that affects the ESG-related category
object: string //an entity related to the esg category that undergoes the predicate}>
““
```

Please output the extracted information in JSON format. Do not output anything except for the extracted information. Do not add any clarifying information. Do not add any fields that are not in the schema. If the text contains attributes that do not appear in the schema, please ignore them. All output must be in JSON format and follow the schema specified above. Wrap the JSON in <json> tags.

Input: In accordance with our ambitious goal, the water withdrawal of the data center decreased remarkably from 3.874 million litres to 2.367 million litres across the past three years.

Output: <json>{"esg\_actions": [{"esg\_category": "Water", "predicate": "Reduction of", "object": "The water withdrawal of the data center by 1.507 million litres"}]}</json>

Input: TotalEnergies introduced an innovative program at its European offices last year to address employees' concerns by creating a dedicated listening space.

Output: <json>{"esg\_actions": [{"esg\_category": "Employee Development", "predicate": "Introduction of", "object": "An innovative program"}]}</json>

Input: In San Antonio, Texas, our company reduced significantly the potable water usage of the data center by around 20% throughout 2020, providing economic and environmental benefits.

Output: <json>{"esg\_actions": [{"esg\_category": "Water", "predicate": "Reduction of", "object": "The data center's potable water usage by around 20%"}]}</json>

Input: In 2019, the ethics training program was completed by over 95% of our employees with outstanding results at our American training centre.

Output: <json>{"esg\_actions": [{"esg\_category": "Employee Development", "predicate": "Completion of", "object": "The ethics training program"}]}</json>

Input: Microsoft has invested €125 million in cutting-edge recycling technologies and smart waste management systems at its offices in Zwijndrecht, Belgium.

Output: <json>{"esg\_actions": [{"esg\_category": "Waste", "predicate": "Investment in", "object": "Cutting-edge recycling technologies and smart waste management systems"}]}</json>

Input: \$INPUT

Output:

**Figure 1:** The full model instruction used in our approach. INPUT represents an ESG-related sentence from which structured data needs to be retrieved. Sentences in the examples (in-context learning) do not represent any real facts.

### 3 Descriptive analysis of the generated triples

ESG category (cat)	Triples (%)	Companies (%)	Sectors (%)	Triples per company sector (%)
Corporate Governance	6.7	94	100	Industrials (7.8%), Financial Services (7.5%), Healthcare (7.4%)
Climate Risk Management	5.1	95	100	Real Estate (7.9%), Financial Services (7.0%), Industrials (5.5%)
Employee Development	4.8	90	100	Healthcare (5.9%), Communication Services (5.7%), Real Estate (5.6%)
Air Emissions	3.7	92	100	Energy (6.9%), Basic Materials (4%), Industrials (3.8%)
Water	3.1	85	100	Consumer Defensive (6.6%), Basic Materials (4.9%), Energy (4.4%)
Green Buildings	1.5	88	100	Real Estate (3.5%), Consumer Cyclical (1.8%), Technology (1.6%)
Packaging	0.9	49	100	Consumer Defensive (4.9%), Consumer Cyclical (1.5%), Technology (0.9%)
Business Ethics	0.3	61	100	Healthcare (0.7%), Utilities (0.5%), Industrials (0.4%),

**Table 4:** Sample of the ESG categories disclosed by companies in their sustainability reports. The table presents the category coverage using three key metrics: the total number of triples (triples), the count of companies that reported it (companies), and the number of company sectors covered (sectors). Additionally, the coverage is assessed by aggregating all the company triples by company sector and the percentage of sector triples concerning a category is reported (Triples per sector).

Action (cat : pred)	Triples (%)	Companies (%)	Sectors (%)	Triples per company sector (%)
AIR EMISSIONS: reduction of	0.6	70	100	Energy (1.2%), Basic Materials (0.8%), Industrials (0.7%)
PHILANTHROPY: donation by	0.4	60	100	Real Estate (0.8%), Communication Services (0.5%), Healthcare (0.5%)
ENERGY: reduction of	0.3	61	100	Real Estate (0.5%), Consumer Cyclical (0.4%), Communication Services (0.4%)
CLIMATE RISK MANAGEMENT: assessment of	0.3	56	100	Real Estate (0.7%), Industrials (0.3%), Financial Services (0.3%)
WATER: reduction of	0.3	50	100	Consumer Defensive (0.4%), Consumer Cyclical (0.4%), Energy (0.4%)
CORPORATE GOVERNANCE: establishment of	0.3	49	100	Healthcare (0.4%), Financial Services (0.4%), Technology (0.4%)
BIODIVERSITY: promotion of	0.2	40	91	Energy (0.5%), Utilities (0.4%), Financial Services (0.4%)
WASTE: reduction of	0.2	39	100	Consumer Defensive (0.5%), Industrials (0.3%), Consumer Cyclical (0.2%)
COMMUNITY AND SOCIETY: engagement in	0.1	37	82	Real Estate (0.5%), Healthcare (0.4%), Basic Materials (0.2%)
EMPLOYEE SAFETY: commitment and involvement with	0.1	25	100	Basic Materials (0.3%), Healthcare (0.1%), Industrials (0.1%)

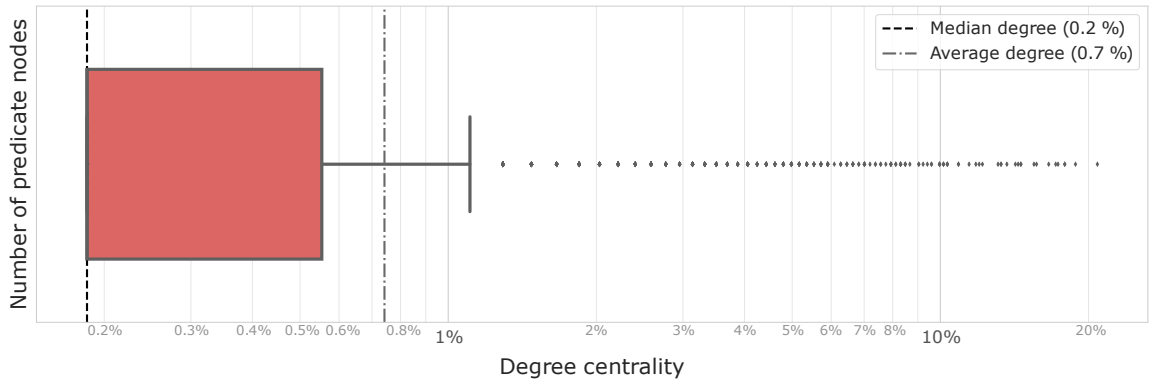
**Table 5:** Sample of the actions disclosed by companies in their sustainability reports. The table presents the action coverage through the same four metrics of the above table.

## 4 Statistics of the bipartite graphs

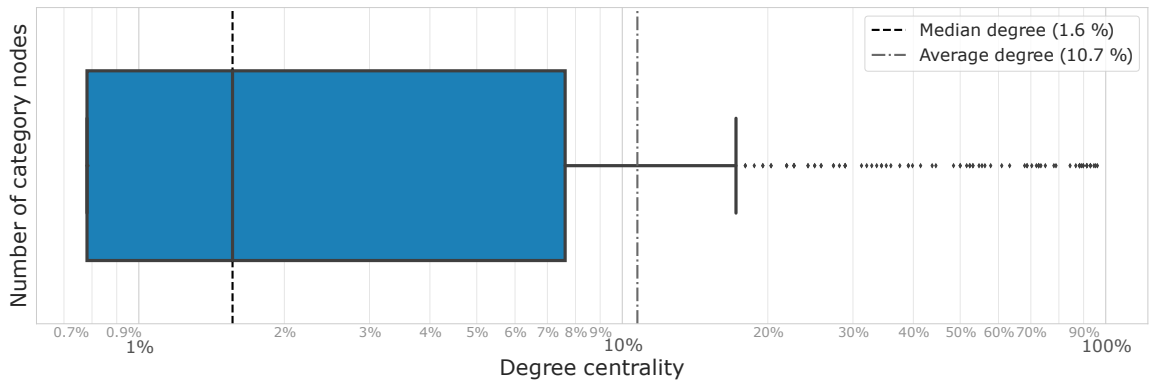
Bipartite graph	Partition A	Partition B	Edges	Density (%)
category-predicate ( $\mathbb{B}_{\text{catpred}}$ )	542	4,864	19,574	0.7
company-category ( $\mathbb{B}_{\text{cocat}}$ )	124	542	7,455	10.8
company-action ( $\mathbb{B}_{\text{coact}}$ )	124	19,574	43,169	1.7

**Table 6:** Statistics of the bipartite graphs obtained from the original KG. The category-predicate and company-action graphs have a low density ( $<1\%$ ), while the company-category graph exhibits a higher density (11%).

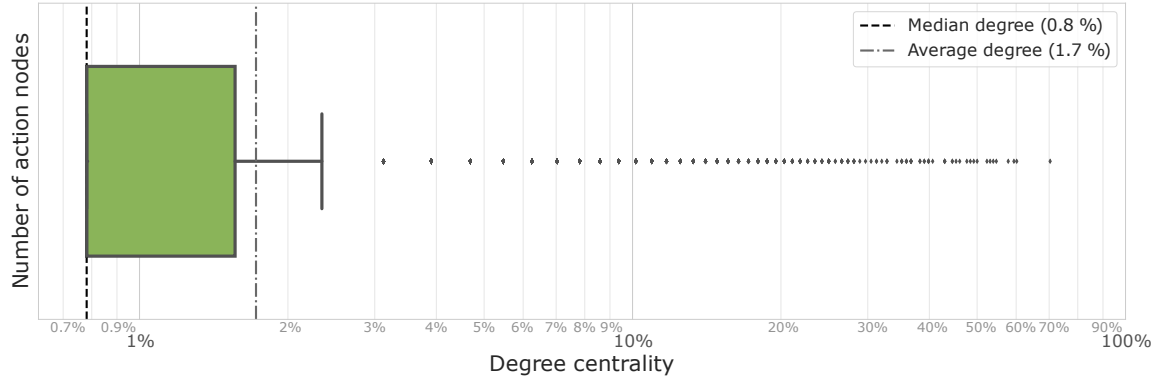
## 5 Degree distributions of the three bipartite graphs



**Figure 2:** Numerical distribution of degree centrality concerning the bipartite graph  $\mathbb{B}_{\text{catpred}}$ .



**Figure 3:** Numerical distribution of degree centrality concerning the bipartite graph  $\mathbb{B}_{\text{cocat}}$ .



**Figure 4:** Numerical distribution of degree centrality concerning the bipartite graph  $B_{coact}$ .

## 6 Descriptive statistics of the three bipartite graphs

Predicate node (pred)	Degree (%)	Closeness (%)	Betweenness (%)
Commitment and involvement with	20.8	89.9	1.5
Advisory support for	18.8	89.0	1.2
Partnership with	17.9	88.2	1.2
Establishment of	17.3	88.3	1.1
Use of	17.2	88.9	1.1
Development and implementation of	16.6	89.0	1.0
Recognition of	13.1	87.6	0.8
Engagement in	12.2	84.2	0.8
Compliance with	10.9	85.3	0.6
Consideration of	10.3	83.9	0.6
Reduction of	10.1	77.3	0.5
Review of	10.0	84.6	0.5
Maintenance of	10.0	84.2	0.5
Launch of	10.0	82.3	0.5
The impacts of	9.6	84.2	0.6
Belief in	9.6	83.2	0.4
Accounting for	1.8	61.2	0.0
Becoming aware of	1.8	60.2	0.0
Re-institution of	0.4	51.0	0.0
Helping consumers achieve	0.2	49.9	0.0
Questioned by	0.2	34.8	0.0
Contesting	0.2	32.2	0.0

**Table 7:** Graph metrics of a sample of all predicate nodes of the bipartite graph  $B_{catpred}$ .

Category node (cat)	Degree (%)	Closeness (%)	Betweenness (%)
Environmental	96.1	99.0	2.3
Supply Chain	95.3	98.9	2.3
Climate Risk Management	94.5	98.5	2.3
Corporate Governance	93.0	98.1	2.1
Energy	93.0	97.9	2.1
Product Sustainability	93.0	97.7	2.2
Community and Society	91.4	98.1	2.1
Board Diversity	89.8	97.4	2.0
Philanthropy	88.3	96.8	2.0
Business Ethics	60.9	85.2	1.0
Anti-Competitive Practices	52.3	79.9	0.7
Packaging	50.0	77.7	0.6
Human Resources	39.8	74.2	0.4
Tax	39.1	74.6	0.4
Lobbying	32.8	71.8	0.3
Whistle-blowing	20.3	64.5	0.1
Recycling	8.6	59.1	0.0
LGBTQ+ Inclusion	4.7	55.4	0.0
Product Responsibility	2.3	53.2	0.0
Anti-Discrimination	2.3	52.1	0.0
Green-washing Behavior	1.6	51.2	0.0
Marketing Responsibly	0.8	47.8	0.0

**Table 8:** Graph metrics of a sample of all the 542 category nodes of the bipartite graph  $\mathbb{B}_{\text{ccat}}$ .

Action node (cat:pred)	Degree (%)	Closeness (%)	Betweenness (%)
AIR EMISSIONS: Reduction of	70.3	87.6	2.1
ENERGY: Reduction of	60.2	83.9	1.5
PHILANTHROPY: Advisory support for	59.4	81.9	1.4
PHILANTHROPY: Donation by	57.8	80.8	1.3
COMMUNITY AND SOCIETY: Advisory support for	54.7	78.2	1.1
CLIMATE RISK MANAGEMENT: Assessment of	53.9	77.2	1.0
BIODIVERSITY: Commitment and involvement with	53.1	77.1	1.1
CORPORATE GOVERNANCE: Commitment and involvement with	52.3	77.1	1.1
WATER: Reduction of	50.0	76.8	1.0
HUMAN RIGHTS: Assessment of	21.9	63.8	0.2
ENVIRONMENTAL: Participation in	21.9	62.4	0.2
EMPLOYEE DEVELOPMENT: Offering	21.9	61.5	0.2
ENVIRONMENTAL: Certification for	21.1	60.7	0.2
HUMAN RIGHTS: Recognition of	20.3	62.6	0.2
WATER: Introduction of	20.3	61.0	0.2
EMPLOYEE SAFETY: Advisory support for	20.3	60.5	0.2
WASTE: Evaluation of	4.7	52.8	0.0
SUPPLIER DIVERSITY: Establishment of	4.7	52.8	0.0
PRODUCT SUSTAINABILITY: Incorporation of	4.7	52.7	0.0
TRANSPORTATION: Opportunity to	0.8	50.4	0.0
INSURANCE: Related to	0.8	39.9	0.0
SUPPLIER ASSESSMENT: Internal risk based	0.8	38.6	0.0

**Table 9:** Graph metrics of a sample of all predicate nodes of the bipartite graph  $\mathbb{B}_{\text{coact}}$ .

## 7 Variability of ESG-related actions

ESG category (cat)	Entropy (nats)	Companies (%)	Category predicates (pred)
Supply Chain	5.72	95	Partnership with (3.7%), Assessment of (2.4%), Engagement in (2.2%)
Biodiversity	5.70	92	Commitment and involvement with (4.7%), Promotion of (4.2%), Increase of (2.1%)
Environmental	5.65	97	Development and implementation of (4.6%), Commitment and involvement with (3.1%), Establishment of (2.1%)
Corporate Governance	5.62	94	Establishment of (4.3%), Commitment and involvement with (3.7%), Overseeing (2.9%),
Human Rights	5.62	90	Commitment and involvement with (5.1%), Respect for (2.6%), Assessment of (2.4%)
Employee Development	5.46	90	Introduction of (4.5%), Provision of (3.9%), Required training (3.3%)
Product Safety	5.27	69	Assessment of (4.4%), Compliance with (2.3%), Development and implementation of (2.3%)
Food Waste	2.44	6	Commitment and involvement with (14.3%), Reduction of (14.3%), Development of (7.1%)
Product Responsibility	1.79	2	Signatory to (16.7%), Commitment and involvement with (16.7%), Includes (16.7%)
Anti-Slavery Practices	1.61	4	Undertaking of (20%), Integration of (20%), Required training (20%)
Clean Energy	0.69	2	Helping clients deploy (50%), Engagement in (50%)
Conflict of Interest	0.00	1	Avoidance of (100%)

**Table 10:** A sample of ESG categories with the computed entropy values. The three most frequent category predicates are reported alongside the percentage of companies disclosing that category.

## 8 Company similarities according to disclosed actions

Company	Top three reported actions	Most similar companies
Sony	PACKAGING: Reduction of (x7) PHILANTHROPY: Advisory support for (x6) CORPORATE GOVERNANCE: Establishment of (x6)	Canon (7%), Tokyo Gas (6%), Hyundai Motor (6%), LG Display (6%), 3M (6%), Toshiba (6%), Kia (6%)
Deutsche Bank	ENERGY: Reduction of (x4) BIODIVERSITY: Promotion of (x4) CORPORATE GOVERNANCE: Establishment of (x4)	Royal Bank of Canada (7%), Banco Santander (6%), UniCredit (6%), Airbus (5%), BPER Banca (5%)
Mastercard	PHILANTHROPY: Advisory support for (x4) HUMAN RIGHTS: Commitment and involvement with (x4) FINANCIAL INCLUSION: Commitment and involvement with (x4)	Visa (6%), Home Depot (5%), American Electric Power Company (5%), BPER Banca (5%), Vodafone (4%)
Moderna	BIODIVERSITY: Commitment and involvement with (x4) PHILANTHROPY: Launch of (x3) PHILANTHROPY: Participation in (x3)	Vertex Pharmaceuticals (7%), AstraZeneca (5%), Broadcom (5%), Franklin Electric (4%), Alcon (4%)
Delta Air Lines	CLIMATE RISK MANAGEMENT: Evaluation of (x4) COMMUNITY AND SOCIETY: Recognition of (x3) EMPLOYEE SAFETY: Introduction of (x3)	American Electric Power Company (6%), Broadcom (6%), PPG Industries (5%), Canadian Pacific Railway (5%)
GlobalFoundries	WATER: Use of (x6) AIR EMISSIONS: Reduction of (x7) PHILANTHROPY: Donation by (x7)	Texas Instruments (7%), PPG Industries (7%), Imperial Oil (6%), Samsung (6%), Visa (6%)
Aluminum Corporation of China	CORPORATE GOVERNANCE: Establishment of (x3) CORPORATE GOVERNANCE: Improving (x3) CORPORATE GOVERNANCE: Organisation around (x3)	China Petroleum Chemical (6%), Geely Automobile (5%), Baidu (5%), United States Steel (4%), Tokyo Gas (4%)
Geely Automobile	PHILANTHROPY: Participation in (x5) SUPPLY CHAIN: Establishment of (x5) CORPORATE GOVERNANCE: Development and implementation of (x5)	China Petroleum Chemical (7%), Baidu (6%), LG Display (6%), Korean Air Lines (6%), Alibaba Group (5%)
STMicroelectronics	AIR EMISSIONS: Reduction of (x4) SUPPLY CHAIN: Assessment of (x4) AUDIT: Conducted (x3)	Texas Instruments (6%), PPG Industries (5%), GlobalFoundries (5%), Monster (5%), TotalEnergies (5%)
Enel	AIR EMISSIONS: Reduction of (x10) COMMUNITY AND SOCIETY: Contribution to (x6) HUMAN RIGHTS: Commitment and involvement with (x6)	Saipem (7%), Snam (6%), Tokyo Gas (6%), TotalEnergies (6%), Banco Santander (6%), UniCredit SpA (5%)
Saudi Aramco	AIR EMISSIONS: Reduction of (x4) BIODIVERSITY: Protection of (x4) ENERGY: Investment in (x4)	Tokyo Gas (5%), Royal Dutch Shell (5%), Yamana Gold (5%), Visa (5%), GlobalFoundries (5%)
Philip Morris	WASTE: Continuous efforts to reduce, reuse, or recycle (x5) BIODIVERSITY: Continuing to set goals and work towards (x4) PHILANTHROPY: Investment in (x3)	3M (4%), Coca Cola (4%), Croda (4%), GlobalFoundries (4%), United States Steel (4%), Samsung (4%)
VMware	CORPORATE GOVERNANCE: Overseeing (x4) PHILANTHROPY: Advisory support for (x4) BIODIVERSITY: Commitment and involvement with (x3)	Visa (5%), Oracle (4%), Mastercard (4%), The Home Depot (4%), Amazoncom (4%), Cisco (4%)

**Table 11:** A company sample with the most reported actions and the most similar companies for each. The company similarity is assessed by computing the Jaccard similarity on the action set of companies.

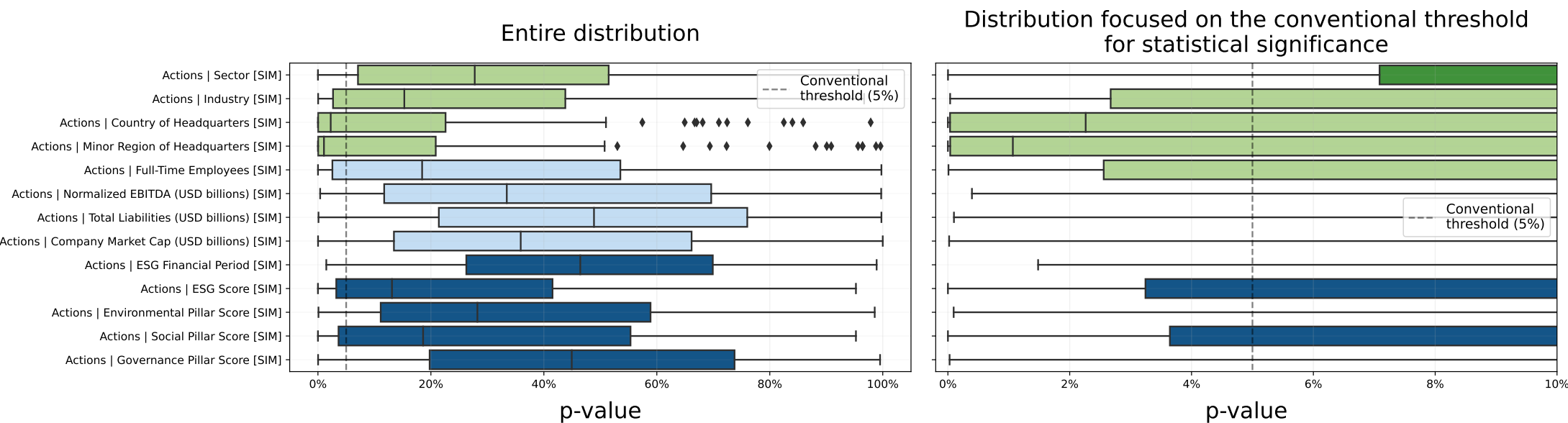


## 9 Features used for the correlation analysis of company similarities

Feature	Similarity Measure
Action Similarity	Jaccard Similarity
Sector Industry Country of Headquarters Minor Region of Headquarters	Cosine Similarity with min-max normalisation
ESG Score Environmental Pillar Score Social Pillar Score Governance Pillar Score ESG Financial Period	Absolute difference with max-min normalisation
Full-Time Employees Normalized EBITDA (USD billions) Total Liabilities (USD billions) Company Market Cap (USD billions)	

**Table 12:** Similarity measures of company features used in the bivariate correlation analysis concerning company similarities. All the measures are defined in the range between zero and one.

# 10 Statistical significance of the pairwise correlations



**Figure 5:** Distribution of the p-values for Kendall’s correlation coefficient computed in the correlation analysis between company similarities in terms of actions disclosed and similarities in other company information. The right graph exhibits the distributions focused around the conventional threshold of statistical significance.

11 Monotonic correlations among company information

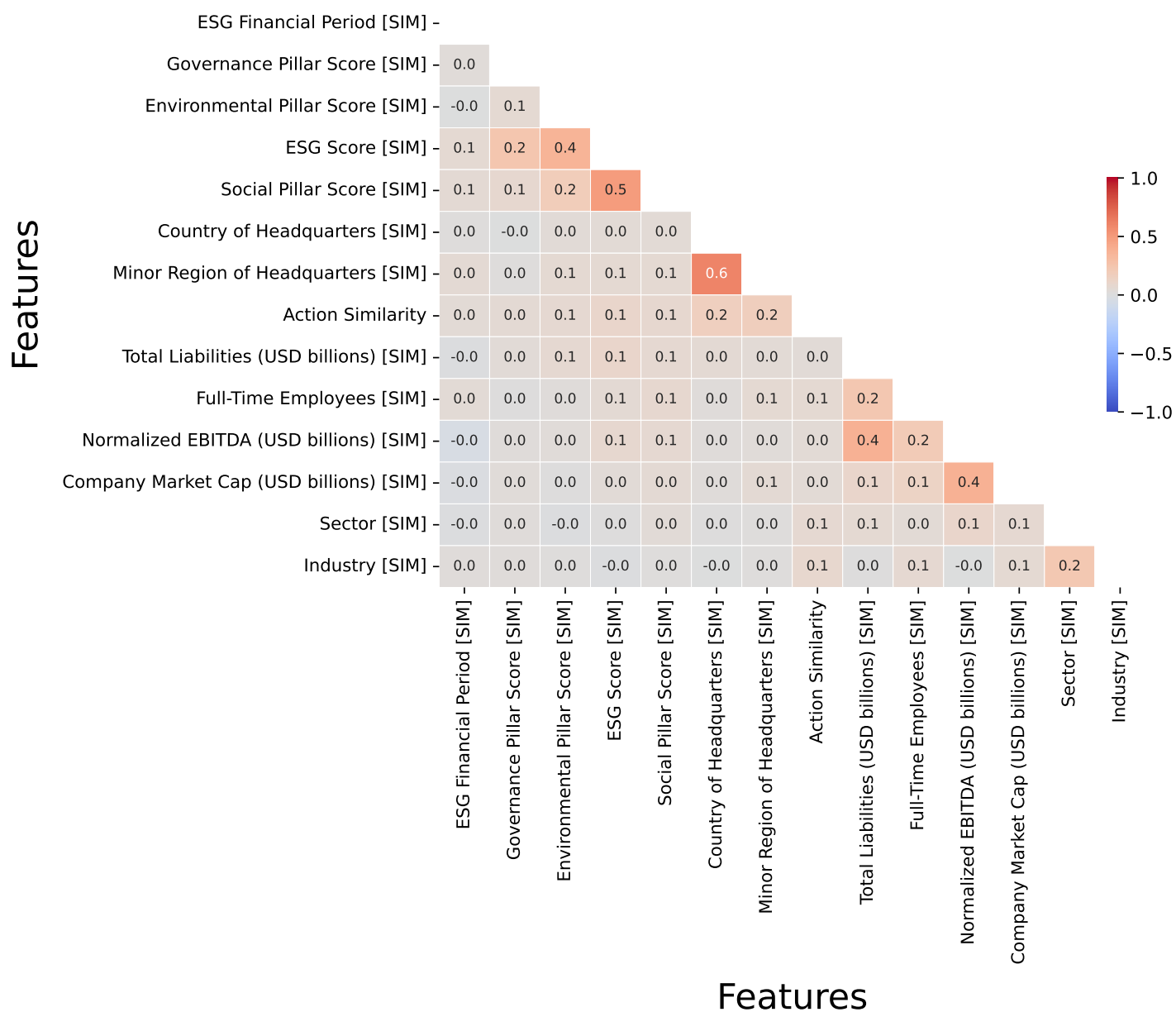


Figure 6: Pairwise monotonic correlations between company similarities.

## 12 Features for the ESG score inference for the interpretability analysis

Feature	Value	Source	Type
Category Entropy	3.6	Sustainability report	Continuous
Action Entropy	1.1		
EBITDA (USD Billions)	16.0	Refinitiv	
Liabilities (USD Billions)	191.8		
Market Cap (USD Billions)	158.6		
Employees	108900	Refinitiv	Discrete
Incorporation Year	1946		
Fiscal Year	2022		
SECTOR: Technology	True	Refinitiv	Binary categorical
SECTOR: Utilities	False		
...	...		
CONTINENT: Asia	True		
CONTINENT: Europe	False		
...	...		
REGION: Eastern Asia	True		
REGION: Northern Europe	False		
...	...		
COUNTRY: Japan	True		
COUNTRY: Canada	False		
...	...	Sustainability report	Percentage
Social Categories	0.27		
Environmental Categories	0.30		
Governance Categories	0.42		
CATEGORY: Corporate Governance	0.09		
CATEGORY: Environmental	0.07		
CATEGORY: Supply Chain	0.06		
CATEGORY: Waste	0.03		
CATEGORY: Board Diversity	0		
...	...		

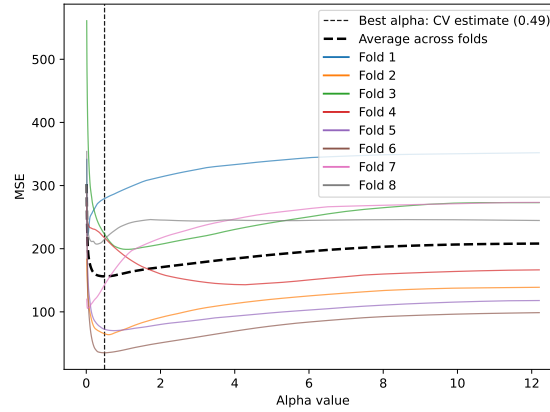
**Table 13:** Example of the features used in the regression model. There are numerical features and categorical features (dummy variables). The data is exhibited beforehand standardization. The observation concerns the fiscal year 2022 of Sony.

## 13 Performance of the OLS model

Metric	Score
Coefficient of Determination ( $R^2$ )	0.707
Mean Absolute Error (MAE)	5.775
Root Mean Square Error (RMSE)	7.761
Weighted Mean Absolute Percentage Error (wMAPE)	0.079

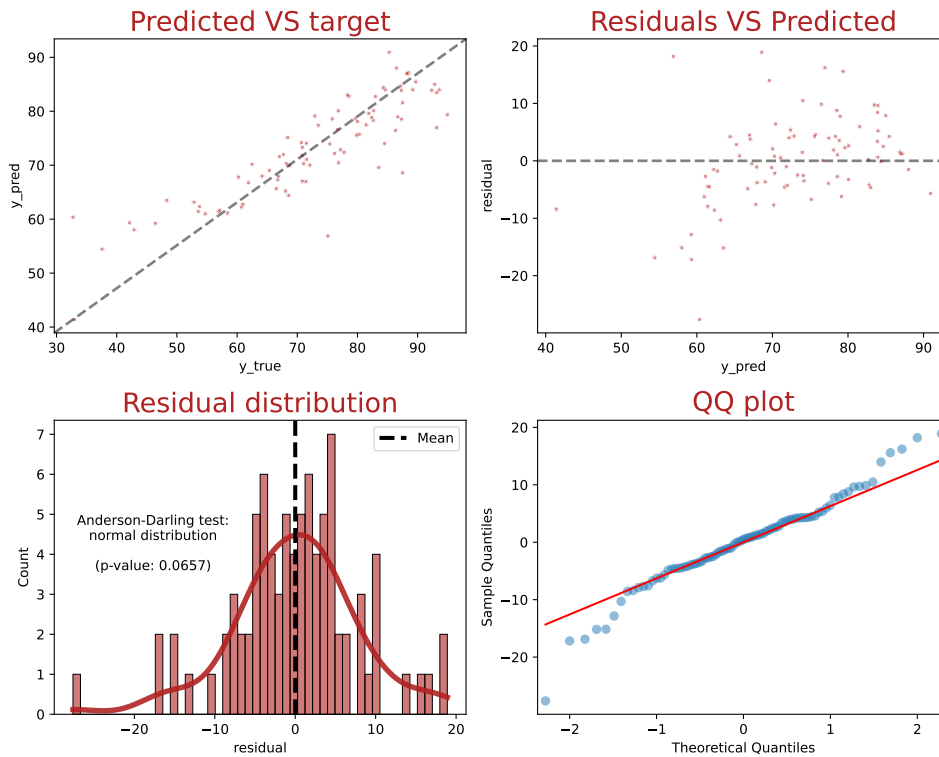
**Table 14:** Performance of the first-order regression model. Elastic Net Regularization was used as a feature selection method with the optimal alpha parametric constant ( $\alpha = 0.493$ ) discovered through a five-fold Cross-Validation approach.

## 14 Training performance of the OLS model using cross-validation



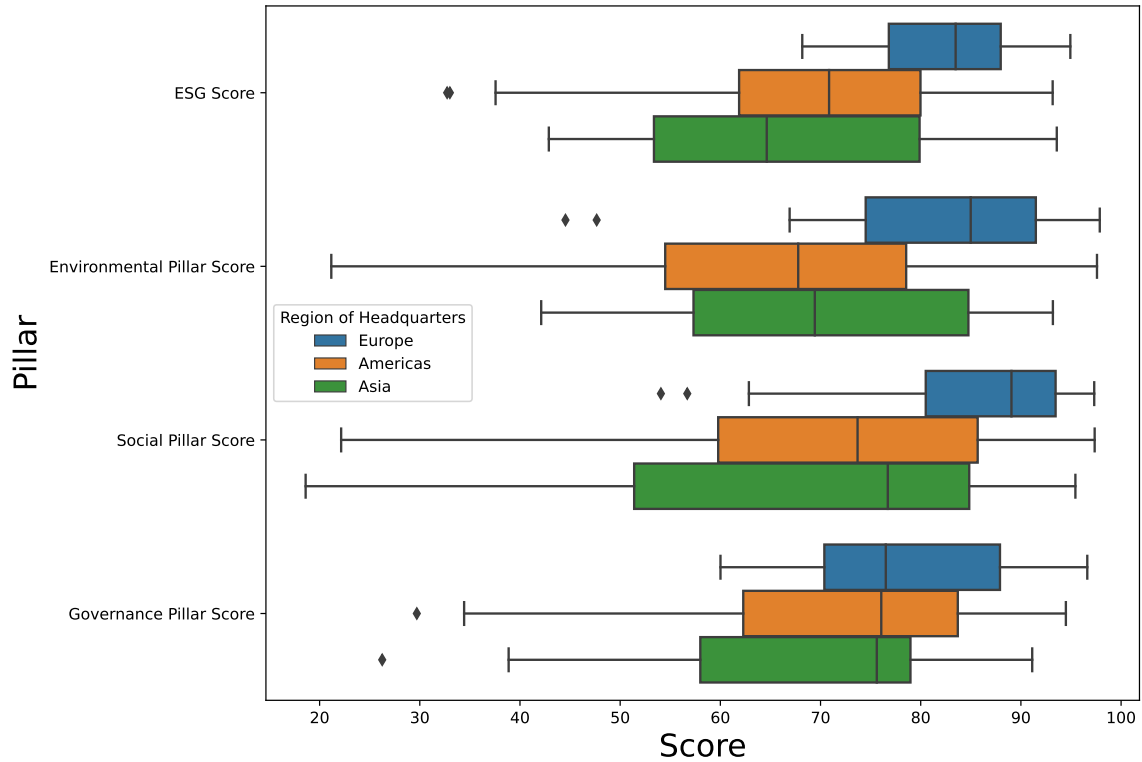
**Figure 7:** Performance of the OLS regression model during training. An eight-fold cross-validation approach, with the Elastic Net cost function based on Mean Squared Error (MSE), was adopted to estimate the optimal  $\alpha$  parameter for the Elastic Net Regularisation.

## 15 Residual analysis of the OLS model



**Figure 8:** Graphical panels exhibiting different residual analyses. The plot at the top left exhibits the predicted scores versus the actual ones, while the top-right graph shows the residuals versus predicted scores in order to check homoscedasticity. At the bottom left, the histogram exhibits the distributions of the residuals alongside the results of the Anderson-Darling test on normal distribution. Lastly, a QQ plot of residuals versus normal distribution is displayed at the bottom right of the graphical panel.

## 16 ESG scores by region



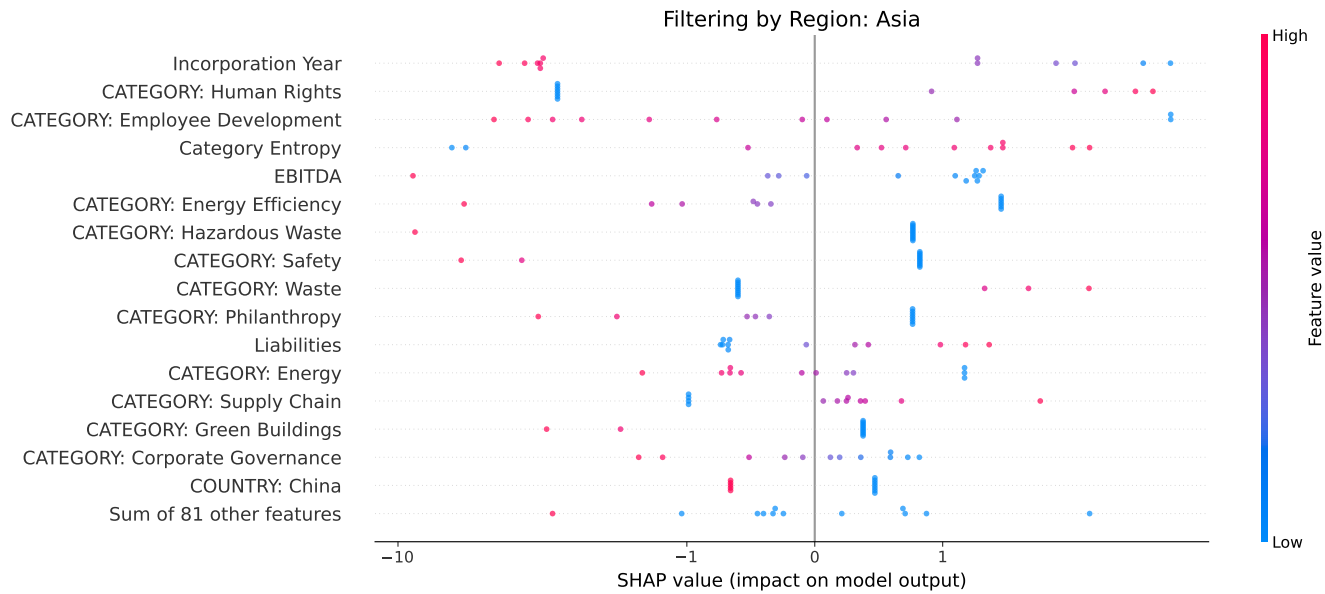
**Figure 9:** ESG scores of companies aggregated by continent. The graphical panel exhibits the E/S/G scores as well as the combined score. European companies have the highest average scores in each pillar.

## 17 Asian companies considered for the region-based ESG score interpretability

Company	Sector	Country	Incorporation Year	Fiscal Year	...	ESG score
Toshiba	Industrials	Japan	1904	2021	...	93.58
Sony	Technology		1946	2022	...	87.51
Toyota Motor	Consumer Cyclical		1937	2020	...	84.51
Tokyo Gas	Utilities		1885	2021	...	78.35
Geely Automobile	Consumer Cyclical	Hong Kong	1946	2022	...	75.42
Daikin Industries	Industrials	Japan	1934	2022	...	66.78
China Petroleum Chemical Corporation	Energy	China	2000	2021	...	62.49
Aluminum Corporation of China	Basic Materials		2001	2019	...	57.14
Baidu	Communication Services		2000	2020	...	53.55
China Evergrande	Real Estate		2006	2020	...	52.85
Alibaba	Consumer Cyclical	Saudi Arabia	1999	2022	...	48.33
Saudi Aramco	Energy		2018	2021	...	42.88

**Table 15:** The twelve Asian companies presented in our data. This subset includes businesses from Japan, China, Hong Kong and Saudi Arabia and was used for the region-based interpretability of ESG scores. The companies are ordered by ESG score.

## 18 Top features impacting the Asian companies' ESG scores



**Figure 10:** Summary of the sixteen primary factors with the greatest influence on determining the ESG score for Asian companies. The features are ordered according to their median shape value. The x-axis represents the degree of a positive and negative impact on model output. Each dot represents a company instance and colours represent the company values of the standardised feature.

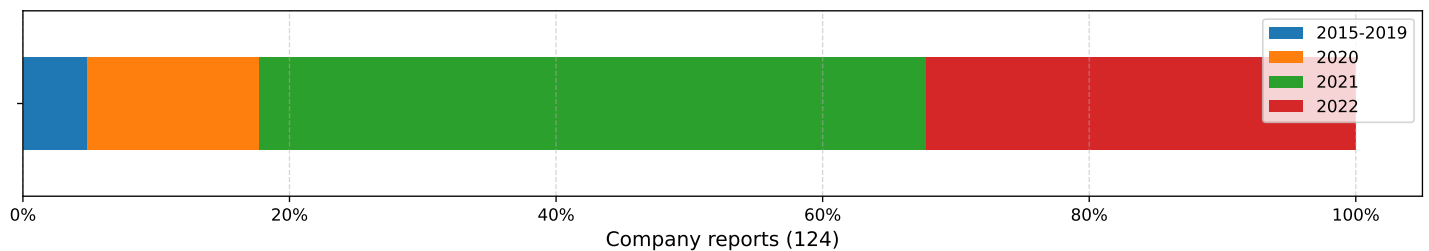
## 19 Categories of the adopted ESG Taxonomy

ESG category	Pillar	Degree centrality (%)
Climate Risk Management	Environmental	94.5
Energy		93.0
Biodiversity		91.4
Waste		88.3
Green Buildings		86.7
Water		84.4
Product Safety		68.8
Forests		53.1
Packaging		50.0
Hazardous Waste		48.4
Animal Welfare		35.9
Green Products		35.2
Environmental Fines		10.8
Environmental Management System		10.8
Environmental Policy		10.8
Environmental Reporting		10.8
GHG Emissions		10.8
GHG Policies		10.8
Non-GHG Air Emissions		10.8
Ozone-Depleting Gases		10.8
Resource Efficiency		10.8
Sustainable Finance		10.8
Electromagnetic Fields		7.0
GMOs		2.3
Toxic Spills		2.3
Supply Chain	Social	95.3
Community and Society		91.4
Employee Development		89.1
Human Rights		88.3
Philanthropy		88.3
Customer Relationship		73.4
Financial Inclusion		70.3

Access to Basic Services		68.8
Access to Healthcare		44.5
Child Labor		33.6
Collective Bargaining		28.9
Employee Turnover		27.3
Public Health		19.5
Diversity		10.8
Health and Safety		10.8
Labor Practices		10.8
Responsible Marketing		10.8
Unions		10.8
Indigenous Rights		8.6
Clinical Trials		7.8
HIV Programs		0.8
Corporate Governance	Governance	93.0
Board Diversity		89.8
Audit		72.7
Remuneration		72.7
Business Ethics		60.9
Anti-competitive Practices		52.3
Lobbying		32.8
Systemic Risk		28.1
Global Compact Membership		21.8
Board		10.8
Chairperson-CEO Separation		10.8
Corruption		10.8
ESG Incentives		10.8
Privacy and IT		10.8
Reporting Quality		10.8
Shareholders		10.8
Taxes		10.8
Site Closure		10.2

**Table 16:** The table exhibits the ESG categories of the ESG taxonomy used in this work. Their degree is derived from our findings, although they represent only a tenth of the categories discovered through our methodology.

## 20 Distribution of the fiscal years of the sustainability reports considered



**Figure 11:** Distribution of the fiscal years of the sustainability reports considered for our study. We picked the latest available reports for each considered company. Almost all reports concern the 2020s (orange, green and red), whereas only 5% refer to fiscal years within the 2015-2019 period (blue).



## 21 Complete list of the company considered

Company	Sector	Industry
3M Corporation	Industrials	Conglomerates
3i Group plc	Financial Services	Asset Management
Activision Blizzard Inc	Communication Services	Electronic Gaming & Multimedia
Adecco Group AG	Industrials	Staffing & Employment Services
Adidas AG	Consumer Cyclical	Footwear & Accessories
Air Canada	Industrials	Airlines
Air Liquide SA	Basic Materials	Specialty Chemicals
Airbus SE	Industrials	Aerospace & Defense
Alcon Inc	Healthcare	Medical Instruments & Supplies
Alibaba Group Holding Limited	Consumer Cyclical	Internet Retail
Alphabet Inc	Communication Services	Internet Content & Information
Aluminum Corporation of China Limited	Basic Materials	Aluminum
Amazoncom Inc	Consumer Cyclical	Internet Retail
American Electric Power Company Inc	Utilities	Utilities: Regulated Electric
Amplifon	Healthcare	Medical Distribution
Apple Inc	Technology	Consumer Electronics
ArcelorMittal SA	Basic Materials	Steel
Assicurazioni Generali SpA	Financial Services	Insurance: Diversified
AstraZeneca PLC	Healthcare	Drug Manufacturers: General
BPER Banca SpA	Financial Services	Banks: Regional
Baidu Inc	Communication Services	Internet Content & Information
Banco Santander SA	Financial Services	Banks: Diversified
Bank of America Corp BofA	Financial Services	Banks: Diversified
Bayer AG	Healthcare	Drug Manufacturers: General
British American Tobacco PLC	Consumer Defensive	Tobacco
British Land Co PLC The	Real Estate	REIT: Diversified
Broadcom Inc	Technology	Semiconductors
Builders FirstSource Inc	Industrials	Building Products & Equipment
CF Industries Holdings Inc	Basic Materials	Agricultural Inputs
Campbell Soup Company	Consumer Defensive	Packaged Foods
Canadian Pacific Railway Limited	Industrials	Railroads
Canon Inc	Technology	Computer Hardware
CarMax Inc	Consumer Cyclical	Auto & Truck Dealerships
China Evergrande Group	Real Estate	Real Estate: Development
China Petroleum Chemical Corporation	Energy	Oil & Gas Integrated
Cisco Systems Inc	Technology	Communication Equipment
Coca Cola	Consumer Defensive	Beverages: Non-Alcoholic
Commonwealth Bank of Australia	Financial Services	Banks: Diversified
Croda International plc	Basic Materials	Specialty Chemicals
Daikin Industries Ltd	Industrials	Building Products & Equipment
Delta Air Lines Inc	Industrials	Airlines
Deutsche Bank AG	Financial Services	Banks: Regional
Deutsche Lufthansa AG	Industrials	Airlines
Deutsche Wohnen	Real Estate	Real Estate: Development
DuPont	Basic Materials	Specialty Chemicals
ENI SpA	Energy	Oil & Gas Integrated
Edison International	Utilities	Utilities: Regulated Electric
Enel SpA	Utilities	Utilities: Diversified
FedEx Corporation	Industrials	Integrated Freight & Logistics
First Republic Bank CA	Financial Services	Banks: Regional
Fox Corporation	Communication Services	Entertainment
Franklin Electric Co Inc	Industrials	Specialty Industrial Machinery
Geely Automobile Holdings Ltd	Consumer Cyclical	Auto Manufacturers
General Motors Co GM	Consumer Cyclical	Auto Manufacturers
GlobalFoundries	Technology	Semiconductors
Goldman Sachs Group Inc The	Financial Services	Capital Markets
Home Depot Inc The	Consumer Cyclical	Home Improvement Retail
Humana Inc	Healthcare	Healthcare Plans
Hyundai Motor Co	Consumer Cyclical	Auto Manufacturers
Imperial Oil Ltd	Energy	Oil & Gas Integrated
Intel Corp	Technology	Semiconductors
Intuitive Surgical Inc	Healthcare	Medical Instruments & Supplies
Iveco Group NV	Industrials	Farm & Heavy Construction Machinery
Johnson Johnson	Healthcare	Drug Manufacturers: General

Kia Corp	Consumer Cyclical	Auto Manufacturers
Korean Air Lines Co Ltd	Industrials	Airlines
Kraft Heinz Co The	Consumer Defensive	Packaged Foods
LG Display Co Ltd	Technology	Consumer Electronics
Leonardo SpA	Industrials	Aerospace & Defense
Lockheed Martin Corp	Industrials	Aerospace & Defense
Mastercard Inc	Financial Services	Credit Services
Meta Platforms Inc	Communication Services	Internet Content & Information
Microsoft Corporation	Technology	Software: Infrastructure
Moderna Inc	Healthcare	Biotechnology
Monster Beverage Corp	Consumer Defensive	Beverages: Non-Alcoholic
NVIDIA Corp	Technology	Semiconductors
National Grid PLC	Utilities	Utilities: Regulated Electric
Nestle SA	Consumer Defensive	Packaged Foods
Netflix Inc	Communication Services	Entertainment
Novo Nordisk A S	Healthcare	Biotechnology
Oracle Corporation	Technology	Software: Infrastructure
PPG Industries	Basic Materials	Specialty Chemicals
Paramount Resources Ltd	Energy	Oil & Gas E&P
Park Hotels Resorts Inc	Real Estate	REIT: Hotel & Motel
PepsiCo Inc	Consumer Defensive	Beverages: Non-Alcoholic
Petroleo Brasileiro SA Petrobras	Energy	Oil & Gas Integrated
Philip Morris International	Consumer Defensive	Tobacco
Poste Italiane	Industrials	Conglomerates
Prologis Inc	Real Estate	REIT: Industrial
Royal Bank of Canada	Financial Services	Banks: Diversified
Royal Dutch Shell PLC	Energy	Oil & Gas Integrated
STMicroelectronics	Technology	Semiconductors
Saipem SpA	Energy	Oil & Gas Equipment & Services
Samsung Electronics Co Ltd	Technology	Consumer Electronics
Saudi Aramco	Energy	Oil & Gas Integrated
Simon Property Group Inc	Real Estate	REIT: Retail
SkyWest Inc	Industrials	Airlines
Sligro Food Group NV	Consumer Defensive	Food Distribution
Snam SpA	Utilities	Utilities: Regulated Gas
Sony Corporation	Technology	Consumer Electronics
Sun Communities	Real Estate	REIT: Residential
Swisscom AG	Communication Services	Telecom Services
Telecom Italia SpA	Communication Services	Telecom Services
Tesco PLC	Consumer Defensive	Grocery Stores
Tesla Inc	Consumer Cyclical	Auto Manufacturers
Texas Instruments Inc	Technology	Semiconductors
Tokyo Gas Co Ltd	Utilities	Utilities: Regulated Gas
Toshiba Corp	Industrials	Conglomerates
TotalEnergies	Energy	Oil & Gas Integrated
Toyota Motor Corp	Consumer Cyclical	Auto Manufacturers
Uber Technologies Inc	Technology	Software: Application
UniCredit SpA	Financial Services	Banks: Regional
Uniper SE	Utilities	Utilities: Independent Power Producers
United States Steel Corp	Basic Materials	Steel
VMware Inc	Technology	Software: Infrastructure
Vertex Pharmaceuticals Inc	Healthcare	Biotechnology
Virgin Atlantic Ltd	Industrials	Transportation & Logistics
Visa Inc	Financial Services	Credit Services
Vodafone Group plc	Communication Services	Telecom Services
WESCO International Inc	Industrials	Industrial Distribution
Walmart Inc	Consumer Defensive	Discount Stores
Walt Disney Co	Communication Services	Entertainment
Washington Real Estate Investment Trust	Real Estate	Commercial Real Estate
Yamana Gold Inc	Basic Materials	Gold

**Companies: 124**

## 22 All the sustainability reports originally gathered

Nationality	Companies	Companies (%)
United States	2059	47.0
Canada	391	8.9
Britain	191	4.4
Brazil	176	4.0
Taiwan	154	3.5
Japan	138	3.2
South Korea	133	3.0
...	...	...
United Arab Emirates	3	0.07
	4,222	100

**Table 18:** Nationality of the companies covered. Results are ordered by cardinality. The majority of the available reports come from North American companies (56%).

Language	Reports	Reports (%)
English (EN)	4,804	94.0
Spanish (ES)	77	1.5
Portuguese (PT)	76	1.5
Chinese (ZH)	61	1.2
Korean (KO)	31	0.6
Japanese (JA)	27	0.5
French (FR)	20	0.4
...	...	...
German (DE)	3	0.1
	6,456	100

**Table 19:** Percentage of the available reports by language. Results are ordered by cardinality. Almost all the non-financial reports are written in English.

Period	Fiscal year	Reports	Reports (%)
2020s	2021	3,483	54.0
	2020	1,588	24.6
	2022	621	9.6
2010s	2019	567	8.8
	2018	131	2.0
	2017	20	0.3
	2015	17	0.3
	...	...	...
	2013	2	0.03
		6,456	100

**Table 20:** Percentage of the available sustainability reports by Fiscal year. The results are ordered by annual reports.