sustainability

MDPI

Article

A Text-Mining Analysis on the Review of the Non-Financial Reporting Directive: Bringing Value Creation for Stakeholders into Accounting

Simona Fiandrino * and Alberto Tonelli

Department of Management, University of Turin, 10134 Turin, Italy; alberto.tonelli@unito.it
* Correspondence: simona.fiandrino@unito.it

Abstract: The recent Review of the Non-Financial Reporting Directive (NFRD) aims to enhance adequate non-financial information (NFI) disclosure and improve accountability for stakeholders. This study focuses on this regulatory intervention and has a twofold objective: First, it aims to understand the main underlying issues at stake; second, it suggests areas of possible amendment considering the current debates on sustainability accounting and accounting for stakeholders. In keeping with these aims, the research analyzes the documents annexed to the contribution on the Review of the NFRD by conducting a text-mining analysis with latent Dirichlet allocation (LDA) probabilistic topic model (PTM). Our findings highlight four main topics at the core of the current debate: quality of NFI, standardization, materiality, and assurance. The research suggests ways of improving managerial policies to achieve more comparable, relevant, and reliable information by bringing value creation for stakeholders into accounting. It further addresses an integrated logic of accounting for stakeholders that contributes to sustainable development.

Keywords: Non-Financial Reporting Directive (NFRD); non-financial information disclosure; text-mining analysis; stakeholder accounting; sustainable development

check for updates

Citation: Fiandrino, S.; Tonelli, A. A Text-Mining Analysis on the Review of the Non-Financial Reporting Directive: Bringing Value Creation for Stakeholders into Accounting. *Sustainability* **2021**, *13*, 763. https://doi.org/10.3390/su13020763

Received: 15 December 2020
Accepted: 11 January 2021
Published: 14 January 2021

1. Introduction

Directive 2014/95/EU on the disclosure of non-financial and diversity information (NFI) has been acknowledged as a historic breakthrough in the transition towards more accountability in and responsiveness to sustainable development [1,2]. In more detail, it amends Directive 2013/34/EU and obliges public interest entities to report "an understanding of the undertaking's development, performance, position, and impact of its activity, relating to environmental, social, and employee matters, respect for human rights, anti-corruption, and bribery matters. This includes a brief description of the undertaking's business model; a description of the policies pursued, the due diligence processes implemented, the outcomes and risks related to those matters and, ultimately, non-financial key performance indicators relevant to the business" [3]. With this regulatory intervention, the disclosure of non-financial information (NFI) has shifted from a primarily voluntary-based approach [4–6] to a mandatory regime. More recently, the Communication on the European Green Deal has stressed the need to strengthen the foundations of sustainable investment by enhancing companies' and financial institutions' sustainability disclosure [7]. The European regulatory setting on the disclosure of NFI has moved a step further in this direction with the Review of the Non-Financial Reporting Directive (NFRD). The consultation document for the Review of the NFRD was available from 20 February 2020 to 11 June 2020, and the European Commission (EC) expects to adopt a proposal regarding the NFRD in the first quarter of 2021 formulated around eight thematic issues: quality and scope, standardization, materiality, assurance, digitalization, location of the reported information and personal scope, simplification, and reduction of administrative burden. The EC has provided a Summary Report of the Public Consultation on the Review of the

NFRD to summarize the survey responses gathered from this consultation and to present related descriptive statistics. The analysis did not, however, deeply examine the annexed documents that some academics and practitioners attached to the questionnaire. These texts constitute valuable units of analysis as they offer well-grounded perspectives on NFI matters, including their relevance, and thus are worthy of further scrutiny. This is the first of two knowledge gaps [8] that this research aims to address. The study, therefore, complements the analysis of the Summary Report of the Public Consultation on the Review of the NFRD by delving into the prevalence of the topics and related relationships.

In the context of NFI disclosure, the engagement of academics in the field of sustainability accounting is remarkable. Academics have extensively analyzed these progressive changes and identified intertwined limitations to further developments. Some scholars have observed high levels of adherence to the new mandatory framework [9,10] in comparison with NFI disclosure under a voluntary basis [11] but negligible quality regarding how NFI relates to the core business and for whom it matters [12,13]. Others have highlighted several guidelines and international frameworks that have led to a heterogeneous landscape of accounting procedures, metrics, and disparate approaches instead of a common language of sustainability reporting [14–16]. Consequently, sustainability reporting is affected by a lack of comparability, reliability, and relevance of NFI [4,17]. This is the second knowledge gap [8] that this research aims to address. Thus, in light of the Review of the NFRD, the study suggests resolutions and areas of possible amendment.

Focusing on these two knowledge gaps, the research has the following objectives. First, it aims to discover which topics have been at the heart of discussions of the Review of the NFRD, how prevalent each topic is, and how the topics relate to each other. Second, it aims to understand which resolutions can be improved to overcome the lack of comparability, reliability, and relevance in NFI.

To pursue these aims, the study applied the latent Dirichlet allocation (LDA) probabilistic topic model (PTM) technique to inductively detect the number and nature of topics included in the annexed documents. This technique has recently been applied to review the literature in various research areas to understand the historical evolution of ideas [18], cultural trends, and the dynamics of literary meanings over time, as well as to derive theory and knowledge from big data [19]. Such a "big data approach" has found applications in accounting history research [20], marketing [21,22], organizational research [23], strategic management [24], and business ethics [18].

This research has both practical and theoretical implications. From a practical perspective, it contributes to the exigence to revise Directive 2014/95/EU by complementing the outlook of the Review of the NFRD with an inductive discovery of topic-based patterns that have been at the heart of the discussion of NFI implementations. Accordingly, regulators may revise the NFRD by considering the ranking of the main issues underlying the annexed documents along with the responses to the questionnaire they have already analyzed. Thus, the study suggests future directions for possible amendments to policy interventions by reflecting upon the challenges ahead to enhance the comparability, reliability, and relevance of NFI. From a theoretical perspective, this research emphasizes the need to adopt an integrated approach to accounting for stakeholders that aligns reporting procedures, strategic objectives with risk assessments of financial and non-financial issues linked to stakeholders' relationships under an integrated thinking logic.

The paper proceeds as follows. Section 2 outlines the developmental pathways of sustainability accounting. Section 3 introduces the LDA PTM technique and describes the method of computing the text-mining analysis. Section 4 presents the findings of the study, showing the prevalence of four topics: quality of NFI, standardization, materiality, and assurance. Section 5 discusses each main topic at stake in the argumentation, and Section 6 concludes by addressing both the theoretical and practical contributions of the study as well as limitations and suggestions for future research avenues.

2. Developmental Pathways of Sustainability Accounting

2.1. The Flourishing of Sustainability Reporting under a Mandatory Regime

Sustainability reporting dates back to the 1970s and derives from the managerial attempts in favor of corporate social responsibility (CSR) practices which exploded in this period [25–27]. This brought a new vocabulary of sustainability accounting, sustainability reporting, CSR reporting, social performance, social disclosure, and accountability [28] with interchangeable meanings and overlaps with one another. The focus of sustainability reporting relates to information directly projected onto interested stakeholders [29]. In this way, economic, social, environmental, and ethical issues may maintain the business continuum in an integrated manner and companies respond to stakeholder interests with accountability [30–33]. Sustainability reporting is of relevance to a plurality of stakeholders and academics have usually likened its analysis to stakeholder engagement practices [34,35]. The idea to include social, environmental, economic, and ethical issues at the core of the business model has become even more rooted in the common language of business after the increasing number of corporate scandals and the stock market collapse during the 2008–2009 Global Financial Crisis. This period of crisis called companies to respond to regulatory changes related to working practices, among the others [36], and to rethink business and ethics as jointly intertwined [37,38]. Moreover, the deficiencies of a short-term view [39] favoring shareholder value maximization [40,41] as the primary and exclusive purpose of business have been questioned. Conversely, the view on how business and ethics need to be integrated with each other has been theorized as a stakeholder approach to value creation [37,38] and has gathered primary importance. The idea of stakeholder theory is that "business can be understood as a set of relationships—which are not reducible to transactions—among groups that have a stake in the activities that make up the business. Business is about how customers, suppliers, employees, financiers (stockholders, bondholders, banks, and so on), communities, and managers interact and create value. To understand a business is to know how these relationships work" [37]. In fact, both investors and shareholders have ignored that "surplus could potentially derive from social and environmental externalities" [42]. In parallel, and more recently, a progressive awareness against grand challenges facing society worldwide, such as environmental concerns, human rights, gender equality, and so forth, has been one of the top priorities in the political and institutional agendas of policymakers. Consequently, on the one hand, businesses have started voluntarily disclosing their objectives, actions, and performances on sustainability issues, thus, sustainability reporting has become institutionalized [43–46]. On the other hand, there has been a rapid explosion of international standards frameworks as well as third-party reporting certification [44–48].

Academics have investigated the effectiveness of sustainability reporting guidelines such as Global Reporting Initiative (GRI) and International Integrated Reporting Council (IIRC). For instance, Higgins, Stubbs and Milne [43] argue that the institutionalization of sustainability rankings and taxonomies of environmental, social, and governance (ESG) metrics lead to business-as-usual practices and greater levels of un-sustainability if those guidelines are implemented into the reporting processes without considering stakeholder involvement and accountability. Similarly, Stubbs and Higgins [49] have analyzed the preferences of users of non-financial reporting for regulatory or voluntary approaches to integrated reporting. Interviews highlighted an inclination toward voluntary approaches to Integrated Reporting (IR) as the more legitimate practices because voluntary sustainability reporting has not increased the quality of reporting. In fact, other academic studies in the field of sustainability reporting under a voluntary basis have investigated the relationship between voluntary NFI disclosure and sustainability performance [50], its linkage with operating profitability, market values, and expected growth rates in cash flows [51] by underlying a positive effect of social disclosure on both profitability and market values.

In modern times, the introduction of Directive 2014/95/EU (NFRD) has signed a progressive movement, from a voluntary regime of disclosure enacted by an international standards framework, to a mandatory regime of disclosure regulated by the law [52,53].

The NFRD introduces the reporting of environmental, social, employee issues, human rights, and bribery and corruption, on an annual basis, and requires large, listed companies, banks, and insurance companies with more than 500 employees to meet these reporting requirements. In 2017, the European Commission published non-binding guidelines to instruct companies towards the disclosure of relevant and comparable information [54].

Several academic notes have deepened the analysis of such disclosure's mandatory requirements, even more adamantly emphasized considering the different transpositions into national laws [55,56] and the managerial discretion the European Commission left to the preparers of NFI [4,5]. The effectiveness of this new regulatory framework has started to be vividly debated in academic terms [56]. Research on sustainability reporting under a mandatory basis has empirically investigated the consequences of a mandatory regime of NFI disclosure [57]. In more detail, this study focuses on the mandatory effects of ESG information by asserting that "even in the absence of a regulation that mandates the adoption of assurance or specific guidelines, firms seek the qualitative properties of comparability and credibility" (p. 1). Moreover, the findings reveal a positive relationship between sustainability disclosure and firms' valuations with Tobin's Q as proxy; namely, companies are increasingly enhancing their disclosure in terms of both quantity and quality to enhance transparency around their societal impacts. Furthermore, other studies have illustrated how different member states have implemented the Directive into national law, and the results have shown different applications around the content requirements, boundaries of reporting, assurance, sanctions for non-compliance level [53]. Other academic studies have highlighted negligible disclosure of NFI, scarce procedures on risk assessment and materiality determination processes [10]. Specific requirements for reporting standards [13,53,58] and the complexity to collect non-financial data and ensure comparability [17] have been acknowledged as primary deficiencies of these new mandatory NFI disclosure. Together with the myriad of international standards frameworks that jeopardize comparability of data, this ill-defined set of non-financial reporting practices may potentially lead to levels of "green-washing" or "un-sustainability" [59,60] that encourage business-as-usual practices with instrumental, merely compliance-driven, and profit-oriented terms.

Moreover, several academic scholars have emphasized uncharted areas. There are two reporting dilemmas that remain unexplored: (1) How do we measure our performance with each stakeholder? (2) How do we measure our performance with our entire set of stakeholders? [60,61]. Relatedly, the stream of stakeholder accounting [15,62–64] identifies lacks into non-financial reporting reports as scholars argue an overall weakness in the underlying premises used to justify these non-financial reports [15]. In greater detail, Mitchell et al. [15] acknowledge that present accounting practice limits the reporting of stakeholder value distribution and does not encompass a proper disclosure of stakeholder value creation. These are related to the weak consideration of the full range of risks borne by stakeholders and inclusive mechanisms to value stakeholder-focused decision-making [15,62–65]. Other research has pointed out the need to better rank the materiality determination process and related stakeholder engagement when defining material issues, as we do not know which comes first and whether there is a participatory involvement when discussing material judgements [65].

Drawing on these considerations so far, on the one hand, it seems to be in a murky setting without a standardization of NFI reporting; on the other hand, the above-mentioned academic open questions along with the EU's commitment toward a greener and more sustainable economy and society have led to the Review of the NFRD [17]. Therefore, in this ever-evolving context, it is of keen focus to outline the new updates on the Consultation of the Review of the NFRD, which are presented in the next section.

2.2. The Challenges Ahead for the Review of the Non-Financial Reporting Directive (NFRD)

The Review of the NFRD aims to overcome the following drawbacks. First, there is inadequate publicly available information about non-financial issues which leads to less comparable, relevant, and reliable NFI. Second, companies incur avoidable costs related to

NFI reporting because of the complexity arising from different disclosure requirements. Thus, the regulator acknowledges that the managerial discretions to reporting NFI need to be revised in order to ensure investors adequate NFI on sustainability-related risks and to ensure civil society organizations, trade unions, and others adequate NFI on companies' impact on society and on the environment, and, ultimately, to simplify and reduce administrative burden for companies [66].

The Summary Report of the Review of the NFRD [67] explained that 588 organizations including companies, non-governmental organizations (NGO), academia, and public authorities have participated to the public consultation on the Review of the NFRD. Twenty-nine percent of respondents were both preparers and users of NFI disclosure, 22% were users of such information, and 21% were only preparers. The consultation consisted of 45 questions divided into eight thematic issues: quality and scope, standardization, materiality, assurance, digitalization, location of the reported information and personal scope, simplification, and reduction of administrative burden.

Considering the quality and scope of NFI, the key underlying concerns raised by respondents included reliability (37% of the total respondents), limited comparability (47%), and limited relevance (36%) of reported information. Respondents suggested further additional content issues with reference to taxonomy regulation related to six environmental objectives: (1) climate change mitigation; (2) climate change adaptation; (3) sustainable use and protection of water and marine resources; (4) transition to a circular economy; (5) pollution prevention and control; (6) protection and restoration of biodiversity and ecosystems. In fact, 69% of respondents agreed that environmental matters need to be revised accordingly. Other content issues refer to governance, supply chain, and intangibles such as intellectual property, software, customer retention, and human capital. Moreover, noteworthy to be included as additional categories are targets and processes, scenario analysis, forward looking information, and contribution to the Sustainable Development Goals (SDGs).

With reference to standardization, 82% of respondents favored the application of a common standard for non-financial reporting with sector-specific elements that will alleviate comparability and reliability issues. Fifty-one percent of respondents agreed with the application of the Global Reporting Initiative as main reporting framework of sustainability-related issues, while 72% of respondents suggested the use of simplified standards as a means to limit the burden for small and medium-sized enterprises (SMEs) regardless of a mandatory regime of application (agreed with by 57% of respondents).

In relation to materiality, the emerging consensus is not clear-cut. While the materiality definition on outside-in risks is proper to determine a company's development, performance, and position (69%), the materiality definition on inside-out risks is not appropriate to understand a company's impact on society and the environment (54%). Relatedly, 72% respondents suggested the further disclosure of the materiality determination process.

With regards to assurance, 66% of respondents raised several differences in the assurance requirements between financial information and NFI; moreover, member states have transposed the assurance requirements differently into national laws. For these reasons, 67% of respondents agreed in favor of stronger assurance legal requirements.

Like both assurance and materiality, digitalization also needs further improvements by tagging of reports containing NFI to make them machine-readable with a proportionate balance of costs and benefits, even if the difficulties of its implementation are acknowledged, especially when considering the primarily qualitative nature of NFI. Moreover, even the location of the NFI reports need to be revised as respondents suggested that publishing non-financial statement as part of a separate report creates problems to find NFI.

Ultimately, possible revisions are taken into account to broaden the scope of the NFRD to additional public interest entities (PIEs). Sixty-two percent of respondents said the scope of the NFRD should be broadened to all large PIEs by removing the 500 employees threshold and aligning with the size definitions used elsewhere in the Accounting Directive.

Drawing on these considerations thus far, the public consultation allowed respondents to attach documents to the responses to support the argumentation. These documents constitute valuable units of analysis; thus, the present study complements the analysis of the Review of the NFRD by delving into the textual understanding of the annexed documents with a text-mining analysis.

In the next section, we focus on the textual analysis technique adopted in this study for rendering the thematic structure of the NFRD: topic modeling.

3. Research Method

The research method employed for this study is based on a text-mining analysis of the documents annexed to the contribution on the Review of the NFRD. The analysis has been conducted with the latent Dirichlet allocation (LDA) probabilistic topic modeling (PTM) technique. We chose this method because it comprehensively assesses extensive textual data sets and inductively sorts out main arguments. Thus, this technique has been employed to address the following research questions:

(1) Which topics have been discussed in the annexed documents to the contribution on the Review of the NFRD?
(2) How prevalent is each topic and how do the topics relate to each other?

The following sub-sections are organized as follows. First, we present the methodological issues involved in conducting a textual analysis with topic model, including its strengths as well drawbacks to reinforce the reasons for applying this method here. Next, we describe the sample selection, and finally, we delve into the procedural steps of the textual analysis with PTM.

3.1. Textual Analysis with Topic Modeling

New methods of textual analysis borrowed from computer science have progressively gained attention. More specifically, computer-based language processing approaches have recently been applied to management disciplines [19,21] as well as social sciences [68] and include content analysis, general natural language processing (NLP) of texts, topic modeling [19], and sentiment analysis [69], among others.

Topic modeling has been applied in the present study because it provides an automatic procedure that scans a large collection of texts and inductively discovers topic-based patterns—namely "it breaks down the corpus into themes" [24]. As stated by the pioneering research of this technique [70] "topic modeling algorithms are statistical methods that analyze the words of the original texts to discover the themes that run through them, how those themes are connected to each other, and how they change over time." In topic modeling, some computational text processing methods are used: latent semantic analysis (LSA) decomposes a corpus of terms into a vector space that records the number of times a term occurs within the document and computes the similarity between terms, while LDA is a probabilistic topic model (PTM) of a collection of texts which represents documents as a "random mixture over latent topics, where each topic is characterized by a distribution over words" [71].

LDA topic modeling inductively discovers underlying topics running through a large collection of texts (i.e., corpus) and infers word probabilities in topics. It relies on two assumptions: first, each document has its own topic distribution in which topics are placed randomly; second, each topic has its own word distribution in which words are placed randomly from the word distribution [72]. In other words, "each word in a document [is modeled] as a sample from a mixture model, where the mixture components are multinomial random variables that can be viewed as representations of 'topics.'" [19]

Under these assumptions, the LDA algorithm considers the documents as "a bag of words" with no syntax and "discovers the topic distribution for each document and the word distribution of each topic iteratively by fitting this two-step generative model to the observed words in the documents until it finds the best set of variables that describe the topic and word distributions" [72]. That is, "words that occur in the same contexts

tend to have similar meanings" [73]. For instance, the algorithm sorts out the following co-occurrence of words within a document: "landform", "hiking", "landscape", "hill", and "views". This means that such repeating words are likely to refer to the topic "mountains". Moreover, "topic models are able to capture polysemy and different uses of a word based on the contexts in which it occurs. For example, the term bank can co-occur with words like money and credit in one topic and river and water in another topic—indicating two very different meanings for the same word" [23]. This means that the meaning surfaces from the linkages between words instead of solely from the words themselves [74].

The main strengths are the following. First, topic modeling comprehensively assesses extensive textual data sets, thus, it has replaced human evaluations of large volumes of documents and words, which are almost impossible to be manually coded by either single researchers or research teams via qualitative data analysis approaches [75]. Other advantages that can be derived from this technique are the objective selection of words from a collection of documents and the capacity to record the weight of each word based on the probabilistic calculations. Thus, the technique overcomes the pre-established clusters of selections and related implicit human biases, because researchers inductively uncover latent topics based on probability-based parameters without imposing categories on the data [18].

Conversely, some drawbacks must be taken into consideration. In more detail, the set of words with the best probability of co-occurrences do not display the name of the topic whose words are being referred to. Taking the abovementioned example of the words "landform", "hiking", "landscape", "hill", and "views", the algorithm is not capable of inferring the topic "mountains" from these words when iterated together; therefore, the joint collaboration of researchers in addition to their understanding of groups of words are crucial for associating related topics.

Despite this, the emerging approach is better suited than other research methods, like manual content analysis, for presenting a more comprehensive picture of emerging issues in a set of documents.

3.2. Sample Selection

The sample for investigation was collected from the annexed documents to the contribution on the Review of the NFRD that can be retrieved at the EC website (at https://ec.europa.eu/12129-Review-of-Non-Financial-Reporting-Directive/public-consultation). The total number of available documents was 128. In order for a document to be included in the sample, we applied the following screening criteria. First, we excluded twelve non-English documents (written in French, Spanish, Italian, German, and other languages) to maintain comparability among the documents. Second, we omitted six duplicated documents to avoid double counting, and we did not include nine documents with just one page or containing just figures and graphs as these could jeopardize the algorithm. Thus, a final total of 101 documents were used. Along with these criteria, the screening procedure is shown in Table 1.

Table 1. Sample selection procedure.

Description	Number of Documents
Total number of documents to the contribution on the Revision of the Non-Financial Reporting Directive (NFRD)	128
• Screening of non-English written documents	12
• Screening of duplicates	6
• Screening of documents with one-written page or that have figures and tables only	9
Final sample of investigation	101

3.3. Data Cleaning, Processing, and Analysis with Topic Model

In accordance with Debortoli et al. [75] and Hannigan et al. [19], we operationalized the topic model using Python in three steps: data cleaning, data processing, and data analysis.

First, with the corpus of data gathered from the 101 documents, we applied data cleaning, as text data are large and contain lots of noise and duplicates, which can undermine the text-mining analysis. Thus, we removed stop words (like 'a', 'the', and 'and'), erased the punctuation, numbers, dots, and non-alphanumeric symbols, and converted all the letters into lower case. Then, we considered variations in word forms with the same meaning, such as 'findings' and 'finding'; thus, we lemmatized words (reduced them to their dictionary form, such as plural to singular for nouns and simple present tense for verbs; see Debortoli et al. [75]). After this screening phase, the base document unit for the analysis with the LDA algorithm was related to each sentence of the 101 collected documents. The total number of collected words was 135,000.

Next, we conducted the data processing to describe the corpus of texts. In doing so, we obtained a visual representation of the most common words by counting the frequency of each term in each document, and we constructed a 'bag of words' using the word cloud package. This pre-processing phase had to be conducted before training the model to ensure that the cleaning phase was properly addressed and to outline the samples of the texts of the analyzed documents. The results are shown below (Section 4.1).

Finally, we processed the LDA model by training the model with the LDA visualization package (pyLDAvis). We analyzed the results by visualizing the relationships between the topics and their interpretability (see below, Section 4.2). To verify whether there was a correspondence with the eight thematic issues suggested by the EC in the consultation, we set this (eight) as the number parameter of topic labels for the LDA model, while the number of words for each topic the model sorted out was set at 10, in line with prior research [18,22,73,74]. In other words, we aimed to understand whether the eight topics proposed by the EC were of equal importance in the discussion in the annexed documents or whether there had developed a concentration on some topics around certain issues. Based on these results, we reviewed the number parameters of topics for the LDA algorithm in order to feature the most relevant topics.

The analysis followed the protocol related to the topic interpretability suggested by Sievert [76] and topic coherence following Röder [77]. The topic interpretability comprehends the understanding of meanings and the prevalence of topics as well as their relationships. The meanings and the prevalence of topics were derived using the pyLDAvis package by counting the topic-specific frequency of terms; when analyzing the relation between the topics, we examined the topics–terms relationship [76]. The topic coherence considers the proximity of topics sorted out in the query words of the document passages [77], exploring the space of topic coherence measures. Thus, we aimed to discern their specificities by analyzing the LDA output as grouped into the words of the main topics and their relations for each topic. We complemented this phase by interpreting the meaning of the topics and by assigning to each group of words the related topic label. In this inductive phase, we also performed an intellectual analysis to rate individual documents within each topic as a validation [24]. In so doing, we coded each group of words autonomously and independently, analyzing the topics of the same data package and considering the literature on the co-occurrence of words and topic model [20]. By way of a final check, we discussed the descriptive labels assigned to each group of words so as to achieve consensus about the thematic pattern of each topic. When divergences occurred, we discussed the matter until we arrived at an agreement on the label related to the topic.

Figure 1 shows the three-part process of data cleaning, data processing, and data analysis as described. The results of the visual representation of most common words, the prevalence of topics, and their relationships are described in the next section.

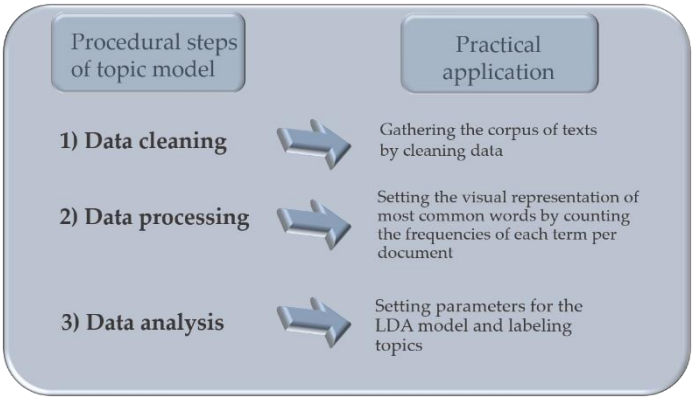

Figure 1. Research method.

4. Results

The results are shown thereafter. First, we display a visual representation of the most common words computed in the data processing phase. Second, we present the LDA model with eight topics to demonstrate the relationships between the topics of the Review of the NFRD, and we inductively identify the main topics at the core of the discussion.

4.1. Visual Representations of the Most Common Words

Figure 2 displays visual evidence of the most common words adopted in the corpus of texts. As already explained in the research method section, this pre-processing outlines the main words with the aim to understand whether the data cleaning has been computed correctly; it does not attribute any kind of quantitative data. Unsurprisingly, the prevailing words are "reporting", "financial", "non-financial", "European", "Directive", "business", "NFRD", "ESG" (which stands for environmental, social, and governance), "standards", "human", and "rights" (interpreted as "human rights").

Figure 2. Word cloud.

In order to quantify the frequency of these words, we plotted the ten most frequent words based on the list of document vectors we created in Python. Figure 3 shows the histogram that graphically displays these ten most common words. We have "reporting" with 1662 occurrences, "information" with 1384, "nonfinancial" with 1177, "financial" with 897, "company" with 751, "business" with 666, "EU" with 635, "rights" with 623, "human"

with 605, and "disclosure" with 603. Each of these words also occur in the word cloud shown above, but they are now displayed in graphical order ranked by their occurrence.

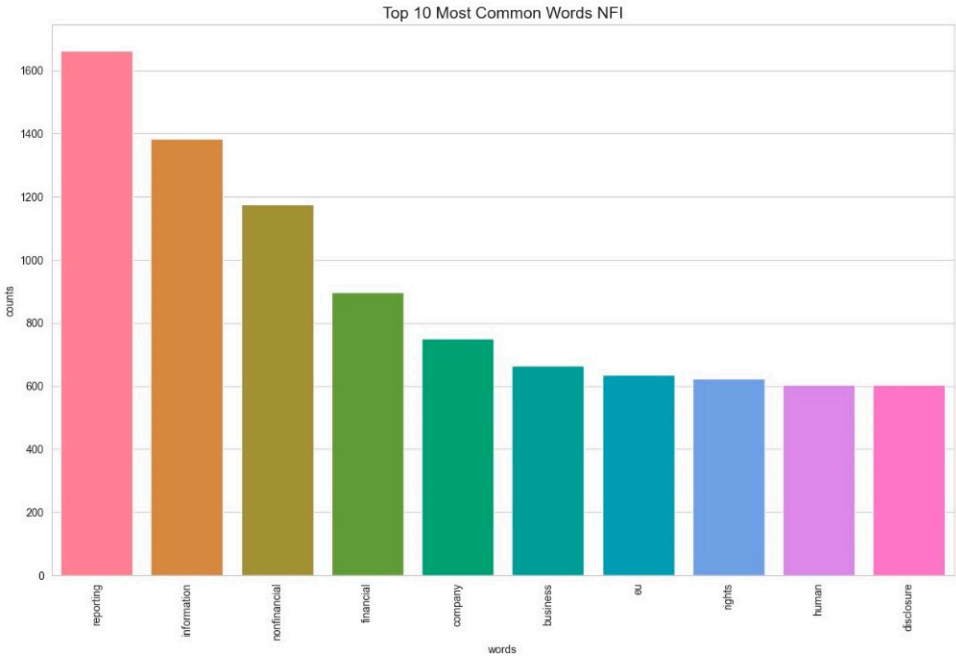

Figure 3. Top ten most common words.

4.2. Inductive Discovery of the Main Topics

We first sought to understand whether there is any correspondence with the eight thematic issues proposed in the consultation of the NFRD. To this aim, we set eight as the main parameter to process the LDA topic modeling using the pyLDAvis package in Python, according with Goloshchapova et al. [78]. pyLDAvis presented a two-dimensional display of the topics, shown in Figure 4.

On the right-hand side, the Top-30 Most Salient Terms identifies the salience of words, namely, the prevalence of words in terms of their overall term frequency. We did not select any specific topic, for instance 1, 2, or 3, thus, the list comprehends the most common words among all the corpus of texts. On the left-hand side, the Intertopic Distance Map presents the relationships between the topics; these relations are reflected in the position and the amount of the overlap. First, we have to consider that all circles are the same size, that means they are of equal importance and none is dominant over another. Second, as we can see, we can clearly distinguish three topics as they are displayed oppositely in the graph, while topics 4, 5, 6, 7, and 8 are overlapping with each other. This means that the word list for the topics is not clear-cut and some words are present both in topic 5, 6, 7, 8. Looking at this Intertopic Distance Map and linking to our research objective, it is possible to draw the following implication. There is not a perfect correspondence with the eight thematic issues proposed by the consultation as the overlap between topics 4, 5, 6, 7, and 8 is evident. Therefore, in order to better delve into the understanding of which topics were the most discussed and what the related argument was about, we hypothesized the presence of four main topics as main issues on debate, then, we ran the LDA algorithm again by setting four topics as the main parameter.

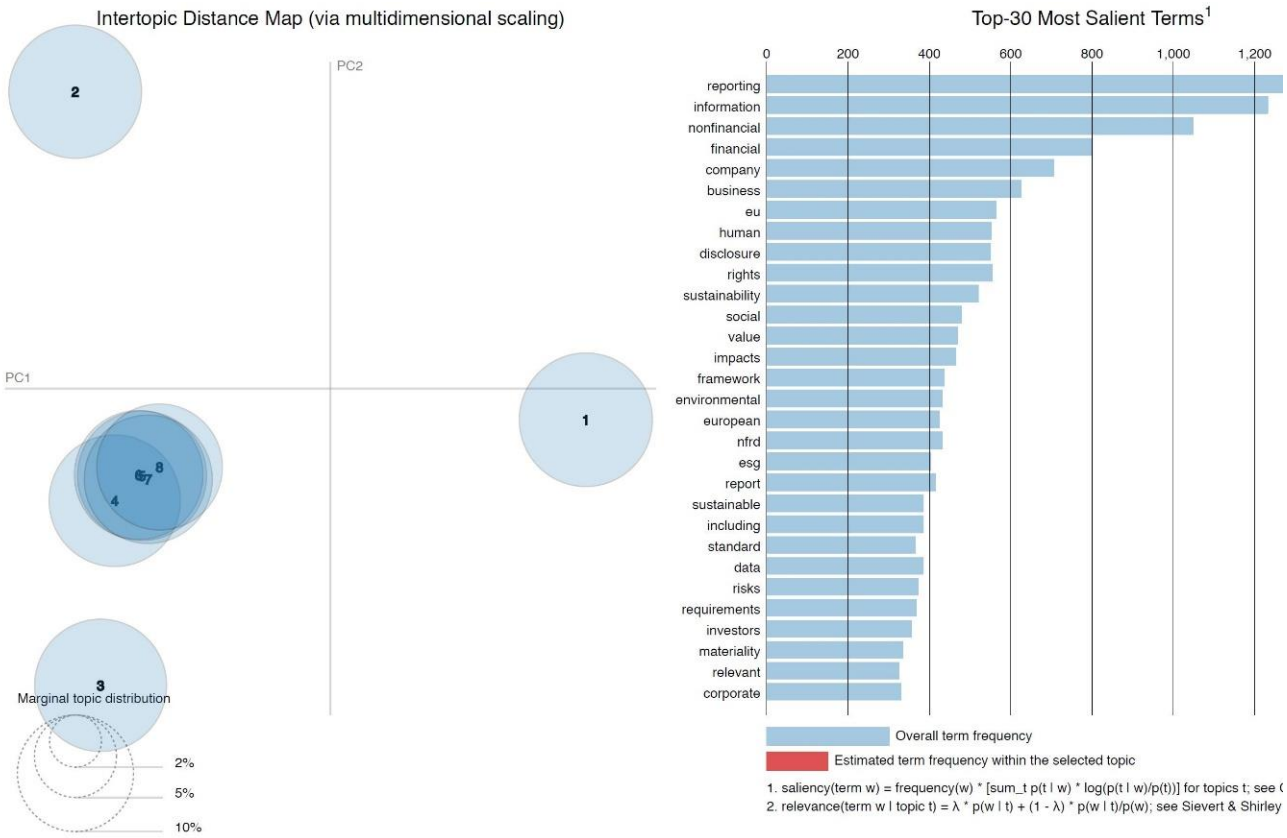

Figure 4. pyLDAvis visualization–eight topics.

Figure 5 plots the distribution of the inductive discovery of four topics when setting four as the parameter on the LDAvis analysis. On the right-hand side, the Top-30 Most Salient Terms are quite similar to before as the selection includes the full dataset. On the left-hand side, however, we now have a clear-cut distinction between the prevailing topics, therefore, we can deduce that there are four main issues at stake in the consultation. As before, all circles are the same size, namely, of equal value. In order to delve into the understanding of each topic, we gathered the LDA results with ten words, in line with prior research that suggests a number ranging from five to ten words as sufficient to decide the content for each topic (see Goloshchapova et al. [78]). In so doing, the algorithm gathers ten words by proximity and granularity; this means that the words are displayed in terms of how closely related they are to the data set.

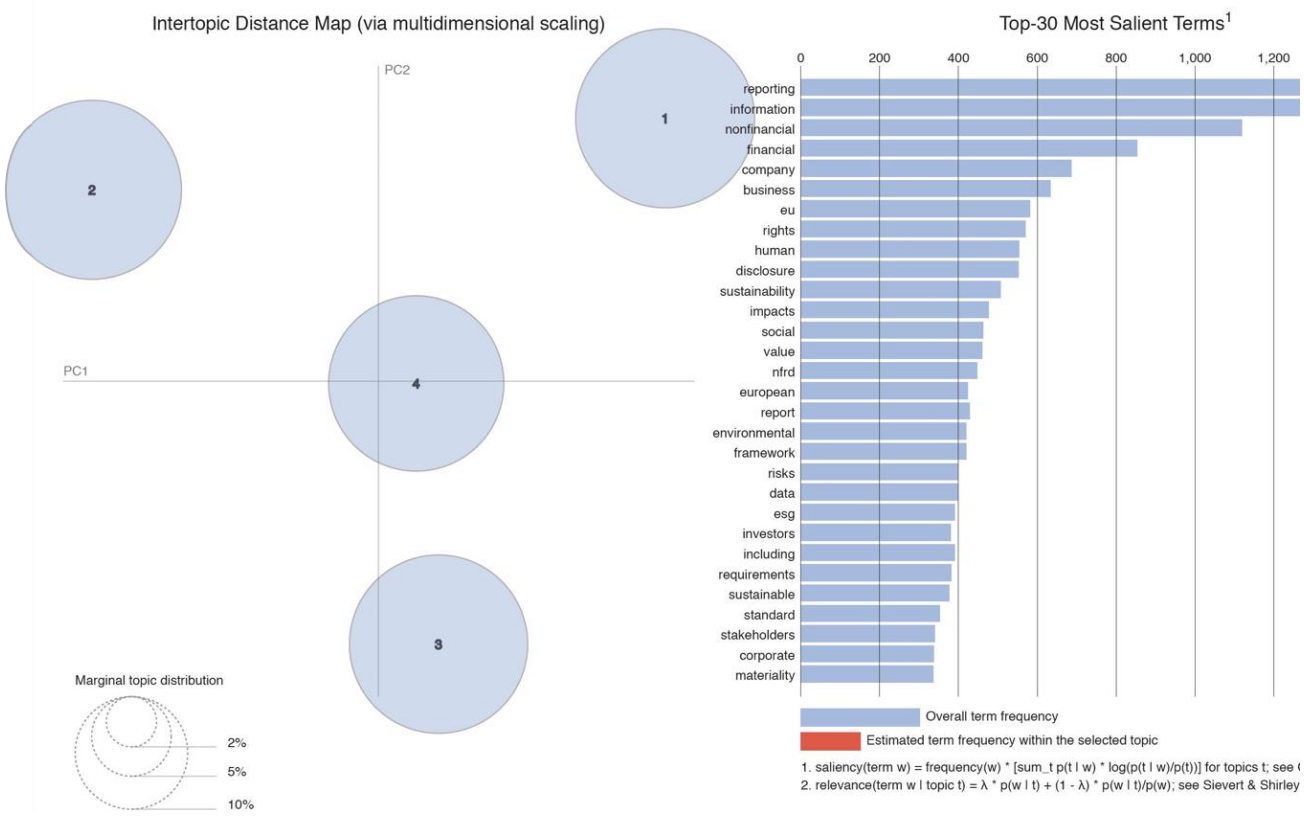

Figure 5. LDAvis visualization–four topics.

Table 2 presents the list of words the algorithm presented and the title of each topic which was inductively derived by the researchers. Topic 2 entails the words "nonfinancial", "financial", "framework", "business", "European", "risks", "stakeholders", "investors", "standard", "public"; thus, it refers to the standardization and the usage of a common framework. Topic 3 has "reporting", "sustainability", "social", "report", "relevant", "including", "sustainable", "corporate", "materiality", "global". Topic 4 includes "Impact", "value", "assurance", "nfrd", "reliability", "data", "esg", "requirements", "society", "process".

Table 2. Four main topics.

Topic Name	10 Words Extracted by the Algorithm
Quality of NFI	Information, company, eu, rights, human, disclosure, management, need, scope, environmental
Standardization	Nonfinancial, financial, framework, business, European, risks, global, investors, standard, public
Materiality	Reporting, sustainability, social, report, relevant, including, sustainable, corporate, materiality, stakeholders
Assurance	Impact, value, assurance, nfrd, reliability, data, esg, requirements, society, process

With this text-mining analysis, we can deduce the prevalence of four main topics (quality of NFI, standardization, materiality, assurance) at stake during discussions between academics, practitioners, and other business organizations. The issues of human rights and environment are the main content issues which respondents suggest as characterization of the quality of NFI. Standardization means the adoption of a framework which facilitates the comparability of data for both investors and stakeholders. Materiality and assurance have always been discussed by academics and practitioners since the first inception of the mandatory regime of NFI disclosure. To further the arguments on these topics in the

Discussion sections, the quotes of some annexed document has been reported to portray the phenomenon together with the theoretical perspectives of academics on these issues.

5. Discussion

Based on the results described above, the discussion delves into the four main topics at stake for the public consultation, which are sorted out from the textual analysis. The four main topics are systematized with the ranking presented in Table 2 and discussed with support for the manual reading of the annexed documents.

NFI quality emerges as the first topic in terms of relevance. There is a common consensus that current non-financial reporting does not offer a holistic view of companies' business models, associated risks (both financial and non-financial), related material issues, a fair taxation of value-added, and does not properly comprehend perceived potential liability or exposure areas. Scholarly academics in sustainability accounting have deepened analysis on NFI disclosure quality regarding the illustration of *what* and *how* companies disclose their sustainability. The *what* dimension of NFI disclosure refers to the policies and practices in outlining social, environmental, and employee matters, as well as human rights and anti-corruption issues related to information completeness. The *how* dimension of NFI disclosure deals with the ways of disclosing, ranging from readability to credibility and from accuracy to the disclosure tone, by including stakeholders' judgments [79,80].

Standardization is the second topic that surfaces from the textual analysis with topic model. In fact, this issue is quite widespread in the stream of NFI disclosure regarding the significant variation across companies and jurisdictions because of managerial discretions left to preparers. The various standards and guidelines should be streamlined to avoid duplication, remove gaps, and ensure that preparers have a clear set of standards. The International Trade Centre has identified nearly 255 standards, codes of conduct, and audit protocols that address sustainability-related information across 80 sectors and 180 countries [81]. Thus, a myriad of international standards frameworks has been developed during the 21st century to improve the reporting of sustainability issues and NFI disclosure, leading to ill-defined and murky procedures that have even emphasized several approaches to assess material topics.

Relatedly, materiality is the other important issue jeopardized by information that is readily available as opposed to relevant matters that are core to the business. For instance, "typical company disclosure is too vague, general, and incomplete to allow for external parties to understand the company's potential and actual human rights risks, impacts and plans to address them." This viewpoint confirms the perspective of Puroila and Mäkelä [65], who argue that the application of the materiality principle and its determination process follows a "tick-the-box" approach. In fact, materiality should be considered as a process that comprehends the interplay between the disclosure of useful information to stakeholders, integrative risk management practices, and ESG performance and evaluations.

Similarly, the assurance and verifiability of information, which is the fourth topic in terms of importance, is demanding, as it has not been applied unanimously in all member states. This procedure will complement the assessment of relevant and reliable information by avoiding "greenwashing" behaviors. In turn, the quality of sustainability reporting will be improved by closing the loop.

Overall, the reporting process should provide a "means whereby the summarized information that produces accountability can be reported in such a way that the collaborating parties receiving the accounting reports can evaluate their risks and apportion rewards" [64]. Businesses must be accountable to a variety of stakeholders and leaders, who are asked to responsibly react to disruptive changes with resource rebalancing, investments in workforce training, and advancements in new models of summarizing sustainability risk-related information for different stakeholders. Harrison and van der Laan Smith [63] maintain that the premises of responsible accounting for stakeholders support information on the risk assessment of financial and non-financial issues linked to stakeholders' relation-

ships and transactions. In such a way, stakeholders can precisely manage risks associated with their investments in both financial and non-financial firm resources.

The accounting method that considers risk assessments of financial and non-financial issues in joint conjunction with stakeholders' relationships will trickle down managerial tools to further develop integrated strategies in line with the core business and stakeholders' interests. Thus, this approach to assess and manage risks considering stakeholders' relationships implies a deep understanding of potential value created and destroyed, both from a financial and non-financial perspective and embraces the value creation/destruction with inclusion/exclusion of any value types (financial, relational, intellectual, or environmental).

6. Conclusions

This research depicted the thematic evolution of the large stream of NFI reporting and delved into analysis of the Review of the NFRD to understand the main topics covered in public consultation. Toward this end, the study tracked 101 annexed documents written in English and performed a textual-mining analysis with the topic model rendering technique. Such a big data approach gathers an extremely high amount of textual data to establish the top themes according to the probabilistic allocation. Topic model is a powerful facilitator for textual analysis, especially when handling a considerable amount of text data, as the algorithm employs rigorous and well-defined control over analysis granularity. However, as with all big data approaches, topic model must be implemented carefully by supporting—and not substituting—the scientific understanding of researchers, who contribute to more thoroughly delineating the boundaries of such a text-mining approach with the ultimate goal of enriching knowledge and generating insights. As such, the study complemented the textual analysis by meticulously reading the text of the annexed documents under analysis. This complementary approach allowed us to extract the following findings. From the eight thematic issues the public consultation focused on, the research found that four main topics were at the stake of discussion. These topics were NFI quality, standardization, materiality, and assurance. Sustainability accounting academics have vividly debated these topics, suggesting a reflection of the premises underlying reporting practices and, more broadly, the purpose of business, which is to create value between and among stakeholders for enhancing sustainable development, namely the Sustainable Development Goals (SDGs), as grand challenges. Drawing on this consideration thus far, this research has the following practical implications for policymakers. First, the research suggests enacting a convergence among standards in favor of more relevant, comparable, and reliable information, both financial and non-financial. Second, the ranking of the four main issues (NFI quality, standardization, materiality, and assurance) that emerged from the annexed documents, together with the questionnaire responses, could be considered by the EU regulator in revising NFRD. Likewise, for what concerns the theoretical perspective, the research implies the need to adopt an integrated approach characterized by the matching of reporting procedures to strategic objectives and governance mechanisms.

Limitations and Avenues for Future Research

This research, like all others, is not without limitations. First, this study does not adopt sentiment analysis, which could complement the understandability of topic meanings and discover the tone of disclosures regarding positive, negative, and neutral word connotations [82,83]. Second, another limitation is related to the topic model methodology. Specifically, since we only used Python to compute the algorithm, our results are deeply linked to Python; however, we could implement the research with other programming languages, such as Matlab. Furthermore, in agreement with other scholars [21,22], this research method is quite new, as is using such a large volume of data. Accordingly, results are subjected to dynamic changes over time, and future algorithms may be refined to facilitate a better understanding of the topics.

Research results could be an important cue for future studies focusing on NFI quality, standardization, materiality, and assurance. Future studies may involve conducting

qualitative research and gathering insights from practitioners working in the field of sustainability reporting to operationalize mechanisms of materiality processes, assurance practices, and NFI standardization to ultimately ensure reliable, relevant, and comparable information. Another line of research could develop literature reviews on such topics using topic model techniques.

To conclude, NFI is gaining ground and promising steps for the NFRD review, along with the International Financial Reporting Standards (IFRS) consultation on sustainability reporting, have started favoring a convergence of both financial and non-financial reporting practices. This needs accounting methods which consider risk assessments and risk management of financial and non-financial issues in joint conjunction with stakeholders' relationships to address stakeholder based-measures and inclusive mechanisms to value stakeholder-focused decision-making.

Author Contributions: Conceptualization, S.F.; methodology, S.F.; validation, S.F. and A.T.; formal analysis, S.F.; investigation, S.F. and A.T.; writing—original draft preparation, S.F.; writing—review and editing, A.T.; supervision, S.F.; project administration, S.F.; funding acquisition, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by SIDREA—Società Italiana dei Docenti di Ragioneria ed Economia Aziendale.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: https://www.mdpi.com/ethics.

Acknowledgments: The corresponding author acknowledges Stefania Fiandrino for her support while using Python for the textual analysis.

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

References

1. Kinderman, D. The challenges of upward regulatory harmonization: The case of sustainability reporting in the European Union. *Regul. Gov.* **2020**, *14*, 674–697. [CrossRef]
2. Howitt, R. The EU law on non-financial reporting-how we got there. *Guardian* **2014**, *16*. Available online: https://www.theguardian.com/sustainable-business/eu-non-financial-reporting-how-richard-howitt (accessed on 13 January 2021).
3. European Parliament. Directive 2014/95/EU of the European Parliament and of the Council—Of 22 October 2014—Amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups. *Off. J. Eur. Union* **2014**, *2014*, 1–9.
4. Fiandrino, S. *Disclosure of Non-Financial Information Disclosure. Evolutionary Paths and Harmonisation to Mandatory Requirements*; G. Giappichelli Editore: Turin, Italy, 2019; ISBN 9788892183414.
5. Schneider, T.; Michelon, G.; Paananen, M. Environmental and Social Matters in Mandatory Corporate Reporting: An Academic Note. *Account. Perspect.* **2018**, *17*, 275–305. [CrossRef]
6. Venturelli, A.; Caputo, F.; Cosma, S.; Leopizzi, R.; Pizzi, S. Directive 2014/95/EU: Are Italian companies already compliant? *Sustainability* **2017**, *9*, 1385. [CrossRef]
7. European Commission. *The European Green Deal*; European Commission: Brussels, Belgium, 2019. [CrossRef]
8. Miles, D.A. A taxonomy of research gaps: Identifying and defining the seven research gaps. *J. Res. Methods Strateg.* **2017**, 1–15. Available online: https://www.academia.edu/35505149/ARTICLE_RESEARCH_A_Taxonomy_of_Research_Gaps_Identifying_and_Defining_the_Seven_Research_Gaps (accessed on 13 January 2021).
9. Matuszak, Ł.; Rózańska, E. CSR disclosure in Polish-listed companies in the light of directive 2014/95/EU requirements: Empirical evidence. *Sustainability* **2017**, *9*, 2304. [CrossRef]
10. Cantino, V.; Devalle, A.; Fiandrino, S.; Busso, D. The level of compliance with the Italian Legislative Decree No. 254/2016 and its determinants: Insights from Italy. *Financ. Rep.* **2019**. [CrossRef]
11. Jonida, C.; Roberto, D.P.; Matteo, M. Mandatory vs voluntary exercise on non-financial reporting: Does a normative/coercive isomorphism facilitate an increase in quality? *Meditari Account. Res.* **2020**. [CrossRef]
12. Pizzi, S.; Venturelli, A.; Caputo, F. The "comply-or-explain" principle in directive 95/2014/EU. A rhetorical analysis of Italian PIEs. *Sustain. Account. Manag. Policy J.* **2020**. [CrossRef]
13. Caputo, F.; Leopizzi, R.; Pizzi, S.; Milone, V. The Non-Financial Reporting Harmonization in Europe: Evolutionary Pathways Related to the Transposition of the Directive 95/2014/EU within the Italian Context. *Sustainability* **2020**, *12*, 92. [CrossRef]

14. Hörisch, J.; Schaltegger, S.; Freeman, R.E. Integrating stakeholder theory and sustainability accounting: A conceptual synthesis. *J. Clean. Prod.* **2020**, *275*. [CrossRef]
15. Mitchell, R.K.; Van Buren, H.J., III; Greenwood, M.; Freeman, R.E. Stakeholder Inclusion and Accounting for Stakeholders. *J. Manag. Stud.* **2015**, *52*, 851–877. [CrossRef]
16. Gray, R.; Adams, C.A.; Owen, D. *Accountability, Social Responsibility and Sustainability. Accounting for Society and the Environment*; Pearson: New York, NY, USA, 2014; ISBN 9780273681380.
17. Venturelli, A.; Pizzi, S.; Caputo, F.; Principale, S. The revision of nonfinancial reporting directive: A critical lens on the comparability principle. *Bus. Strateg. Environ.* **2020**, *29*, 3584–3597. [CrossRef]
18. Liu, Y.; Mai, F.; MacDonald, C. A Big-Data Approach to Understanding the Thematic Landscape of the Field of Business Ethics, 1982–2016. *J. Bus. Ethics* **2019**, *160*, 127–150. [CrossRef]
19. Hannigan, T.R.; Haan, R.F.J.; Vakili, K.; Tchalian, H.; Glaser, V.L.; Wang, M.S.; Kaplan, S.; Jennings, P.D. Topic modeling in management research: Rendering new theory from textual data. *Acad. Manag. Ann.* **2019**, *13*, 586–632. [CrossRef]
20. Ferri, P.; Lusiani, M.; Pareschi, L. Accounting for Accounting History: A topic modeling approach (1996–2015). *Account. Hist.* **2018**, *23*, 173–205. [CrossRef]
21. Reisenbichler, M.; Reutterer, T. Topic modeling in marketing: Recent advances and research opportunities. *J. Bus. Econ.* **2019**, *89*, 327–356. [CrossRef]
22. Amado, A.; Cortez, P.; Rita, P.; Moro, S. Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis. *Eur. Res. Manag. Bus. Econ.* **2018**, *24*, 1–7. [CrossRef]
23. Schmiedel, T.; Müller, O.; vom Brocke, J. Topic Modeling as a Strategy of Inquiry in Organizational Research: A Tutorial With an Application Example on Organizational Culture. *Organ. Res. Methods* **2019**, *22*, 941–968. [CrossRef]
24. Pröllochs, N.; Feuerriegel, S. Business analytics for strategic management: Identifying and assessing corporate challenges via topic modeling. *Inf. Manag.* **2020**, *57*, 103070. [CrossRef]
25. Deegan, C. Twenty five years of social and environmental accounting research within Critical Perspectives of Accounting: Hits, misses and ways forward. *Crit. Perspect. Account.* **2017**, *43*, 65–87. [CrossRef]
26. Adam, C.A.; McNicholas, P. Making a difference: Sustainability reporting, accountability and organizational change. *Account. Audit. Account. J.* **2007**, *20*, 382–402. [CrossRef]
27. Hahn, R.; Kühnen, M. Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *J. Clean. Prod.* **2013**, *59*, 5–21. [CrossRef]
28. Gray, R. Thirty years of social accounting, reporting and auditing: What (if anything) have we learnt? *Bus. Ethics A Eur. Rev.* **2003**, *10*, 9–15. [CrossRef]
29. Hess, D. The three pillars of corporate social reporting as new governance regulation: Disclosure, dialogue, and development. *Bus. Ethics Q.* **2008**, *18*, 447–482. [CrossRef]
30. Haslam, C.; Tsitsianis, N.; Andersson, T.; Gleadle, P. Accounting for Business Models and Increasing the Visibility of Stakeholders. *J. Bus. Model.* **2015**, *3*, 62–80.
31. Zadek, S. Balancing Performance, Ethics, and Accountability. *J. Bus. Ethics* **1998**, *17*, 1421–1442. [CrossRef]
32. Herremans, I.M.; Nazari, J.A.; Mahmoudian, F. Stakeholder Relationships, Engagement, and Sustainability Reporting. *J. Bus. Ethics* **2016**, *138*, 417–435. [CrossRef]
33. Devalle, A.; Gromis di Trana, M.; Fiandrino, S.; Vrontis, D. Integrated thinking rolls! Stakeholder engagement actions translate integrated thinking into practice. *Meditari Accountancy Research* **2020**. [CrossRef]
34. Manetti, G. The quality of stakeholder engagement in sustainability reporting: Empirical evidence and critical points. *Corp. Soc. Responsib. Environ. Manag.* **2011**, *18*, 110–122. [CrossRef]
35. Bellantuono, N.; Pontrandolfo, P.; Scozzi, B. Capturing the Stakeholders' View in Sustainability Reporting: A Novel Approach. *Sustainability* **2016**, *8*, 379. [CrossRef]
36. Polak, P.; Masquelier, F.; Michalski, G. Towards treasury 4.0/the evolving role of corporate treasury management for 2020. *Management* **2018**, *23*, 189–197. [CrossRef]
37. Freeman, R.E.; Harrison, J.S.; Wicks, A.C.; de Colle, S. *Stakeholder Theory. The State of the Art*; Cambridge University Press: Cambridge, UK, 2010; ISBN 9780521190817.
38. Freeman, R.E. *Strategic Management: A Stakeholder Approach*; Pitman Publishing Inc.: San Francisco, CA, USA, 1984; Volume 1, ISBN 0-273-01913-9.
39. Dallas, L.L. Short-Termism, the Financial Crisis, and Corporate Governance. *J. Corp. Law* **2011**, *37*, 264.
40. Jensen, M.C. Value maximization, stakeholder theory, and the corporate objective function. *J. Appl. Corp. Financ.* **2001**, *14*, 8–21. [CrossRef]
41. Stout, L.A. *The Shareholder Value Myth: How Putting Shareholders First Harms Investors, Corporations, and the Public*; Berrett-Koehler Publishers: San Francisco, CA, USA, 2012; ISBN 1605098167.
42. Gray, R. Social, environmental and sustainability reporting and organisationa value creation? Whose value? Whose creation? *Account. Audit. Account. J.* **2006**. [CrossRef]
43. Higgins, C.; Stubbs, W.; Milne, M. Is Sustainability Reporting Becoming Institutionalised? The Role of an Issues-Based Field. *J. Bus. Ethics* **2015**, 309–326. [CrossRef]

44. Adams, C.A.; Whelan, G. Conceptualising future change in corporate sustainability reporting. *Account. Audit. Account. J.* **2009**, *22*, 118–143. [CrossRef]
45. Silva, S.; Nuzum, A.K.; Schaltegger, S. Stakeholder expectations on sustainability performance measurement and assessment. A systematic literature review. *J. Clean. Prod.* **2019**, *217*, 204–215. [CrossRef]
46. De Villiers, C.; Alexander, D. The institutionalisation of corporate social responsibility reporting. *Br. Account. Rev.* **2014**, *46*, 198–212. [CrossRef]
47. Perego, P.; Kennedy, S.; Whiteman, G. A lot of icing but little cake? Taking integrated reporting forward. *J. Clean. Prod.* **2016**, *136*, 53–64. [CrossRef]
48. Buhr, N.; Gray, R.; Milne, M. Histories, rationales, voluntary standards and future prospects for sustainability reporting: CSR, GRI, IIRC and beyond. In *Sustainability Accounting and Accountability*; Routledge: Abingdon, UK, 2014; pp. 69–89.
49. Stubbs, W.; Higgins, C. Stakeholders' Perspectives on the Role of Regulatory Reform in Integrated Reporting. *J. Bus. Ethics* **2018**, *147*, 489–508. [CrossRef]
50. Rezaee, Z.; Tuo, L. Voluntary disclosure of non-financial information and its association with sustainability performance. *Adv. Account.* **2017**, *39*, 47–59. [CrossRef]
51. Qiu, Y.; Shaukat, A.; Tharyan, R. Environmental and social disclosures: Link with corporate financial performance. *Br. Account. Rev.* **2016**, *48*, 102–116. [CrossRef]
52. Camilleri, M.A. Environmental, social and governance disclosures in Europe. *Sustain. Account. Manag. Policy J.* **2015**, *6*, 224–242. [CrossRef]
53. Manes-Rossi, F.; Tiron-Tudor, A.; Nicolò, G.; Zanellato, G. Ensuring more sustainable reporting in europe using non-financial disclosure-de facto and de jure evidence. *Sustainability* **2018**, *10*, 1162. [CrossRef]
54. European Commission. *Guidelines on Non-Financial Reporting (Methodology for Reporting Non-Financial Information)*; EC: Brussels, Belgium, 2017.
55. Jeffery, C. Comparing the Implementation of the EU Non-Financial Reporting Directive. 2017. Available online: tps://ssrn.com/abstract=3083368 (accessed on 14 December 2020).
56. Doni, F.; Martini, S.B.; Corvino, A.; Mazzoni, M. Voluntary versus mandatory non-financial disclosure. *Meditari Account. Res.* **2019**, *28*, 781–802. [CrossRef]
57. Ioannou, I.; Serafeim, G. The Consequences of Mandatory Corporate Sustainability Reporting. 2014. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1799589 (accessed on 13 January 2021).
58. La Torre, M.; Sabelfeld, S.; Blomkvist, M.; Tarquinio, L.; Dumay, J. Harmonising non-financial reporting regulation in Europe: Practical forces and projections for future research. *Meditari Account. Res.* **2018**, *26*, 598–621. [CrossRef]
59. Milne, M.J.; Gray, R. W(h)ither Ecology? The Triple Bottom Line, the Global Reporting Initiative, and Corporate Sustainability Reporting. *J. Bus. Ethics* **2013**, *118*, 13–29. [CrossRef]
60. Taïbi, S.; Antheaume, N.; Gibassier, D. Accounting for strong sustainability: An intervention-research based approach. *Sustain. Account. Manag. Policy J.* **2020**. [CrossRef]
61. Fuller, M. Letters from the top: A comparative control group study of CEO letters to stakeholders. *Int. J. Corp. Soc. Responsib.* **2018**, *3*, 1–14. [CrossRef]
62. Greenwood, M.; Kamoche, K. Social accounting as stakeholder knowledge appropriation. *J. Manag. Gov.* **2013**, *17*, 723–743. [CrossRef]
63. Harrison, J.S.; van der Laan Smith, J. Responsible Accounting for Stakeholders. *J. Manag. Stud.* **2015**, *52*, 935–960. [CrossRef]
64. Freeman, R.E. Five Challenges to Stakeholder Theory: A Report on Research in Progress. In *Stakeholder Management Business and Society 360*; Emerald Publishing Limited: Bentley, UK, 2017; Volume 1, pp. 1–20. ISBN 2514-1759.
65. Puroila, J.; Mäkelä, H. Matter of opinion: Exploring the socio-political nature of materiality disclosures in sustainability reporting. *Account. Audit. Account. J.* **2019**. [CrossRef]
66. European Commission. *Inception Impact Assessment*; EC: Brussels, Belgium, 2020.
67. European Commission. *Summary Report of the Public Consultation on the Review of the Non-Financial Reporting Directive*; EC: Brussels, Belgium, 2020.
68. Valdez, D.; Pickett, A.C.; Goodson, P. Topic Modeling: Latent Semantic Analysis for the Social Sciences. *Soc. Sci. Q.* **2018**, *99*, 1665–1679. [CrossRef]
69. Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [CrossRef]
70. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]
71. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
72. Huang, A.H.; Lehavy, R.; Zang, A.Y.; Zheng, R. Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Manag. Sci.* **2018**, *64*, 2833–2855. [CrossRef]
73. Turney, P.D.; Pantel, P. From Frequency to Meaning: Vector Space Models of Semantics. *J. Artifical Intell. Res.* **2010**, *37*, 141–188. [CrossRef]
74. DiMaggio, P.; Nag, M.; Blei, D. Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics* **2013**, *41*, 570–606. [CrossRef]
75. Debortoli, S.; Müller, O.; Junglas, I.; vom Brocke, J. Text mining for information systems researchers: An annotated topic modeling tutorial. *Commun. Assoc. Inf. Syst.* **2016**, *39*, 110–135. [CrossRef]

76. Sievert, C.; Shirley, K. LDAvis: A method for visualizing and interpreting topics. In Proceedings of the Workshop Oninteractive Language Learning, Visualization, and Interfaces, Baltimore, MD, USA, 27 June 2014; pp. 63–70.
77. Röder, M.; Both, A.; Hinneburg, A. Exploring the space of topic coherence measures. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 399–408. [CrossRef]
78. Goloshchapova, I.; Poon, S.-H.; Pritchard, M.; Reed, P. Corporate social responsibility reports: Topic analysis and big data approach. *Eur. J. Financ.* **2019**, *25*, 1637–1654. [CrossRef]
79. Helfaya, A.; Whittington, M. Does designing environmental sustainability disclosure quality measures make a difference? *Bus. Strateg. Environ.* **2019**, *28*, 525–541. [CrossRef]
80. Melloni, G.; Caglio, A.; Perego, P. Saying more with less? Disclosure conciseness, completeness and balance in Integrated Reports. *J. Account. Public Policy* **2017**, *36*, 220–238. [CrossRef]
81. van Steenis, H. Defective data is a big problem for sustainable investing. *Financial Times*, 21 January 2019; 7–9.
82. Loughran, T.I.M.; Mcdonald, B. When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks Journal of Finance, forthcoming. *J. Financ.* **2011**, *66*, 35–65. [CrossRef]
83. Melloni, G.; Stacchezzini, R.; Lai, A. The tone of business model disclosure: An impression management analysis of the integrated reports. *J. Manag. Gov.* **2016**, *20*, 295–320. [CrossRef]