Investigating Environmental, Social, and Governance (ESG) Discussions in News: A Knowledge Graph Analysis Empowered by AI

Simone Angioni[1], Sergio Consoli[2], Danilo Dessí[3], Francesco Osborne[4], Diego Reforgiato Recupero[1] and Angelo Salatino[4]

[1]*Department of Mathematics and Computer Science, University of Cagliari, Cagliari, Italy*
[2]*European Commission, Joint Research Centre (DG JRC), Ispra, Italy*
[3]*Knowledge Technologies for Social Sciences Department, GESIS Leibniz Institute for the Social Sciences, Cologne, Germany*
[4]*Knowledge Media Institute, The Open University, Milton Keynes, UK*

Abstract

This paper explores the growing importance of Environmental, Social, and Governance (ESG) criteria in financial assessments and conducts an AI-driven analysis of ESG concepts' evolution from 1980 to 2022. Focusing on media sources from the United States and the United Kingdom, the study utilizes the Dow Jones News Article dataset for a comprehensive analysis focused on the environmental domain. The research introduces an innovative information extraction technique, transforming extracted data into a knowledge graph. Key findings highlight recent trends in ESG aspects, with a notable emphasis on climate change, renewable energy sources, and biodiversity conservation in the environmental dimension.

Keywords

ESG, Knowledge Graph, Monitoring Tool, Extraction Pipeline

1. Introduction

Over the past few years, there has been a growing significance in using Environmental, Social, and Governance (ESG) criteria for assessing financial investments[1]. The European Parliament has recognized the importance of ESG ratings in its legislative endeavors to foster an economy that truly serves the interests of the people. This recognition has resulted in specific initiatives,

SemTech4STLD '24: Second International Workshop on Semantic Technologies for Scientific, Technical and Legal Data, May 26th, 2024, Hersonissos, Greece.
*Corresponding author.
†These authors contributed equally.
simone.angioni@unica.it (S. Angioni); sergio.consoli@ec.europa.eu (S. Consoli); danilo.dessi@gesis.org (D. Dessí); francesco.osborne@open.ac.uk (F. Osborne); diego.reforgiato@unica.it (D. R. Recupero); angelo.salatino@open.ac.uk (A. Salatino)
0000-0002-6682-3419 (S. Angioni); 0000-0001-7357-5858 (S. Consoli); 0000-0003-3843-3285 (D. Dessí); 0000-0001-6557-3131 (F. Osborne); 0000-0001-8646-6183 (D. R. Recupero); 0000-0002-4763-3943 (A. Salatino)

[1]Environmental, Social and Governance (ESG) rating activities - https://www.europarl.europa.eu/legislative-train/theme-an-economy-that-works-for-people/file-esg-rating

including the implementation of the *EU taxonomy for sustainable activities*[2], a resource aiming to define a set of ESG standards for organizational conduct. It serves as a valuable tool for socially conscious investors assessing potential investments.

Monitoring and analyzing the portrayal and evolution of ESG-related concepts is crucial for assessing changing perceptions in media and public opinion on sustainability and diversity [1]. While various news monitoring tools, such as Brandwatch[3], Brand24[4], Repustate[5], Cision Communication Cloud[6], SentiOne[7], and Meltwater[8] are available for news analysis, current systems lack a sufficient representation of the nuanced dynamics of discourse. This deficiency hinders their capability to support advanced queries related to entities mentioned in news articles, limiting their ability to perform a comprehensive analysis of ESG discourse.

To address this constraint, researchers have proposed different methods to create structured, interconnected, and machine-readable data frameworks for analyzing news [2, 3]. Over the past few years, knowledge graphs (KGs) have gained growing recognition for their capacity to structure data in a semantically meaningful manner, offering valuable assistance to diverse AI systems across domains like medicine, research, education, robotics, manufacturing, social media, and beyond [4]. Large-scale knowledge graphs are often created through a process that combines both structured and unstructured data, which is partially automated. When dealing with extensive textual data, these methods commonly employ a range of natural language processing techniques to create triples that capture essential concepts within a specific domain [5] and optionally refined using a variety of link prediction techniques [6]. This approach has been applied across various fields, producing a variety of knowledge graph of research articles [7, 8], medical data [9], tourism-related information [10], educational materials [11], and social media posts [12]. These knowledge graphs are capable of facilitating a variety of intelligent services, such as conversational agents [13] and analytical dashboards [14, 15], in addition to supporting extensive domain analysis [16, 17, 18].

In this paper, we present an AI-powered examination of ESG concepts and their development spanning 1980 to 2022. The focus is on media outlets in the United States and the United Kingdom, encompassing notable publications like The Guardian, The New York Times, and The Times. The primary dataset employed for this investigation is the Dow Jones News Article dataset[9], recognized for its extensive and high-quality compilation of news articles.

Our approach utilizes advanced information extraction techniques to condense relevant information from articles into structured statements, represented as triples (`<subject, predicate, object>`). The operational pipeline developed for this process is versatile. It can be implemented on a standard server, eliminating the need for extensive computational resources typically required by current large-scale language models for processing vast data sets. The primary advantage of this innovative approach lies in its capacity to analyze various entity types (e.g.,

[2]EU taxonomy for sustainable activities - https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en
[3]Brandwatch - https://www.brandwatch.com/
[4]Brand24 - https://brand24.com/
[5]Repustate - https://www.repustate.com/
[6]Cision Communication Cloud - https://www.cision.com/
[7]SentiOne - https://sentione.com/
[8]Meltwater - https://www.meltwater.com/
[9]Dow Jones News Article dataset - https://developer.dowjones.com/datasets/details/news

organizations, persons, topics) while establishing meaningful relationships between entities based on predicates extracted from the articles. Consequently, it serves as an effective tool for analyzing substantial volumes of news content, gaining insights into key concepts, and comprehending the evolution of discourse over time.

In detail, our paper contributes in the following ways:

- We offer an AI-driven analysis of the news discourse on ESG concepts spanning from 1980 to 2022.
- We introduce a comprehensive and automated pipeline designed for creating a Knowledge Graph (KG) from a collection of news documents.
- We provide various analytics on ESG concepts, delving into entities and statements derived from a KG extracted from the news.

Section 2 explores prior works related to KGs in news. Section 3 outlines the general pipeline utilized for KG generation. In Section 4, the data source is outlined, and an overview of the resultant KG centered on ESG aspects is provided. Section 5 delves into the analysis results. Finally, Section 6 concludes the paper.

2. Related Work

The central aim of incorporating Knowledge Graphs (KGs) into news analysis is to depict and establish connections among diverse entities within the news domain, encompassing individuals, locations, events, topics, and factual information. This systematic representation enables a more insightful examination of shifts in discourse over time. For example, Al-Obeidat et al. [19] constructed a KG focused on COVID-19-related news, providing a platform for researchers and data analysts to address the challenges posed by the pandemic. Gangopadhyay et al. [20] analyzed a knowledge graph of online claims showing how misinformation can spread on the web and the extensive work for verifying online discussed facts. Tan et al. [3] concentrated on electronics and supply-chain industry news to develop a KG emphasizing causal relations, aiding companies in informed decision-making. Liu et al. [21] proposed a KG-based news recommendation system, incorporating topic context, user interactions, and relevant relations. Rospocher et al. [22] established an event-centric knowledge graph rooted in news sources, emphasizing a temporal dimension for comprehensive entity histories. Fu et al. [23] devised a multi-domain KG for enhanced fake news detection, leveraging semantic links and background information. This tool surpasses existing techniques by effectively generalizing across single, mixed, and multiple domains. Opdahl et al. [24] conducted a survey on research methods utilizing semantic KGs for news production, distribution, and consumption.

In contrast, our paper introduces a KG as a pivotal element in our AI-driven analysis of ESG sectors. Uniquely, our KG is crafted to specifically facilitate the exploration of the evolution of ESG discourse over time, providing a detailed representation of various entity types.

3. The Adopted Pipeline

The constructed pipeline comprises two main phases. In its initial stage, a *Text Parsing Module* is employed to extract entities and their relationships from a set of news articles. The subsequent stage involves a three-step process to generate the knowledge graph. Firstly, the *Entity Extraction Module* identifies crucial entities and classifies them by type. Subsequently, the *Relationship Extraction Module* discerns relationships among these entities from the news articles. Lastly, the *Triple Refinement Module* concludes the process by refining the resulting triples, yielding the finalized knowledge graph.

The Text Parsing Module relies on the Stanford CoreNLP[10] suite, an extensive collection of natural language processing tools developed in Java. Utilizing the Part-of-Speech (PoS) Tagger, this module assigns tags to each word in the provided text, identifying and classifying tokens based on their grammatical categories (e.g., preposition (PRP), verb (VB), noun (NN), adjective (JJ), etc.). Additionally, the module constructs a dependency tree for each sentence.

Within the Entity Extraction Submodule, nominal phrases are identified as entities for the knowledge graph. Nominal phrases constitute word groups with a noun or pronoun as the primary word, accompanied by modifiers, determiners, and complements offering additional information about the noun (e.g., 'long news article'). On the other hand, the Relationship Extraction Submodule identifies connections between entities. For every sentence s, all the shortest paths of the dependency tree between each pair of entities $\langle e_i, e_j)|e_i, e_j \in E_s$ containing a verb are selected. This process yields various types of paths between entities, and the analysis of these paths determines the most suitable ones for identifying relationships in the given context. The paths used in [25] were employed for this purpose.

The Triple Extraction Submodule performs three primary tasks: 1) relation refinement, 2) entity refinement, and 3) triple refinement. The set of triples T, generated in the preceding step, may include triples with similar meanings but expressed through different verbs, for example, `<company, build, 200-unit motel>`, `<company, construct, buildings>`, `<craftsmen, create, accommodation>`. Relation refinement aims to identify the most suitable predicate label r for each relation verb v in a triple $< e_m, v, e_n >$ and map v to r in the resulting triple. This phase reduces the space of possible relationships by analyzing the resultant verbs and clustering them into a more concise set of well-defined relationships [25]. The method was applied to the 393 verbs found in all the triples extracted from the ESG news dataset, resulting in a final set of 57 predicates.

The Entity Refinement module establishes an index based on the tokens contained within the entities. This index links each token to all entities that contain it. For instance, the token *Obama* is linked to entities such as *Barack Obama*, *President Obama*, *former President Barack Obama*, *Barack Obama's Administration*, *Michelle Obama*, and so on. Entities e_i and $e_j \in E$ are compared if they share at least one token. This comparison is executed using the state-of-the-art framework SentenceTransformers[11], encoding the entities with the *all-mpnet-base-v2*[12] transformer model. If the cosine similarity between entity e_i and e_j exceeds 0.9, they are grouped into the same cluster. The reader notices that this value has been selected based on an empirical analysis that considered the values 0.7, 0.8, 0.85, 0.9, and 0.95.

For the Triple Refinement module, akin to the entity refinement step, a sentence transformer model is used to detect and merge triples with the same meaning. As a final step, the resulting

---

[10]Stanford CoreNLP - https://stanfordnlp.github.io/CoreNLP/
[11]SentenceTransformers - https://huggingface.co/sentence-transformers
[12]all-mpnet-base-v2 - https://huggingface.co/sentence-transformers/all-mpnet-base-v2

triples are linked to the original papers and employed to construct the knowledge graph. Each triple is associated with its *support*, indicating the number of news articles from which it was extracted. To evaluate the pipeline's accuracy, a sample of triples underwent assessment by three reviewers, considering both the triple and the original sentences from which it originated. The reviewers marked the triple as 1 if it accurately reflected the news articles' content and 0 if it did not. The average agreement between annotators was 0.89, indicating substantial consensus. The pipeline's accuracy, evaluated against the majority vote of the three annotators for 200 statements, was 0.85, with individual rater estimates ranging from 0.85 to 0.93, demonstrating the pipeline's ability to extract triples with high accuracy. More information about this pipeline can be found in [26].

4. The Data Source and the Generated ESG Knowledge Graph

The Dow Jones News Datasets encompass an extensive compilation of 15, 105, 283 news articles spanning various languages. The dataset incorporates 13 English sources, including renowned ones like The Wall Street Journal, New York Times, and The Guardian, contributing to a total of 7.3 million distinct news items. In assembling a repository of news articles concerning ESG, we considered all news from 1980 to 2022 containing keywords related to Environmental, Social, and Governance, either within the text body or metadata fields. The ultimate collection comprises approximately 850, 000 news articles, distributed as 500, 000 on environmental topics, 290, 000 on social issues, and 60, 000 on governance. The pipeline detailed in Section 3 was applied to this set of 850, 000 ESG news articles, resulting in a KG that includes over 7.2M statements and 4M entities.

For structuring the statements, we utilized a lightweight ontology tailored to the primary purpose of aiding news analysis. The ontology defines four main classes: i) aggregated statement, ii) fine-grained statement, iii) News, and iv) Entity. It also specifies 57 object properties derived from the predicates outlined in Section 3. Furthermore, the ontology maps the statements using the original verb alongside their version utilizing the 57 predicates obtained by clustering them.

Each statement in ESG-KG incorporates: - *rdf:subject*, *rdf:predicate*, and *rdf:object*, providing the reification of triples within an rdf:Statement; - *provo:wasDerivedFrom*, supplying provenance information and listing the DNA-IDs of the news from which the statement is derived; - *esg-kg:statement_negated*, a boolean indicating whether the statement was derived from a negative sentence (True) or not (False); - *esg-kg:original_triple*, listing the fine-grained versions of the statement.

Additionally, each news ID is linked to *xsd:date*, offering the news publishing date, and *esg-kg:source*, providing the original journal source name.

5. Exploration of the News Discourse on ESG

The analysis presented in this section relies on various analytics derived from the knowledge graph.

Figure 1 illustrates the distribution of the three primary topics (Environmental, Social, and Governance) over time. Historically, environmental issues have consistently dominated, rep-

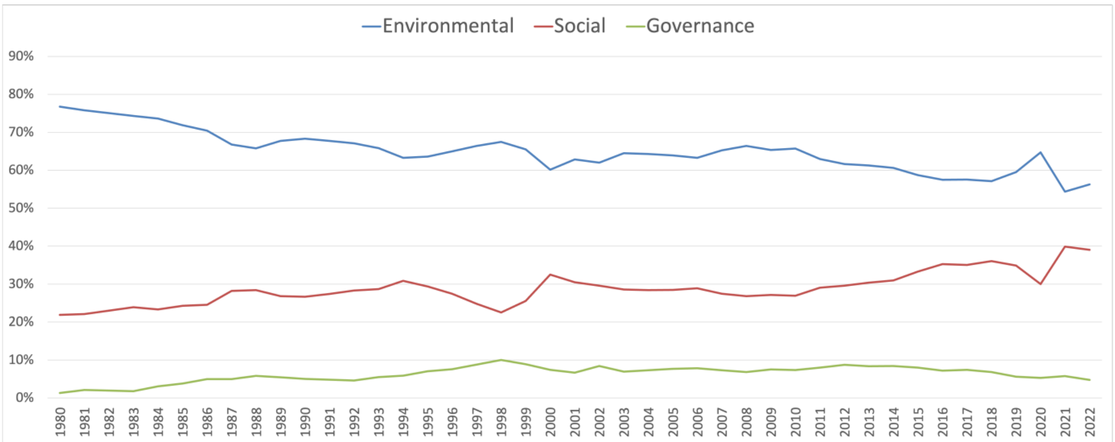

Figure 1: News Distribution for Year

Table 1
Categories of entities and their occurrence. GPE denotes Geopolitical Entities, while NORP encompasses nationalities, religious, and political groups.

Entity Type	**# Entities**
PERSON	519K
ORG	359K
CARDINAL	244K
GPE	232K
NORP	116K
DATE	101K
LOC	32K
ORDINAL	30K
PERCENT	19K
TIME	16K
MONEY	7.2K
EVENT	2K
OTHER	2.2M

resenting 55% to 75% of coverage. However, a noticeable trend shows an increasing focus on social issues, growing from approximately 20% in 1980 to nearly 40% by 2022. This shift seems to be propelled by heightened interest in subjects such as ethics, racism, gender identity, and global human rights.

The environmental component encompasses 2 million entities, the social aspect covers 361,000 entities, and the governance section comprises 209,000 entities. These entities are categorized based on the Named Entity Recognition (NER) tool provided by Spacy, and their types and frequencies are detailed in Table 1.

The KG includes 3.8 million statements: 3 million statements about environmental topics, 600,000 related to social issues, and 236,000 concerning governance.

We present the foremost ten political groups (Table 2), geopolitical entities (Table 3), prominent individuals (Table 4), and organizations (Table 5) for each category.

An impactful observation underscores the USA's significant role in ESG discourse, evident in the top three groups comprising Democrats, Republicans, and Americans. Moreover, frequently mentioned individuals center around US Presidents, encompassing figures like Bush, Obama, Clinton, and Trump. The United States emerges as the most cited country in articles related to

Table 2
Top 10 NORP Entities.

environmental		social		governance	
Democrats	7457	**Democrats**	1699	**Democrats**	434
Republicans	6818	**Republicans**	1435	**Republicans**	369
Americans	5111	**Americans**	1324	**Americans**	321
Russians	1273	**Jews**	460	**European union**	118
Democrat	901	**Palestinians**	405	**Japanese corporation**	86
Palestinians	894	**Chinese government**	376	**European commission**	65
Europeans	741	**Russians**	345	**Chinese authority**	48
Germans	495	**Muslims**	317	**European community**	35
British government	249	**African**	291	**British company**	25
Italians	214	**Islamic State**	247	**German company**	23

Table 3
Top 10 Geopolitical Entities.

environmental		social		governance	
United States	13587	**United States**	5469	**China**	2078
China	7600	**China**	4310	**United States**	1355
California	5029	**Russia**	1822	**Russia**	466
Russia	4527	**U.S.**	1811	**Beijing**	423
U.S.	4358	**Israel**	1588	**Japan**	272
Washington	3877	**Britain**	1434	**America**	261
America	3308	**Washington**	1369	**Britain**	246
New York	3005	**America**	1166	**India**	214
Israel	2989	**UK**	932	**UK**	191
Japan	2729	**France**	712	**Australia**	146

Table 4
Top 10 Person Entities.

environmental		social		governance	
Bush	4727	**Biden**	1209	**Trump**	221
Obama	3478	**Trump**	1007	**Bush**	180
Clinton	2162	**Bush**	890	**Obama**	162
Johnson	1812	**Johnson**	688	**Johnson**	159
Brown	1709	**Obama**	667	**Clinton**	156
Trump	1705	**Clinton**	595	**Brown**	85
Mr. Reagan	1466	**Brown**	455	**Greg Abbott**	78
President Reagan	947	**Putin**	444	**Mrs. Clinton**	72
McCain	792	**Jackson**	421	**Mr. Smith**	70
Miller	704	**Harry**	306	**Mr. Dimon**	66

Environmental and Social aspects, underlining its perceived leadership status in these domains. Conversely, China takes the lead in discussions on Governance, with Beijing and Japan often mentioned in conversations about employee rights, securing the fourth and fifth positions among the most referenced countries for Governance.

Our exploration then delves into the Environmental domain, scrutinizing the evolving trends of key entities. This involves computing the annual frequency of each entity and applying

Table 5
Top 10 Organization Entities.

environmental		social		governance	
Congress	12828	**Congress**	2807	**Congress**	1431
EPA	6491	**Taliban**	963	**Microsoft**	1212
White House	4243	**White House**	924	**Apple**	882
Senate	4062	**Senate**	849	**Google**	819
House	3585	**United Nations**	804	**SEC**	706
NASA	2649	**Supreme Court**	707	**White House**	417
Environmental Protection Agency	2242	**House**	661	**Justice Department**	396
Fed	2148	**State Department**	628	**Intel**	395
Ford	1810	**EU**	612	**Facebook**	390
Dow Jones industrial average	1765	**Facebook**	523	**Fed**	385

linear regression to discern the trajectory of these entities' yearly distributions. The slope of the regression line serves as an indicator of the trend's momentum, with a steeper slope signifying a more rapid surge in media coverage for the specified entity. This analytical technique, commonly employed to detect key trends, finds application in areas such as research topics [27]. Results are presented in Table 6, where *slope_10* denotes the trend over the last 10 years, and *slope_5* indicates the trend over the last 5 years. To accentuate common themes, related entities have been manually clustered and highlighted in the same color.

Table 6
Environmental entities

Category	Entity	Freq.	Slope 10 Years	Slope 5 Years
Climate & Carbon Emissions	climate change	4818	58.01	99.20
	temperature	4047	20.47	51
	carbon emission	910	5.33	16.40
	coal	1492	2.56	11
	greenhouse gas emission	919	3.56	9.60
	global warming	2831	1.76	7.20
	natural gas	2508	−9.63	5.10
	rising temperature	184	2	4.70
	oil industry	585	−1.84	4.40
	greenhouse gas	959	0.07	1.90
	global temperature	145	0.38	0.70
	air pollution	1073	4.57	0.30
Renewable Energy	solar panel	834	2.13	8.90
	wind power	130	0.42	2.30
	wind farm	199	0.62	1.90
	solar farm	43	0.97	1.60
	energy efficiency	444	−1.07	1.30
	solar energy	251	0.22	1
	solar power	253	−0.06	1
	wind energy	45	−0.29	0.40
Biodiversity & Land Use	deforestation	223	1.68	4
	fertiliser	22	0.94	2.30
	biodiversity	249	3.03	2.20

The environmental facet of ESG centers on assessing a company's influence on the natural world and its approach to managing environmental risks. Table 6 presents entities that have exhibited notable increases in mentions over the past decade and the last five years.

We examined the network of entities and their statements, identifying three key topics that have gained prominence in recent years:

- **Climate** and **Carbon Emissions** 6: Central to discussions involving the measurement of carbon footprints, implementation of initiatives to reduce greenhouse gas emissions, and the formulation of strategies to mitigate the effects of climate change [28]. The increased visibility of these entities in news narratives underscores the growing importance of taking concrete measures to combat climate change.
- **Renewable Energy**: Positive trends suggest a rising emphasis on the use of renewable energy sources and the adoption of energy-saving practices in public discourse. This shift towards energy efficiency reflects broader societal and economic recognition of the benefits associated with sustainable energy practices. With the increasing urgency to address climate change, the push for more efficient energy usage and the transition to renewables becomes a central theme in policy, industry, and community conversations [29].
- **Biodiversity** and **Land Use**: The upward trend over the past five years underscores the significance of these issues. Specifically, the focus on deforestation and biodiversity highlights the media's growing concern and interest in the conservation of natural habitats, ecosystems, and biodiversity. This focus aligns with global efforts to achieve biodiversity conservation targets and sustainable development goals [30], emphasizing the need for a holistic approach to environmental stewardship that includes protecting diverse ecosystems and ensuring responsible land use.

6. Conclusions and Future Works

In this manuscript, we provide a preliminary examination of ESG concepts and their evolutionary trajectory spanning from 1980 to 2022, concentrating on news content sourced from the United States and the United Kingdom. Employing the Dow Jones Article dataset, our investigation encompasses news articles from well-known newspapers like The Guardian, The New York Times, and The Times. To execute this analysis, we initially applied an extraction pipeline to the news articles, involving the organization of extracted data into a Knowledge Graph (KG). The methodology employed advanced information extraction techniques to distill pertinent information from articles into structured statements represented as triples. These triples underwent aggregation, and verification, and were used in constructing a comprehensive knowledge graph. The implemented pipeline is versatile, applicable across domains, and facilitates the analysis of various entity types while establishing semantic relationships between them based on information extracted from news articles. The information extraction pipeline underwent rigorous evaluation by three annotators, achieving an accuracy of 0.85. Subsequently, the resulting knowledge graph was utilized to scrutinize the three core components of Environmental, Social, and Governance (ESG), with a specific focus on the environmental domain. In future work, we plan to overcome some limitations that we have encountered to further improve the generation pipeline. First, we would like to experiment and develop a novel model to merge entity mentions that refer to the same entity (for example, in Table 4 *Mr. Reagan* and *President Reagan* are not merged). Second, we would like to assign news into categories to make it simple the explore both the news and the KG content. Finally, we intend to explore additional datasets beyond Dow Jones Article dataset to enrich our analysis and validate the findings across different sources, thereby enhancing the robustness and generalizability of our research.

Acknowledgements

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - Call for tender No.3277 published on December 30, 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union – NextGenerationEU. Project Code ECS0000038 – Project Title eINS Ecosystem of Innovation for Next Generation Sardinia – CUP F53C22000430001- Grant Assignment Decree No. 1056 adopted on June 23, 2022 by the Italian Ministry of University and Research (MUR).

References

[1] LISI, UK and EU – An analysis of ESG reporting requirements and trends, Technical Report, 2023. URL: https://www.lisi-law.eu/resources/uk-and-eu-an-analysis-of-esg-reporting-requirements-and-trends.

[2] A. L. Opdahl, T. Al-Moslmi, D.-T. Dang-Nguyen, M. Gallofré Ocaña, B. Tessem, C. Veres, Semantic knowledge graphs for the news: A review, ACM Comput. Surv. 55 (2022). URL: https://doi.org/10.1145/3543508. doi:10.1145/3543508.

[3] F. A. Tan, D. Paul, S. Yamaura, M. Koji, S.-K. Ng, Constructing and interpreting causal knowledge graphs from news, 2023. arXiv:2305.09359.

[4] C. Peng, F. Xia, M. Naseriparsa, F. Osborne, Knowledge graphs: opportunities and challenges, Artificial Intelligence Review (2023) 1–32.

[5] D. Dessì, F. Osborne, D. R. Recupero, D. Buscaldi, E. Motta, Generating knowledge graphs by employing natural language processing and machine learning techniques within the scholarly domain, Future Generation Computer Systems 116 (2021) 253–264.

[6] M. Nayyeri, G. M. Cil, S. Vahdati, F. Osborne, M. Rahman, S. Angioni, A. Salatino, D. R. Recupero, N. Vassilyeva, E. Motta, et al., Trans4e: Link prediction on scholarly knowledge graphs, Neurocomputing 461 (2021) 530–542.

[7] D. Dessí, F. Osborne, D. Reforgiato Recupero, D. Buscaldi, E. Motta, Cs-kg: A large-scale knowledge graph of research entities and claims in computer science, in: International Semantic Web Conference, Springer, 2022, pp. 678–696.

[8] S. Angioni, A. Salatino, F. Osborne, D. R. Recupero, E. Motta, Aida: A knowledge graph about research dynamics in academia and industry, Quantitative Science Studies 2 (2021) 1356–1398.

[9] F. Michel, F. Gandon, V. Ah-Kane, A. Bobasheva, E. Cabrio, O. Corby, R. Gazzotti, A. Giboin, S. Marro, T. Mayer, et al., Covid-on-the-web: Knowledge graph and services to advance covid-19 research, in: The Semantic Web–ISWC 2020: 19th International Semantic Web Conference, Athens, Greece, November 2–6, 2020, Proceedings, Part II 19, Springer, 2020, pp. 294–310.

[10] A. Chessa, G. Fenu, E. Motta, F. Osborne, D. R. Recupero, A. Salatino, L. Secchi, Data-driven

methodology for knowledge graph generation within the tourism domain, IEEE Access (2023).
[11] M. Rizun, et al., Knowledge graph application in education: A literature review, Acta Universitatis Lodziensis. Folia Oeconomica 3 (2019) 7–19.
[12] A. Tchechmedjiev, P. Fafalios, K. Boland, M. Gasquet, M. Zloch, B. Zapilko, S. Dietze, K. Todorov, Claimskg: A knowledge graph of fact-checked claims, in: The Semantic Web–ISWC 2019: 18th International Semantic Web Conference, Auckland, New Zealand, October 26–30, 2019, Proceedings, Part II 18, Springer, 2019, pp. 309–324.
[13] A. Meloni, S. Angioni, A. Salatino, F. Osborne, D. R. Recupero, E. Motta, Integrating conversational agents and knowledge graphs within the scholarly domain, Ieee Access 11 (2023) 22468–22489.
[14] T. M. Nguyen, H.-W. Chun, M. Hwang, L.-N. Kwon, J.-M. Lee, K. Park, J. J. Jung, Sociopedia+: a visual analytics system for social knowledge graph-based event exploration, PeerJ Computer Science 9 (2023) e1277.
[15] S. Angioni, A. Salatino, F. Osborne, A. Birukou, D. R. Recupero, E. Motta, Leveraging knowledge graph technologies to assess journals and conferences at springer nature, in: International Semantic Web Conference, Springer, 2022, pp. 735–752.
[16] A. ZEKIYE, A. ALPKOCAK, Extracting and analyzing covid-19-related entities and relationships: A knowledge graph approach, in: Proceedings of 14th Turkish Congress of Medical Informatics, volume 16, 2023.
[17] D. Dessí, R. Dessí, Diving into knowledge graphs for patents: Open challenges and benefits (2022).
[18] P. Manghi, A. Mannocci, F. Osborne, D. Sacharidis, A. Salatino, T. Vergoulis, New trends in scientific knowledge graphs and research impact assessment, 2021.
[19] F. Al-Obeidat, O. Adedugbe, A. B. Hani, E. Benkhelifa, M. Majdalawieh, Cone-kg: A semantic knowledge graph with news content and social context for studying covid-19 news articles on social media, in: 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS), 2020, pp. 1–7. doi:10.1109/SNAMS52053.2020.9336541.
[20] S. Gangopadhyay, K. Boland, D. Dessí, S. Dietze, P. Fafalios, A. Tchechmedjiev, K. Todorov, H. Jabeen, Truth or dare: Investigating claims truthfulness with claimskg (2023).
[21] D. Liu, T. Bai, J. Lian, G. Sun, W. X. Zhao, J. rong Wen, X. Xie, News graph: An enhanced knowledge graph for news recommendation, in: KaRS@CIKM, 2019. URL: https://api.semanticscholar.org/CorpusID:204777685.
[22] M. Rospocher, M. Van Erp, P. Vossen, A. Fokkens, I. Aldabe, G. Rigau, A. Soroa, T. Ploeger, T. Bogaard, Building event-centric knowledge graphs from news, Journal of Web Semantics 37 (2016) 132–151.
[23] L. Fu, H. Peng, S. Liu, Kg-mfend: an efficient knowledge graph-based model for multi-domain fake news detection, The Journal of Supercomputing (2023) 1–28.
[24] A. L. Opdahl, T. Al-Moslmi, D.-T. Dang-Nguyen, M. Gallofré Ocaña, B. Tessem, C. Veres, Semantic knowledge graphs for the news: A review, ACM Comput. Surv. 55 (2022). URL: https://doi.org/10.1145/3543508. doi:10.1145/3543508.
[25] D. Dessí, F. Osborne, D. R. Recupero, D. Buscaldi, E. Motta, Scicero: A deep learning and nlp approach for generating scientific knowledge graphs in the computer science domain,

Knowledge-Based Systems 258 (2022) 109945.
[26] S. Angioni, S. Consoli, D. Dessí, F. Osborne, D. R. Recupero, A. Salatino, Exploring environmental, social, and governance (esg) discourse in news: An ai-powered investigation through knowledge graph analysis, IEEE Access (2024) 1–1. doi:10.1109/ACCESS.2024.3407188.
[27] A. A. Salatino, F. Osborne, E. Motta, How are topics born? understanding the research dynamics preceding the emergence of new areas, PeerJ Computer Science 3 (2017) e119.
[28] M. Kabir, U. E. Habiba, W. Khan, A. Shah, S. Rahim, P. R. D. los Rios-Escalante, Z.-U.-R. Farooqi, L. Ali, M. Shafiq, Climate change due to increasing concentration of carbon dioxide and its impacts on environment in 21st century; a mini review, Journal of King Saud University - Science 35 (2023) 102693. URL: https://www.sciencedirect.com/science/article/pii/S1018364723001556. doi:https://doi.org/10.1016/j.jksus.2023.102693.
[29] A. Slameršak, G. Kallis, D. W. O'Neill, Energy requirements and carbon emissions for a low-carbon energy transition, Nature Communications 13 (2022) 6932. URL: https://doi.org/10.1038/s41467-022-33976-5. doi:10.1038/s41467-022-33976-5.
[30] A. Opoku, Biodiversity and the built environment: Implications for the sustainable development goals (sdgs), Resources, Conservation and Recycling 141 (2019) 1–7. URL: https://www.sciencedirect.com/science/article/pii/S0921344918303768. doi:https://doi.org/10.1016/j.resconrec.2018.10.011.