sustainability

Review

Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review

Wayne Moodaley * and Arnesh Telukdarie

Johannesburg Business School, University of Johannesburg, Johannesburg 2092, South Africa
* Correspondence: wayne.moodaley@jbs.ac.za

Abstract: The rise of stakeholder interest globally in sustainable business practices has resulted in a rise in demands from stakeholders that companies report on the environmental and social impacts of their business activities. In certain cases, however, companies have resorted to the practice of providing inaccurate disclosures regarding sustainability as part of their corporate communications and sustainability reporting—commonly referred to as "greenwashing". Concurrently, technological improvements in artificial intelligence have presented the means to rapidly and accurately analyze large volumes of text-based information, such as that contained in sustainability reports. Despite the possible impacts of artificial intelligence and machine learning on the fields of greenwashing and sustainability reporting, no literature to date has comprehensively and holistically addressed the interrelationship between these three important topics. This paper contributes to the body of knowledge by using bibliometric and thematic analyses to systematically analyze the interrelationship between those fields. The analysis is also used to conjecture a conceptual and thematic framework for the use of artificial intelligence with machine learning in relation to greenwashing and company sustainability reporting. This paper finds that the use of artificial intelligence in relation to greenwashing, and greenwashing within sustainability reporting, is an underexplored research field.

Keywords: greenwashing; sustainability reporting; artificial intelligence; machine learning; sustainability

check for updates

Citation: Moodaley, W.; Telukdarie, A. Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review. *Sustainability* **2023**, *15*, 1481. https://doi.org/10.3390/su15021481

Academic Editor: Evangelos Katsamakas

Received: 22 November 2022
Revised: 6 January 2023
Accepted: 10 January 2023
Published: 12 January 2023

1. Introduction

Sustainability as both a practice and concept is complex, particularly in business. Environmental sustainability implies the stewardship of resources by companies for current and future generations, which in turn involves appropriate measurement and evaluation of such impacts [1]. Social sustainability incorporates considerations regarding people, communities, and broader society. While clear consensus on the definition of social sustainability is not present, in the context of business it is frequently defined in relation to minimizing the negative, and maximizing the positive impacts of a company's operations [2]. Economic sustainability, in turn, reflects economic progress which is not at the expense of the environment, people, or society [3]. Jurisdictions such as the European Union have recognized the importance of these elements of sustainability and adopted initiatives such as the EU Circular Economy Action Plan [4].

The practice of sustainability reporting reflects increasing public awareness regarding the importance of sustainability, and an increase in stakeholder expectations and pressure for companies to disclose the impact of their operations on the environment and society. In that context, sustainability reporting has become increasingly important for companies to signal their commitment to stakeholders regarding environmental, social, and governance (ESG) factors [5].

The Global Reporting Initiative (GRI) states that the "foundation of sustainability reporting is for an organization to identify and prioritize its impacts on the economy, environment, and people to be transparent about their impacts" [6]. Many companies now choose to report on ESG factors in sustainability reports [7].

Within that context, the use of artificial intelligence (AI) and machine learning (ML) has fundamentally changed the landscape of business and society in recent years [8]. The progression in the field of sustainability reporting has coincided with the rise in interest in, and use of, artificial intelligence—both within business and society, as well as academic literature. The term "Artificial Intelligence" stems from the 1956 Dartmouth summer research project proposal on artificial intelligence, in which the authors, John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, proposed a study on "the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" [9].

Since then, the field has experienced a dramatic evolution, and the last decade has seen significant improvements and breakthroughs relating to data, computing technology, and algorithms [10]. Goodell et al. [8] reference this evolution and indicate that the evolution relates to technology that is able to tackle tasks at the level of difficulty at which humans are able to operate. A steady increase in academic literature relating to AI has accompanied this evolution [11].

With the increased interest in sustainability reporting and the impact of company actions on ESG factors, has come an increase in the prevalence of misleading information regarding such impact—commonly referred to as greenwashing [5]. In and Schumacher [12] contend that "companies have misaligned incentives to deliberately or selectively communicate information not matched with actual environmental impacts or make largely unsubstantiated promises around future ambitions".

The term greenwashing was coined by Jay Westerveld in 1986 [13]. Westerveld first used the term in relation to hotel signs requesting that guests hang up used towels for environmental reasons, and posited that these requests related more to the hotels' desire to reduce the costs of laundry, rather than reducing the impact on the environment. He christened this misleading form of communication "greenwashing" [14].

Large bodies of academic literature are devoted to sustainability reporting, greenwashing, and AI with ML. Certain recent studies have sought to link two of these respective fields to each other, while others link one of these fields to other disciplines. None, however, consider the intersection of all three within a systematic or bibliometric literature review. This paper considers extant literature reviews relating to each intersection below.

Goodell et al. [8] conduct a bibliometric study of the application of artificial intelligence and machine learning in the field of finance. Lombardi and Secundo [15] conduct a systematic literature review to analyze the intersection of reporting, and digital and smart technologies.

Beltrami et al. [16] conduct a systematic literature review to evaluate the link between Industry 4.0 and sustainability, and develop a framework that highlights relationships between those fields—both of which are significantly broader than the fields of sustainability reporting and AI and ML.

Crucially, none of the abovementioned reviews specifically or explicitly study the link between sustainability reporting and AI with ML.

While there is a lack of literature reviews relating to the intersection of sustainability reporting and AI with ML, it is worthwhile noting that recent advancements in ML and natural language processing (NLP) have highlighted the ability of ML and NLP tools in assessing large volumes of complex textual information, such as that contained in sustainability reports [17,18]. Kotzian [19] proposes methods for applying AI with ML to the Corporate Social Responsibility (CSR) domain to detect CSR non-compliance.

Other studies relating to sustainability reporting or greenwashing have incorporated AI with ML as methodological tools. Examples include the use of AI with ML in analyzing sustainability reporting using NLP [20,21], applying machine learning algorithms for topic modeling to analyze firm sustainability reporting disclosures available online [22], measuring the readability of sustainability reports [23], and measuring corporate alignment with the UN Sustainable Development Goals (SDGs) [24].

Similar to the intersection for sustainability reporting and AI with ML, limited systematic literature or bibliometric reviews have been conducted to study the interplay between greenwashing and sustainability reporting. Velte [25] conducts a systematic literature review of research on integrated reporting in relation to the "business case for integrated reporting" and creates a differentiation within the corpus reviewed which the author indicates "is crucial, to stress the challenges of greenwashing policies and information overload". The study is, however, focused on integrated reporting and does not specifically review literature on the intersection between sustainability reporting and greenwashing.

Systematic literature reviews or bibliometric reviews relating to the intersection of greenwashing and AI with ML have not been found.

Given the paucity of comprehensive systematic or bibliometric reviews of literature described above, it is clear that comprehensive analysis of the intersections between greenwashing, sustainability reporting, and AI with ML in academic literature is not present.

As described above, the previous literature relating to sustainability reporting, greenwashing, and AI with ML, have linked one or the other to adjacent fields, or dealt with subsets of each field. Other literature reviews have a much wider purview and different aims, such as literature dealing with the broader topic of Industry 4.0, of which AI with ML are subsets.

The aim of this paper is to infer and develop a clear thematic and conceptual framework relating to the intersection of AI with ML, greenwashing, and sustainability reporting, and specifically, for the use of AI with ML in relation to greenwashing and sustainability reporting. This paper fills the gap in the academic literature by applying bibliometric and thematic analysis to analyze the interrelationship between these fields within that context and identify specific thematic trends.

The remainder of the paper is structured as follows. Section two sets out an overview of each of these fields. Section three details the research methodology, while section four presents a bibliometric and thematic analysis and results. Lastly, section five sets out the conclusions, implications of the research, and limitations of the study.

2. Overview

2.1. Sustainability Reporting

Sustainability reporting is synonymous with a number of different types of reporting, including, inter alia, corporate social responsibility (CSR) reporting, greenhouse gas emissions (GHG) reporting, and UN Sustainable Development Goals (SDG) reporting [26]. In academic literature and practice, the term is also often used interchangeably with corporate social responsibility (CSR) reporting [27], and triple-bottom-line (TBL) reporting [28].

Various bodies globally have responded to stakeholder calls for reliable sustainability reporting standards by developing a plethora of guidance documents and frameworks on disclosures relating to ESG and sustainability. These have included initiatives and guidelines published by, inter alia, the Global Reporting Initiative (GRI), the International Financial Reporting Standards (IFRS) Foundation, the European Financial Reporting Advisory Group (EFRAG), the Sustainability Accounting Standards Board (SASB), the International Integrated Reporting Council (IIRC), and the European Commission [29,30]. Other initiatives include those by the Task Force on Climate Related Financial Disclosure (TFCD), the G20, G7, and World Economic Forum [31].

The diversity in frameworks, standards, and practices relating to sustainability reporting reflects the complexity of sustainability itself, which involves multiple actors, stakeholders, and interests [2]. However, at the heart of sustainability reporting is the goal of holding companies accountable for the environmental, social, and economic impact of their operations [2].

2.2. Greenwashing

Numerous approaches have been followed in an attempt to define what greenwashing is. De Freitas Netto et al. [5] present three forms in terms of the approach followed in academic literature to define greenwashing. The first, greenwashing as selective disclosure, refers to retaining “the disclosure of negative information related to the company’s environmental performance and expose positive information regarding its environmental performance” [5]. The second, “greenwashing as decoupling”, refers to the decoupling of the reality of company behavior on sustainability issues from its communications through the use of, inter alia, “symbolic actions” [5]. The third identifies the use of legitimacy theory and signaling in relation to the practice of greenwashing [5].

Tateishi [32] defines greenwashing in relation to “communication that misleads people (e.g., consumers and stakeholders) regarding environmental performance/benefits”. Lyon and Maxwell [33] define greenwashing more broadly to include social elements, i.e., “selective disclosure of positive information about a company’s environmental or social performance, without full disclosure of negative information on these dimensions, so as to create an overly positive corporate image”.

Greenwashing is therefore, “an umbrella term for a variety of misleading communications and practices that intentionally or not, induce false positive perceptions of an organization’s environmental performance” [34].

2.2.1. Impact of Greenwashing

Testa et al. [35] state that greenwashing undermines “corporate accountability toward stakeholders and the credibility of environmental initiatives”. Various industry and regulatory bodies have acknowledged the negative impact of greenwashing on sustainability efforts.

The European Commission [4] states that greenwashing “misleads market actors and does not give due advantage to those companies that are making the effort to green their products and activities. It ultimately leads to a less green economy”. The former Commissioner of the USA’s Securities and Exchange Commission (SEC), Allison Herron Lee, states that “Greenwashing can mislead investors as to the true risks, rewards, and pricing of investment assets” [36]. The United Kingdom’s Financial Conduct Authority (FCA) states that “Greenwashing misleads consumers and erodes trust in all ESG products” [37].

One of the more prominent greenwashing examples relates to Volkswagen publishing misleading and incorrect information relating to its vehicles’ emissions, which led to significant reputational damage, as well as financial losses, and a drastic decline in the company’s share price [38]. Greenwashing may also lead to a decline in consumer confidence in a company’s brand and products [39].

In response, the EU has proposed changes to the Unfair Commercial Practices Directive to combat greenwashing and misleading green claims [4]. Similarly, the U.S. Securities and Exchange Commission has proposed rules to enhance and standardize climate related disclosures to investors [40]. In her support of the proposal, former SEC Commissioner Allison Herren Lee states the need for “consistent, comparable, and reliable information—information to help protect investors from “greenwashing” or exaggerated or false claims about ESG practices” [40].

In the context of sustainability reporting, greenwashing represents a threat to the accuracy, reliability, and transparency of such reporting, because at the heart of greenwashing is the difference between what a company chooses to disclose (and signal) to stakeholders regarding its performance on ESG and climate-related factors, and its actions in relating to such factors. Steiner et al. [41] express this as the “incongruence between the reputational intention and the actual, real sustainability performance of the company”. Greenwashing is therefore detrimental to the interests of multiple stakeholders, including investors, consumers, and other market actors.

2.2.2. Increase in Greenwashing

Delmas and Burbano [42] reported an increase in the practice of greenwashing by firms. Lyon and Montgomery [43] conduct a review of the greenwashing literature and indicate that both green claims and “the incidence of greenwash” have increased rapidly in recent years. Nemes et al. [44], in their more recent study state the issue more bluntly when they contend that despite "growing awareness" regarding greenwashing, the phenomenon is still "widespread".

Regulators and industry bodies have affirmed this trend. A screening by the European Commission focusing on greenwashing analyzed online claims [45]. The findings indicate that "authorities had reason to believe that in 42% of cases the claims were exaggerated, false or deceptive and could potentially qualify as unfair commercial practices under EU rules", and stated that greenwashing "has increased as consumers increasingly seek to buy environmentally sound products" [45]. Varying definitions of greenwashing make the identification of greenwashing more complex [46].

2.2.3. Increase in Greenwashing Academic Literature

Besides the increasing trend in the practice of greenwashing, recent years have seen a significant increase in the academic literature relating to the topic of greenwashing. Montero-Nevarro et al. [47] analyze academic literature on greenwashing in the agriculture, food, and retail industries, and identify a significant increase in greenwashing literature from 2016. They identify three distinct periods for this trend, with the last being a period of significant growth between 2016 and 2020. The increase reflects growing interest in greenwashing as a topic of research [47].

Lyon and Montgomery [43], in their review of the greenwashing literature, identify a "sharp increase in articles since 2011". Lu Zhang et al. [48] state that "references to greenwashing in the literature has increased rapidly and the types and consequences of greenwashing have become a research hotspot".

This trend is supported by the work of Yang et al. [49] who identify an increasing trend in academic literature relating to greenwashing from 2010, and Andreoli et al. [39] in their bibliometric analysis of the field. De Freitas Netto et al. [5], based on a systematic review of the literature, identify a rise in academic interest in greenwashing. Pope and Wæraas [50] contend that that the growth in academic discourse relating to greenwashing has been significant.

2.3. Artificial Intelligence and Machine Learning

AI presents multiple opportunities and possibilities to change society, whether in the production of goods or services, in broader business, or in shaping and changing approaches to environmental and social issues [11,51]. The field is becoming increasingly relevant in terms of both its impacts on, and transformations of, various fields [11,52]. Within the context of sustainability, AI may have both positive and negative impacts on societal, economic, and environmental outcomes [53] which reflects its importance in relation to the field of sustainability, and concomitantly, sustainability reporting.

A subfield of AI that has emerged is machine learning (ML). ML uses algorithms to recognize patterns, make decisions, and imitate the way that humans learn and solve problems [7,8,52]. ML, like AI, has become increasingly prominent in the academic literature, as reflected by increases in the occurrences of the term as well as its evolution toward being a term that is sometimes viewed as being distinct or autonomous from AI within academic literature [11].

Natural language processing (NLP), itself a subset of AI and ML, uses algorithms and ML to analyze text, and allows analysis of large quantities of data much faster than manual analysis and review [54]. Recent bibliometric studies identify NLP as an emerging topic requiring more research in future [11], while scholarly tools, such as VosViewer, employ NLP to conduct bibliometric analyses [55].

Text mining is related to NLP as it "is the process of transforming unstructured text into structured data for easy analysis", and "uses natural language processing tools to interpret the human language and process text documents automatically" [56].

3. Materials and Methods

The methodology applied is designed to ensure transparency and replicability of the study. To achieve those aims, this paper follows a bibliometric analysis approach, combined with thematic analysis. Bibliometric analysis as a research method has become more prominent in academic research and allows for the analysis of large quantities of bibliometric information to identify themes and trends [8,26]. Thematic analysis is the process of "identifying, analyzing, and reporting patterns (themes) within data" [57].

A five-step process was followed for the bibliometric analysis, presented graphically in Figure 1 below:

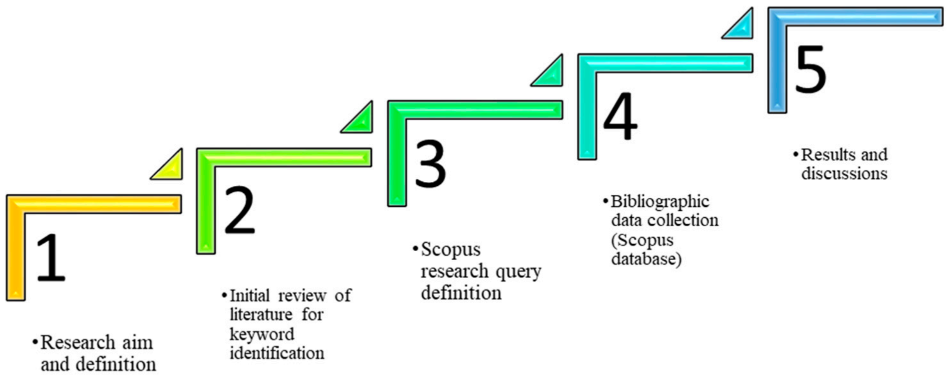

Figure 1. Analysis process.

The bibliometric process incorporates an approach recommended by Donthu et al. [58] for bibliometric analysis, which begins with the definition of the research aim. This paper applies RStudio and Biblioshiny, which are powerful bibliometric tools for analyzing bibliographic data to identify themes and core relationships [59].

3.1. Research Aim Definition

As a first step, the aim of the review is defined. The aim of the review is to use bibliometric mapping and analysis techniques, as well as thematic analysis, to infer and develop a clear thematic and conceptual framework relating to the intersection of AI with ML, greenwashing, and sustainability reporting, and specifically, for the use of AI with ML in relation to greenwashing and sustainability reporting.

The scope of the bibliometric review is sufficiently large [58] when taking into account the impact of AI with ML as fields within the literature and practice, the importance of sustainability and sustainability reporting, and the potential negative impact of greenwashing on both sustainability and sustainability reporting.

Thematic analysis is used to generate additional insights for the corpora analyzed. Similarly, where the results of a defined query return a literature corpus that is too small for bibliometric analysis, thematic analysis is applied.

3.2. Initial Review of Literature for Keyword Identification

As a second step, in August 2022, a review of papers was conducted relating to:

- sustainability reporting;
- AI with ML;
- greenwashing.

This was done to gain a better understanding and overview of each of these fields within the research topic context, with the aim of constructing a robust research search query.

To identify keywords used in search queries in academic literature databases, extant literature often relies on author perceptions, author brainstorming [8], or limited initial investigations. Such analyses also typically do not provide information on how initial literature was selected, providing a lack of transparency on the identification of keywords, as well as whether search terms, which directly affect bibliometric analysis results, are comprehensive.

The goal is to conduct an initial review in a transparent and replicable manner that provides a high probability of identifying the correct corpus of literature to be analyzed as part of the bibliometric analysis. This approached is referred to as 'purposeful': the review and search methodology is a purposeful review, the selection of keywords and sample selection is a purposeful selection, and the review of the results of the bibliometric procedures focuses on the research aim. This is because the goal is purposeful navigation of literature to ensure the research aim is met.

To begin the purposeful review, bibliometric review literature was identified relating to each field separately to ensure a more global overview of the existing literature in each field. Data were obtained from Scopus, considered to be the most comprehensive database of peer-reviewed literature in these fields [60].

The following initial search queries were constructed and run in the Scopus database, using the operator "TITLE-ABS-KEY" to identify the search terms in the titles, abstracts, or keywords of documents in the database:

The titles and abstracts in the search results returned by the search strings in Table 1 below were then reviewed to select a sample of appropriate and relevant bibliometric or systematic literature reviews in each field. The focus of the initial review was on the most recent literature reviews to ensure an up-to-date view of each field, as well as on reviews that were more global, encapsulating a broad range of bibliometric data, rather than reviews that focused on a specialized discipline or niche (e.g., AI in the field of oncology). The literature selected for review consists of nine papers, and is presented in Table 2 below.

Table 1. Initial search queries.

Topic	Search String
Artificial intelligence	TITLE-ABS-KEY (("artificial intelligence" AND "machine learning") AND ("bibliometric"))
Sustainability reporting	TITLE-ABS-KEY (("sustainability report*" OR "csr report*") AND ("literature review" OR "bibliometric"))
Greenwashing	TITLE-ABS-KEY (("greenwashing*" OR "greenwash*" OR "green claim ") AND ("literature review" OR "bibliometric"))

Table 2. Initial literature selection.

Bibliometric Literature: Artificial Intelligence and ML (Title, Reference)	Bibliometric Literature: Sustainability Reporting (Title, Reference)	Bibliometric Literature: Greenwashing (Title, Reference)
Global bibliometric mapping of the frontier of knowledge in the field of artificial intelligence for the period 1990–2019, [11]	Providing a Roadmap for Future Research Agenda: A Bibliometric Literature Review of Sustainability Performance Reporting (SPR), [2]	Concepts and forms of greenwashing: a systematic review, [5]
Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis, [8]	Mapping the literature on sustainability reporting: A bibliometric analysis grounded in Scopus and web of science core collection, [61]	What has been (short) written about greenwashing: A bibliometric research and a critical analysis of the articles found regarding this theme, [39]
Approaching Artificial Intelligence in business and economics research: a bibliometric panorama (1966–2020), [62]	Non-financial reporting research and practice: Lessons from the last decade, [26]	Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review, [49]

Each of these literature reviews were read to identify keywords commonly used and relevant to each field and the aim of the study. Tables 3–5 below illustrate the keywords identified for each field, with the letter "X" used to indicate keywords identified for a specific paper within the selected literature for each field:

Table 3. Identified keywords: AI with ML literature.

Reference	Artificial Intelligence	Big Data	Machine Learning	Natural Language Processing
[11]	X	X	X	X
[8]	X	X	X	X
[62]	X	X	X	

Table 4. Identified keywords: Sustainability reporting literature.

Reference	CSR Reporting	Sustainability Reporting	Non-Financial Reporting	ESG Reporting
[2]	X	X		
[61]		X		
[26]	X	X	X	X

Table 5. Identified keywords: Greenwashing literature.

Reference	Greenwashing	Green Claim
[5]	X	X
[39]	X	
[49]	X	

The identified keywords then informed the construction of three combined intermediate Scopus search queries to return literature for the intersections of each of the relevant fields i.e., the intersections of:

1. Sustainability reporting and AI with ML
2. Sustainability reporting and greenwashing
3. Greenwashing and AI with ML

These intersections are depicted graphically in Figure 2 below:

Queries were defined for these intersections to better understand the interplay of literature relating to each one.

3.2.1. Intersection 1: Sustainability Reporting and AI with ML

The search for this intersection was limited to results from 2018–2022, given the rapidly evolving nature of each field, and to ensure identification of the most up-to-date keywords. The search was also limited to articles, conference papers, and reviews in English. The intermediate query applied in Scopus for the intersection of sustainability reporting and AI with ML is as follows, and is presented in the format proposed by Goodell et al. [8] in Table 6 below:

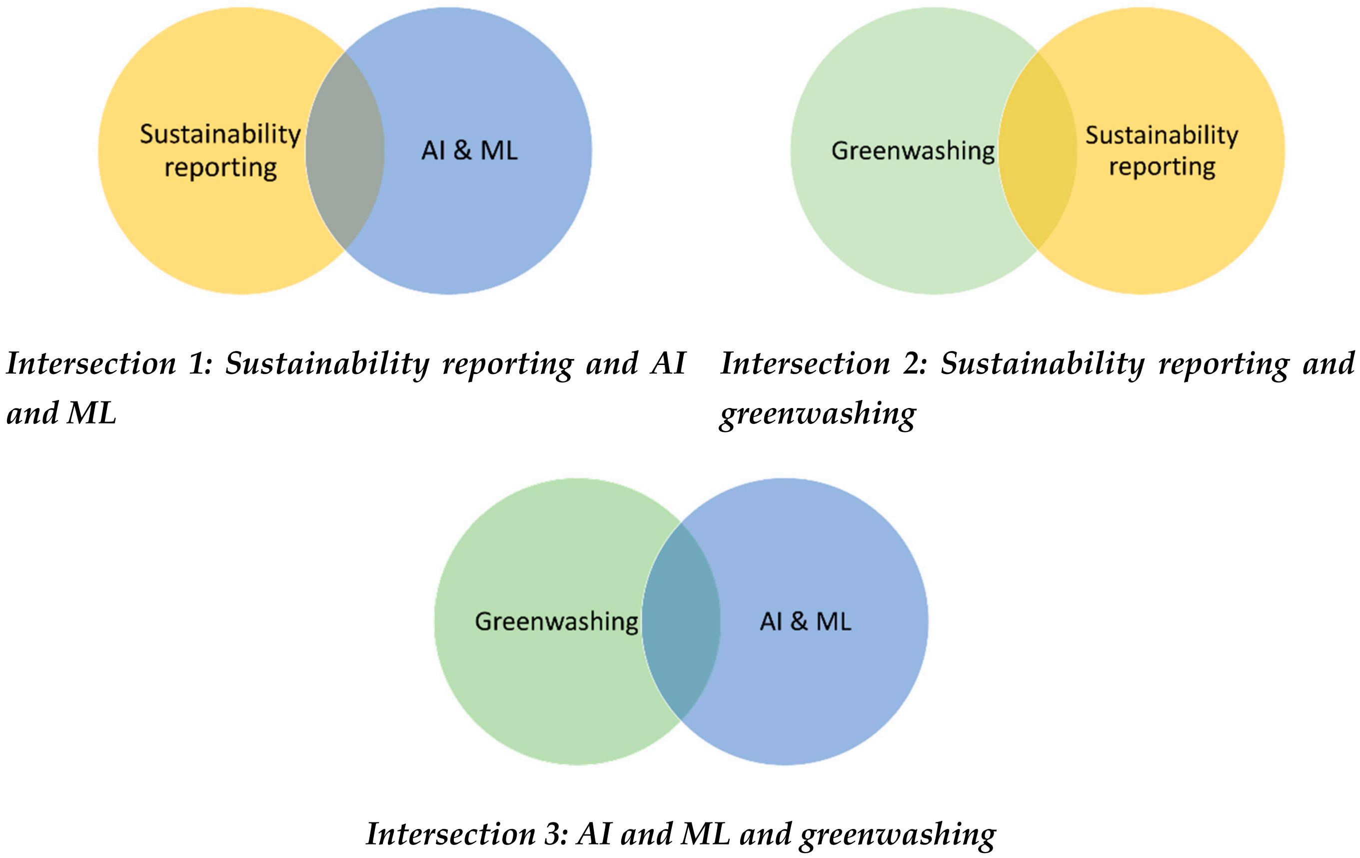

Figure 2. Graphical depiction of field intersections.

Table 6. Intermediate search criteria: Intersection 1.

Filtering and Search Criteria
Database: Scopus Search date: 14 October 2022 Search term: TITLE-ABS-KEY ((“machine learning” OR “AI” OR “artificial intelligence”) AND (“sustainability report*” OR “non-financial report*” OR “ESG report*” OR “CSR report”)) Year: 2018–2022 Document type: “Articles”, “Conference paper”, “Review” Language screening: Only English language documents included Relevance screening: Articles selected for inclusion only where “Titles, abstracts, and keywords” are relevant to review aim (i.e., sustainability reporting and AI) Number of items returned by query: 20 Number of items selected for review: 14

Of the twenty results returned by the query, a sample of fourteen was selected based on a review of each title and abstract for relevance. Each of these fourteen items were then read in order to identify additional relevant keywords relating to the review aim. Table 7 below reflects the relevant keywords identified in those fourteen items relating to the intersection of the fields of sustainability reporting and AI. The letter “X” within the table denotes keywords identified for a particular title.

Table 7. Intermediate search keywords identified: Intersection 1.

	Title	Artificial Intelligence	Big Data	Machine Learning	Natural Language Processing	Text Mining	Deep Learning	CSR Reporting	Sustainability Reporting	Non-Financial Reporting	ESG Reporting
1	Performance Evaluation of the Implementation of the 2013/34/EU Directive in Romania on the Basis of Corporate Social Responsibility Reports							X	X		
2	Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research	X	X	X					X		
3	A framework for evaluating and disclosing the ESG related impacts of AI with the SDGs	X	X	X	X			X	X		
4	Society 5.0 as a Contribution to the Sustainable Development Report	X	X	X					X		
5	Artificial intelligence activities and ethical approaches in leading listed companies in the European Union	X	X	X			X	X	X	X	
6	Who Are the Intended Users of CSR Reports? Insights from a Data-Driven Approach	X	X	X	X	X		X	X		
7	Identifying Corporate Sustainability Issues by Analyzing Shareholder Resolutions: A Machine-Learning Text Analytics Approach	X			X	X	X		X		
8	Interrelation between the climate-related sustainability and the financial reporting disclosures of the European automotive industry				X	X			X		
9	Fundamental ratios as predictors of ESG scores: a machine learning approach			X			X				X
10	Sentiment analysis of CSR disclosures in annual reports of EU companies			X				X	X		
11	Natural Language Processing Methods for Scoring Sustainability Reports—A Study of Nordic Listed Companies	X		X	X	X		X	X		
12	Incorporating ESG in Decision Making for Responsible and Sustainable Investments using Machine Learning	X		X				X	X	X	X
13	Develop CSR Themes using Text-Mining and Topic Modelling Techniques			X	X	X		X	X		
14	Classification of CSR Using Latent Dirichlet Allocation and Analysis of the Relationship Between CSR and Corporate Value			X				X	X		

3.2.2. Intersection 2: Sustainability Reporting and Greenwashing

The process was then repeated for each of the other intersections. That systematic process is summarized in Tables 8–10 below:

Table 8. Intermediate search criteria: Intersection 2.

Filtering and Search Criteria
Database: Scopus Search date: 14 October 2022 Search term: (TITLE-ABS-KEY((“sustainability report*” OR “ESG report*” OR “CSR report” OR “non-financial report*”) AND (“greenwash*” OR “green claim”)) Year: 2018–2022 Document type: “Articles”, “Conference paper”, “Review” Language screening: Only English language documents included Relevance screening: Articles selected for inclusion only where “Titles, abstracts, and keywords” are relevant to review aim (i.e., sustainability reporting and greenwashing) Number of items returned by query: 39 Number of items selected for review: 15

Table 9. Intermediate search keywords identified: Intersection 2.

Item	Title	Sustainability Disclosure	ESG Disclosure
1	CSR Performance and the Readability of CSR Reports: Too Good to be True?		
2	Greenwashing in environmental, social and governance disclosures	X	X
3	Is corporate social responsibility reporting a tool of signaling or greenwashing? Evidence from the worldwide logistics sector	X	
4	Green brand of companies and greenwashing under sustainable development goals		
5	CSR achievement, reporting, and assurance in the energy sector: Does economic development matter?	X	
6	Corporate social responsibility disclosure level, external assurance and cost of equity capital	X	
7	Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research		X
8	The relationship between non-financial reporting, environmental strategies and financial performance. Empirical evidence from Milano stock exchange		
9	The greenwashing triangle: adapting tools from fraud to improve CSR reporting		
10	Is Femvertising the New Greenwashing? Examining Corporate Commitment to Gender Equality		
11	Mapping corporate climate change ethics: Responses among three Danish energy firms		
12	Does internal control contribute to a firm’s green information disclosure? Evidence from China	X	
13	Green, blue or black, but washing–What company characteristics determine greenwashing?	X	
14	Representative account or greenwashing? Voluntary sustainability reports in Australia’s mining/metals and financial services industries	X	
15	Sustainable grocery retailing: Myth or reality?—A content analysis	X	

Table 10. Intermediate search criteria: Intersection 3.

Filtering and Search Criteria
Database: Scopus Search date: 14 October 2022 Search string: TITLE-ABS-KEY (“greenwash*” OR “green claim”) AND (“machine learning” OR “AI” OR “artificial intelligence”) Year: ALL Document type: “Articles”, “Conference paper”, “Review” Language screening: Only English language documents included Relevance screening: Articles selected for inclusion only where “Titles, abstracts, and keywords” are relevant to review aim (i.e., greenwashing and AI) Number of items returned by query: 9 Number of items selected for review: 9

Of the thirty-nine results returned by the query, a sample of fifteen was selected based on a review of each title and abstract for relevance. Each of these fifteen items were then read in order to identify additional relevant keywords relating to the review aim. Table 9 below reflects the relevant keywords identified in those fifteen items relating to the intersection of the fields of sustainability reporting and greenwashing. It is important to note that the table illustrates only additional keywords identified in the fifteen items selected for review. Where keywords were previously identified for example, from the initial search queries, or the intermediate search query relating to sustainability reporting and AI described above, these are not duplicated here. The letter “X” within the table denotes keywords identified for a particular title.

3.2.3. Intersection 3: Greenwashing and AI with ML

Of the nine results returned by the query shown in Table 10, nine were selected based on a review of each title and abstract for relevance. Each of these nine items were then read in order to identify additional relevant keywords relating to the review aim. No new keywords were identified within the review.

3.3. Scopus Search Query Definition

The third step involved defining the final search queries to be used within the bibliometric analysis, for each field intersection. Based on the initial literature review and keyword analysis a list of keywords was created.

The identified list of keywords was further enriched by analyzing the queries used in the following bibliometric literature identified in the initial review:

- Turzo et al. [26] relating to non-financial and sustainability reporting;
- Goodell et al. [8] relating to artificial intelligence and machine learning;
- De Freitas Netto et al. [5] relating to greenwashing.

Each query was optimized based on Scopus guidelines to consider the impact of Boolean operators, punctuation, wildcards, and braces in arriving at the final search query.

The query for each binary intersection was defined sufficiently broadly, based on the initial review, to make provision for differences in terminology, synonyms, and definitions used and identified in the initial literature review. Consistent with the research aim, a final query for the intersection of all three fields was also defined. The defined search query for each intersection is shown in Table 11 below:

Table 11. Final search queries.

#	Intersection Description	Intersection Defined Query
1	Sustainability reporting and AI with ML	("non financ* report*" OR "sustainab* disclo*" OR "environment* report*" OR "CSR report*" OR "CSR disclo*" OR "corporate social responsibility report*" OR "corporate social responsibility disclo*" OR "sustainab* report*" OR "responsib* report*" OR "ESG report*" OR "TBL report*" OR "triple* report*" OR "GHG report*" OR "greenhouse gas report*" OR "integr* report*" OR "corporate citizenship report*" OR "SDG* report*" OR "sustainable development goal* report*" OR "carbon report*" OR "social report*") AND ("machine learning" OR "AI" OR "artificial intelligence" OR "natural language processing" OR "text mining" OR "algorithm" OR "soft computing" OR "data mining" OR "big data" OR "robot" OR "automation" OR "analytics" OR "deep learning")
2	Sustainability reporting and greenwashing	("non financ* report*" OR "sustainab* disclo*" OR "environment* report*" OR "CSR report*" OR "CSR disclo*" OR "corporate social responsibility report*" OR "corporate social responsibility disclo*" OR "sustainab* report*" OR "responsib* report*" OR "ESG report*" OR "TBL report*" OR "triple* report*" OR "GHG report*" OR "greenhouse gas report*" OR "integr* report*" OR "corporate citizenship report*" OR "SDG* report*" OR "sustainable development goal* report*" OR "carbon report*" OR "social report*") AND ("greenwashing*" OR "greenwash*" OR "green claim")
3	AI with ML and greenwashing	("machine learning" OR "AI" OR "artificial intelligence" OR "natural language processing" OR "text mining" OR "algorithm" OR "soft computing" OR "data mining" OR "big data" OR "robot" OR "automation" OR "analytics" OR "deep learning") AND ("greenwashing*" OR "greenwash*" OR "green claim")
4	Greenwashing, sustainability reporting, and AI with ML	("greenwashing*" OR "greenwash*" OR "green claim") AND ("non financ* report*" OR "sustainab* disclo*" OR "environment* report*" OR "CSR report*" OR "CSR disclo*" OR "corporate social responsibility report*" OR "corporate social responsibility disclo*" OR "sustainab* report*" OR "responsib* report*" OR "ESG report*" OR "TBL report*" OR "triple* report*" OR "GHG report*" OR "greenhouse gas report*" OR "integr* report*" OR "corporate citizenship report*" OR "SDG* report*" OR "sustainable development goal* report*" OR "carbon report*" OR "social report*")

3.4. Bibliographic Data Collection

For the fourth step, the final defined query for each intersection was run in the Scopus database.

Intersection 1: Sustainability Reporting and AI with ML

Table 12 illustrates the structured procedure followed to select the population of documents for each intersection.

Table 12. Document selection process per intersection.

Intersection.	1		2		3		4	
Filtering and Search Criteria	**Reject**	**Accept**	**Reject**	**Accept**	**Reject**	**Accept**	**Reject**	**Accept**
Database: Scopus Search date: 14 October 2022								
Search string results:		296		89		17		2
Year: 2003–2023	10	286	1	88	0	17	0	2
Document type: "Articles", "Conference paper", "Review"	27	259	12	76	1	16	0	2

Table 12. *Cont.*

Intersection.	1		2		3		4	
Filtering and Search Criteria	**Reject**	**Accept**	**Reject**	**Accept**	**Reject**	**Accept**	**Reject**	**Accept**
Language screening: Only English language documents included	4	255	0	76	0	16	0	2
Relevance screening: Articles selected for inclusion only where "Titles, abstracts, and keywords" are relevant to intersection and review aim	94	160	0	76	0	14	0	2

4. Results and Discussion

4.1. Summary Information

Summary information provides an overview of the literature analyzed, which is an important element in understanding the corpus. The basic statistics returned relating to the distinct literature items analyzed, for each of the three intersections, are presented in Table 13 below.

Table 13. Summary statistics per intersection.

Description	Intersection 1 Sustainability Reporting and AI and ML	Intersection 2 Greenwashing and Sustainability Reporting	Intersection 3 Greenwashing and AI and ML	Intersection 4 Greenwashing, Sustainability Reporting and AI and ML
MAIN INFORMATION ABOUT DATA				
Timespan	2004:2022	2003:2022	2016:2022	2022:2022
Sources (Journals, Books, etc.)	121	61	15	2
Documents	160	76	16	2
Annual Growth Rate %	14.25	14.45	38.31	0.00
Document Average Age	4.28	3.51	1.25	0.00
Average citations per doc	11.66	31.97	8.06	8
References	8098	5691	1127	228
DOCUMENT CONTENTS				
Keywords Plus (ID)	876	156	143	17
Author's Keywords (DE)	528	278	90	19
AUTHORS				
Authors	417	179	45	10
Authors of single-authored docs	16	18	4	0
AUTHORS COLLABORATION				
Single-authored docs	16	18	4	0
Co-Authors per Doc	2.96	2.54	2.94	5.00
International co-authorships %	19.38	28.95	18.75	50.00
DOCUMENT TYPES				
article	100	72	10	2
conference paper	54	2	4	0
review	6	2	2	0

The literature identified relating to Intersection 4 is too small a corpus for bibliometric analysis, and is therefore analyzed separately in 4.5 below using thematic analysis.

4.2. Evolution—Publications

Figure 3 below illustrates a summary graph of the publication trends for the three intersections:

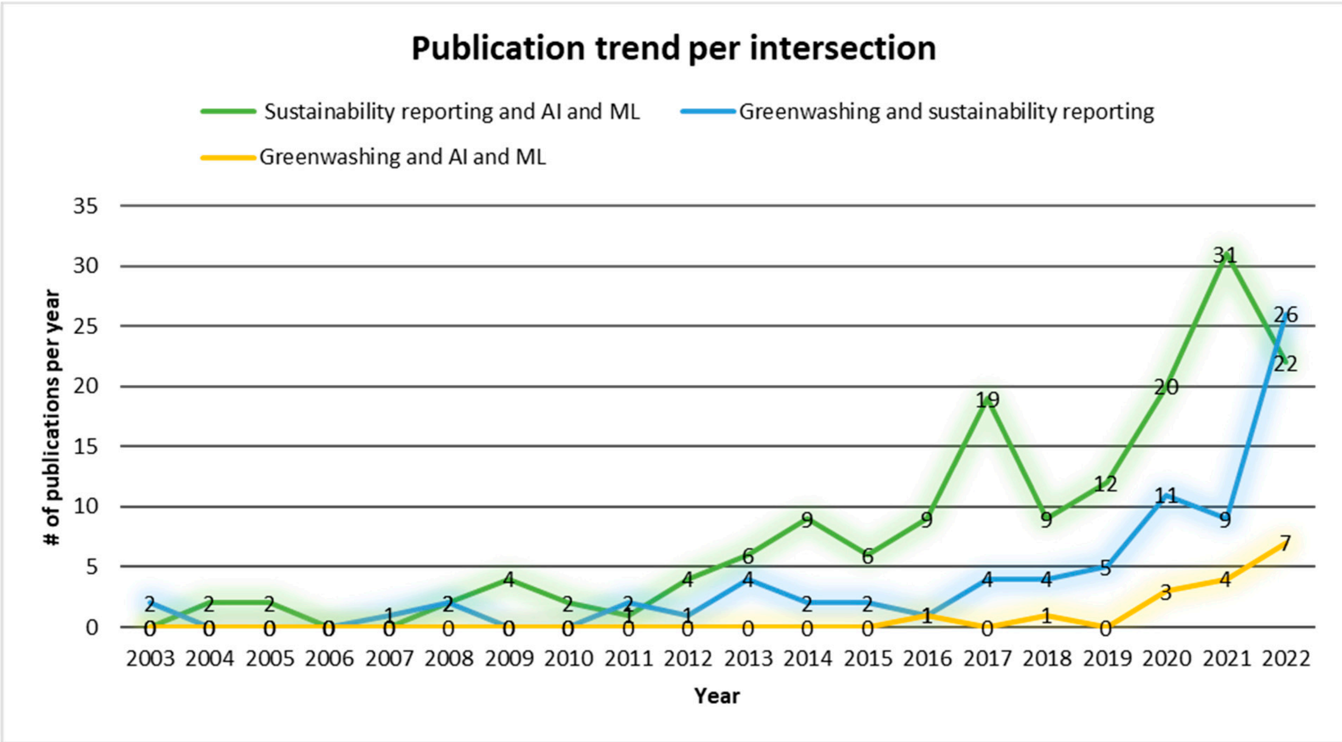

Figure 3. Publication trend per intersection.

4.2.1. Intersection 3

The highest growth rate is found in the literature for Intersection 3, the greenwashing and AI with ML intersection, with a 38.31% annual average growth rate. This corpus covers a shorter time period, however, with publications for this intersection only occurring from 2016, and is a smaller corpus, with only 16 total articles during that timespan. The relative size and timespan of the literature in this corpus reflects the emerging and nascent nature of research pertaining to artificial intelligence and ML within the field of greenwashing.

4.2.2. Intersection 2

Within Intersection 2, the intersection of greenwashing and sustainability reporting, an upward trend is also observed, with some fluctuations during the time period analyzed. This intersection reflects an annual average growth rate of 14.45%, over a much longer timespan than the greenwashing and AI intersection, from 2003 to 2022.

The number of publications relating to this research field intersection has increased consistently since the beginning of the analysis period. This upward trend may be understood with reference to three distinct periods.

The first, for the period 2003 to 2016, saw a limited increase in the literature in the research field, with only 17 publications within 14 years, an average of only 1.21 publications per year. The second period, from 2017 to 2019 saw faster growth, with 13 publications in 3 years at an average of 4.33 publications per year, 3.5 times more than the prior period. The third period has seen the most rapid growth, with 46 publications from 2020 to October 2022. The highest number of publications in a specific period to date occurs in 2022.

Comparing the number of publications from the first to the second period illustrates that the average number of publications per year was 1.21 in the first period, 4.33 in the second, and 15.33 in the last. When considering the upward trend from 2017 onward, this may be understood with reference to the Paris Agreement on Climate Change, which concluded in late 2015, with the increase in academic literature relating to this intersection then trending upward from 2017, as academic literature relating to both sustainability reporting and greenwashing was published.

4.2.3. Intersection 1

The Intersection 1 publication trend illustrates a high average annual growth rate of 14.25%, close to that of Intersection 2 (14.45%). This trend is for the period from 2004 to 2022, whereas the trend for Intersection 2 is for slightly longer, from 2003 to 2022.

Though Intersection 1′s growth rate is marginally lower than that of Intersection 2, it relates to a much larger corpus of academic literature than both Intersection 2 and Intersection 3, with 160 documents published during that time period. This upward trend may be also understood with reference to three distinct periods.

The first, for the period 2004 to 2016, saw a limited increase in literature in the research field, with only 47 of the total 160 publications occurring within those 13 years, an average of only 3.62 publications per year. The second period, from 2017 to 2019 saw faster growth, with 40 publications in only 3 years. The third period has seen the most rapid growth, with 73 publications from 2020 to October 2022. The highest number of publications in a specific period to date occurs in 2021.

Comparing the number of publications from the first to the second period illustrates that the average number of publications per year was 3.62 in the first period, 13.33 in the second, and 24.33 in the last.

The key publication trend statistics across each intersection are summarized in Table 14 below.

Table 14. Publication trend summary statistics.

#	Intersection	Average Annual Growth Rate	Size of Corpus (# of Documents)	Timespan	Document Average Age
1	Sustainability reporting and AI with ML	14.25%	160	2003:2022	4.28
2	Greenwashing and sustainability reporting	14.45%	76	2003:2022	3.51
3	Greenwashing and AI with ML	38.31%	16	2016:2022	1.25

The relative timespan for literature for each intersection reflects that research within Intersection 3 is an emerging field, with research in this field only being published since 2016. The larger corpus of literature relating to Intersection 1, for the timespan from 2004 to 2022, reflects that research within the intersection is more mature and established than that both Intersections 2 and 3. This is supported by Intersection 1 having the highest average document age amongst the three intersections.

The above analysis reflects the relative maturity of the academic literature relating to Intersections 1 and 2 in relation to Intersection 3. It also reveals that literature relating to the use of AI with ML within the field of greenwashing, Intersection 3, has only emerged recently and is in its nascent stages both in terms of age and volume of publications.

4.3. Evolution—Journal Publications

Figures 4–6 below show top journal sources for Intersections 1 to 3:

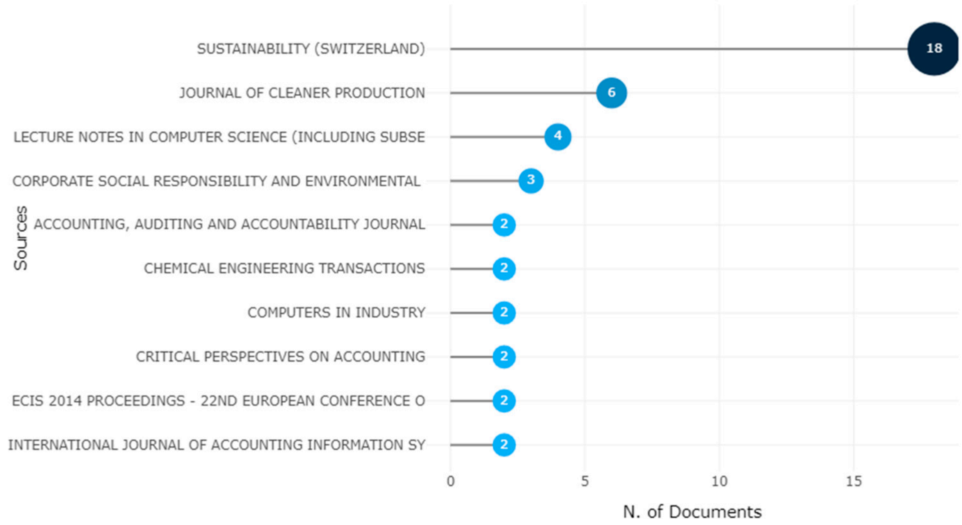

Figure 4. Top sources: Sustainability reporting and AI with ML.

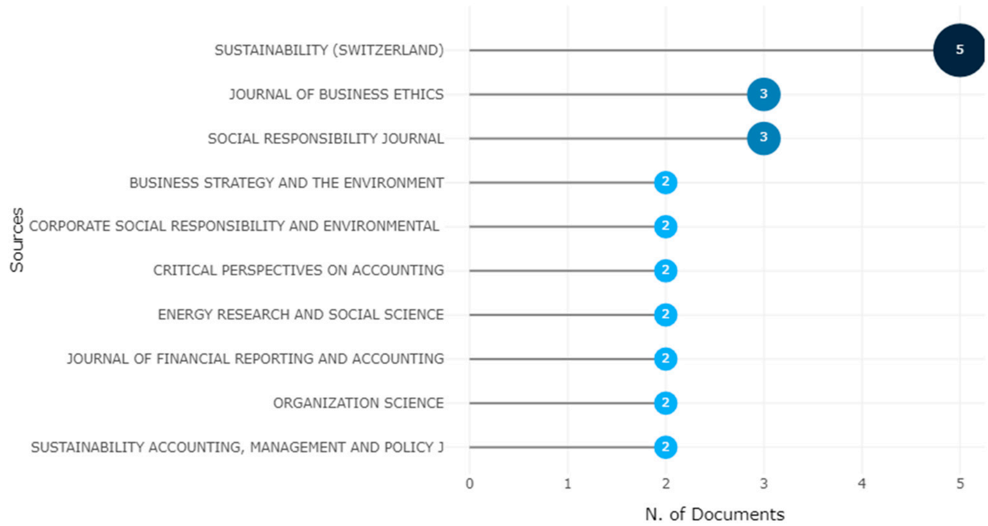

Figure 5. Top sources: Sustainability reporting and greenwashing.

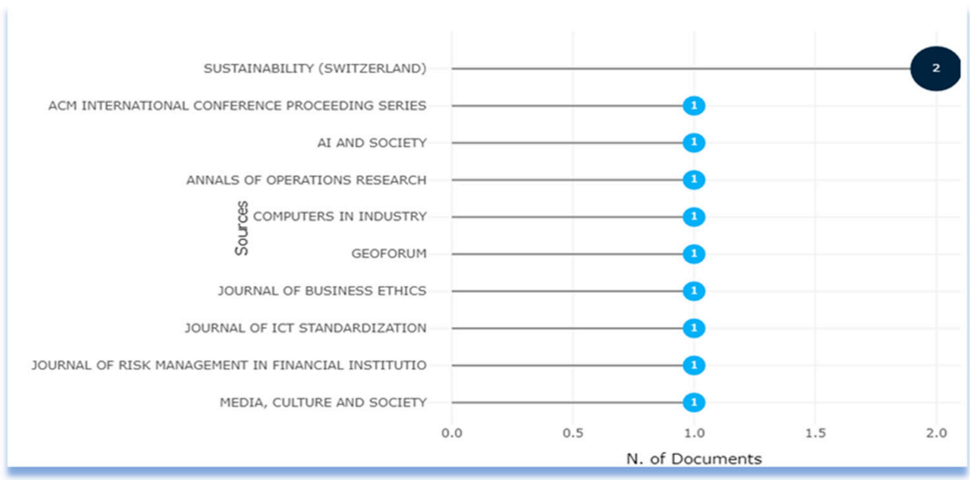

Figure 6. Top sources: Greenwashing and AI with ML.

Across all of the intersections, the journal "Sustainability (Switzerland)" is the top source journal for documents.

4.4. Research Trends and Themes

In order to identify research trends, Biblioshiny's document keyword functionality is used, based on "Keywords Plus". Keywords Plus are "words or phrases that frequently appear in the titles of the article's references and not necessarily in the article's title or as Author Keywords" [60]. Zhang et al. [63] suggest a number of advantages to using Keywords Plus and recommend that Keywords Plus be used for bibliometric analyses.

4.4.1. Highest Occurrence Keywords

Intersection 1: Sustainability Reporting and AI with ML

Figure 7 below illustrates the top 10 keywords, based on Keywords Plus and ranked according to frequency of occurrence for Intersection 1. The high number of occurrences of the terms 'data mining' (32), 'text mining' (17), and 'big data' (11) reflects the relevance of these methods within both sustainable development and sustainability as broader themes, as well as within sustainability reporting and corporate social responsibility, which are also keywords identified within the ranking.

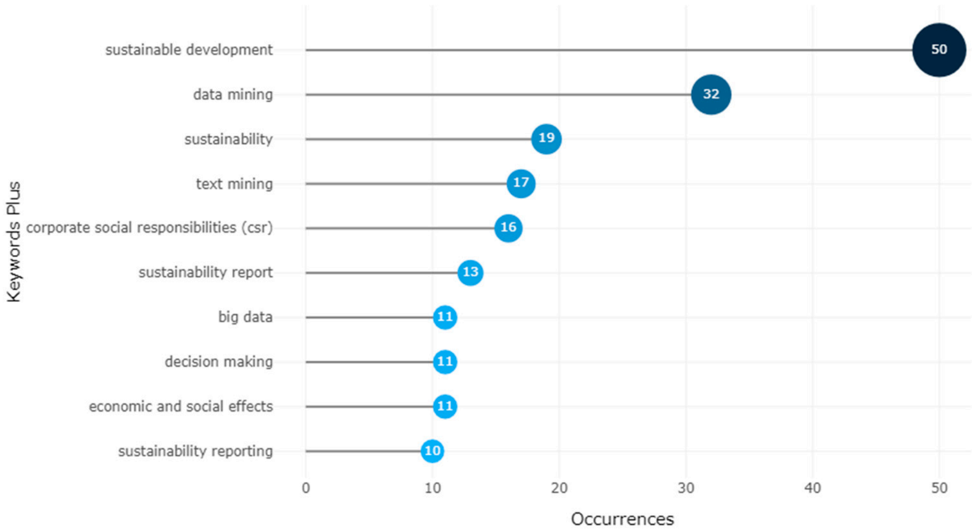

Figure 7. Intersection 1 keyword ranking based on Keywords Plus and number of occurrences.

Intersection 2: Greenwashing and Sustainability Reporting

The keyword ranking for Intersection 2 shown in Figure 8 below provides similar insights, showing within the smaller corpus the highest occurrence of terms relates to sustainability, sustainable development, and CSR, with greenwashing also identified within the ranking.

Intersection 3: Greenwashing and AI with ML

Within Intersection 3, 'data mining', 'regression analysis', and 'artificial intelligence' are highly-ranked keywords, as are 'greenwashing', 'environmental monitoring', and related terms such as 'air pollution' and 'anomaly detection', as shown in Figure 9 below.

4.5. Thematic Analysis

As the research aims relates to understanding the use of AI with ML in relation to sustainability reporting and greenwashing, a thematic analysis was conducted to consider the use of AI with ML within the literature identified. This was done by reviewing the titles and abstracts within each intersection corpus to identify documents which indicate the use of AI or ML tools within their research methodologies.

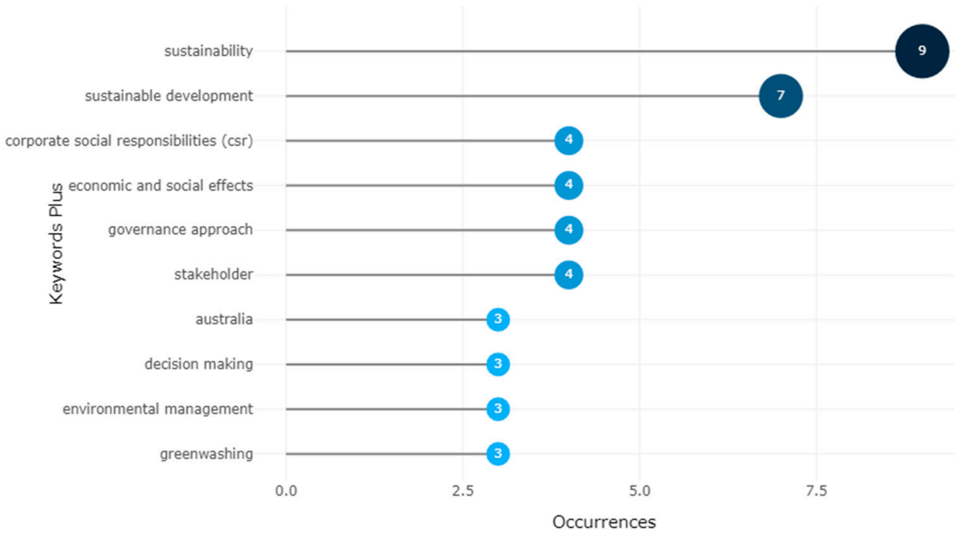

Figure 8. Intersection 2 keyword ranking based on Keywords Plus and number of occurrences.

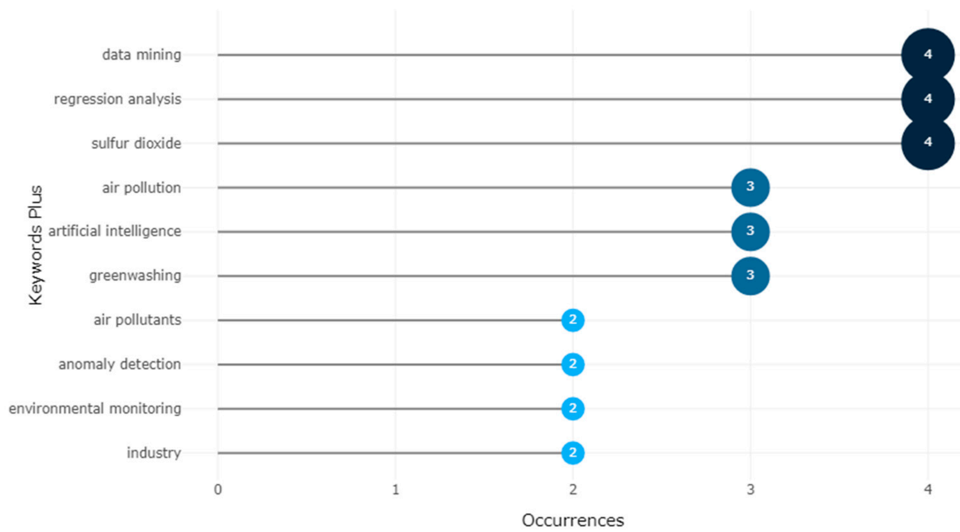

Figure 9. Intersection 3 keyword ranking based on Keywords Plus and number of occurrences.

4.5.1. Intersection 1: Sustainability Reporting and AI with ML

The thematic analysis of this intersection reveals that 82 of the 160 documents (51.25%) within this corpus use AI or ML as a methodological tool. From this result, it may be inferred that a significant amount of research within this field applies AI with ML as a methodological tool. Below is a list of AI with ML methodological tool descriptions, as identified within those 82 documents:

- Algorithms;
- Artificial intelligence;
- Automated content analysis;
- Bibliometric analysis;
- Big data analysis;
- Data mining;
- Machine Learning;
- Machine Learning clustering algorithm;
- Machine learning using Multivariate Discriminant Analysis (MDA);
- Natural Language Processing;
- Natural Language Processing using Latent Dirichlet Allocation (LDA);

- Random forest;
- Topic modeling;
- Text mining.

Note that within the list, bibliometric analysis is considered to be an AI with ML tool, as this type of analysis relates to the use of machine learning within research [64].

These identified methodologies predominantly relate to the use of NLP in the analysis of textual information relating to sustainability reporting. This reflects the maturity in the use of NLP techniques in the analysis of sustainability reporting information.

4.5.2. Intersection 2: Greenwashing and Sustainability Reporting

The thematic analysis of this intersection reveals one article that uses AI or ML within this corpus. However, this article is an article common to Intersections 1, 2, and 3. This exemplifies the limited use of AI with ML, and NLP, as a methodological tool relating to the research on greenwashing within the field of sustainability reporting. The identified article is shown in Table 15 below:

Table 15. Intersection 2 common literature.

Title	Year
Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research	2022

4.5.3. Intersection 3: Greenwashing and AI with ML

Similar to Intersection 2, only one article that uses AI or ML as a methodological tool is found within this corpus, and is the same article common to common to Intersections 1, 2, and 3. This further exemplifies the limited use of AI with ML as a methodological tool in relating to the research on greenwashing. The identified article is shown in Table 16 below:

Table 16. Intersection 2 common literature.

Title	Year
Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research	2022

The thematic analysis above illustrates the relative maturity of the use of AI with ML as methodological tools within the field of sustainability reporting (Intersection 1). However, when looking at the corpus of the field of greenwashing within sustainability reporting (Intersection 2), or AI with ML in relation to greenwashing (Intersection 3), limited use of AI with ML as a methodological tool is found.

4.5.4. Intersection 4: Greenwashing, Sustainability Reporting and AI with ML

Intersection 4 reflects the intersection of greenwashing, sustainability reporting, and AI with ML. In order to conduct the thematic analysis, the literature within this corpus which consisted of two documents, shown in Table 17 below, was read.

Table 17. Intersection 4 common literature.

Item #	Title	Year
1	Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research	2022
2	Unsupervised neural network-enabled spatial-temporal analytics for data authenticity under environmental smart reporting system	2022

Item 1 above is found within the corpora of Intersections 1 to 4. Item 1 shows the use of AI with ML as a methodological tool. The aim of this item is a major review to provide an overview of the field of sustainable finance by using machine learning [64]. This item therefore, when considering the research aim, does not relate to the use of AI with ML in relation to greenwashing or sustainability reporting.

Item 2, besides presenting a reporting system, also evaluates "the authenticity of the data collected from IoT devices, considering human-made counterfeits on measuring instruments for greenwashing" [65]. This article does therefore relate to the use of AI with ML in relation to greenwashing or sustainability reporting.

The use of AI with ML in relation to greenwashing and sustainability reporting is found in one article within Intersection 4. While that article accounts for half of this corpus of two documents, identifying only a single instance of the use of AI with ML in relation to greenwashing within sustainability reporting reflects that such use is underexplored in the literature.

5. Conclusions

In this paper, a purposeful review of the literature was conducted using bibliometric and thematic analysis of the intersections of the fields of greenwashing, sustainability reporting, and AI with ML, respectively. The foundation of the review is bibliometric analyses of the binary combinations of these fields, which we combine with a thematic analysis of each of those intersections and of the intersection of all three fields. We introduce additional insights by a thematic analysis of the methodological tools applied.

This paper presents two foundational contributions. The first is the application of bibliometric and thematic analysis to comprehensively and holistically address the interrelationship between greenwashing, sustainability reporting, and AI with ML, an analysis which is not present within extant literature. The second is the conjecture of a conceptual and thematic framework for the use of artificial intelligence and machine learning in relation to greenwashing and company sustainability reporting.

The systematic methodological process followed for the analysis is designed to ensure transparency and replicability, in a manner that provides scope for future use of a similar process when analyzing the interrelationships between multiple fields.

Through the purposeful review, a number of trends and themes are identified.

First, the review identifies the emerging and nascent nature of the research relating to the use of AI with ML within the field of greenwashing. This is apparent when considering multiple maturity measures, such as the number of documents that relate to the intersection of those two fields, as well as the average age and timespan of the documents within the corpus. These maturity measures are in stark contrast to those for the literature analyzed for the intersections of AI with ML and sustainability reporting, and greenwashing and sustainability reporting, respectively. Both of those intersections reflect a body of research that is significantly more mature for each maturity measure namely, corpus size, document average age, and timespan. The corpus for the intersection of AI with ML and sustainability reporting is both the most mature and largest amongst the intersections.

Second, from a thematic perspective, clear keyword trends are identified relating to foundational themes such as sustainability and sustainable development.

Third, important trends are observed relating to the high keyword occurrences of AI with ML-related tools and techniques, such as data mining, text mining, big data, and artificial intelligence. Those trends reflect the significance of AI with ML within the fields of greenwashing and sustainability reporting, respectively.

Fourth, the significance of the use of AI with ML in the field of sustainability reporting is supported by the thematic analysis which identifies maturity in the use of AI with ML techniques in that field. The wide application of AI with ML tools such as NLP within the field of sustainability reporting reflects its perceived usefulness by researchers.

Last, the mature or wide application of AI with ML techniques within the other intersections is not found, and it may therefore be inferred that the use of AI with ML in relation to greenwashing, and in relation to greenwashing within sustainability reporting, is an underexplored research field. This finding is significant given the negative impacts of greenwashing on sustainability and sustainable reporting, and the potential of AI and ML to ameliorate such impacts.

5.1. Implications for Future Research

This paper provides a number of implications for future research.

Given the significance of the business and societal impact of greenwashing, future research could explore the use of ML and NLP tools and techniques in relation to greenwashing generally or greenwashing within sustainability reporting specifically. The use of such tools presents significant opportunities for identifying or combating greenwashing. Such a study would likely contribute significant business and societal value in terms of environmental, social, and economic impacts.

Future research may also relate to how insights from other fields that are more mature may contribute to greenwashing research, given the emergent nature of research into the intersection of AI and greenwashing. Such research could consider to what extent theories, practices, themes, tools, taxonomies, or analyses from the fields of AI with ML or sustainability reporting may have relevance to the study of greenwashing and greenwashing behaviors. These may include, for example, research into more data-driven approaches to sourcing and reporting on environmental information, or data authentication relating to green claims and reported sustainability information.

5.2. Limitations

The study reflects a number of inherent limitations. We limit the study to academic databases and academic literature. Practice-based literature may reflect different themes, practices, or conclusions. Given the nature of the three fields analyzed, practice-based literature may either lag or be ahead of academic literature in inferring a thematic or conceptual structure for these intersections. We also note the rapidly evolving nature of each of the three fields studied, which may affect the currency of the findings

Author Contributions: W.M.: Conceptualization; Methodology; Validation; Formal analysis; Investigation; Data Curation; Writing—Original Draft; Writing—Review and Editing; Visualization; Project administration. A.T.: Conceptualization; Methodology; Validation; Formal analysis; Investigation; Data Curation; Writing—Original Draft; Writing—Review and Editing; Visualization; Supervision; Project administration. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: The authors would like to acknowledge the professional support of the University of Johannesburg's Johannesburg Business School.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Commission on Environment and Development. Our Common Future. Available online: https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf (accessed on 9 October 2022).
2. Osobajo, O.A.; Oke, A.; Lawani, A.; Omotayo, T.S.; Ndubuka-McCallum, N.; Obi, L. Providing a roadmap for future research agenda: A bibliometric literature review of sustainability performance reporting (SPR). *Sustainability* **2022**, *14*, 8523. [CrossRef]
3. van Niekerk, A.J. Inclusive economic sustainability: SDGs and global inequality. *Sustainability* **2020**, *12*, 5427. [CrossRef]
4. European Commission. Initiative on Substantiating Green Claims. Available online: https://ec.europa.eu/environment/eussd/smgp/initiative_on_green_claims.htm (accessed on 6 October 2022).
5. de Freitas Netto, S.V.; Sobral, M.F.F.; Ribeiro, A.R.B.; Soares, G.R.d.L. Concepts and forms of greenwashing: A systematic review. *Environ. Sci. Eur.* **2020**, *32*, 19. [CrossRef]
6. Global Reporting Initiative. A Short Introduction to the GRI Standards. Available online: https://www.globalreporting.org/media/wtaf14tw/a-short-introduction-to-the-gri-standards.pdf (accessed on 8 October 2022).
7. Macpherson, M.; Gasperini, A.; Bosco, M. Implications for Artificial Intelligence and ESG Data. Available online: https://ssrn.com/abstract=3863599 (accessed on 4 October 2022).
8. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial Intelligence and Machine Learning in Finance: Identifying Foundations, Themes, and Research Clusters from Bibliometric Analysis. *J. Behav. Exp. Financ.* **2021**, *32*, 100577. [CrossRef]
9. McCarthy, J.; Minsky, M.L.; Rochester, N.; Shannon, C.E. A proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Mag.* **2006**, *27*, 12. [CrossRef]
10. Annoni, A.; Benczur, P.; Bertoldi, P.; Delipetrev, P.; de Prato, G.; Feijoo, C.; Fernandez-Macias, E.; Gomez, E.; Iglesias, M.; Junklewitz, H.; et al. *Artificial Intelligence—A European Perspective*; Craglia, M., Ed.; EUR 29425 EN; Publications Office of the European Union: Luxembourg, 2018. [CrossRef]
11. de la Vega Hernández, I.M.; Urdaneta, A.S.; Carayannis, E. Global Bibliometric Mapping of the Frontier of Knowledge in the Field of Artificial Intelligence for the Period 1990–2019. *Artif. Intell. Rev.* **2022**. [CrossRef]
12. In, S.Y.; Schumacher, K. Carbonwashing: A New Type of Carbon Data-Related ESG Greenwashing Working Paper. 2021. Available online: https://ssrn.com/abstract=3901278 (accessed on 14 December 2022).
13. Guo, R.; Zhang, W.; Wang, T.; Li, C.B.; Tao, L. Timely or considered? Brand Trust repair strategies and mechanism after greenwashing in China—From a legitimacy perspective. *Ind. Mark. Manag.* **2018**, *72*, 127–137. [CrossRef]
14. Baumgarth, C.; Binckebank, L. Building and Managing CSR Brands—Theory and Applications. Available online: https://www.researchgate.net/publication/283350902_Building_and_managing_CSR_brands_-_Theory_and_applications (accessed on 28 October 2022).
15. Lombardi, R.; Secundo, G. The digital transformation of corporate reporting—A systematic literature review and avenues for future research. *Meditari Account. Res.* **2020**, *29*, 1179–1208. [CrossRef]
16. Beltrami, M.; Orzes, G.; Sarkis, J.; Sartor, M. Industry 4.0 and Sustainability: Towards Conceptualization and Theory. *J. Clean. Prod.* **2021**, *312*, 127733. [CrossRef]
17. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186. [CrossRef]
18. Huang, A.; Wang, H.; Yang, Y. FinBERT—A Large Language Model for Extracting Information from Financial Text. *Contemp. Account. Res.* **2022**. [CrossRef]
19. Kotzian, P. Applying Machine Learning and Artificial Intelligence to CSR-Compliance. A Conceptual Framework with Illustrations. Available online: https://ssrn.com/abstract=3788977 (accessed on 5 October 2022).
20. Kang, H.; Kim, J. Analyzing and visualizing text information in corporate sustainability reports using natural language processing methods. *Appl. Sci.* **2022**, *12*, 5614. [CrossRef]
21. Luccioni, A.; Baylor, E.; Duchene, N. Analyzing Sustainability Reports Using Natural Language Processing. *arXiv* **2020**. [CrossRef]
22. Ning, X.; Yim, D.; Khuntia, J. Online sustainability reporting and firm performance: Lessons learned from text mining. *Sustainability* **2021**, *13*, 1069. [CrossRef]
23. Smeuninx, N.; de Clerck, B.; Aerts, W. Measuring the readability of sustainability reports: A corpus-based analysis through standard formulae and NLP. *Int. J. Bus. Commun.* **2020**, *57*, 52–85. [CrossRef]
24. Amel-Zadeh, A.; Chen, M.; Mussalli, G.; Weinberg, M. NLP for SDGs: Measuring Corporate Alignment with the Sustainable Development Goals. Available online: https://ssrn.com/abstract=3874442. (accessed on 20 September 2022).
25. Velte, P. Archival research on integrated reporting: A systematic review of main drivers and the impact of integrated reporting on firm value. *J. Man. Gov.* **2022**, *26*, 997–1061. [CrossRef]

26. Turzo, T.; Marzi, G.; Favino, C.; Terzani, S. Non-financial reporting research and practice: Lessons from the last decade. *J. Clean. Prod.* **2022**, *345*, 131154. [CrossRef]
27. Uyar, A.; Karaman, A.S.; Kilic, M. Is corporate social responsibility reporting a tool of signaling or greenwashing? Evidence from the worldwide logistics sector. *J. Clean. Prod.* **2020**, *253*, 119997. [CrossRef]
28. Dienes, D.; Sassen, R.; Fischer, J. What are the drivers of sustainability reporting? A systematic review. *Sustain. Account. Manag. Policy J.* **2016**, *7*, 154–189. [CrossRef]
29. Afolabi, H.; Ram, R.; Rimmel, G. Harmonization of sustainability reporting regulation: Analysis of a contested arena. *Sustainability* **2022**, *14*, 5517. [CrossRef]
30. Tettamanzi, P.; Venturini, G.; Murgolo, M. Sustainability and financial accounting: A critical review on the ESG dynamics. *Environ. Sci. Pollut. Res.* **2022**, *29*, 16758–16761. [CrossRef]
31. JSE Limited. Leading the Way for a Better Tomorrow JSE Sustainability Disclosure Guidance. Available online: https://www.jse.co.za/sites/default/files/media/documents/JSE%20Sustainability%20Disclosure%20Guidance%20June%202022.pdf (accessed on 28 October 2022).
32. Tateishi, E. Craving gains and claiming "Green" by Cutting Greens? an exploratory analysis of greenfield housing developments in Iskandar Malaysia. *J. Urban Aff.* **2017**, *40*, 370–393. [CrossRef]
33. Lyon, T.P.; Maxwell, J.W. Greenwash: Corporate environmental disclosure under threat of audit. *J. Econ. Manag. Strategy.* **2011**, *20*, 3–41. [CrossRef]
34. Climate Social Science Network. CSSN Working Paper 2021:1 An Integrated Framework to Assess Greenwashing. Available online: https://cssn.org/wp-content/uploads/2021/09/CSSN-Working-Paper-2021-on-Assessing-Greenwashing-1.pdf (accessed on 21 March 2022).
35. Testa, F.; Boiral, O.; Iraldo, F. Internalization of environmental practices and institutional complexity: Can stakeholders pressures encourage greenwashing? *J. Bus. Ethics* **2015**, *147*, 287–307. [CrossRef]
36. U.S. Securities and Exchange Commission. It's Not Easy Being Green: Bringing Transparency and Accountability to Sustainable Investing. Available online: https://www.sec.gov/news/statement/lee-statement-esg-052522 (accessed on 12 October 2022).
37. Financial Conduct Authority. FCA Proposes New Rules to Tackle Greenwashing. Available online: https://www.fca.org.uk/news/press-releases/fca-proposes-new-rules-tackle-greenwashing (accessed on 28 October 2022).
38. Pimonenko, T.; Bilan, Y.; Horák, J.; Starchenko, L.; Gajda, W. Green brand of companies and greenwashing under sustainable development goals. *Sustainability* **2020**, *12*, 1679. [CrossRef]
39. Andreoli, T.P.; Crespo, A.; Minciotti, S. What Has Been (Short) written about greenwashing: A bibliometric research and a critical analysis of the articles found regarding this theme. *RGSA* **2017**, *11*, 54–72. [CrossRef]
40. U.S. Securities and Exchange Commission. SEC Proposes Rules to Enhance and Standardize Climate-Related Disclosures for Investors. Available online: https://www.sec.gov/news/press-release/2022-46 (accessed on 6 October 2022).
41. Steiner, G.; Geissler, B.; Schreder, G.; Zenk, L. Living Sustainability, or Merely Pretending? From Explicit Self-Report Measures to Implicit Cognition. *Sustain. Sci.* **2018**, *13*, 1001–1015. [CrossRef]
42. Delmas, M.A.; Burbano, V.C. The drivers of greenwashing. *Calif. Manage. Rev.* **2011**, *54*, 64–87. [CrossRef]
43. Lyon, T.P.; Montgomery, A.W. The means and end of greenwash. *Organ. Environ.* **2015**, *28*, 223–249. [CrossRef]
44. Nemes, N.; Scanlan, S.J.; Smith, P.; Smith, T.; Aronczyk, M.; Hill, S.; Lewis, S.L.; Montgomery, A.W.; Tubiello, F.N.; Stabinsky, D. An Integrated Framework to Assess Greenwashing. *Sustainability* **2022**, *14*, 4431. [CrossRef]
45. European Commission. Screening of Websites for 'Greenwashing': Half of Green Claims Lack Evidence. Available online: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_269 (accessed on 13 October 2022).
46. Ruiz-Blanco, S.; Romero, S.; Fernandez-Feijoo, B. Green, blue or black, but washing–What company characteristics determine greenwashing? *Environ. Dev. Sustain.* **2022**, *24*, 4024–4045. [CrossRef]
47. Montero-Navarro, A.; González-Torres, T.; Rodríguez-Sánchez, J.L.; Gallego-Losada, R. A bibliometric analysis of greenwashing research: A closer look at agriculture, food industry and food retail. *Br. Food J.* **2021**, *123*, 547–560. [CrossRef]
48. Zhang, L.; Li, D.; Cao, C.; Huang, S. The influence of greenwashing perception on green purchasing intentions: The mediating role of green word-of-mouth and moderating role of green concern. *J. Clean. Prod.* **2018**, *187*, 740–750. [CrossRef]
49. Yang, Z.; Nguyen, T.T.H.; Nguyen, H.N.; Nguyen, T.T.N.; Cao, T.T. Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review. *J. Bus. Econ. Man.* **2020**, *21*, 1486–1507. [CrossRef]
50. Pope, S.; Wæraas, A. CSR-washing is rare: A conceptual framework, literature review, and critique. *J. Bus. Ethics* **2016**, *137*, 173–193. [CrossRef]
51. Hopkins, E. Machine learning tools, algorithms, and techniques in retail business operations: Consumer perceptions, expectations, and habits. *J. Self-Gov. Manag. Econ.* **2022**, *10*, 43–55. [CrossRef]
52. Feuerriegel, S.; Raj Shrestha, Y.; von Krogh, G.; Zhang, C. Bringing artificial intelligence to business management. *Nat. Mach. Intell* **2022**, *4*, 611–613. [CrossRef]
53. Vinuesa, R.; Azizpour, H.; Leite, I.; Balaam, M.; Dignum, V.; Domisch, S.; Felländer, A.; Langhans, S.D.; Tegmark, M.; Fuso Nerini, F. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* **2020**, *11*, 233. [CrossRef]
54. Yim, W.W.; Yetisgen, M.; Harris, W.P.; Sharon, W.K. Natural language processing in oncology: A review. *JAMA Oncol.* **2016**, *2*, 797–804. [CrossRef] [PubMed]

55. Jan van Eck, N.; Waltman, L. VOSviewer Manual. Available online: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.8.pdf (accessed on 28 October 2022).
56. Civitani, M.M. X-Ray Optics through the Text Mining of Five Decades of Conference Proceedings. In Proceedings of the Optics for EUV, X-Ray, and Gamma-Ray Astronomy X, Proc. SPIE 11822, San Diego, CA, USA, 7 September 2021. [CrossRef]
57. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
58. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [CrossRef]
59. Aria, M.; Cuccurullo, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]
60. Ampah, J.D.; Yusuf, A.A.; Afrane, S.; Jin, C.; Liu, H. Reviewing two decades of cleaner alternative marine fuels: Towards imo's decarbonization of the maritime transport sector. *J. Clean. Prod.* **2021**, *320*, 128871. [CrossRef]
61. Pasko, O.; Chen, F.; Oriekhova, A.; Brychko, A.; Shalyhina, I. Mapping the Literature on Sustainability Reporting: A Bibliometric Analysis Grounded in Scopus and Web of Science Core Collection. *Eur. J. Sustain. Dev.* **2021**, *10*, 303. [CrossRef]
62. Yang, D.; Zhao, W.G.; Du, J.; Yang, Y. Approaching Artificial Intelligence in business and economics research: A bibliometric panorama (1966–2020). *Technol. Anal. Strateg. Manag.* **2022**. [CrossRef]
63. Zhang, J.; Yu, Q.; Zheng, F.; Long, C.; Lu, Z.; Duan, Z. Comparing keywords plus of wos and author keywords: A case study of patient adherence research. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 967–972. [CrossRef]
64. Kumar, S.; Sharma, D.; Rao, S.; Lim, W.M.; Mangla, S.K. Past, Present, and Future of Sustainable Finance: Insights from Big Data Analytics through Machine Learning of Scholarly Research. *Ann. Oper Res.* **2022**. [CrossRef]
65. Wu, W.; Chen, W.; Fu, Y.; Jiang, Y.; Huang, G.Q. Unsupervised Neural Network-Enabled Spatial-Temporal Analytics for Data Authenticity under Environmental Smart Reporting System. *Comput. Ind.* **2022**, *141*, 103700. [CrossRef]