CLIMATE CHANGE AND CLIMATE-RELATED FINANCIAL DISCLOSURES IN THE BANKING SECTOR

Dario Aversa *

* Department of Social Sciences and Economics, Sapienza University of Rome, Rome, Italy
Contact details: Department of Social Sciences and Economics, Sapienza University of Rome, 113, Via Salaria, 00198 Rome, RM, Italy

OPEN ACCESS

How to cite this paper: Aversa, D. (2023). Climate change and climate-related financial disclosures in the banking sector. *Risk Governance and Control: Financial Markets & Institutions, 13*(1), 70–94. https://doi.org/10.22495/rgcv13i1p6

ISSN Online: 2077-4303
ISSN Print: 2077-429X

Received: 12.12.2022
Accepted: 28.03.2023

JEL Classification: G10, G20, G30, K32
DOI: 10.22495/rgcv13i1p6

Abstract

The aim of the paper is to analyze sustainability report disclosures (Task Force on Climate-related Financial Disclosures [TCFD], 2017a, 2017b, 2019, 2020a, 2020b, 2021, 2022; AlHares & Al-Hares, 2020; Lagasio, 2019; Lucchese, 2020; International Sustainability Standards Board [ISSB], 2022) of the listed banks on FTSE Italia All-Share index of Borsa Italiana through text analytics (Giuliano, 2004). The research questions tend to verify: how and whether physical risk (acute and chronic) is reported; how and whether transition risk (legal, technology, market, and reputational) is reported; how and whether scenario analysis (The Bank of England, 2022; Rogelj et al., 2018) is conducted. Using Iramuteq (www.iramuteq.org) and SAS Viya (www.sas.com), the research combines unsupervised learning (Reinert, 1990) and supervised techniques (SAS, 2019) pointing out the inadequacy, the lack of transparency, and the lack of comparability of the sustainability reports that may increase the potential for uncertainty and financial instability. Disclosing climate information on a mandatory basis allows an increase in the quantity and quality of climate-related reporting, an increase in transparency, and comparability accountability, and provides clearer disclosures to investors and regulators.

Keywords: Climate Change, Disclosure, Transition Risks, Physical Risks, Scenario Analysis, Text Mining, Text Analytics, Unsupervised Learning, Supervised Learning

Authors' individual contribution: The Author is responsible for all the contributions to the paper according to CRediT (Contributor Roles Taxonomy) standards.

Declaration of conflicting interests: The Author declares that there is no conflict of interest.

Acknowledgements: The Author wants to express his sincere gratitude to Fiorenza Deriu for her helpful tips on how to utilize the software and for her text-mining suggestions, which raised the quality of the article. The Author sincerely wishes to thank Mariantonietta Fiore and Nino Adamashvili for their constant confidence in him and assistance.

1. INTRODUCTION

Climate change[1] (Hwang et al., 2021) is a matter of science, a systemic risk (Ramani, 2020; Stern et al., 2022) that produces uncertainty (Minenna & Aversa, 2019; U.S. Governing Publishing Office, 2021), affects financial stability (Battiston et al., 2021) and

[1] Climate change refers to a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. Climate change may be due to natural internal processes or external forcings such as modulations of the solar cycles, volcanic eruptions and persistent anthropogenic changes in the composition of the atmosphere or in land use. Note that the Framework Convention on Climate Change (UNFCCC), in its Article 1, defines climate change as: "a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods". The UNFCCC thus makes a distinction between climate change attributable to human activities altering the atmospheric composition and climate variability attributable to natural causes.

manifests itself in financial impacts on banks and corporations balance sheet (TCFD, 2017b, 2019, 2020[2], Eccles & Krzus, 2018; Aversa & Cincinelli, 2019). The investigation of the disclosures[3] (Moreno & Caminero, 2022) in sustainability reports (Directive 2014/95/EU on non-financial information (NFI)) of firms constitute a way to face the problem (El-Hage, 2021) and the dire consequences of this *anthropogenic*, *multifaceted*, *non-linear*, *complex* and *interconnected* urgency of our time (Aversa et al., 2022). The analysis of disclosures and the provision of future affordable scenario leads to a better valuation of financial impacts on corporations, entail better management of the implications on the business, and allow to focus on the most relevant effects for investors, stakeholders, and regulators (TCFD, 2020b, 2022; Feridun & Güngör, 2020). TCFD-aligned disclosures are a crucial part of managing these impacts, they are a means of inclusive capitalism and a way of communication for the stakeholder audience. The aim of the paper is to analyze sustainability report (EU non-financial reporting) disclosures of the listed bank on the FTSE Italia All-Share index of Borsa Italiana through multivariate analysis of text to assess how and whether they disclose transition and physical risks, and how and whether they carry out scenario analysis to tackle the urgency of climate change. With regard to methodology, the research combines *unsupervised learning* using Iramuteq (www.iramuteq.org) (Souza et al., 2018) and the technique of *information extraction* with SAS Viya (www.sas.com); the first part of the analysis includes *lexicometric measures*, specificity, and *correspondence analysis*, *similarities* tools, and *visual graphic representations*, *clustering* with the Reinert's method; the second part uses concepts and categories nodes in SAS Viya for Text Analytics pipeline adopting language interpretation and text interpretation (LITI) coding language. For the *conceptualisation* and *operationalisation* of the concepts, the definitions of the TCFD (2017b) and The Bank of England (2022) are used, dividing transition risk into policy and legal, technology, market, reputation, physical risks into acute and chronic, and applying the scenario definition. The TCFD was launched at the Paris COP21 in 2015 by the Financial Stability Board (FSB), and it is commonly recognized as the *benchmark* for financial and non-financial corporations; the TCFD in June 2017 also sets out a framework of four core elements (sometimes referred to as "pillars"), supported by eleven recommended disclosures, therefore the paper, as a corollary, verifies the alignment to TCFD of the bank reports for the risk management and metric sections. The research investigates the climate change disclosures included in the sustainability reports of listed banks on the FTSE All-share Index of Borsa Italiana, using text analytics (Moreno & Caminero, 2022) and combining *unsupervised learning* with *supervised* techniques. Text analytics helps our analysis to extract meanings, patterns, and structures hidden in textual data by adopting lexicometric measures, analysis of specificity and lexical correspondences, similarity, and cluster analysis. The literature shows that banks' sustainability reports (GARP, 2022) are characterised by a wealth of information in terms of social and environmental disclosure but they are opaque on climate change risks and do not prioritise climate information, the paper attempts to fill this gap by combining various analysis techniques and proposes solutions to increase market transparency and firms accountability.

The general structure of the article is as follows: an introduction is presented in Section 1. Section 2 provides the literature review and Section 3 describes the research methodology. Section 4, the results, is divided into data and lexicometric measures, corpus analysis, specificities, and lexical correspondence analysis, similarities analysis, cluster analysis, and analysis with SAS Viya[4] Text Mining. Section 5 provides a conclusion adding some consideration about limitations and future research directions.

2. LITERATURE REVIEW

Climate change is a matter of urgency, a systemic risk (Ramani, 2020) in a world of high uncertainty (Minenna & Aversa, 2019) an interconnected (Pörtner et al., 2022) and global issue (Stranadko, 2022), a planetary process in which its elements interact in "non-linear and complex"[5] way requiring a system thinking approach (Johnson et al., 2019; Arnold & Wade, 2015) to address the solutions. As a matter of emergency, a problem of immense risk (Stern et al., 2022), uncertainty (Chenet et al., 2021; Eory et al., 2018), financial risk and stability (Ebner, 2018), as well as the not-linear[6] and not incremental global process, entail the evaluation of risks and opportunities (Rising et al., 2022) calling for solutions through innovation and technology and better information. Climate change is anthropogenic, it is a change in the state of the climate and could be better expressed as "a change in the mean and (or) variability of the properties of this state" measured in a multidecadal period (GARP, 2022). The transition from a traditional to a low-emission economy brings with it risks (TCFD, 2019, 2020, 2020b, 2021) that can have a financial or reputational impact on banks and corporations. To assess and mitigate climate change the research refers to the TCFD framework. The TCFD is "an industry-led group which helps investors understand their financial exposure to climate risk and works with companies to disclose this information in a clear and consistent way". The TCFD was launched at the Paris COP21 in 2015 by the Financial Stability Board (FSB) and it is commonly recognized as the benchmark for financial and non-financial corporations.

The TCFD divided climate-related risks into two major categories:

- risks related to the *transition* to a lower-carbon economy (transition risks);
- risks related to the *physical* impacts of climate change (physical risks).

[2] The complete list of TCFD publications is available at https://www.tcfdhub.org and https://www.fsb-tcfd.org/publications/.
[3] An innovative approach to mitigating climate change beyond the international negotiations and hard-law approaches is governing by disclosure — the acquisition and dissemination of information to influence the behavior of particular actors.
[4] This part uses *concepts* and *categories nodes* in SAS Viya for Text Analytics pipeline adopting LITI coding language.
[5] This means that "an event of one element can involve numerous others".
[6] The key distinguish feature of the non-linear relationship together with globalization determines the greater interdependence of different systems (Arnold & Wade, 2015).

Transition risk is divided into four categories: policy and legal, technological, market, and reputation while the physical risk is split into acute and chronic. The political and legal risks (the latter is also named litigation risk) are characterized by measures aimed at limiting the negative effects or at adapting to climate changes (e.g., carbon tax, energy, and water efficiency, rationale use of the soil). The technological risk connected to renewable energies and efficient energy is conditioned by the timing of the development and diffusion of technology. Market risk includes the changes in demand and supply of products and services while reputation risk consists of the change in the perception by customers or the community. The acute physical risks (acute events) are extreme meteorological events: cyclones, hurricanes, or floods, and the chronic physical risks that are a change of long term in the climatic model ("long-term shifts in climate change (e.g., sustained higher temperatures)"), these types of risks may have financial implications, direct and indirect damage in the distribution and supply chain.

The assessment of climate and chronic risks occurs by employing scenario analysis which describes its content. Therefore, addressing climate change requires the use of scenario analysis[7] (TCFD, 2017b, 2021[8]) that is a planning tool, a narrative on the "potential future state of the world" and a conditional prediction that appeared academically in 1950 and requires some decisions about the setting of *parameters* and (or) *assumptions*, the choice of *analytical tools* and *outputs*. The use is for strategy and stakeholder communication to investors and regulators, it responds to both related physical impacts of climate change (physical risks) and the zero-carbon economic transition (transition risks); the two main types of risk that describe climate change.

Scenario analysis varies quite significantly between both risks and also for this reason we use "global reference scenarios"[9] that is, "agreed and widely used projections of future emissions" sometimes with socio-economic narratives attached. The most universal scenario is from the Intergovernmental Panel on Climate Change (IPCC)[10] and includes "representative concentration pathways"[11] also named "RCPs" (van Vuuren, 2011), and "shared socioeconomic platforms" (SSPs) incorporating socioeconomic projections up to 2100 (Riahi et al., 2017), "they are used to derive GHG emissions scenarios with different climate policies" (Rogelj et al., 2018). Other important providers of reference scenarios are the International Energy Agency (IEA), Greenpeace, IRENA, and the Network for Greening Financial System (NGFS).

Transition risk scenario analysis can make use of integrated assessment models (IAMs), "economic models that also include representations of societal and environmental phenomena and sector-specific decarbonization pathways" (GARP, 2022), physical risk scenario analysis instead uses physical climate models, but it also benefits from resilience planning. Additionally, the TCFD recommended scenario analysis as a way to "*enhance critical strategic thinking*" for companies using plausible, distinctive, consistent, relevant, and challenging scenarios. From a "nice to have" instrument to a best practice, scenario analysis[12] is increasingly being implemented by, and even mandated by, regulators. The latter adopted a stress test[13] in the wake of the financial crisis of 2008 and they model "the reaction of both a financial system as a whole and an individual institutions' balance sheet to a hypothetical shock" relying on scenario analysis (Ebner, 2018). The Final Report of Recommendations was published by the TCFD in June 2017 and sets out a framework of four core elements (sometimes referred to as "pillars"), supported by eleven recommended disclosures[14] for corporations (Figure 1).

The disclosure information helps assess the extent of the impact that climate change will have on them (Attenborough, 2022) and "how and when it should be reflected in the financial statements as well as the narrative disclosures [...] the physical risks (including direct damage to assets and operational disruption) and transition risks (such as policy, legal, technology, and market changes); short, medium, and long-term implications; and the implications for others in the same supply chain".

There are growing calls from investors, regulators, and other stakeholder groups for better information on how these issues (climate change and scenario analysis) have been "communicated" in the sustainability reports[15], but they should also consider whether further disclosures are necessary to enhance transparency in order to understand the financial impact of climate change on the balance sheet of companies and to do that the landmark is both TCFD and Net Zero Strategy.

[7] This part of the research is based on the paper "Scenario Analysis and Climate Change: A Literature Review via Text Analytics" which was submitted to the British Food Journal (revised and resubmitted — second round).
[8] https://www.tcfdhub.org/
[9] Global references scenario is a set of standard and cross-comparable scenarios.
[10] The Intergovernmental Panel on Climate Change (IPCC) is the United Nations body for assessing the science related to climate change (https://www.ipcc.ch/).
[11] RCPs is a "grenhouse gas concentration (not emissions) trajectory adopted by the IPCC" (van Vuuren, 2011).

[12] Starting April 6, 2022, the UK will become the first G20 country that enforced in law mandatory requirements to "largest business and financial institution" to disclose climate-related risk and opportunities aligning with TCFD.
[13] Various central banks, including the European Central Bank (ECB), do stress tests with varying degrees of granularity and distinct climate and political scenarios.
[14] The TCFD also releases status reports that highlight the latest reporting trends in relation to the four core elements and eleven recommended disclosures and include helpful published examples of reporting. The full suite of TCFD documents can be found at https://www.fsb-tcfd.org/publications/.
[15] The term sustainability report is interchangeable with the term "integrated reporting" used in the European Directive as the European Commission has never clarified the differences due to the adoption of this term.

Figure 1. TCFD pillars and recommendation

Governance	Strategy	Risk Management	Metrics and Targets
Disclose the organization's governance around climate-related risks and opportunities.	Disclose the actual and potential impacts of climate-related risks and opportunities on the organization's businesses, strategy, and financial planning where such information is material.	Disclose how the organization identifies, assesses, and manages climate-related risks.	Disclose the metrics and targets used to assess and manage relevant climate-related risks and opportunities where such information is material.
Recommended Disclosures	**Recommended Disclosures**	**Recommended Disclosures**	**Recommended Disclosures**
a) Describe the board's oversight of climate-related risks and opportunities.	a) Describe the climate-related risks and opportunities the organization has identified over the short, medium, and long term.	a) Describe the organization's processes for identifying and assessing climate-related risks.	a) Disclose the metrics used by the organization to assess climate-related risks and opportunities in line with its strategy and risk management process.
b) Describe management's role in assessing and managing climate-related risks and opportunities.	b) Describe the impact of climate-related risks and opportunities on the organization's businesses, strategy, and financial planning.	b) Describe the organization's processes for managing climate-related risks.	b) Disclose Scope 1, Scope 2, and, if appropriate, Scope 3 greenhouse gas (GHG) emissions, and the related risks.
	c) Describe the resilience of the organization's strategy, taking into consideration different climate-related scenarios, including a 2°C or lower scenario.	c) Describe how processes for identifying, assessing, and managing climate-related risks are integrated into the organization's overall risk management.	c) Describe the targets used by the organization to manage climate-related risks and opportunities and performance against targets.

Source: TCFD (2017a, p. 14).

3. RESEARCH METHODOLOGY

The aim of the research is to analyze the disclosures information within the sustainability reports (Directive 2014/95/EU) of the listed banks on the FTSE Italia All-Share index of Borsa Italiana with the use of quantitative tools for qualitative analysis (data mining for textual data) (Moreno & Caminero, 2022) approaching the topic with mixed research method and design.

The research consists of three research questions that tend to verify the environmental disclosures of the bank's reporting as follows:

RQ1: How and whether physical risk (acute and chronic) is reported.

RQ2: How and whether transition risk (legal, technology market, and reputational) is reported.

RQ3: How and whether scenario analysis is conducted.

Iramuteq (*Interface de R pour les Analyzes Multidimensionnelles de Textes et de Questionnaires*) was developed by the French programmer Pierre Ratinaud (Aversa et al., 2022; Souza et al., 2018). It is an open-source, multifunctional, and statistical software for the analysis of textual data.

Its algorithms are based on Python programming language and statistical software R (Camargo & Justo, 2016), with statistical tools and through a recursive approach, IRaMuTeQ 0.7 performs text analysis such as similarity analysis, specificity analysis and lexical correspondence analysis (LCA) for text data which are specified step by step in the analysis sections of the paper. Iramuteq was chosen for the textual analysis because, compared to other software, it shows a high processing capacity and allows us to arrive at detailed reports with quality and highly reliable results. In addition, unlike other software of the same quality, it is open source while maintaining the same standards.

According to the definition of the Italian banking act (TUB), banks are companies that carry out the exercise of credit and the collection of savings from the public (Legislative Decree of September 1, 1993, no. 385 in force in Italy since January 1, 1994).

Table 1. The FTSE Italia All-Share banks

Bank name	*Index*	*Coding*	*Year*
UniCredit Bank	FTSE MIB	****001 *bank_unicredit *index_ftsemib *year_2020	2020
Mediobanca	FTSE MIB	****002 *bank_mediobanca *index_ftsemib *year_2020	2020
Intesa Sanpaolo Bank	FTSE MIB	****003 *bank_intesasanpaolo *index_ftsemib *year_2020	2020
BPER Banca	FTSE MIB	****004 *bank_bper *index_ftsemib *year_2020	2020
Banca BPM	FTSE MIB	****005 *bank_bpm *index_ftsemib *year_2020	2020
Banca Mediolanum	FTSE MIB	****006 *bank_mediolanum *index_ftsemib *year_2020	2020
Banca Generali	FTSE MIB	****007 *bank_generali *index_ftsemib *year_2020	2020
Banca Monte dei Paschi di Siena	FTSE Italia Mid Cap	****008 *bank_mps *index_ftseitaliamidcap *year_2020	2020
Banca Popolare di Sondrio	FTSE Italia Mid Cap	****009 *bank_popolaresondrio *index_ftseitaliamidcap *year_2020	2020
Banca Credito Emiliano	FTSE Italia Mid Cap	****010 *bank_creditoemiliano *index_ftseitaliamidcap *year_2020	2020
Banca Credito Valtellinese	FTSE Italia Mid Cap	****011 *bank_creditovaltellinese *index_ftseitaliamidcap *year_2020	2020
Banca di Desio e Brianza	FTSE Italia Small Cap	****012 *bank_desiobrianza *index_ftseitaliasmallcap *year_2020	2020

The corpus (see Table 1) consists of the sustainability reports of 12 banks (partition variables) belonging to the FTSE Italia All-Share index which consists of the aggregation of all the elements of the FTSE MIB, FTSE Italia Mid Cap, and FTSE Italia Small Cap indices. FTSE MIB is the most significant share index of the Italian Stock Exchange, it includes the shares of the 40 companies with the greatest capitalization, free float, and liquidity that represent over 80% of the total capitalization and almost 90% of the turnover, FTSE Italia mid-cap is the stock market index that includes the top 60 companies by capitalization that does not belong to the FTSE MIB index. FTSE Italia Small Cap is the basket of low-cap equities representing 4% of Borsa Italiana's capitalization, 1% of total daily turnover, and 6% of total contracts of an average session.

The contents were saved in text format (UTF-8 coding) for automatic analysis, each text was accompanied by the text partition key variables indicated below:

a) identifier: **** number;
b) bank: * bank_bank name;
c) index: * index_index name;
d) year: * year_number.

Example of a trace: **** 001 * bank_unicredit * index_ftsemib * year_2020.

The techniques used are text analytics (*data mining for textual data*) and combine *unsupervised learning* with supervised techniques (Moreno & Caminer, 2022; Kalamara et al., 2020) it was not possible to implement the analysis with machine learning tools due to the dimensions of the corpus. Text analytics helps our analysis to extract meanings, patterns, and structure hidden (Anderson et al., 2021) in unstructured textual data using advanced software like Iramuteq, and SAS Viya for Text Analytics, and R-Studio for the vocabulary growth curve.

By the use of Iramuteq, the following elements will be implemented:

- Descriptive analysis: type-token ratio (TTR), TTR made using lemmas (L-TTR), Guiraud index, and Herdan index, Zipf's curve and its slope to verify lexical richness and justify the automatic and semi-automatic treatment of the data);
- Specificities analysis and correspondence analysis;
- Similarities analysis;
- Cluster analysis with Reinert's method.

The techniques use an *unsupervised learning* process (data-driven). The use of similarities analysis will be dealt with from a "frequentistic"[16] point of view (number of words and/or boolean presence or absence of the word in an array) to a relational approach (the corpus is represented in a contingency table and seen like a network of words). To complete the analysis we will use SAS Viya, especially for information extraction and information retrieval through programming in LITI language on categories and concept nodes[17].

4. RESULTS

4.1. Data and lexicometric measures

The treatment of the qualitative data of the collection of documents is carried out with multivariate techniques, it begins with the lexicographic analysis (*lexicometric measures*) of the *corpus* (homogeneous, monolingual Italian and referred to the period 2020–2019), subsequently the analysis of specificity and lexical correspondences, similarity and cluster analysis by means of the Reinert's method as previously written.

The construction of the validation of the corpus took place with an iterative process through three phases:

a) application of multivariate techniques on the original corpus (cluster analysis);
b) lexicalization (to identify lexias);
c) application of multivariate techniques after lexicalization.

In phase (a), the lexicographic analysis carried out with Iramuteq, before the lemmatization and after the lexicalization, identifies 12 texts in the corpus, 42,045 tokens, 5,811 types and a number of hapaxes equal to 2708 in absolute terms, corresponding to 6.44% of occurrences and 46.60% of forms (see Table 2). "In the practice of content analysis, a corpus consisting of 15,000 occurrences is considered to be small in size, between 50,000 and 100,000 medium-sized, while above 200,000 it is large" (Giuliano, 2004).

The details of the descriptive aspects of the dataset are summarized in the following abstract.

Table 2. Corpus abstract after/before lexicalization

Parameters	*Values*
Corpus abstract after lexicalization	
Number of texts	12
Number of text segments	1,166
Number of occurrences	42,045
Number of forms	5,811
Number of hapaxes	2,708 (46.60% of forms – 6.44% of occurrences)
Corpus abstract before lexicalization	
Number of texts	12
Number of occurrences	43,012
Number of forms	5,775
Number of hapaxes	2,702 (46.79% of forms – 6.28% of occurrences)

Source: Author's elaboration using Iramuteq.

In phase (b) the lexicalization generated 40 lexias between compound and complex, i.e., sets of two or more words that take on a different meaning (see Table 2).

The measures that, together with the previous ones, determine the descriptive aspects of the corpus, are represented/illustrated in Table 3:

- TTR;
- Guiraud index;
- Herdan index;
- Token ratio after lemmatization type;
- Zipf's curve;
- Zipf's slope curve;
- Curve of vocabulary growth;
- Abstract after lemmatization and L-TTR.

[16] Our adaptation of the term.
[17] The size of the corpus did not allow the application of machine learning techniques and made the topic not very robust detection and sentiment analysis unreliable.

Table 3. Lexicalization

Lexicalized word	*No.*	*Lexicalized word*	*No.*	*Lexicalized word*	*No.*	*Lexicalized word*	*No.*
1. impatti ambientali	24	1. consumo di carta	17	1. energia elettrica	114	1. intesa sanpaolo	57
2. iso 14001	17	2. carta riciclata	15	2. fonti rinnovabili	43	2. circular economy	36
3. efficientamento energeticc	16	3. banca generali	11	3. gas naturale	32	3. green bond	16
4. efficienza energetica	15	4. materiali utilizzati	11	4. consumi energetici	30	4. economia circolare	14
5. sistema di gestione	14	5. totale carta	11	5. emissioni indirette	30	5. climate change	9
6. covid-19	12	6. raccolta differenziata	9	6. consumo di energia	25	6. miliardi di euro	9
7. gestione ambientale	12	7. carta e cartone	8	7. t co2e	22	7. transizione verso un economia	9
8. banco desio	10	8. carta e toner	8	8. energia elettrica acquistat	21	8. green deal	8
9. sistema di gestione ambiental	10	9. metodo di smaltiment	8	9. market based	20	9. circular economy	8
10. consumi energetici	9	10. per mq	8	10. emissioni derette	19	10. ambiente e climate	7

Table 4. Lexicometric measures

Lexicometric measures	*Formula*	*Value*
TTR	$(V/N)*100$	$(5811/42045)*100 = 13.82\%$
% Hapax	$(V_1/V)*100$	$(2708/5811)*100 = 46.60\%$
Zipf's slope curve	$(\log N/\log V)$	$(\log 42045/\log 5811) = 4.62/3.76 = \|1.22\|$
Guiraud index	$V/\sqrt{N}$	$5811/\sqrt{42045} = 5811/205.04 = 28.34$
Herdan index	$(\log V/\log N)$	$\log 5811/\log 42045 = 3.76/4.62 = 0.81$

Type-token ratio (TTR): The first quantitative measure, the TTR is 13.82 expressed in percentage terms, it is given by the ratio between the number of different words (*type*), equal to 5,811, and the total number of words (*occurrences or token*) of 42,045 lexical units; this ratio generates a value of less than 20%, a figure of a valid lexicographic measure that allows us to consider the adequate corpus for an automatic or semi-automatic treatment.

The TTR is therefore a relationship between the width of the vocabulary (*V*) and the size of the corpus (*N*). It is an index that is sensitive to the size of the corpus, its limitation lies in the fact that as the number of occurrences increases its value tends to decrease and therefore fall below the 20% threshold since the graphic forms tend to repeat themselves. As mentioned, if its value is less than 20%, the corpus is considered adequate for a lexicometric-type treatment. Later in the analysis, the lemmas-based version *(L-TTR) = Lemma/N* was also used.

The absolute number of hapax, 2,708 units, corresponds to 46.60% of forms and is below the limit threshold of 50%; this measure confirms the correct size and breadth of the corpus as a whole.

Guiraud index: To continue validating the corpus, the Guiraud index is used, given by the ratio between the number of forms (*V*) and the square root of the occurrences (*N*). It indicates a certain lexical richness as it is higher than the minimum limit of 22 in absolute terms (see Table 4).

Herdan index: The Herdan index, expressed by the ratio between the logarithm of the types and the logarithm of the occurrences (*logV/logN*), is confirmatory for the considerations relating to lexical richness (see Table 4).

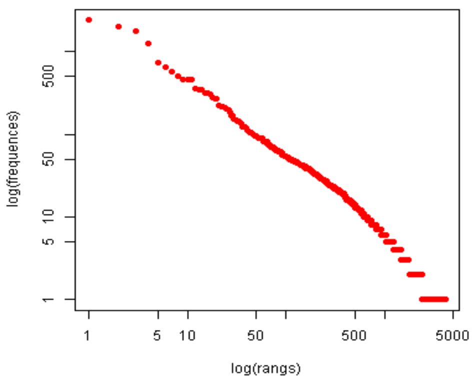
Abstract (after lemmatization) and L-TTR:

- Number of texts: 12;
- Number of occurrences: 42,045;
- Number of forms (lemmas): 4,192;
- Number of hapaxes: 1,714 (4.08% of occurrences – 40.89% of forms);
- The mean of occurrences by text: 3503.75.

The L-TTR index is 9.97% and obviously consists of an improvement in the TTR to the extent that it is determined by a reduction in the absolute value of the numerator which results in a lower output.

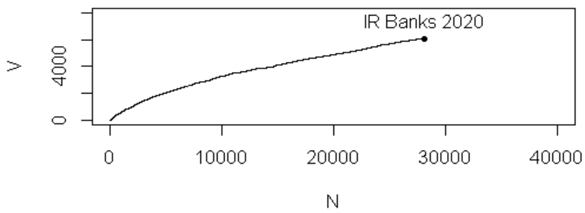
Zipf's curve and Zipf's slope curve: The words of a vocabulary always have a well-defined distribution in terms of occurrences; its form is known and was identified by the linguist George Kingsley Zipf[18] (Zipf's law) whereby frequency is inversely proportional to rank.

Figure 2. Zipf's curve



Source: Author's elaboration using Iramuteq.

Figure 3. Vocabulary growth curve



Source: Author's elaboration using R.

The Zipf's curve, expressed by the logarithm of the rank on the abscissa by the logarithm of the frequency on the ordinate, takes the form in Figure 2 above; it shows that the frequency is

[18] Zipf's law provides a probability distribution for the frequency of words in text. It is like a discrete version of the Pareto distribution. A feature of Zipf's law is that a plot of the frequency of words versus the rank of the word on a log scale will be approximately linear. Perl can be used to tabulate the frequencies of words in a document or database to see if they follow Zipf's law.

inversely proportional to the rank: the higher the rank, the lower the frequency. This is a rule to be understood in a statistical sense on average because it is necessary to take an average value of occurrence of the words belonging to a certain neighborhood of the considered rank (there are not all the possible frequencies and then there are the *ex aequo*).

Zipf's slope curve: The coefficient of the curve is the slope of a line on a graph with logarithmic coordinates — the abscissa (x-axis) indicates the logarithm of the rank and the ordinate (y-axis) the log of the frequency and represents the lexical richness of a vocabulary: the proportion of different words. The slope, on the logarithmic coordinate graph, is well approximated and its value must be around 1.3 in absolute terms. In this case, the slope has a value of |1.22| and represents a good lexical richness of the corpus, while by lexical richness we mean the proportion between different words. The vocabulary growth curve, created with R-Studio, shows an ever-increasing trend: as the number of occurrences increases, the graphic forms grow.

The summary structure of the statistics shows that the text as a whole can be subjected to automatic or semi-automatic processing and that the quantitative statistical characteristics of the corpus validation are robust, documenting the possibility of lexicometric processing of the text data.

4.2. Corpus analysis

The study of the reference corpus, of the distribution of words in the text, as they are processed by the software, takes place by opting for three tools:

- a table of the frequency range (high, medium, low);
- a table and a Pareto chart ordered according to decreasing frequencies,
- a word cloud.

Table 5 depicts the frequency range divided into high, medium, and low for the first five nouns, adjectives, and verbs.

The high-frequency range of words incorporates the subject words and is composed solely of nouns and adjectives.

The middle range includes keywords as well as nouns and adjectives and three verbs.

Taken together, these two "aggregates" show the efforts of banks towards a low environmental impact economy.

Table 5. Frequency ranges (high, medium, low)

Noun			*Adjective*			*Verb*		
High frequency ranges (257 to 96)								
emissione	257	nom	ambientale	158	adj	---		
consumo	234	nom	totale	149	adj			
gruppo	219	nom	energetico	115	adj			
energia	144	nom	rinnovablle	115	adj			
carta	124	nom	---					
Medium frequency ranges (95 to 52)								
obiettivo	89	nom	rinnovabile	88	adj	utilizzare	61	ver
impianto	83	nom	sostenibile	83	adj	acquistare	59	ver
utilizzo	82	nom	aziendale	74	adj	prevedere	57	ver
prodotto	81	nom	proprio	71	adj	---		
ambiente	80	nom	relativo	65	adj			
Low frequency ranges (51 to 1)								
sede	52	nom	europeo	49	adj	riciciare	49	ver
acqua	52	nom	specifico	46	adj	considerare	36	ver
societa	50	nom	principale	44	adj	ridurre	35	ver
linea	50	nom	primo	44	adj	destinare	35	ver
intensita	50	nom	immobile	41	adj	derivare	34	ver

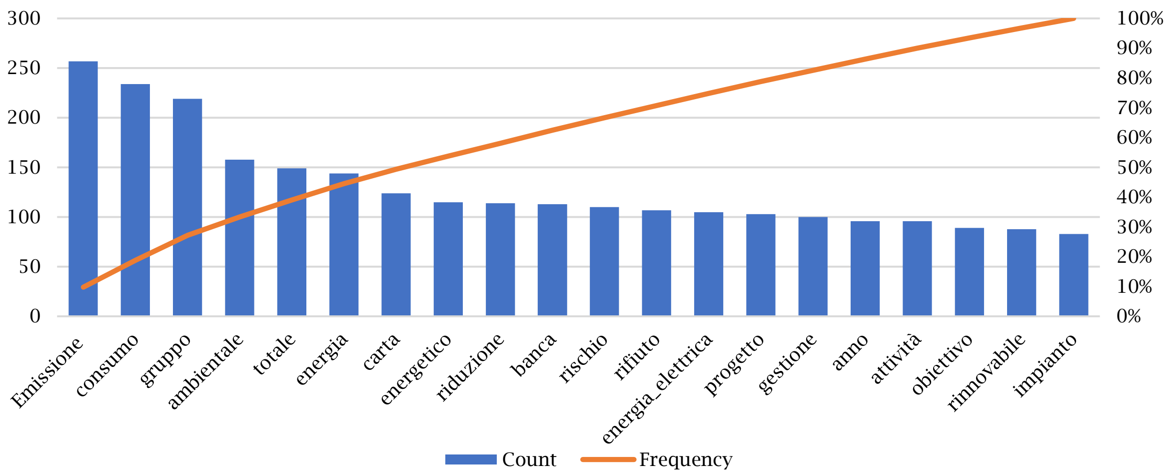
The frequency range between 257 and 96 features the keywords, that is the words that represent the main meaning. In this case, 5 nouns and 4 adjectives are used. In the middle range (frequencies between 95 and 52), the keywords of the analyzed text are included: 5 nouns, 5 adjectives, and 3 verbs.

The first 20 words (see Table 6) of the lexical matrix (*terms documents matrix*), represented in absolute frequencies, include referential words that are mainly in favor of an orientation towards the mitigation of *climate impacts change*, the companies analyze, express, and communicate greater attention to the processes of transition and reduction of carbon emissions, in line with the international agreements and the requirements of the main corporate reporting standards; this is a distinctive feature that can be found in all the banks subjected to this analysis.

Table 6. Frequency of the first twenty words

Words (Italian/English)		*Count*	*Part of speech*
emissione	emission	257	nom
consumo	consumption	234	nom
gruppo	group	219	nom
ambientale	environmental	158	adj
totale	total	149	adj
energia	power	144	nom
carta	paper	124	nom
energetico	energetic	115	adj
riduzione	reduction	114	nom
banca	bank	113	nom
rischio	risk	110	nom
rifiuto	rejection	107	nom
energia elettrica	electric energy	105	nr
progetto	project	103	nom
gestione	management	100	nom
anno	year	96	nom
attività	activity	96	nom
obiettivo	objective	89	nom
rinnovabile	renewable	88	adj
impianto	plant	83	nom

Figure 4. Pareto graph (frequency of the first twenty words)



The most frequently used words, organized according to the decreasing order of frequency, indicate that the first five items ("*consumption*", "*emission*", "*energy*", "*group*", "*environmental*", and "*total*") are needed to induce to evaluate in terms of the centrality the company activities aimed at the interest "in supporting the transition to a net zero emissions economy over time" and the treatment of environmental risks, in particular those of transition, while the words referring to adaptation activities as a response to climate change are less considered.

In this sense, some analyses of the concordance are reported (these are words that precede or follow a key term):

******011 *bank_creditovaltellinese *index_ftseitaliamidcap *year_2020**

il sensibile calo dei **consumi** rispetto_al 2019 sconta anche il fatto che per il 2020 è stato adottato un diverso metodo di calcolo dei **consumi** in accordo con le linee guida abi

The lower presence of personnel, due to COVID-19, implies that a minor use of the heating and cooling systems has led to a reduction in consumption in Banca BPM.

******005 *bank_bpm *index_ftsemib *year_2020**

che hanno richiesto un maggiore numero di ricambi d aria e la necessità di riscaldare e raffrescare gli ambienti a prescindere dal numero di addetti presenti in ufficio in linea con i **consumi** sono diminuite anche le emissioni

Consumption refers to the various sources of energy used, both traditional and renewable, but also water and the supply of raw materials such as paper and certified paper, and the emissions of Scope 1, Scope 2, and Scope 3 are reported for most banks in line with the greenhouse gas (GHG) Protocol.

******002 *bank_mediobanca *index_ftsemib *year_2020**

le emissioni indirette derivanti dal **consumo** di energia elettrica acquistata da terzi scopo 2 market based e location based nonché dalle trasferte dei nostri collaboratori in treno e in aereo scopo 3

The pandemic has resulted in less presence in the branches and therefore less consumption in-house of paper and toner.

******005 *bank_bpm *index_ftsemib *year_2020**

la pandemia ha comportato effetti anche in termini di **consumo** dei materiali l allargamento della platea di smart worker e la minore presenza della clientela in filiale hanno infatti contribuito alla riduzione dei **consumi** inhouse di carta e toner

The consumption of resources also refers to more aware practices of the circular economy, such as more efficient management of consumption and separate collection; smart working has decreased the use of paper.

The substantive emissions have similar indications, and the consumption of the whole dataset refers in particular to the methodologies and reporting measures adopted (tons of CO_2 equivalent).

******001 *bank_unicredit *index_ftsemib *year_2020**

nelle filiali temporaneamente chiuse gli impianti sono_stati spenti con un risparmio energetico di circa 1 800 gj pari a circa 170 tonnellate di **emissioni** di co2 rispetto_al 2019

Basically, there is a decrease in energy consumption and greenhouse gas emissions, COVID-19 has contributed to this regard.

The parts of speech of Table 6 are made up of 15 nouns (nom) and 4 adjectives (adj), they do not present interpretative ambiguities as they have a well-identified semantic area, especially within the type of topic being addressed.

A single exception, in terms of ambiguity, could be given by the word "*total*" but it originates from the summary values of the corpus tables obtained, with the cleaning of the respective numerical values (eliminated from the document because they would have been considered as empty forms).

The remaining 5 words, less frequent than the previous ones, namely ("*paper*", "*energy*", "*reduction*", "*bank*" and "*renewable*"), confirm the orientation of banks to the effort to reduce the impact of climate change and it is believed that they are used in this sense because they are aimed at providing a representation which is consistent with the reporting standards (the companies analyzed mainly used the Global Reporting Initiative (GRI) criteria for reporting) adopted and commonly used by financial institutions.

The word risk is present with an absolute frequency of 110 and the analysis of concordances starting from the matrix of active forms highlights its proximity to nouns and adjectives including currency, COVID-19 pandemic, corruption, environmental, the climate in general and in terms more specific to portfolio risk management and portfolio selection (portfolio, loans, investments, financing) to a concentrated extent in very few banks.

Worthy of note are the concordances of the word risk in Banca BPER, they are immediately connected to the physical adjective and the nouns transition and scenario.

Subsequently, the word risk and its many forms will be the object of the construction of a model in SAS Viya using information extraction, through grouping variables.

The analysis of the concordances of the word energy identifies the different types of energy (electrical, thermal but also renewable) and is consistent with the scientific considerations that underline how any activity limiting the effects of climate change is connected to the type, production, and use of energy resources, and is also consistent with the reporting frameworks adopted and in most cases of international value.

Energy is a word that occurs 260 times in the singular form and 15 times in the plural form and also shows collaborative approaches to tackle the problem, considered by the literature to be of particular importance, as in the following example:

******005 *bank_bpm *index_ftsemib *year_2020**

la prosecuzione del dialogo con fire federazione italiana per l uso razionale dell **energia** nell ambito del progetto europeo goesi

The study of the linguistic data conducted so far, for the research questions of our study, shows overall a report aimed at making the impact of climate change less serious through the prevention or reduction of GHG emissions in the atmosphere, through the share of renewable energy use, a less impactful corporate mobility system; mitigation is achieved by reducing the sources of these greenhouse gases.

In terms of document matrix for analysis, there is less attention to anticipating the adverse effects of climate change, in other words: the adoption of adequate measures to prevent or minimize damage or take advantage of the opportunities that may arise; in other words, adaptation can be understood as the process of adapting to the current and future effects of climate change[19].

The word cloud has been set on Iramuteq with reference to the first 150 most frequent words in the corpus, this does not alter the final result but improves its readability.

Figure 5 is the third way in which we intended to represent the textual terms of the lexical matrix.

The word consumption ("*consume*") is very frequent in the placement of texts, it is associated with efficiency and reduction of emissions in almost all banks, and also with graphic forms of energy and energy consumption. Its significant presence is also linked to direct and indirect adjectives which together, with the type of consumption (water, paper, gas, petrol, etc.) illustrate the reporting methods.

Figure 5. Word cloud

Source: Authors' elaboration using Iramuteq

4.3. Specificities and lexical correspondence analysis

The analysis of the specificities allows verifying if a word is specific or not of a single text/category of texts. A word is considered "specific" if it appears in a document to a greater extent than expected (the reverse indicates a rare word). In fact, the observed frequency and the expected frequency of a word within a document are compared, and through a calculation of probabilities based on hypergeometric distribution specificities are identified.

[19] https://www.eea.europa.eu/it/help/domande-frequenti/qual-e-la-differenza-tra Although the operational definitions for the following research derive from the TCFD, it seems useful to allow an understanding of these technical concepts with the definitions of the European Environment Agency (EEA). We have also put definitions of IPCC and a part of TCFD.

The following tables show the positive specificities of all the banks referring to the first 19 words found in the specificities tables by Iramuteq. Our representation is a graphic simplification that facilitates reading.

The specific words sorted by decreasing score identify the areas of greatest interest for each bank.

Table 7. Specificities (BPER, BPM, and Credito Emiliano)

X.bank_bper		*X.bank_bpm*		*X.bank_creditomiliano*	
bper	55	raccogliere	14	credem	49
scenario	23	clientela	13	carbone	16
alto	23	contrasto	12	target	16
portafoglio	22	a4	12	reconducibile	15
intensita	14	smaltire	8	fisico	13
ets	14	social	8	foresta	11
risparmiare	13	contratto	8	percorrenza	10
modena	11	comunicazione	7	edificio	9
ftv	11	importante	7	trasferta	7
basso	10	carta	6	aziendale	7
Co2e	10	singolo	6	razionalizzazione	6
rischio	10	cambiamento_climatico	6	albero	6
stima	10	materiale	6	immobiliare	5
espozione	9	variazone	5	finanlizzare	5
settore	9	ottico	4	working	5
medio	9	quantitativo	4	serra	5
kwh	8	avvio	4	persona	4
potenza	7	utilizzare	4	change	4
firma	7	mln	4	remoto	4

Table 7 is interpreted by reading the individual words and with an overview and comparison with the terms and other variables investigated.

Based on the specificities, it should be noted that an element that accompanies the financial statements of these banks is the fact that they are strongly focused on the practices and concrete objectives of reducing environmental impact.

- *Banca BPER* (see Table 7) uses the word "*scenario*" as a specific word, it connotes itself as a financial institution of interest to verify the research questions of studies. In fact, in its text, there are also physical and transition risks and their association with the portfolio, credits, investments, and financing.

There are also a series of words that define the units of measurement adopted to report the reduction of impacts.

Compared to other banks, Banca BPER identifies the word "*risk*" very well and makes a portfolio assessment of it, and speaks of scenario analysis with respect to plausible future states (the word plausible is an attribute explicitly called into question by the TCFD). It refers to the business as scenario usual (BAU) in which a strong increase in temperature is assumed and climatic, physical and transition risk scenarios, in which the assumptions of the BAU scenario are significantly altered. Banca BPER also talks about transition risk on the loan portfolio. The "*scenario*" is a word that appears with very low absolute frequency, but as we will see in the analysis with SAS Viya it is concentrated in a single document.

In addition, Banca BPER makes a specific reference to the TCFD benchmark.

******004 *bank_bper *index_ftsemib *year_2020**

l analisi di **scenario** è dunque uno strumento importante e utile per un organizzazione al_fine_di valutare le potenziali implicazioni di business dei rischi legate al clima e per informare gli stakeholder su come l organizzazione si sta posizionando alla luce di questi rischi

******004 *bank_bper *index_ftsemib *year_2020**

la task force on climate related financial disclosure tcfd raccomanda l utilizzo di un analisi di scenario proprio al_fine_di valutare i rischi legati al **cambiamento_climatico**

- *Banca BPM* (see Table 7) has positive specificities: climate change, disposal, and paper and pollution reduction measures due to the use of these materials and is characterized by showing their green practice with a view to a circular economy.

Climate change appears 25 times in the corpus in general, in BPM it is also linked to the offer of products with low environmental impact and the awareness of energy supply:

******005 *bank_bpm *index_ftsemib *year_2020**

questa offerta rappresenta di fatto un_potenziale green per tutti i nuovi mutui ipotecari e per gli immobili di tutte le classi energetiche e tiene conto dell importanza della riqualificazione energetica del patrimonio immobiliare esistente anche_come contrasto all aumento della cementificazione dei centri urbanie della conseguente minore resilienza di_fronte agli effetti del **cambiamento_climatico**

******005 *bank_bpm *index_ftsemib *year_2020**

energetiche e di contrasto al **cambiamento_climatico** approvvigionamento di energia elettrica da fonti rinnovabili iniziative di efficienza e monitoraggio dei consumi utilizzo di materiali riciclati raccolta differenziata e recupero di materiale gestione centralizzata dei consumi di carta e toner sicurezza ambientale degli

- *Banca Credito Emiliano* (see Table 7) identifies among its peculiarities the way of rationalizing the pollution activities of the group due to the mobility of the workforce, as well as to the efficiency of the facilities.

The next set analyzed consists of the group shown in Table 8.

Table 8. Specificities (Credito Valtellinese, Desio e Brianza, and Generali)

X.bank_creditovaltellinese		*X.bank_desiobrianza*		*X.bank_generali*	
associare	10	desio	38	generale	82
riportare	10	mq	20	quintale	31
sito	8	espresso	16	capire	27
elettrico	8	tempo	15	acqua	24
benzina	8	dipendente	13	banca	21
impianto	7	banco	11	m3	20
peso	7	apparecchiatura	10	tabella	19
milano	7	riferimento	9	ripartizione	19
fsc	6	relativo	7	sede	13
condominiale	6	segnalare	7	variazione	12
ctntrale	6	termico	6	treno	11
fonte	6	riutiltzzare	6	km	11
totale	6	toner	6	defra	10
teleriscaldamento	5	indicatore	6	rifiuto	9
equivalente	4	soluzione	6	aereo	9
riscaldamento	4	gruppo	6	operativo	9
derivare	4	consumo	5	scope_3	8
tco2	4	gi	5	performance	8
gasolio	4	sede	5	rocorso	7

- In *Banca Credito Valtellinese* (see Table 8), the consumption activity of the plants, the measures associated with the reduction of carbon dioxide (greenhouse gas), and a series of words such as source, total, and heating are used as specific characteristics. In it, there is tCO2 (tons of carbon dioxide equivalent) which is the measure adopted for the disclosures referable to the GHG Protocol.

******011 *bank_creditovaltellinese *index_ftseitaliamidcap *year_2020**

emissioni di gas serra dirette **tco2** equivalenti da gas naturale per riscaldamento da impianti autonomi da gasolio per riscaldamento da gpl per riscaldamento da autoproduzione da cogenerazione totale emissioni della flotta aziendale dirette **tco2** equivalenti aziendali diesel benzina elettrico benzina auto a noleggio

- In *Banca Desio e Brianza* (see Table 8), the specificities are all connected to the reporting methods for activities aimed at reducing pollution.
- In *Banca Generali* (see Table 8), we note the presence of the term "Scope 3", generally used as a metric for reporting carbon emissions, and among the specific words the units of measurement adopted, the various references to the areas of intervention where an attempt was made to manage the process towards net zero transition.

Table 9. Specificities (Intesa San Paolo, Mediobanca, and Mediolanum)

X.bank_intesa		*X.bank_mediobanca*		*X.bank_mediolanum*	
sanpaolo	64	scopo	55	mediolanum	171
intendere	50	diesel	23	donare	109
circular	44	indiretto	15	euro	59
economy	42	unita	11	fondo	55
green	27	fonte	10	erogazione	47
circolarre	20	based	9	fondazione	17
plafond	16	rinnovabile	8	emergenza	13
offerta	13	derivare	7	molano	12
impresa	13	idrico	7	raccolta	11
finanziamento	13	dirigere	5	sostenere	9
miliardo	12	market	5	collaborazione	8
linked	11	location	5	distinare	8
esg	11	aziendale	5	impegnare	8
economia	11	termico	4	covid	7
bond	11	treno	4	acquisto	7
studio	10	aero	4	ricerca	6
europeo	10	naturale	4	strumento	5
euro	10	benzina	4	acquistare	4
dedicare	9	consumo	4	associazione	4

- *Banca Intesa* (see Table 9) is particularly connected to the principles and general arguments of the circular economy, also with reference to the actions that govern the fight against climate change. The following are present in the specific discourse: the reference to environment, social and governance (ESG) criteria, sustainable activities finance (linked, financing, plafond, bond) customer services, and investments adopted to be a green company.
- *Mediobanca* (see Table 9) is a large bank of the FTSE MIB that uses disclosures relating to the measures adopted to combat climate change and transition from a carbon-intensive economy to one with a low environmental impact. The terms of the specificity table mainly refer to concrete practices

and actions of ecological transition (emission sources, renewable, thermal), in this sense the "purpose" form should be disambiguated, which is probably to be understood as a measurement metric and reporting criterion of the GRI framework used.

- *Banca Mediolanum* (see Table 9) mainly exhibits the social practices connected to the theme of the environment with positive specificities: disbursement, fund, support, commit, and allocate, also practiced on the territory. In detail, the report summarizes the social practices that have an inclusive meaning with the collaboration and support of scientific activities with the word research.

Table 10. Specificities (Monte dei Paschi di Siena, Popolare di Sondrio, and UniCredit)

X.bank_mps		*X.bank_popolaresondrio*		*X.bank_Unicredit*	
gri	29	Popolare	49	capitale	24
strumentale	15	Sondrio	45	finanza	12
parametro	11	Udm	42	naturale	8
coefficiente	11	Tco2e	15	efficienza	8
gasolio	10	Mc	14	disponibile	8
noto	10	Diminuzione	11	informasione	8
promiscuo	10	quota	10	fornitore	7
metodologia	9	Impiego	9	Min	7
litro	8	Acquistare	9	maggiore	6
uso	8	Rendicontazione	8	sostenibile	6
guida	7	Ronnovablile	8	intraprendere	4
ghg	7	Maggiore	8	integrara	4
pericoloso	7	Incremento	8	finanziamento	4
gi	7	Effetto	7	approvvigionamento	3
linea	6	condominiale	7	documento	3
tonnoliata	6	teleriscaldamento	7	azione	3
informativo	6	kwh	6	programma	3
abi	6	standard	5	sostegno	3
precendente	6	perimetro	5	prestito	3

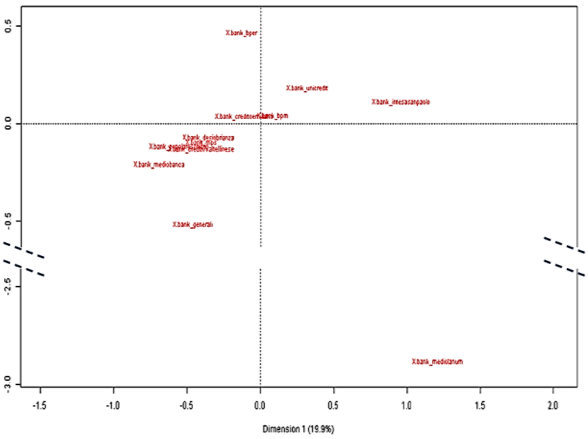
- *Banca Monte dei Paschi di Siena* (see Table 10) is extremely focused on data and measures, with the indication of the methodology adopted, on the international GRI framework, and with the explicit reference to GHG. It is consistent with the calculation of emissions according to this protocol (even if not clearly specified).
- *Banca Popolare di Sondrio* (see Table 10) concretely relates the strategies to contain emissions, and the strategies to contain the impact, allowing us to evaluate the green effort even if with an orientation to past activities.
- *Banca UniCredit* (see Table 10) refers to major scenarios and major "principles" such as sustainability, integration, support, linked to the negative effects of climate change, and capital and finance for sustainable finance and green products for customers.

4.4. Lexical correspondence analysis (LCA)

The analysis of lexical correspondences (LCA) integrates with that of specificities and both identify the peculiar characteristics of a text. LCA is an application to textual data of simple correspondence analysis (a multivariate statistical technique belonging to the group of factorial analyses[20]).

[20] The large category of factorial analyses takes a very large data matrix and reduces it in dimensionality by projecting the large amount of data into a small subspace. This process is accomplished through the creation of factors — algebrically constructed variables — starting from the original data. The first two factors are taken; they are nothing more than the latent dimensions on which the data are projected, reproducing the maximum possible amount of variability that is contained in the original matrix. Factor analysis par excellence is the analysis of the principal components that works on quantitative variables; it creates the principal components (factors) — which are precisely a linear combination of the original variables; the use of the first two factors allows the representation in two dimensions. The analysis of lexical correspondences has the same objective as the analysis of the principal components: it synthesizes the content of many (qualitative) variables in a reduced subspace, and then on the factors we project our information.

Figure 6. Lexical correspondence analysis (LCA)



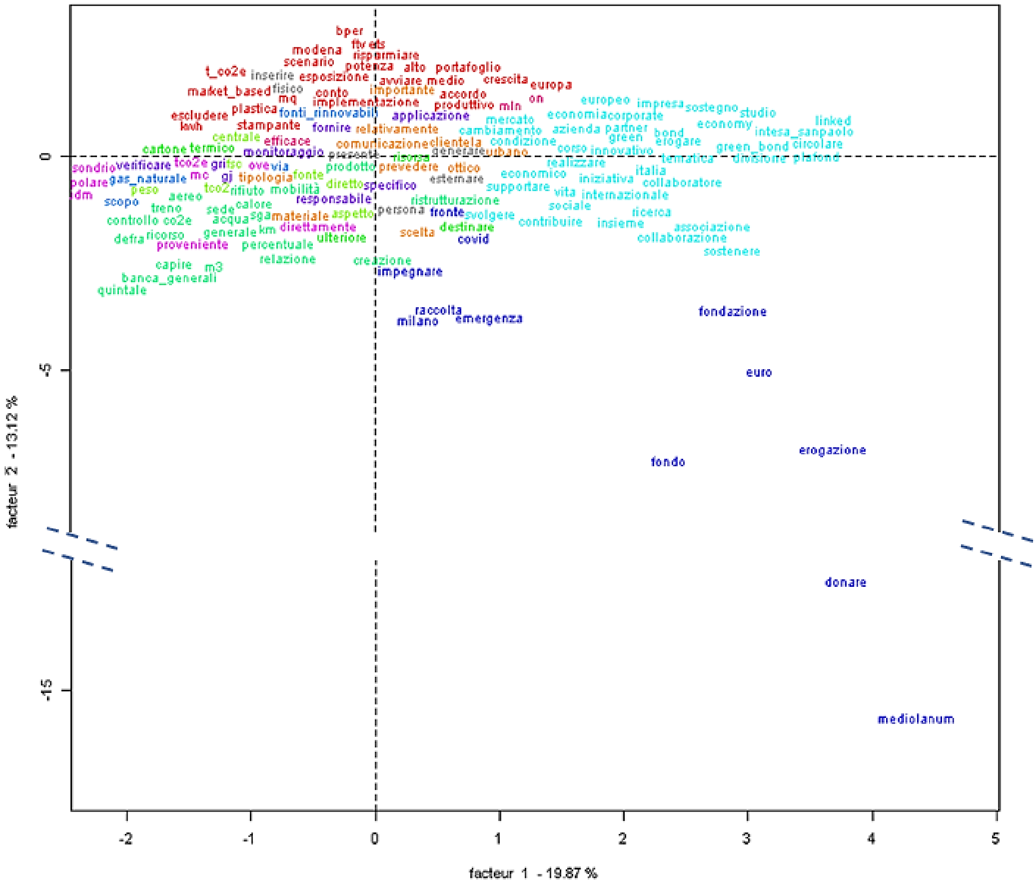
Source: Author's elaboration using Iramuteq.

After examining the two overlapping factorial levels of the cross-analysis of the lexicon used by the banks, the following considerations can be made: the first dimension (first factor, 19.9% of total inertia represented), consists of the difference between lexicon which describes the effective actions to reduce consumption and energy impact, on the left, and sustainable finance policies on the right.

On the right side of the axis (see Figure 6), we find Intesa San Paolo and UniCredit, which, therefore, use a partially assimilable lexicon, and on the part of the first factorial dimension we find all the other financial companies, apart from BPER and Mediolanum, which, instead, characterize the second factorial axis (second factor, 13.12% of total variance represented).

The second factorial dimension presents polarization between the lexicon that describes financial sustainability (the words "*portfolio*", "*estimate*", and "*power*", located at the top of Figure 7, in red, which represent the color of the typical words of BPER), and the lexicon that describes social investment with the words "*disbursement*", "*fund*", "*donate*", "*foundation*", "*support*", "*collaboration*", along the positive semi-axis characterized by the blue color typical of Mediolanum.

Figure 7. LCA: Polarization between the lexicon



Source: Author's elaboration using Iramuteq.

4.5. Similarities analysis

Similarity analysis, using graph theory, identifies and represents co-occurrences within text segments, allows us to study how speech is constructed in a text and, therefore, how words co-occur in small segments (in Iramuteq there are 40 occurrences by default), making explicit the words which are most associated with each other.

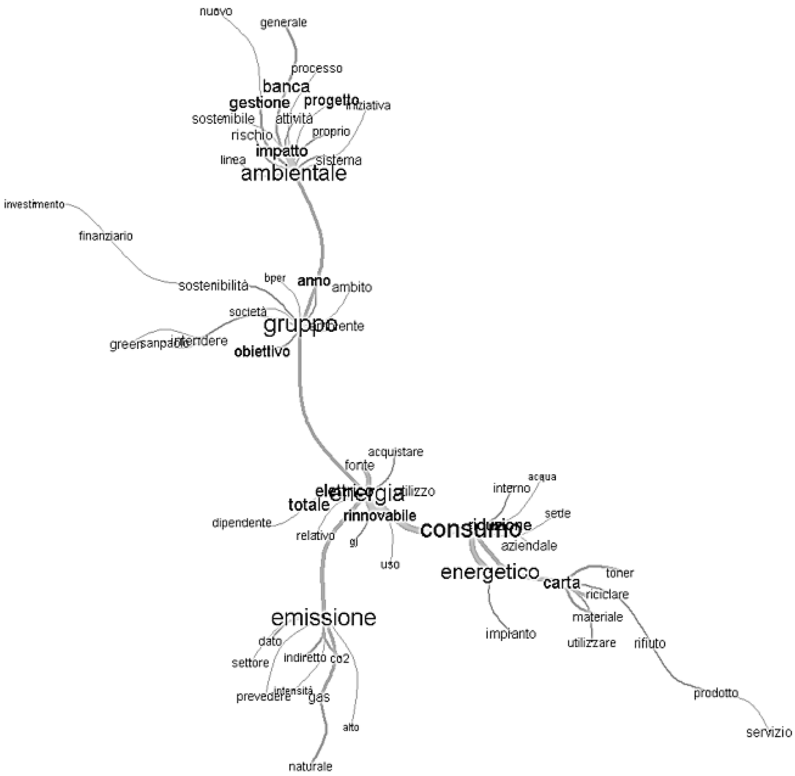
The analysis of networks within the textual analysis is a fairly recent approach in the literature and with it, we substantially pass from a frequentist to a relational approach.

In the frequentist approach, the text is encoded in a frequency table (in the cells of the matrix there is the absolute frequency or the presence/absence of the term), in the relational one, the text becomes a network: the lexical matrix is transformed into a network through the passage to a contingency matrix that has the same variables or the same cases on both axes.

And so we can imagine that words have a relationship with each other like documents, obviously, these relationships are built starting from co-occurrences.

The main words of the speech, written larger and located at the intersection points, are energy, consumption, environmental, group, and emission. The terms related to the word energy, consumption, and emission, appear with greater probability within the text segment. Excluding the graphic form called group, the study of the sequences clearly shows the concrete commitment to the reduction of emissions and the transition towards a low environmental impact economy. The construction of the discourse is entirely centered on the low carbon strategy and the reconstruction of the topic is unequivocally the ecological transition, in fact in the text segment almost exclusively terms referring to the fight against climate change are needed. The substantive group that highlights the management of the process at a central level, the impact of climate change, and its risks deserve attention.

Figure 8. Similarities analysis



Source: Author's elaboration using Iramuteq.

4.6. Cluster analysis

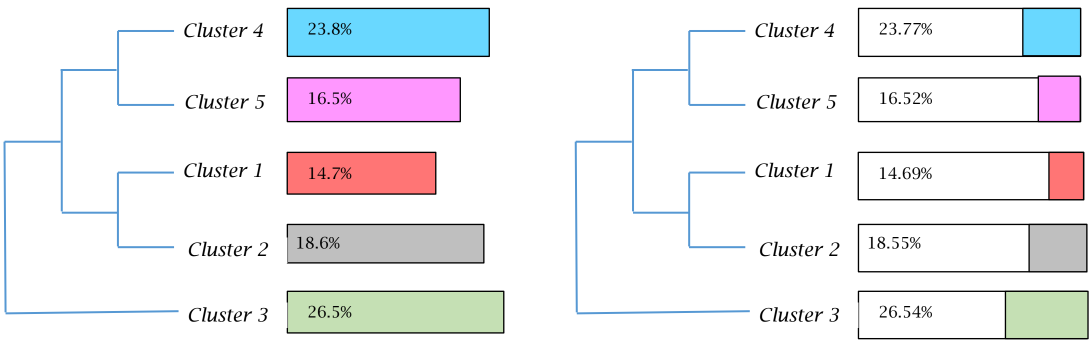
Cluster analysis, a method of unsupervised data mining, divides a set of statistical units so that the resulting subdivision has some properties considered desirable (Sokal & Rohlf, 1962), that is, to form groups of units so that the units that are assigned to the same group are homogeneous with respect to the characters considered, and each unit of the collection is contained in one and only one group (Gordon, 1999).

Cluster analysis is a set of relatively recent techniques that have drawn statisticians' interest since about 1960, also thanks to the parallel evolution of automatic calculation tools, which has made it possible to tackle computational complexities without difficulty.

Iramuteq uses Reinert's method (Reinert, 1990) a descending agglomerative hierarchical method (hierarchical method) whose application creates five groupings. Clustering is represented by:

- Number of texts: 12;
- Number of text segments: 1,166;
- Number of forms: 5,811;
- Number of occurrences: 42,045;
- Number of lemmas: 4,192;
- Number of active forms: 3,629;
- Number of supplementary forms: 524;
- Number of active forms with a frequency ≥ 3: 1,542;
- The mean of forms by segment: 36.059177;
- Number of clusters: 5;
- 1,035 segments classified on 1,166 (88.77%).

Figure 9. Dendrogram and cluster coverage



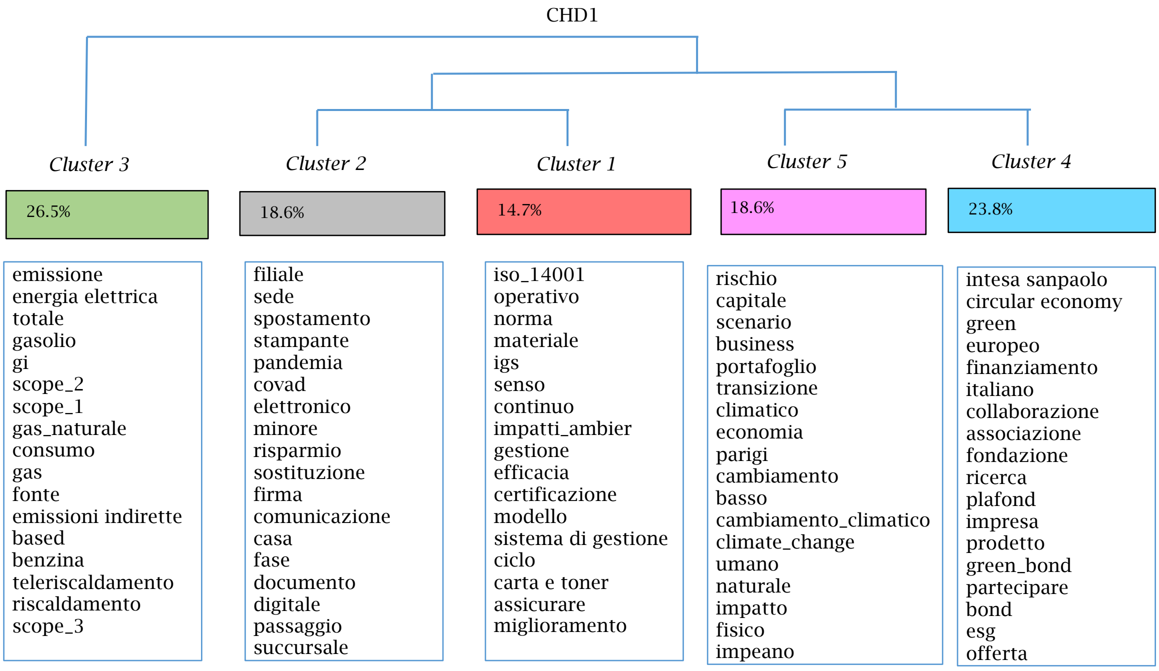
Dendrogram starting from CHD1 | Dendrogram CHD1 - phylogram

The cluster identified five lexical classes covering approximately 85% above the variance (the minimum acceptable threshold is 75%).

The steps of the descending reclassification algorithm that led to the identification of the final five clusters are represented by the dendrogram (see Figure 9).

- Cluster 1 classified 152 text segments out of a total of 1,035 (14.69%).
- Cluster 2 classified 192 text segments out of a total of 1,035 (18.55%).
- Cluster 3 classified 274 text segments out of a total of 1,035 (26.47%).
- Cluster 4 classified 246 text segments out of a total of 1,035 (23.77%).
- Cluster 5 classified 171 text segments out of a total of 1,035 (16.52).

Figure 10. Cluster topic



Dendrogram CHD1 - phylogram

Cluster 1 (red) refers to the regulations, in particular, the voluntary ones, linked to the reduction of environmental impact, on the internal management systems and therefore the implementation of regulations and company requirements.

Cluster 2 (grey) specifies the concrete measures adopted at each organizational level and reported as good practices for reducing emissions.

Cluster 3 (green) is a universe of meaning linked to the practices of reducing emissions and environmental impact.

Cluster 4 (blue) refers to the founding principles of actions to contain emissions and environmental impact.

Cluster 5 (viola) shows the perception of climate risk in relation to the offer of green products to customers.

4.7. Analysis with SAS Viya Text Mining

This part of the research analyses SAS Viya for the correct identification and categorization of textual data.

The pipeline includes five analysis nodes named: *concepts*, *parsing*, *sentiment*, *topics*, and *categories* to allow the identification of relevant textual data and building "concept and categorization models" (SAS, 2019).

It was not possible to apply machine learning techniques since the small size of the corpus does not allow a correct structured flow of analysis nor was it possible to identify whether the document expressed positive, neutral, or negative attitudes for the same reason (impossibility of executing the sentiment analysis)

The "*concepts*" node allows the creation of additional concepts to extract knowledge from the corpus, discover in a document, or a set of documents. A concept is a useful property for analyzing information in context and for extracting useful information. A set has been created with the LITI programming language and the construction rules are illustrated further on.

The "*topics*" analysis node groups the "similar documents in themes or topics", it is an automatic grouping of important terms that occur in the documents and represent them; (first 5 terms that frequently appear in the topic).

The "*categories*" analysis node labels documents based on their content and by creating thought constructions that match research interests.

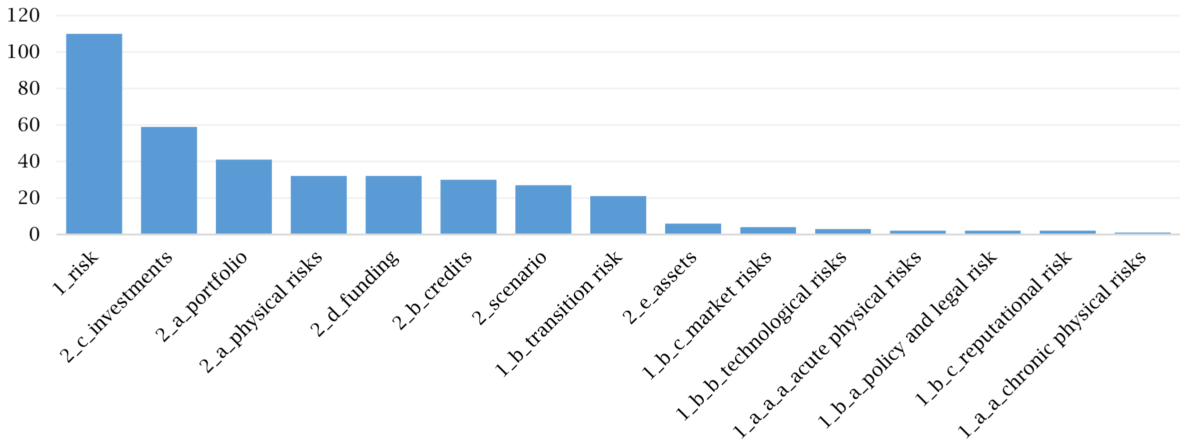
The construction of the aforementioned concepts allows for contextual analysis and is useful for extracting information and information retrieval consistent with the research questions.

Appendix provides a list of concepts and categories created using the LITI programming language: in addition to the extraction of information with "*classifier*" code, 8 concepts have been identified with the relative modalities and a window length of 20 words, and 16 categories with the relative sub-variables.

The concept node analysis extracts specific information from the corpus on SAS Viya for Text Analytics.

Figure 11 expresses the number of matches per concept that is how many times the concept appears in absolute terms in the text.

Figure 11. Number of matches per concept



Words (English/Italian)		Words (English/Italian)		Words (English/Italian)	
risk	rischio	credits	crediti	technological risks	rischi tecnologici
investments	investimenti	scenario	scenario	acute physical risks	rischi fisici acuti
portfolio	portafoglio	transition risk	rischio transizione	policy and legal risk	rischio policy and legal
physical risks	rischi fisici	assets	asset	reputational risk	fischi reputazionali
funding	finanziamenti	market risks	rischi mercato	chronic physical risks	rischi fisici cronici

The word "*risk*" has an absolute frequency of 110, "*transition risk*" 21, "*physical risk*" 32, and "*scenario*" 27, these words with their respective frequencies represent a general, but also generic focus on risk management that is not specific for our analysis. In fact, the partition variables of transition and physical risk have a very low presence: "*market risk*" 4, "*technological risk*" 3, "*policy and legal risk*" 2, "*reputational risk*" 2, "*chronic physical risk*" 1, and "*acute physical risk*" 1.

The additional concepts used outline a focus on *portfolio selection* and *portfolio management* in terms of products offered to customers and investments; the terms investment, portfolio, financing, and credit have the following frequencies: 58, 41, 32, and 30, respectively. Based on these data, it is believed that banks' product offerings are sensitive to climate change criteria, although it should be noted that no clear disclosure is given with respect to stranded assets and investments of this type.

Figure 12. Number of documents by the concept

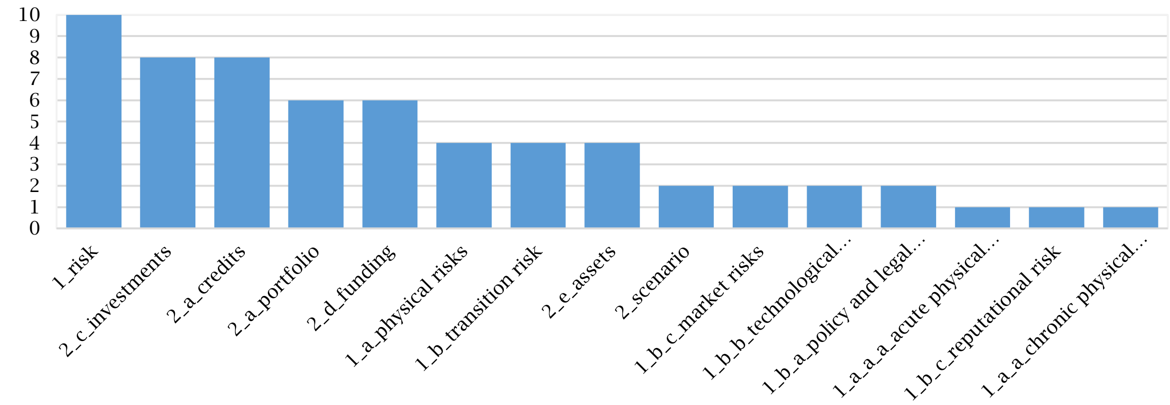


Figure 12 shows that the concept of "*risk*", the one with the highest frequency, is however present in 10 sustainability reports out of a total of 12, "*transition risk*" in only 4, and the same for "*physical risk*". The corpus is characterized by a very scarce presence of risk types such as market, technological, policy, legal, and even the term scenario. In fact, they are present in two documents only, and reputational risk is found in only one sustainability report.

Figure 13. The average number of matches per document

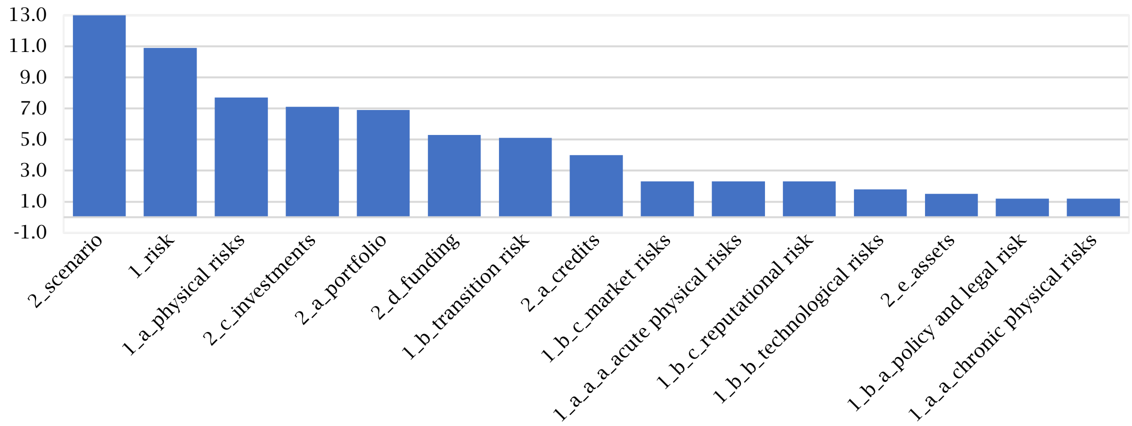
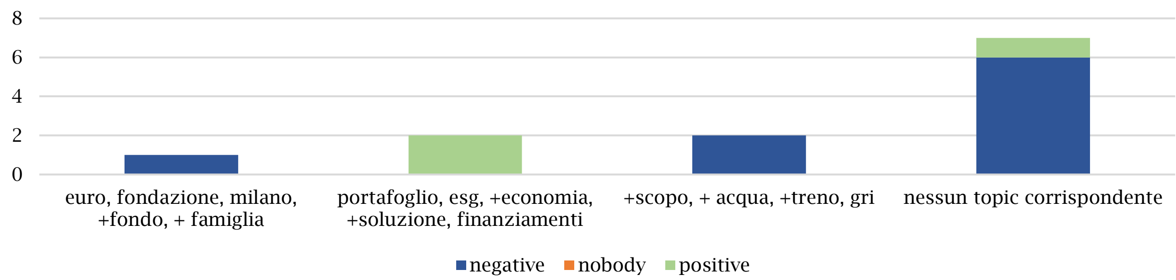


Figure 13 takes each term of corpus and divides it by the number of documents: the average values using Figure 11 and Figure 12. The word "*scenario*" only occurs in two documents and has the highest average value.

Figure 14. Topic detection

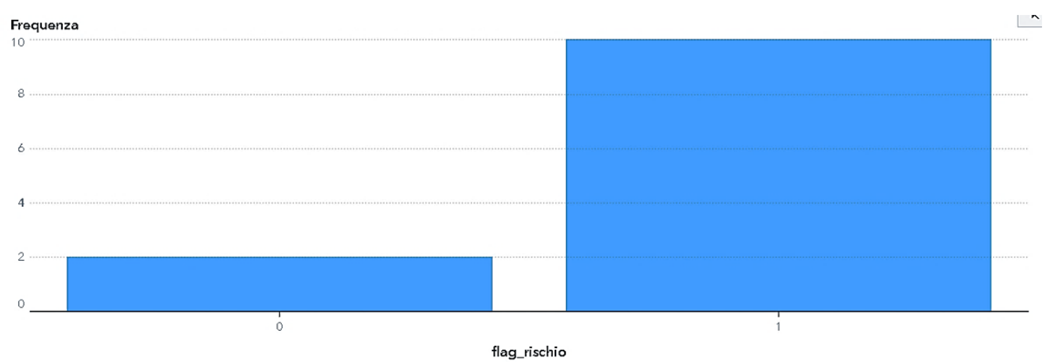


Note: The color green indicates a positive sentiment, while the color blue is the opposite.

In Figure 14, the topic detection shows only the grouping "*portfolio + esg + economy + reduction + financing*" as positive, outlining the banks' propensity to activate ESG-oriented products and a circular economy approach for only two documents in the corpus, while no topic detection can be found for about 6 documents; there is no corresponding topic. Figure 14 is unreliable because topic detection requires a lot of training data that are not available.

The analysis node called "*categories*" labels documents according to their content, i.e., it classifies documents by subject, a value of zero is interpreted as the absence of the category, and a value of 1 as the presence of the category (*dummy variable*); it also allows to focus on new categories that correspond to the research interests.

Figure 15. Risk category



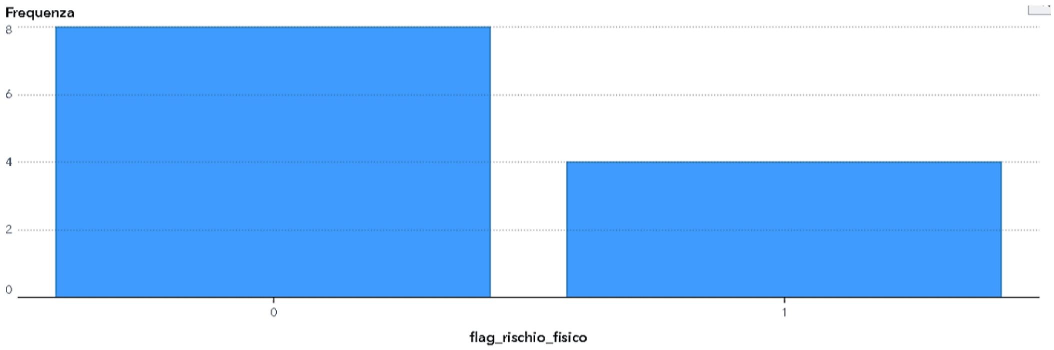
flag_rischio ▲	Frequenza
0	2
1	10

As seen above, the term "*risk*" is present in 10 documents and has a frequency equal to 110.

The category "*physical risk*" appears in only 4 documents and is not present in 8, it defines a lack of attention to this concept of climate change, the banks concerned are BPER, BPM, Credem, and Credito Valtellinese.

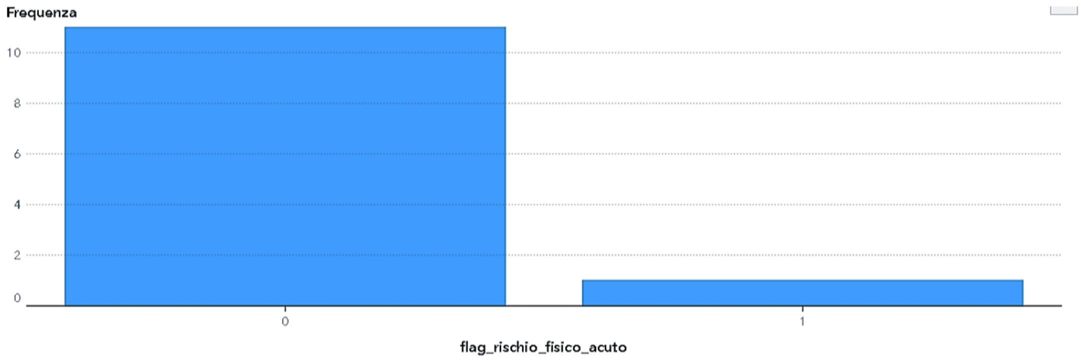
Confirming the chart illustrated above, the category pertaining to "*acute physical risk*" appears in only one document, specifically in the Credem report, while it is not present in the remaining 11.

Figure 16. Physical risk category



flag_rischio_fisico ▴	Frequenza ▾
0	8
1	4

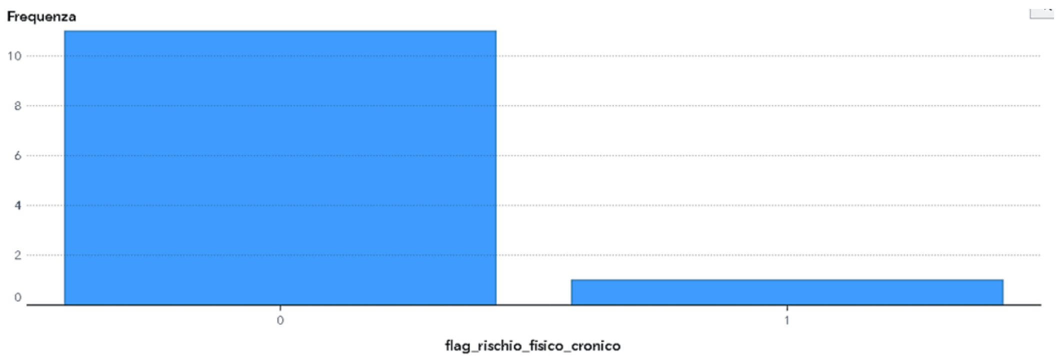
Figure 17. Acute physical risk category



flag_rischio_fisico_acuto ▴	Frequenza ▾
0	11
1	1

A similar consideration for "*chronic physical risk*" is only encountered in Credem, therefore it is evident that banks do not disclose these two important aspects (Figure 18).

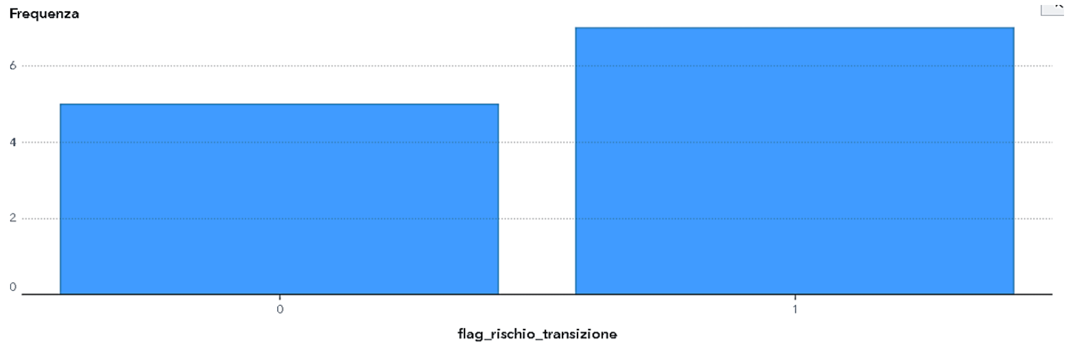
Figure 18. Chronic physical risk category



flag_rischio_fisico_cronico ▴	Frequenza ▾
0	11
1	1

Things are better for “*transition risk*” (Figure 19), which is identifiable in 7 out of 12 sustainability reports, but does not meet Banca Desio e Brianza, Mediobanca, Mediolanum, Monte dei Paschi di Siena, and Popolare Sondrio.

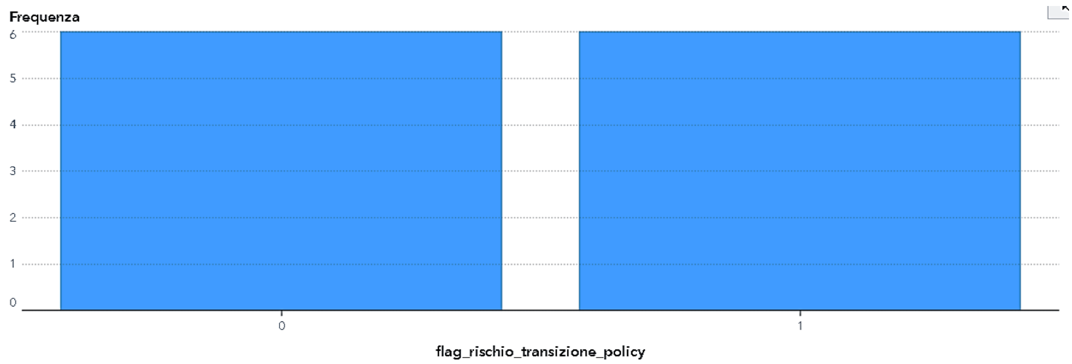
Figure 19. Transition risk category



flag_rischio_transizione ▲	Frequenza ▲
0	5
1	7

The breakdown of “*transition risk*” in its policy aspect (Figure 19) shows a positive response in 50% of the documents: BPER, BPM, Credem, Credito Valtellinese, Generali, and UniCredit.

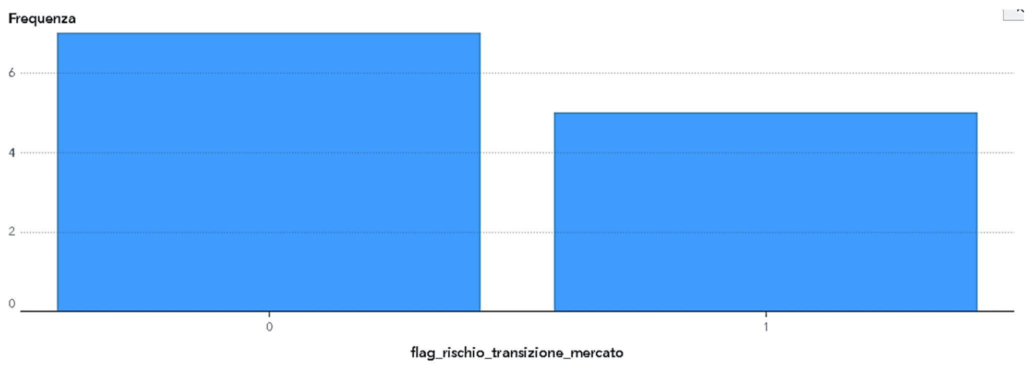
Figure 20. Policy transition risk category



flag_rischio_transizione_policy ▲	Frequenza ▲
0	6
1	6

The words “*transition*” and “*market*” (Figure 21), conceptualized in a single variable, do not appear in 7 documents but only in 5, those in which they are found are BPER, BPM, Credem, Intesa San Paolo, and UniCredit.

Figure 21. Market transition risk category

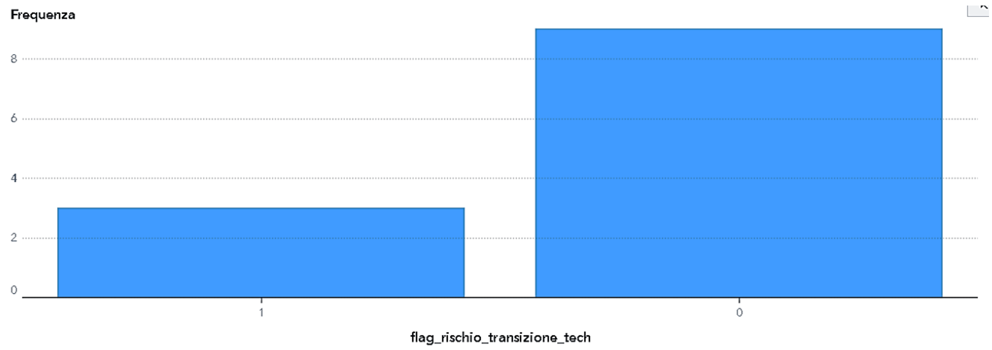


flag_rischio_transizione_mercato ▲	Frequenza
0	7
1	5

The aspect of "*technology-related transition risk*" (Figure 22) is encountered in only 3 documents and is absent in the remaining 9, it is presented in Credem, Generali, and Intesa.

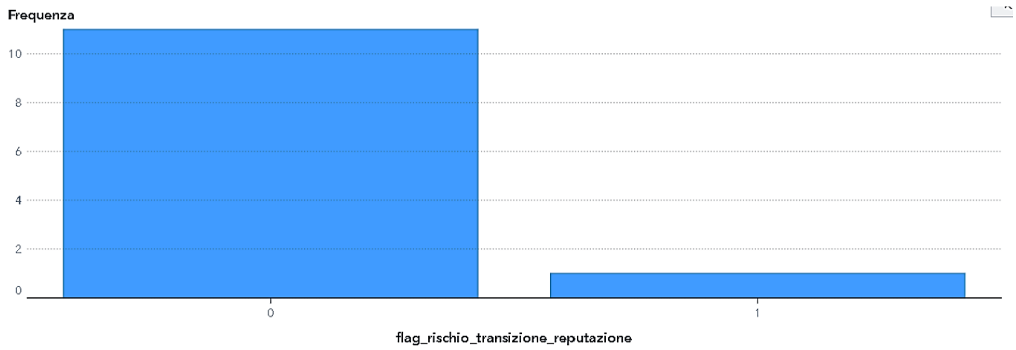
"*Transition risk*" in its reputational articulation (Figure 23) is not present in as many as 11 documents.

Figure 22. Technology-related transition risk category



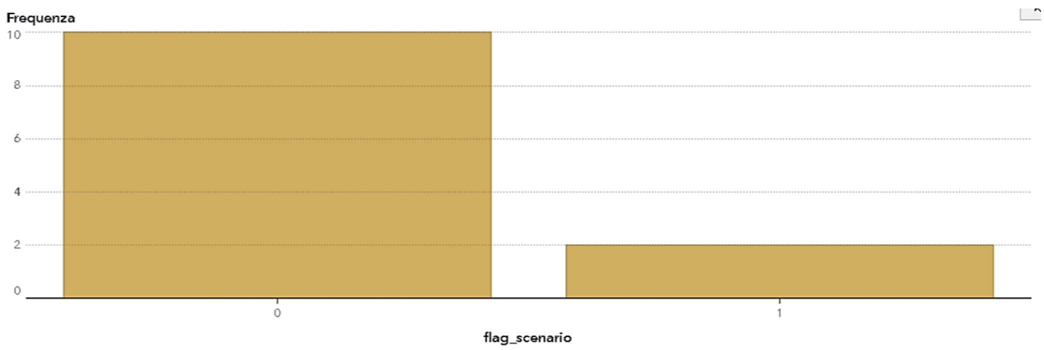
flag_rischio_transizione_tech	Frequenza
1	3
0	9

Figure 23. Reputational transition risk category



flag_rischio_transizione_reputazione	Frequenza
0	11
1	1

Figure 24. Scenario analysis category



flag_scenario	Frequenza
0	10
1	2

The absence in 10 reports of the scenario flag (Figure 24) shows the lack of interest in a forecasting tool within the individual bank balance sheets; the term is only found in BPER and UniCredit.

The key findings are the inadequacy and incompleteness of sustainability reports to disclose forward-looking and firm-specific information for investors, stakeholders, and regulators. The not common data information and the scarce completeness of them lead to incomparability and reduced transparency of the ESG disclosures. In addition, there is a pronounced opacity in the assessment of physical risks compared to transitional risks, which are, however, partially and selectively disclosed in an attempt to provide adequate information in terms of risk management. Most of the disclosures do not provide "a complete and understandable picture", increasing the potential for uncertainty in the markets and affecting financial stability. With regard to alignment to the pillars and recommendations of the TCFD, only very few reporting categories are fairly well covered while others are not, with variation between companies.

5. CONCLUSION

The results derive from the use of various multivariate techniques and lead to the insufficient disclosure of climate change information in banks' sustainability reports, requiring an effort to adapt the in-house organization of risk management functions and a path of integration with TCFD frameworks.

A limit of the research is the size of the corpus analysis that does not match with the application of machine learning and data mining techniques, but at the same time points to research directed towards the use of mixed methods of research, adding machine learning and supervised learning methods, and, as a result, the elaboration of better classifications and predictive models.

A clear and understandable picture of climate change disclosures helps investors, regulators, and stakeholders to better assess the companies and potentially reduce uncertainty and improve financial stability but they have to prioritise climate change information in the structure of sustainability reports.

This study promotes a recent research trend: the use of text analytics in economics and finance, and fosters its use among public institutions.

In conclusion, it is recommended that regulators and governments enforce in law TCFD disclosures by making them mandatory and aligned with this standard. Disclosing climate-related financial information on a mandatory basis allows to increase in the quantity and quality of climate-related reporting, set out the emission reduction plans and sustainability credentials, helps investors and businesses to better understand the financial impact of their exposure, assess climate-related risks more accurately, overcome short-terminism to favor long-term strategies, increase transparency and more comparability on corporate sustainability reports, also improve accountability, and provide clearer disclosures to actual and potential investors, lenders and regulators.

Researchers are calling for innovative methodological perspectives mixing quantitative and qualitative methods. The new research directions according to greater availability of data could be Machine Learning (ML) and especially ML techniques in text analytics (e.g., Glove and Word2Vec) for combating climate change by enabling the analysis of complex relationships, the combination of qualitative and quantitative variables in the same model, and better predictions.

As written in a recent paper under publication, implications for companies are the need to improve areas of risk management using scenario analysis, bearing in mind that combating climate change also requires the sound application of accounting disclosures; at the same time, this approach is useful for assessment by regulators.

A limit of the article can be found in the impossibility of applying ML techniques due to the small size of the corpus.

REFERENCES

1. Agrawal, A., & Hockerts, K. (2021). Impact investing: Review and research agenda Informa UK Limited. *Journal of Small Business & Entrepreneurship, 33*(2), 153–181. https://doi.org/10.1080/08276331.2018.1551457
2. AlHares, A., & Al-Hares, O. M. (2020). The impact of corporate governance mechanisms on risk disclosure. *Corporate Ownership and Control, 17*(4, Special Issue), 292–307. https://doi.org/10.22495/cocv17i4siart7
3. Anderson, G., Galang, P., Gamba, A., Medina, L., & Zheng, T. (2021). *How have IMF priorities evolved? A text mining approach* (IMF Working Paper 2021/159). International Monetary Fund. https://doi.org/10.5089/9781513573458.001
4. Arnold, R. D., & Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science, 44*, 669–678. https://doi.org/10.1016/j.procs.2015.03.050
5. Attenborough, D. (2022). Corporate disclosures on climate change: An empirical analysis of FTSE all-share British fossil fuel producers. *European Business Organization Law Review, 23*, 313–346. https://doi.org/10.1007/s40804-022-00246-2
6. Aversa, D., & Cincinelli, P. (2019). The contribution of Italian banking system to the Italian economic and social growth. In D. Aversa (Ed.), *Politics of banking: Regulatory capture, thin political markets, ideology and beyond* (pp. 53–84). Aracne.
7. Aversa, D., Adamashvili, N., Fiore, M., & Spada, A. (2022). Scoping review (SR) via text data mining on water scarcity and climate change. *Sustainability, 15*(1), Article 70. https://doi.org/10.3390/su15010070
8. Battiston, S., Dafermos, Y., & Monasterolo, I. (2021). Climate risks and financial stability. *Journal of Financial Stability, 54*, Article 100867. https://doi.org/10.1016/j.jfs.2021.100867
9. Camargo, B. V., & Justo, A. M. (2016). IRaMuTeQ: R interface for multidimensional analysis of texts and questionnaires. *Um Software Gratuito para Análise de Dados Textuais.Temas em Psicologia, 21*(2), 513–518. http://www.iramuteq.org/documentation/fichiers/IRaMuTeQ%20Tutorial%20translated%20to%20English_17.03.2016.pdf
10. Chenet, H., Ryan-Collins, J., & van Lerven, F. (2021). Finance, climate-change and radical uncertainty: Towards a precautionary approach to financial policy. *Ecological Economics, 183*, Article 106957. https://doi.org/10.1016/j.ecolecon.2021.106957
11. Ebner, A. (2018). The financial stability aspects of the EU-wide stress test. *Journal of Financial Regulation, 4*(2), 326–336. https://doi.org/10.1093/jfr/fjy009
12. Eccles, R. G., & Krzus, M. P. (2018). Implementing the task force on climate-related financial disclosures recommendations: An assessment of corporate readiness. *Schmalenbach Business Review, 71*(2), 287–293. https://doi.org/10.1007/s41464-018-0060-4

13. El-Hage, J. (2021). Fixing ESG: Are mandatory ESG disclosures the solution to misleading ESG ratings? *Fordham Journal of Corporate & Financial Law, 26*(2), 359–390. https://ir.lawnet.fordham.edu/jcfl/vol26/iss2/2/
14. Eory, V., Topp, C. F. E., Butler, A., & Moran, D. (2018). Addressing uncertainty in efficient mitigation of agricultural greenhouse gas emissions. *Journal of Agricultural Economics, 69*(3), 627–645. https://doi.org/10.1111/1477-9552.12269
15. Feridun, M., & Güngör, H. (2020). Climate-related prudential risks in the banking sector: A review of the emerging regulatory and supervisory practices. *Sustainability, 12*(13), Article 5325. https://doi.org/10.3390/su12135325
16. GARP. (2022). *SCR certificate sustainability and climate risk exam.* GARP.
17. Giuliano, L. (2004). *L'analisi automatica dei dati testuali: Software e istruzioni per l'uso.* LED Edizioni Universitarie.
18. Gordon, A. D. (1999). *Classification* (2nd ed.). CRC Press.
19. Huiskamp, U., ten Brinke, B., & Kramer, G. J. (2022). The climate resilience cycle: Using scenario analysis to inform climate-resilient business strategies. *Business Strategy and the Environment, 31*(4), 1763–1775. https://doi.org/10.1002/bse.2982
20. Hwang, H., An, S., Lee, E., Han, S., & Lee, C.-h. (2021). Cross-societal analysis of climate change awareness and its relation to SDG 13: A knowledge synthesis from text mining. *Sustainability, 13*(10), Article 5596. https://doi.org/10.3390/su13105596
21. International Sustainability Standards Board (ISSB). (2022). *Exposure draft IFRS® sustainability disclosure standard: [Draft] IFRS S2 climate-related disclosures.* International Financial Reporting Standards (IFRS) Foundation.
22. Johnson, A., Papi-Thornton, D., & Stauch, J. (2019). *Student guide to mapping a system* (3rd ed.). Skoll Centre for Social Entrepreneurship. https://www.mtroyal.ca/nonprofit/InstituteforCommunityProsperity/_pdfs/ssdata_icp_mts_2020.pdf
23. Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., & Kapadia, S. (2020). *Making text count: Economic forecasting using newspaper text* (Bank of England Working Paper No. 865). https://doi.org/10.2139/ssrn.3610770
24. Lagasio, V. (2019). Corporate governance in banks: Systematic literature review and meta-analysis. *Corporate Ownership and Control, 16*(1–1), 113–126. https://doi.org/10.22495/cocv16i1c1art1
25. Lucchese, M. (2020). Country-specific institutional effects on non-financial disclosure level: Evidence from European listed banks. *Corporate Ownership and Control, 17*(4), 166–182. https://doi.org/10.22495/cocv17i4art14
26. Minenna, M., & Aversa, D. (2019). A revised European stability mechanism to realize risk sharing on public debts at market conditions and realign economic cycles in the Euro area. *Economic Notes, 48*(1), Article 12118. https://doi.org/10.1111/ecno.12118
27. Moreno, A.-I., & Caminero, T. (2022). Application of text mining to the analysis of climate-related disclosures. *International Review of Financial Analysis, 83*, Article 102307. https://doi.org/10.1016/j.irfa.2022.102307
28. Network for Greening the Financial System (NGFS). (2022). *Credit ratings and climate change — Challenges for central bank operations.* https://www.ngfs.net/sites/default/files/medias/documents/credit_ratings_and_climate_change_-_challenges_for_central_bank_operations.pdf
29. Pörtner, H.-O., Roberts, D. C., Tignor, M. M. B., Poloczanska, E., Mintenbeck, K., Alegria, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). *Climate change 2022: Impacts, adaptation and vulnerability* (IPCC Sixth Assessment Report). Intergovernmental Panel on Climate Change.https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/
30. Ramani, V. (2020). *Addressing climate as a systemic risk: A call to action for financial regulators.* Ceres. https://www.ceres.org/sites/default/files/reports/2020-06/Financial%20Regulators%20FULL%20FINAL.pdf
31. Reinert, M. (1990). Alceste une méthodologie d'analyse des données textuelles et une application: Aurelia de gerard de nerval. *Bulletin De Méthodologie Sociologique, 26*(1), 24–54. https://doi.org/10.1177/075910639002600103
32. Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., … Tavoni, M. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change, 42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009
33. Rising, J. A., Taylor, C., Ives, M. C., & Ward, R. E. T. (2022). Challenges and innovations in the economic evaluation of the risks of climate change. *Ecological Economics, 197*, Article 107437. https://doi.org/10.1016/j.ecolecon.2022.107437
34. Rogelj, J., Popp, A., Calvin, K. V., Luderer, G., Emmerling, J., Gernaat, D., Fujimori, S., Strefler, J., Hasegawa, T., Marangoni, G., Krey, V., Kriegler, E., Riahi, K., van Vuuren, D. P., Doelman, J., Drouet, L., Edmonds, J., Fricko, O., Harmsen, M., Havlík, P., … Tavoni, M. (2018). Scenarios towards limiting global mean temperature increase below 1.5 C. *Nature Climate Change, 8*(4), 325–332. https://doi.org/10.1038/s41558-018-0091-3
35. SAS. (2019). *Exploring SAS Viya: Visual analytics, statistics, and investigations.* https://support.sas.com/content/dam/SAS/support/en/books/free-books/exploring-sas-viya-va-statistics-investigations.pdf
36. Sokal, R. R., & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *Taxon, 11*(2), 33–40. https://doi.org/10.2307/1217208
37. Souza, M. A. R. d., Wall, M. L., de Morais Chaves Thuler, A. C., Lowen, I. M. V., & Peres, A. M. (2018). O uso do software IRAMUTEQ na análise de dados em pesquisas qualitativas [The use of IRAMUTEQ software for data analysis in qualitative research]. *Revista Da Escola De Enfermagem Da USP, 52*, Article e03353. https://doi.org/10.1590/S1980-220X2017015003353
38. Stern, N., Stiglitz, J., & Taylor, C. (2022). The economics of immense risk, urgent action and radical change: Towards new approaches to the economics of climate change. *The Journal of Economic Methodology, 29*(3), 181–216. https://doi.org/10.1080/1350178X.2022.2040740
39. Stranadko, N. (2022). Global climate governance: Rising trend of translateral cooperation. *International Environmental Agreements: Politics, Law and Economics, 22*, 639–657. https://doi.org/10.1007/s10784-022-09575-6
40. Task Force on Climate-related Financial Disclosures (TCFD). (2017a). *Recommendations of the Task Force on Climate-related Financial Disclosure* (Final report). https://assets.bbhub.io/company/sites/60/2020/10/FINAL-2017-TCFD-Report-11052018.pdf

41. Task Force on Climate-related Financial Disclosures (TCFD). (2017b). *The use of scenario analysis in disclosure of climate-related risks and opportunities* (Technical supplement). https://assets.bbhub.io/company /sites/60/2020/10/FINAL-TCFD-Technical-Supplement-062917.pdf
42. Task Force on Climate-related Financial Disclosures (TCFD). (2019). *2019 Status report.* https://assets.bbhub.io /company/sites/60/2020/10/2019-TCFD-Status-Report-FINAL-053119.pdf
43. Task Force on Climate-related Financial Disclosures (TCFD). (2020a). *Forward-looking financial sector metrics consultation.* https://www.fsb.org/wp-content/uploads/P291020-4.pdf
44. Task Force on Climate-related Financial Disclosures (TCFD). (2020b). *2020 Status report.* https://www.fsb.org /wp-content/uploads/P291020-1.pdf
45. Task Force on Climate-related Financial Disclosures (TCFD). (2021). *2021 Status report.* https://www.fsb.org /wp-content/uploads/P141021-1.pdf
46. Task Force on Climate-related Financial Disclosures (TCFD). (2022). *2022 Status report.* https://assets.bbhub.io /company/sites/60/2022/10/2022-TCFD-Status-Report.pdf
47. The Bank of England. (2022). *Results of the 2021 climate biennial exploratory scenario (CBES).* https://www.bankofengland.co.uk/stress-testing/2022/results-of-the-2021-climate-biennial-exploratory-scenario
48. U.S. Governing Publishing Office Washington. (2021). *Addressing climate as a systemic risk.* U.S. Government Publishing Office.
49. van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T. Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: An overview. *Climatic Change, 109*(1–2), 5–31. https://doi.org/10.1007/s10584-011-0148-z

APPENDIX

A.1. Concepts

Rischio Risk
CLASSIFIER:rischio
CLASSIFIER:rischi

Rischio Fisico Physical risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "fisici")
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "fisico")
CONCEPT_RULE:(DIST_20, "_c{danni}", "fisici")
CONCEPT_RULE:(DIST_20, "_c{danno}", "fisico")

Rischio fisico cronico Chronic physical risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "cronico")
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "cronici")
CONCEPT_RULE:(DIST_20, "_c{1_a_Rischi_Fisici}", "cronico")
CONCEPT_RULE:(DIST_20, "_c{1_a_Rischi_Fisici}", "cronici")

Rischio fisico acuto Acute physical risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "acuto")
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "acuti")
CONCEPT_RULE:(DIST_20, "_c{1_a_Rischi_Fisici}", "acuto")
CONCEPT_RULE:(DIST_20, "_c{1_a_Rischi_Fisici}", "acuti")

Rischio transizione Transition risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "transizione")

Policy and legal risk
CONCEPT_RULE:(DIST_20, "_c{1_b_Rischi_transizione}", "normativo")
CONCEPT_RULE:(DIST_20, "_c{1_b_Rischi_transizione}", "policy")

Technology risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "tecnologico")
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "tecnologici")
CONCEPT_RULE:(DIST_20, "_c{technological}", "risk")
CONCEPT_RULE:(DIST_20, "_c{technology}", "risk")

Market risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "mercato")
CONCEPT_RULE:(DIST_20, "_c{market}", "risk")

Reputation risk
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "reputazionale")
CONCEPT_RULE:(DIST_20, "_c{1_Rischio}", "reputazionali")

Mitigazione Mitigation
CLASSIFIER:mitigazione
CLASSIFIER:mitigare

Adattamento Adaptation
CLASSIFIER:adattamento

Resilienza Resilience
CLASSIFIER:resilience
CLASSIFIER:resilient
CLASSIFIER:resiliente
CLASSIFIER:resilienza

Covid
CLASSIFIER:covid
CLASSIFIER:Covid
CLASSIFIER:COVID

Scenario Scenario
CLASSIFIER:scenario

A.2. Categories

Flag_Ambiente Environment
(OR,"Ambiente","ambiente","Ambientale","ambientale","Environment","environment","Environmental","environmental")

Flag_Clima CLimate
(OR,"clima","Clima","climatico","Climatico","climate","Climate")

Flag_Covid
(OR,"covid","Covid","COVID")

Flag_Emission
(OR,"Emission","emission","Emissions","emissions")

Flag_Mitigazione_Adattamento Mitigation, Adaptation
(OR,"mitigazione","mitigare","adattamento","adapt")

Flag_Resilienza Flag resilience
(OR, "resilienza","resiliente","resilience","resilient")

Flag_Rischio Flag risk
(OR, "rischio","rischi","risk")

Flag_Rischio_Fisico Flag physical risk
(AND,(OR,"danni","danno","rischio","rischi"),(OR,"fisici","fisico"))

Flag_Rischio_Fisico_Acuto (tradurre i seguenti come i precedenti)
(AND,(OR,"danni","danno","rischio","rischi"),(OR,"acuti","acuto"))

Flag_Rischio_Fisico_Cronico
(AND,(OR,"danni","danno","rischio","rischi"),(OR,"cronici","cronico"))

Flag_Rischio_Transizione
(AND,(OR,"rischio","rischi"), (OR,"transizione"))

Flag_Rischio_Transizione_Mercato
(OR, (AND, (AND,(OR,"rischio","rischi"),"transizione"), "mercati"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "mercato"))

Flag_Rischio_Transizione_Policy
(OR, (AND, (AND,(OR,"rischio","rischi"),"transizione"), "policy"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "normativo"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "legale"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "legali"))

Flag_Rischio_Transizione_Reput
(OR, (AND, (AND,(OR,"rischio","rischi"),"transizione"), "reputazionali"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "reputazionale"))

Flag_Rischio_Transizione_Tech
(OR, (AND, (AND,(OR,"rischio","rischi"),"transizione"), "tecnologici"), (AND, (AND,(OR,"rischio","rischi"), "transizione"), "tecnologico"))

Flag_Scenario
(OR, "scenario", "scenari")