



**UNIVERSITÀ  
DI TORINO**

## **Università degli Studi di Torino**

*Corso di Laurea in Innovazione sociale, Comunicazione e  
Nuove tecnologie*

### **Titolo**

**Social Media e Cambiamento Climatico: Topic  
Modeling e Sentiment Analysis di Discussioni Online  
sull'Ambiente**

**Relatore/Relatrice**

Mensa Enrico

**Correlatore/Correlatrice**

Villata Sabrina

**Candidato/a**

**Cognome Nome**

Matricola 952248

Lazzarotto Francesco

Anno Accademico 2023/2024

## **ABSTRACT**

Con l'aumento degli eventi meteorologici estremi e dei disastri ambientali, cresce sempre di più l'attenzione verso le problematiche legate al clima, mettendo in evidenza l'importanza di comprendere l'opinione pubblica e le dinamiche discorsive legate a questa tematica, diventata ormai cruciale. Questo studio mira a esaminare le opinioni relative al cambiamento climatico attraverso tecniche di text-mining applicate a testi provenienti da social media come Reddit e Twitter. La prima fase dell'analisi consiste nell'identificare e raccogliere i dati disponibili in rete, selezionando dataset e testi pertinenti all'obiettivo di ricerca. Una volta raccolti, i testi vengono elaborati e preparati per l'analisi attraverso procedure di pulizia e normalizzazione. Segue una parte di Topic Modeling mirata a scoprire i principali argomenti trattati e le parole chiave ad essi collegate. Infine, l'analisi del sentiment permette di contestualizzare le emozioni associate ai testi e alle opinioni espresse. Comprendere gli argomenti trattati e le percezioni del pubblico riguardo al cambiamento climatico è quindi d'aiuto per fornire un quadro informativo più chiaro e approfondito, contribuendo così alla consapevolezza e a sviluppare possibili strategie di azione per affrontare questa sfida, ora più che mai, globale.

# Indice

<b>1</b>	<b>Introduzione</b>	<b>7</b>
1.1	Introduzione . . . . .	7
1.2	Domande e principi della ricerca . . . . .	8
1.3	Struttura della ricerca . . . . .	10
<b>2</b>	<b>Stato dell'arte</b>	<b>12</b>
<b>3</b>	<b>NLP e Text-Mining</b>	<b>17</b>
3.1	Elaborazione del linguaggio naturale . . . . .	17
3.1.1	Storia . . . . .	17
3.1.2	Il text-mining . . . . .	18
3.1.3	Passaggi ed estrazione di informazioni . . . . .	19
3.1.4	Applicazioni pratiche . . . . .	25
<b>4</b>	<b>Metodologia</b>	<b>27</b>
4.1	Strumenti utilizzati . . . . .	27
4.1.1	Linguaggio di programmazione e librerie . . . . .	27
4.1.2	Hardware . . . . .	28
4.2	Raccolta dei dati e dataset . . . . .	28
4.3	Pre-processing dei dati . . . . .	29
4.4	Topic Modeling . . . . .	32
4.5	Modello BERTopic . . . . .	34
4.6	Applicazione pratica . . . . .	35
4.6.1	Creazione degli embeddings . . . . .	35
4.6.2	Clustering . . . . .	35
4.6.3	UMAP . . . . .	36
4.6.4	HDBSCAN . . . . .	36
4.6.5	c-TF-IDF . . . . .	37
4.6.6	Addestramento del modello . . . . .	38
4.6.7	Valutazione modello . . . . .	39
4.7	Analisi del sentiment . . . . .	41
4.7.1	Distribuzione sentiment . . . . .	41
4.7.2	Parole chiave per polarità . . . . .	42
4.7.3	Analisi basata su elenchi di parole . . . . .	44
4.7.4	Analisi delle emozioni . . . . .	45
4.7.5	Analisi delle emozioni per topic . . . . .	46
4.8	Data Visualization . . . . .	48

<b>5 Risultati e Discussione</b>	<b>51</b>
5.1 Analisi dei topic . . . . .	52
5.1.1 Cambiamento Climatico e Impatto Umano . . . . .	59
5.1.2 Cooperazione Internazionale e Clima . . . . .	62
5.1.3 Cambiamento Climatico e Politica . . . . .	65
5.1.4 Inquinamento ed Ecosistemi . . . . .	68
5.1.5 Crisi, Sanità e Sicurezza pubblica . . . . .	73
5.1.6 Industria, Inquinamento e Impatto Ambientale . . . . .	77
5.1.7 Riscaldamento Globale . . . . .	81
5.1.8 Il Cambiamento Climatico fra Bufala e Realtà . . . . .	86
5.1.9 Impatto e Controversie della Società Exxonmobil . . . . .	91
5.1.10 Consumo di Carne, Allevamenti e Impatto . . . . .	96
5.2 Analisi del sentimento . . . . .	101
5.2.1 Distribuzione e parole chiave . . . . .	101
5.2.2 Analisi delle emozioni complessiva . . . . .	105
5.2.3 Analisi delle emozioni per topic . . . . .	108
<b>6 Conclusioni</b>	<b>111</b>
6.1 Approccio innovativo . . . . .	111
6.2 Limitazioni e aree di miglioramento . . . . .	111
6.3 Conclusioni e risultati chiave . . . . .	112

# Elenco delle figure

1.1	Procedura di ricerca . . . . .	10
2.1	Credenza Pubblica " <i>Peoples' climate vote - fig.2</i> " [11] . . . . .	12
2.2	Andamento della consapevolezza generale. [7] . . . . .	13
3.1	Processo Text-mining . . . . .	19
3.2	Bag of Words . . . . .	21
3.3	Formule metriche di misura [29] . . . . .	23
3.4	Silhouette Score [30] . . . . .	24
3.5	Applicazione pratiche del text-mining . . . . .	25
4.1	Architettura completa dei transformatori(Fig 1 e 2 in Vaswani, et al., 2017.) [25] . . . . .	32
4.2	Funzione softmax [44] . . . . .	33
4.3	Processo di BERTopic [45] . . . . .	34
4.4	Clustering [46] . . . . .	36
4.5	c-TF-IDF [45] . . . . .	37
4.6	bm25 [45] . . . . .	38
4.7	Ruota delle emozioni [50] . . . . .	44
5.1	Rete semantica . . . . .	51
5.2	Distribuzione dei documenti per topic . . . . .	53
5.3	Cloudword delle 1000 parole più frequenti . . . . .	55
5.4	Grafico a dispersione di similarità. . . . .	58
5.5	Deviazione livelli di CO <sub>2</sub> nell'atmosfera . . . . .	59
5.6	Livelli di anidride carbonica . . . . .	60
5.7	Keyword-Peso Cambiamento Climatico e impatto umano . . . . .	61
5.8	Urgenza di risposta da parte delle istituzioni tra le persone che credono nell'emergenza climatica - Fig. 4 IPCC [11] . . . . .	63
5.9	Urgenza di risposta da parte delle istituzioni tra le persone che credono nell'emergenza climatica per gruppi di paesi - Fig 5 IPCC [11] . . . . .	63
5.10	Keyword-Peso Cooperazione internazionale e Clima . . . . .	64
5.11	Keyword-Peso Cambiamento climatico e Politica . . . . .	66
5.12	Il cambiamento climatico deve essere una priorità primaria per il Presidente e il Congresso. - Global Warming as a Priority Issue: Political Trends [61] . . . . .	67
5.13	Produzione globale di plastica [65] . . . . .	69
5.14	Livello del mare [66] . . . . .	69
5.15	Livelli di pH nell'oceano per anno [70] . . . . .	71
5.16	Keyword-Peso Inquinamento ed Ecosistemi . . . . .	72

5.17 Percentuale di persone che credono, che a lungo termine, il cambiamento climatico sia una crisi seria quanto quella da Covid 19 [71] . . . . .	74
5.18 Impatto dei pericoli del cambiamento climatico sulle malattie [72] . . . . .	75
5.19 Keyword-Peso Crisi, Sanità e sicurezza pubblica . . . . .	76
5.20 Emissioni globale per settore [73] . . . . .	78
5.21 Keyword-Peso Industria inquinamento e impatto ambientale . . . . .	79
5.22 Capacità energia solare [75] . . . . .	80
5.23 Anomalie della temperatura globale terrestre e oceanica su base annuale [76] . . . . .	82
5.24 Influenze umane e naturali sulla temperatura globale [79] . . . . .	83
5.25 Keyword-Peso Riscaldamento Globale . . . . .	85
5.26 Dove vivono i negazionisti del cambiamento climatico [85] . . . . .	87
5.27 Indice di preoccupazione per il cambiamento climatico per affiliazione religiosa [86] . . . . .	88
5.28 Keyword-Peso Cambiamento climatico fra bufala e realtà . . . . .	90
5.29 Confronto osservazioni e proiezioni di temperatura [88] . . . . .	91
5.30 Confronto emissioni di gas serra 2009-2012 . . . . .	93
5.31 Keyword-peso Impatto e controversie Exxonmobil . . . . .	94
5.32 Emissione di gas serra per chilogrammo di cibo prodotto [94] . . . . .	97
5.33 Emissioni di gas serra nella produzione di cibo [95] . . . . .	98
5.34 Keyword-Peso Consumo di carne, allevamenti e impatto . . . . .	99
5.35 Grafico a Torta distribuzione sentiment . . . . .	102
5.36 20 Parole più comuni per ogni polarità . . . . .	103
5.37 Word Cloud delle parole più frequenti per ogni polarità . . . . .	104
5.38 Grafico a torta della distribuzione delle emozioni nel dataset . . . . .	105
5.39 Conteggio testi per emozione . . . . .	106
5.40 Distribuzione delle emozioni per polarità . . . . .	107
5.41 Distribuzione delle emozioni per topic . . . . .	108



# Capitolo 1

## Introduzione

### 1.1 Introduzione

Negli ultimi decenni, il mondo ha assistito ad un aumento senza precedenti degli eventi climatici estremi e dei disastri ambientali. Eventi recenti come la temperatura più calda mai registrata al Polo Antartico, gli incendi australiani e californiani, l'incremento delle temperature medie, l'innalzamento del livello del mare e le forti siccità presenti in diverse aree del pianeta hanno portato ad una crescente attenzione sulla crisi climatica. Riconoscendo l'entità del fenomeno, numerosi studi si sono prestati ad esaminare i fattori che contribuiscono a questo cambiamento. Il report annuale dell'Intergovernmental Panel on Climate Change's (IPCC) [1], riporta diverse statistiche e i dati che sottolineano l'urgenza del problema. Inoltre, diverse ricerche sul cambiamento climatico indicano che questi eventi meteorologici, siano essi naturali o antropologici, tendono a innescare la discussione sulla crisi climatica in atto [2, 3]. È, oltre quanto già detto, essenziale considerare gli impatti socio-economici derivanti dai cambiamenti climatici. Le regioni più vulnerabili, spesso situate nei paesi in via di sviluppo, sono quelle che subiscono le conseguenze più gravi, includendo perdita di mezzi di sussistenza, insicurezza alimentare e migrazioni forzate. Uno studio del World Bank ha evidenziato che, senza interventi adeguati, il cambiamento climatico potrebbe spingere oltre 100 milioni di persone in condizioni di estrema povertà entro il 2030 [4]. Questo sottolinea il carattere globale di questa sfida in virtù delle sue ramificazioni multidimensionali e degli impatti su scala planetaria. Tale fenomeno presenta quindi una natura intrinsecamente interconnessa e interdisciplinare, richiedendo un approccio olistico e collaborativo per affrontare le sue complesse implicazioni e per garantire la sostenibilità delle risorse ambientali per le generazioni presenti e future.

Con l'avvento dell'era digitale, i social media sono diventati un luogo di scambio di opinioni e di dibattito pubblico [5], innescando la necessità di sviluppare tecnologie in grado di comprendere in maniera sufficientemente precisa le dinamiche di questi scambi. In questo scenario, diventa fondamentale capire come la tecnologia possa aiutare a risolvere queste problematiche e come possa la ricerca comprendere e incoraggiare la consapevolezza delle persone. *L'elaborazione del linguaggio naturale*, in inglese *Natural Language Processing* (NLP), è un campo dell'*intelligenza artificiale* (AI) e della *linguistica* che si occupa dell'interazione tra computer e linguaggio umano. Coinvolge lo sviluppo di algoritmi e modelli per consentire ai computer di comprendere, interpretare e generare il linguaggio umano in modo significativo e

---

contestualmente appropriato. Il *Text-Mining*, sotto-branca dell’NLP volta a estrarre informazioni utili dai dati linguistici, può perciò avere un ruolo importante nel comprendere l’opinione pubblica, le percezioni e i comportamenti linguistici legati al cambiamento climatico. Questo studio vuole quindi esplorare i principali argomenti d’interesse delle persone, cercando di estrarre informazioni da un corpus di testi attraverso due fra le principali tecniche di *text-mining*: il *Topic Modeling* e la *Sentiment Analysis*.

## 1.2 Domande e principi della ricerca

Delle domande di ricerca ben definite sono fondamentali nell’ambito della ricerca scientifica poiché forniscono un quadro chiaro e strutturato per l’indagine e l’analisi di un determinato argomento [6]. Esse fungono da guida per l’intero processo di ricerca, orientando la raccolta e l’interpretazione dei dati e contribuendo così a raggiungere gli obiettivi prefissati. In particolare, nel contesto di questo studio, le domande di ricerca diventano fondamentali per individuare i principali temi e argomenti trattati, valutare la loro frequenza e distribuzione, esplorare le relazioni tra di essi e comprendere le emozioni e il sentimento associati. Queste domande non sono statiche, ma possono essere plasmate e riformulate in base al contesto di ricerca e ai risultati ottenuti durante l’analisi dei dati. Ad esempio, l’individuazione di nuovi pattern o relazioni durante l’analisi potrebbe suggerire la necessità di aggiungere domande supplementari o di modificare quelle esistenti per approfondire ulteriormente determinati aspetti dell’argomento. Infine, le domande di ricerca possono essere influenzate dai risultati emersi durante lo studio, adattandosi di conseguenza per rispondere in modo più accurato alle sfide e alle questioni che emergono durante il processo di ricerca. Il contesto dello studio, che si focalizza sull’analisi dei social media come Reddit e Twitter, rende ancora più cruciale lo sviluppo efficace di tali domande di ricerca. La natura dinamica e la vastità dei contenuti generati dagli utenti su queste piattaforme richiedono una formulazione precisa e rigorosa delle domande di ricerca per garantire un’analisi accurata e pertinente. Nella presente ricerca, l’argomento principale è il cambiamento climatico. Tuttavia, vengono studiati diversi temi specifici all’interno di questo vasto argomento. Ad esempio, quando si parla di frequenza dei temi, ci si riferisce alla frequenza dei vari sotto-argomenti specifici che emergono nelle discussioni online sul cambiamento climatico.

- **Quali sono i principali argomenti o temi discussi riguardo al cambiamento climatico sui social media?**

Questa domanda mira a identificare i temi prevalenti nelle discussioni sul cambiamento climatico. L’obiettivo è comprendere quali aspetti del cambiamento climatico attraggono maggiormente l’attenzione del pubblico e quali temi vengono trattati più frequentemente. Il cambiamento climatico rappresenta l’argomento principale, ma vengono esplorati diversi sotto-argomenti, cioè i topic, all’interno di questo vasto ambito. Ad esempio, si possono individuare discussioni su eventi meteorologici estremi, politiche ambientali, innovazioni tecnologiche, energie rinnovabili e movimenti sociali per l’ambiente. Quando parliamo di frequenza dei temi, ci riferiamo alla frequenza con cui questi vari temi emergono nelle discussioni online relative al cambiamento climatico.

- 
- **Con quale frequenza e distribuzione vengono discussi questi argomenti all'interno dei tweet e i post reddit?**

Questa domanda cerca di quantificare la frequenza e la distribuzione delle discussioni sugli argomenti identificati. Si analizzano i dati per determinare quanto spesso vengono trattati specifici temi e come questa frequenza varia.

- **Come sono correlati gli argomenti tra loro e quali sono le relazioni tra i diversi temi?**

Questa domanda esplora le relazioni e le correlazioni tra i diversi temi discussi. Utilizzando tecniche di analisi della similarità e delle reti, è possibile individuare come diversi argomenti si intrecciano e si influenzano reciprocamente. Ad esempio, si potrebbe scoprire che le discussioni sulle politiche ambientali sono spesso correlate con quelle sull'innovazione tecnologica, suggerendo una connessione percepita tra regolamentazioni governative e soluzioni tecniche.

- **Quali sono le parole chiave più frequentemente associate a ciascun argomento?**

Questa domanda identifica le parole chiave che emergono con maggiore frequenza in relazione a ciascun argomento. Questo permette di creare un vocabolario tematico specifico e di capire meglio il contenuto e l'orientamento delle discussioni.

- **Quali sono e dove si trovano le discrepanze tra l'importanza percepita per certe problematiche espresse sui social e la loro rilevanza scientifica?**

Questa domanda mira a identificare eventuali disallineamenti tra la percezione pubblica delle problematiche legate al cambiamento climatico e la loro effettiva rilevanza scientifica. Utilizzando dati climatici e scientifici, si confronta la frequenza e l'intensità delle discussioni pubbliche con l'importanza attribuita dagli esperti. Ad esempio, potrebbe emergere che la plastica nei mari è un tema molto discusso pubblicamente, mentre la fusione dei ghiacciai, sebbene scientificamente più critica, riceve meno attenzione.

- **Quali sono le emozioni associate ai tweet e ai post e ai rispettivi argomenti e con quale distribuzione si presentano?**

Questa domanda esplora le emozioni espresse nei testi in relazione ai vari argomenti. Utilizzando tecniche di analisi del sentimento, si possono identificare le emozioni predominanti, come paura, speranza, rabbia o tristezza, e osservare come queste emozioni si distribuiscono tra i diversi argomenti. Questo aiuta a comprendere meglio la risposta emotiva del pubblico alle varie tematiche climatiche.

- **Qual è la distribuzione del sentimento all'interno dei testi e quali sono le parole chiave associate a ogni polarità?**

Questa domanda analizza la distribuzione del sentimento (positivo, negativo, neutro) nei testi e identifica le parole chiave associate a ciascuna polarità del sentimento. Questo permette di comprendere quali termini sono più spesso utilizzati in contesti positivi o negativi e come le opinioni del pubblico si distribuiscono rispetto ai temi trattati.

## 1.3 Struttura della ricerca

La struttura dello studio, rappresentato in Fig. 1.1, è la seguente: nel Capitolo 1, ossia l'introduzione alla ricerca, sono stati presentati i concetti fondamentali e il contesto della ricerca, fornendo una panoramica delle questioni principali che verranno approfondite nei capitoli successivi. Nel Capitolo 2 viene illustrata una revisione della letteratura per contestualizzare il lavoro nello scenario accademico e scientifico attuale. Il Capitolo 3 si compone di una panoramica della materia, passando dalla storia ai principali campi di utilizzo. Nel Capitolo 4 vengono descritte nel dettaglio le metodologie e gli strumenti utilizzati. Il Capitolo 5 è composto dall'analisi dei risultati ottenuti attraverso l'applicazione delle tecniche di Topic Modeling e Sentiment Analysis ai testi raccolti. Vengono presentati e discussi i principali temi emersi, insieme alle emozioni e alle variazioni di sentimento riscontrate nei testi analizzati e dove possibile, verrà presentata una correlazione con i dati climatici. Per concludere, con il Capitolo 6, vengono fornite alcune considerazioni finali, che sintetizzano i principali risultati ottenuti e discutono le implicazioni e i limiti della ricerca, insieme a possibili direzioni future per lo sviluppo di ulteriori studi in questo campo.

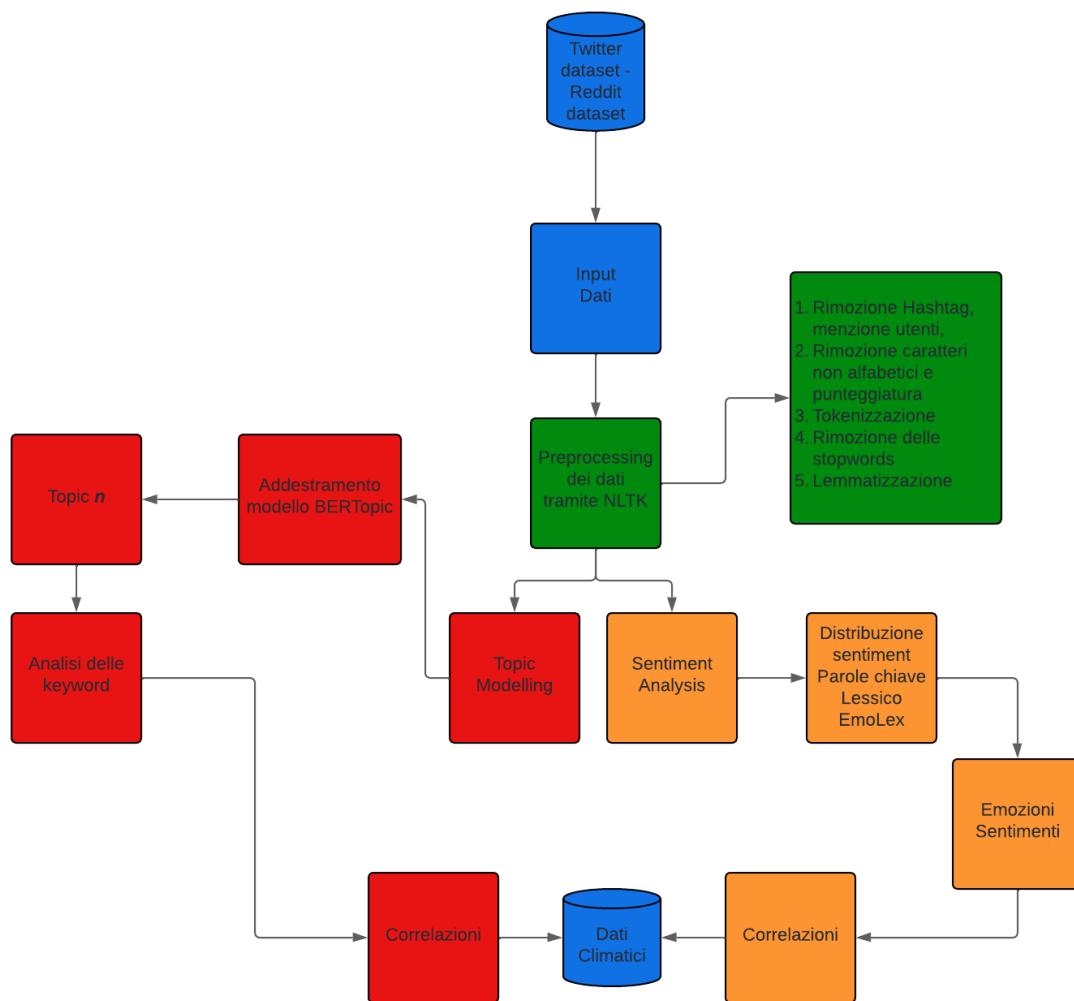


Figura 1.1: Procedura di ricerca



# Capitolo 2

## Stato dell'arte

L'opinione pubblica sul cambiamento climatico è influenzata da una vasta gamma di fattori, tra cui quelli sociodemografici, politici, culturali, economici e ambientali. A livello internazionale, la maggioranza della popolazione considera la crisi climatica un'emergenza [7]. In questo contesto, i sondaggi rappresentano uno strumento utile per comprendere le considerazioni sul clima, arricchendo il complesso quadro dell'opinione pubblica e contribuendo ad aumentare l'impegno politico verso azioni più concrete per affrontare questa sfida globale. Dal 1980 sono stati condotti numerosi sondaggi per valutare l'opinione pubblica, inizialmente focalizzati sulla consapevolezza, e, successivamente, sulle sfaccettature del cambiamento climatico e sulle possibili risposte da parte della politica [8, 9, 10]. Negli ultimi anni, i sondaggi a livello globale forniscono dati sempre più accurati, ad esempio, nel Gennaio 2021, il Programma di sviluppo delle Nazioni Unite (UNDP) ha condotto il *Peoples' Climate Vote* [11], la più ampia indagine sul clima mai realizzata, coinvolgendo 1,2 milioni di persone in 50 paesi. L'obiettivo principale del sondaggio è raccogliere una vasta gamma di opinioni e prospettive sulla crisi climatica da parte delle persone di tutto il mondo. Attraverso questo sondaggio, l'UNDP cerca di valutare il grado di consapevolezza della popolazione mondiale riguardo al cambiamento climatico, nonché il livello di preoccupazione e l'urgenza per affrontare la questione. I risultati hanno rivelato che in media, come mostrato dalla Figura 2.1, il 64% degli intervistati considera il cambiamento climatico come un'emergenza, e che la conservazione delle foreste e del territorio rappresentano le soluzioni più popolari.

**Figure 2. Public Belief in the Climate Emergency, by Region**

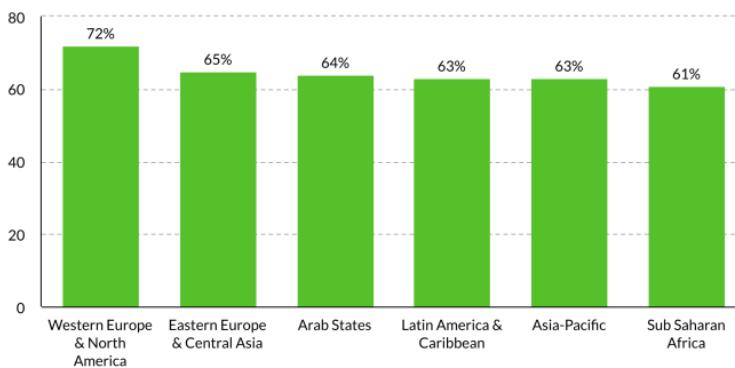


Figura 2.1: Credenza Pubblica "Peoples' climate vote - fig.2" [11]

Secondo un articolo del 2015 basato su una revisione della letteratura che ha esaminato migliaia di articoli provenienti da oltre duecento studi condotti tra il 1980 e il 2014, si è osservato un aumento della consapevolezza pubblica sui cambiamenti climatici negli anni '80 e '90, seguito da un periodo di crescente preoccupazione, accompagnato da posizioni contrastanti, alla fine degli anni '90 e all'inizio degli anni 2000 [12]. Successivamente, si è verificato un periodo di diminuzione della preoccupazione pubblica e crescente scetticismo tra la metà e la fine degli anni 2000. Dal 2010 al 2014, viene mostrata una possibile stabilizzazione della preoccupazione pubblica, dovuta anche all'aumento delle politiche ambientali e alle campagne pubbliche di consapevolezza.

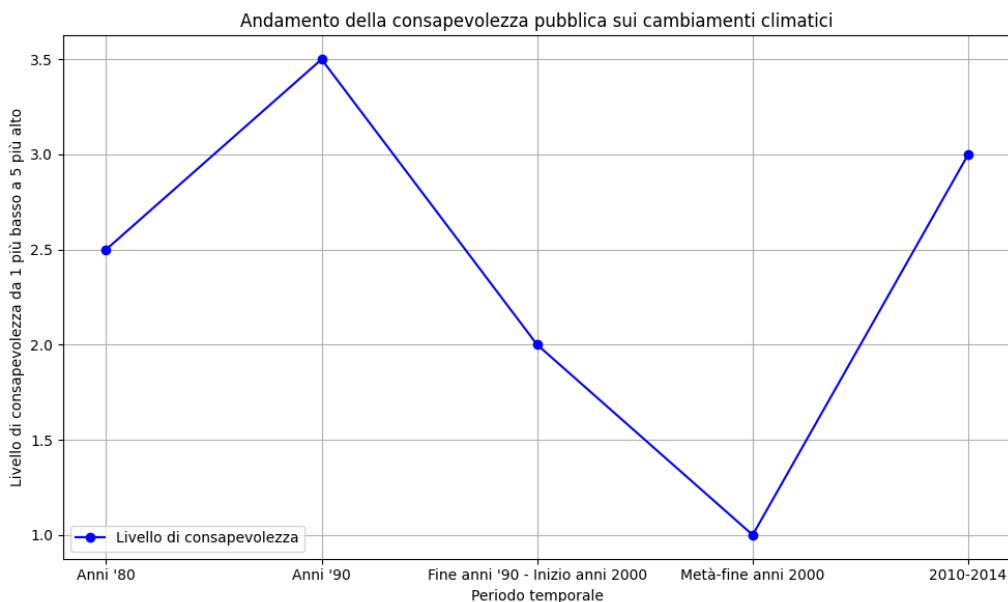


Figura 2.2: Andamento della consapevolezza generale. [7]

Dai risultati dei sondaggi emerge chiaramente l'importanza della consapevolezza pubblica e di come venga espressa l'opinione pubblica sui cambiamenti climatici. Tuttavia, è importante riconoscere che i sondaggi hanno dei limiti e potrebbero non catturare completamente tutte le sfumature di opinione e percezione. Alcune fra le problematiche maggiori associate ai sondaggi sono [13]:

- *Selezione del campione non rappresentativo:* se il campione non è rappresentativo della popolazione di interesse, i risultati del sondaggio potrebbero essere distorti e non riflettere accuratamente le opinioni della popolazione nel loro insieme.
- *Bias di risposta:* alcune persone potrebbero essere più propense a rispondere ai sondaggi rispetto ad altre, introducendo un bias nei risultati. Questo può essere dovuto alle loro forti opinioni sull'argomento o al contrario all'indifferenza verso di esso.
- *Errore di misurazione:* gli errori di misurazione possono verificarsi quando le risposte dei partecipanti non riflettono accuratamente le loro opinioni o comportamenti reali, a causa di fraintendimenti delle domande o risposte socialmente desiderabili.

- 
- *Effetto dell'intervistatore*: in alcuni tipi di sondaggi, l'effetto dell'intervistatore può influenzare le risposte dei partecipanti attraverso il tono della voce, l'espressione facciale o il linguaggio del corpo, introducendo un bias nei risultati. Nei sondaggi telematici questa problematica si può riflettere sull'ente che effettua la ricerca.
  - *Cambiamenti nel tempo delle opinioni e dei comportamenti*: le opinioni e i comportamenti delle persone possono cambiare nel tempo, rendendo difficile ottenere una misurazione accurata e aggiornata attraverso i sondaggi.

Osservando i limiti dei sondaggi può venir d'aiuto l'elaborazione del linguaggio naturale, attraverso questo campo di ricerca, è possibile esaminare grandi quantità di dati testuali provenienti dai social media per identificare tendenze, opinioni e sentimenti nascosti che potrebbero non emergere dai sondaggi tradizionali. Si pone quindi come uno strumento utile ad arricchire la comprensione delle modalità con la quale questo argomento è percepito e discusso, in particolar modo si differenza dei sondaggi tradizionali, che spesso si basano su risposte predefinite e campionamenti limitati, poichè consente di accedere a una vasta gamma di espressioni spontanee e autentiche degli utenti online. Questo approccio permette di cogliere in tempo reale le dinamiche delle percezioni pubbliche e di identificare temi emergenti che potrebbero non essere stati previsti. Inoltre, la granularità dei dati analizzati dall'NLP permette di segmentare le opinioni in base a variabili demografiche, geografiche e temporali, fornendo una comprensione più dettagliata e mirata dei sentimenti collettivi. Con l'avvento dei social media, essi sono diventati uno dei luoghi principali di scambio delle opinioni e possono essere utilizzati come fonte di dati per analizzare il pensiero pubblico da un punto diverso rispetto quello che viene tratto dai sondaggi. Nel contesto dell'analisi del cambiamento climatico e delle dinamiche sociali ad esso collegate, diverse ricerche precedenti hanno contribuito a delineare il panorama complesso delle opinioni pubbliche, delle discussioni online e delle reazioni alle politiche ambientali. Diversi studi si concentrano sull'esame delle percezioni pubbliche riguardanti il cambiamento climatico, utilizzando l'analisi dei dati provenienti da varie piattaforme di social media. Queste ricerche condividono l'obiettivo di identificare i principali temi di discussione e le emozioni prevalenti espresse dagli utenti in relazione al cambiamento climatico. Spesso, l'analisi si focalizza su piattaforme online come Twitter, che forniscono un ampio campione di dati testuali spontanei. Gli studi tendono a utilizzare metodologie simili, impiegando tecniche avanzate di elaborazione del linguaggio naturale (NLP) per estrarre e classificare i temi e le emozioni dai dati testuali. L'approccio generale di queste ricerche consente di comprendere meglio le preoccupazioni e le priorità del pubblico riguardo al cambiamento climatico nel loro complesso. Per esempio, la Latent Dirichlet Allocation (LDA), una tecnica di identificazione dei topic utilizzata in ambito NLP, viene spesso utilizzata per effettuare un'analisi dell'opinione pubblica sul cambiamento climatico [14, 15]. Nello specifico l'algoritmo assegna casualmente una pluralità di topic a ciascun documento e successivamente riassegna iterativamente i topic in base alla distribuzione delle parole nei vari topic. Altre metodologie spesso utilizzate per effettuare questa tipologia di analisi si basano sul modello Bidirectional Encoder Representations from Transformers (BERT), verrà successivamente approfondito poichè è il modello sulla quale si basa BERTopic, la tecnica di topic modeling utilizzata nello studio [16]. Al contrario delle analisi che hanno alla loro base un approccio più genera-

---

le sono presenti altri studi che si compongono di analisi più specifiche che mirano ad analizzare da più vicino il rapporto fra cambiamento climatico e altri argomenti esterni ma connessi a quello principale, come per esempio le elezioni politiche, l'inquinamento industriale e le politiche cosiddette verdi [17, 18]. *Effrosynidis et al.* [19], hanno influenzato in modo significativo la definizione degli obiettivi dello studio, fornendo un quadro dettagliato delle tematiche e delle metodologie pertinenti da esplorare. Utilizzando un vasto corpus di dati provenienti da Twitter, cioè un insieme strutturato di testi linguistici o dati linguisticamente annotati che sono raccolti e conservati per scopi di analisi linguistica, questo studio esplora le differenze comportamentali e di percezione tra i negazionisti e coloro che credono al cambiamento climatico, nonché la correlazione tra il sentimento umano e le variazioni dalla temperatura storica. È proprio questo approccio di ricerca delle correlazioni che ha ispirato la formulazione delle domande di ricerca, orientate a esplorare non solo i temi principali discussi, ma anche le interconnessioni tra essi e i cambiamenti ambientali. Per concludere, tramite varie metodologie e approcci, queste pubblicazioni contribuiscono alla costruzione di un quadro più completo e articolato della situazione scientifica riguardante l'atteggiamento e le percezioni della popolazione nei confronti di questo tema. L'eterogeneità delle metodologie adottate, insieme alla diversità delle popolazioni e dei contesti considerati, arricchisce il dibattito scientifico e offre una panoramica più ricca e dettagliata delle dinamiche sociali e culturali correlate al cambiamento climatico.



# Capitolo 3

## NLP e Text-Mining

### 3.1 Elaborazione del linguaggio naturale

L'elaborazione del linguaggio naturale è una disciplina che si situa al crocevia tra linguistica, informatica e intelligenza artificiale. Si concentra sul capire come creare modelli informatici per gestire e processare grandi quantità di dati sotto forma di linguaggio naturale. Lo scopo di questa disciplina è rendere la tecnologia in grado di comprendere in modo più approfondito il contenuto dei documenti analizzati, in modo tale che possa estrarre informazioni, classificare e categorizzare il linguaggio umano [20].

#### 3.1.1 Storia

Nonostante l'NLP sia entrato nel panorama informatico e di ricerca solamente negli ultimi decenni, la sua storia pone le radici nella seconda metà del Novecento. Questo campo infatti nasce nei primi anni '50 con la pubblicazione da parte di Alan Turing dell'articolo: *"Computing Machinery and Intelligence"* [21, 22]. Nasce in questi anni il famoso "Test di Turing" usato come criterio per determinare se una macchina possa pensare o no. Il test mira a testare la capacità di una macchina di essere percepita come umana nel corso di una conversazione. Nel 1954, l'esperimento di traduzione automatica Georgetown-IBM segnò un punto di svolta per l'NLP, quando un computer IBM tradusse con successo oltre 60 frasi dal russo all'inglese [23]. In questo periodo, l'approccio alla NLP era basato su regole e logica e i programmi venivano scritti per eseguire specifiche operazioni di elaborazione del linguaggio naturale, come la traduzione di frasi da una lingua all'altra [24]. I primi approcci dell'NLP erano per lo più rule-based, basati cioè su regole formalizzate che venivano date ai computer sotto forma di istruzioni. A questo approccio, troppo rigido e poco adatto a gestire la grande ambiguità e variabilità del linguaggio umano, è stato sostituito, a partire soprattutto dagli anni '90, un approccio di NLP basato sulla statistica o su corpora [22]. L'approccio statistico permette alla macchina di 'imparare' autonomamente sulla base delle strutture che incontra in corpora di grandi dimensioni, permettendo una maggiore robustezza, generalizzazione e gestione dell'ambiguità. Parallelamente, l'uso di corpora implica l'analisi di grandi raccolte di testi autentici per trarre inferenze linguistiche, consentendo una comprensione più realistica e contestuale del linguaggio. Nei primi anni '2000 si ha poi l'ascesa del *Deep Learning* e delle *Reti Neurali*, Lo sviluppo di mo-

---

delli come le Reti Neurali Convoluzionali (CNN) e le Reti Neurali Ricorrenti (RNN) sottolinea l'importanza dell'apprendimento automatico. Le CNN, originariamente sviluppate per il riconoscimento di immagini, sono in grado di identificare pattern complessi attraverso l'uso di convoluzioni, che permettono di estrarre caratteristiche rilevanti da grandi quantità di dati visivi. Questo approccio è stato adattato anche per l'elaborazione del linguaggio naturale, migliorando compiti come la classificazione di testo e l'analisi del sentimento. D'altra parte, le RNN sono progettate per gestire dati sequenziali, come il testo o il parlato, poiché mantengono una memoria dei passaggi precedenti nella sequenza. Questo le rende particolarmente efficaci per compiti come la traduzione automatica, il riconoscimento vocale e la generazione di testo. Entrambi questi tipi di reti neurali dimostrano come l'apprendimento automatico possa affrontare la complessità e la variabilità del linguaggio umano, offrendo soluzioni più flessibili e adattabili rispetto agli approcci tradizionali. Oggi, l'evoluzione di questo campo continua a progredire, grazie a modelli di linguaggio basati su *transformer* [25], come Bidirectional Encoder Representations from Transformers (BERT) e Generative Pre-trained Transformer (GPT). Questo nuovo tipo di modello di Deep Learning permette di concentrarsi su diverse parti di una sequenza di input contemporaneamente e di pesare le diverse parti dell'input dandone un'importanza, elaborando successivamente le diverse sequenze in modo parallelo [25]. Concludendo, la tendenza verso il transfer learning e la condivisione di modelli volti ad analisi basate su fine-tuning dei modelli stessi, ha reso possibile il trasferimento di conoscenze, consentendo agli sviluppatori di creare soluzioni NLP altamente performanti con meno risorse computazionali e dati di addestramento. L'NLP continua quindi ad acquisire importante crescita in molti settori, apre nuove frontiere nelle applicazioni pratiche e nella comprensione dell'interazione tra linguaggio umano e tecnologia. Con possibili ulteriori innovazioni e scoperte in arrivo, il futuro dell'NLP promette di essere ancora più entusiasmante e pieno di opportunità.

### 3.1.2 Il text-mining

Il text-mining è una sotto-branca dell'elaborazione del linguaggio naturale che ha come obiettivo l'analisi e l'estrazione di informazioni utili da grandi quantità di testo non strutturato, trasformandolo in un formato strutturato utilizzabile per applicare, successivamente, tecniche analitiche e di apprendimento automatico. I dati strutturati sono organizzati in un formato predefinito, come tabelle di database, dove le informazioni sono disposte in righe e colonne con campi ben definiti. Questo tipo di dati è facilmente interpretabile da algoritmi e strumenti di analisi. D'altra parte, i dati non strutturati, come documenti di testo, immagini, video, e post sui social media, non seguono uno schema rigido e richiedono processi più complessi di elaborazione per estrarre e organizzare le informazioni rilevanti in modo che possano essere analizzate efficacemente. Il text mining è definito da *Palwinder Kaur Manigate e Dr. Kamaljit Singh Saini* come:

*"Text mining is the data mining technique or process which discovers earlier unfamiliar and valuable information from a huge quantity of unstructured text data. This knowledge is then analyzed and processed for operators, so they can receive valid knowledge"*[26]

Il text-mining è quindi un'applicazione pratica dell'elaborazione del linguaggio naturale, e poichè si stima che entro il 2025 circa l'80% dei dati sarà di tipo non strutturato [27], questa tecnica diventa di fondamentale importanza per la ricerca, le aziende e all'interno delle organizzazioni. Negli ultimi anni, l'esplosione del web e l'aumento degli utenti online portano con sè una generazione di dati non strutturati senza precedenti. Sempre più persone si stanno avvicinando allo studio del text mining perché si crede che abbia un potenziale commerciale, di ricerca e analitico superiore rispetto all'estrazione dei dati, poiché la maggior parte delle informazioni è memorizzata sotto forma di testo [28]. Basti pensare che la maggior parte delle comunicazioni online, come chat e e-mail sono espressi in linguaggio naturale, mentre solo una piccola percentuale è in un formato strutturato e quindi processabile subito da un algoritmo. Con questo scenario in mente, il text mining diventa cruciale per migliorare l'analisi automatica dei documenti, poiché è l'unica tecnologia in grado di esaminare profondamente il "significato" del testo. Il text-mining si divide quindi in diversi passaggi, fortemente dipendenti fra di loro e dalla quale dipendono i risultati delle task o dell'analisi che si sta svolgendo.



Figura 3.1: Processo Text-mining

### 3.1.3 Passaggi ed estrazione di informazioni

In questo campo, i dati, come detto precedentemente, per essere sottoposti a tecniche analitiche devono essere trasformati in un formato strutturato. Questo processo viene definito come *preprocessing*, cioè il processo di pulizia, trasformazione e preparazione del testo grezzo per le successive analisi. Questo non è l'unico passaggio però, l'estrazione di informazioni dai dati sotto forma di linguaggio naturale si compone di alcuni fondamentali procedimenti che permettono; una volta ripuliti i dati, di essere utilizzati e di estrarre della conoscenza da essi.

- *Selezione dei dati:* La selezione e la raccolta dei dati è il primo step fondamentale nel processo di text-mining, i dati devono essere testi, estratti principalmente da: motori di ricerca, social media, documenti digitali e trascrizioni. La selezione online avviene principalmente tramite *API* (Application Programming Interface) cioè quell'insieme di metodi informatici volti a facilitare la comunicazione e l'interazione fra diversi software. Fungendo da collegamento fra le diverse applicazioni consentendo loro l'interoperabilità e lo scambio efficiente di dati. I dati inoltre devono essere coerenti con l'obiettivo di ricerca, con i task che si vogliono eseguire e il più privi di bias possibile.
- *Annotazioni:* I testi del corpus vengono etichettati da annotatori umani, ad ogni testo viene assegnata una label che indichi il contenuto presente, questo processo può rendere il corpo di dati più facilmente interpretabile e utilizzabile per l'analisi. Inoltre, nei processi di analisi supervisionati, dove l'annotazione

---

avviene necessariamente, si permette una valutazione del modello e un calcolo dell'efficienza dello stesso.

- *Preprocessing*: Il terzo passo è quello con cui si inizia la pulizia del testo, i dati presentano molto spesso rumori ed elementi non fondamentali per le successive analisi, è necessario quindi rimuoverli e modificarli per ottenere una versione pulita del testo, un'estrazione di informazioni coerente con l'obiettivo di ricerca e l'anonimizzazione dei dati.

I principali passaggi del preprocessing sono:

- Tokenizzazione: Questo passaggio consiste nel suddividere il testo in parole o token individuali.
- Rimozione di caratteri speciali, punteggiatura e elementi di disturbo: In questa fase vengono rimossi dalla frase tutti i caratteri speciali, come simboli di punteggiatura e altri elementi non alfabetici, che potrebbero essere considerati rumore nei dati.
- Conversione in minuscolo: Qui le parole vengono trasformate tutte in minuscolo. Questo assicura che le parole con la stessa radice, ma con lettere maiuscole e minuscole diverse, vengano trattate allo stesso modo.
- Rimozione delle stopwords: Le stopwords sono parole comuni e ricorrenti che non portano significato specifico al contesto del testo e quindi potrebbero aumentare il rumore nei dati.
- Lemmatizzazione: Quest'ultimo processo consiste nel ridurre le parole alla loro forma di base o lemma. Questo significa che le parole vengono convertite nella loro forma canonica, che rappresenta il loro significato fondamentale. Questo aiuta a ridurre la complessità dei dati e a trattare parole simili, verbi coniugati in diversi tempi verbali, allo stesso modo.
- *Vettorizzazione*: Dopo aver preparato e pulito i dati, nonché averli normalizzati, il passaggio successivo consiste nella trasformazione di tali dati da un formato non strutturato a uno strutturato e comprensibile dal computer. Questo processo, comunemente noto come vettorizzazione, è cruciale per la conversione dei dati in un formato numerico, interpretabile da modelli di machine learning e reti neurali. È fondamentale chiarire che un formato strutturato non implica necessariamente un formato numerico e viceversa. I dati strutturati sono organizzati secondo schemi definiti, come nelle tabelle di database o nei file CSV, mentre i dati non strutturati, quali per esempio testi o immagini, non seguono una struttura predeterminata. La vettorizzazione si occupa di trasformare questi dati non strutturati in vettori numerici, rendendoli adatti per l'analisi mediante algoritmi di machine learning. La rappresentazione vettoriale dei dati, all'interno di uno spazio vettoriale, offre numerosi vantaggi sostanziali nell'esecuzione delle analisi. Lo spazio vettoriale è un concetto matematico che consente di rappresentare i dati in una forma geometrica, dove ogni punto dello spazio corrisponde a un vettore. Questa rappresentazione permette di applicare misure di similarità, come la similarità del coseno, che quantifica la somiglianza tra due vettori basandosi sull'angolo tra essi. Tale capacità è fondamentale per molte applicazioni di machine learning, come il clustering e il retrieval di informazioni. Fra i principali metodi di vettorizzazione utilizzati nel text-mining sono presenti:

- Attributi Semplici: La vettorizzazione tramite attributi semplice mira ad associare delle caratteristiche numeriche sotto forma di attributi semplici al testo, per esempio la lunghezza, la presenza di determinate parole chiave, il numero di occorrenze dei termini, presenza di punti interrogativi o esclamativi ma anche presenza o meno di determinati caratteri.
- Dizionari e lessici: Una raccolta di parole o frasi associate a significati specifici. Ad ogni parola viene quindi assegnata un’informazione significativa, rientrando nel lessico di quell’informazione assieme ad altre parole.
- Bag of Words: La bag of words (BoW) può essere definito come un borsa-dizionario (bag) contenente tutte le parole contenute nel corpus al netto delle ripetizioni. Non considera quindi l’ordine o la struttura del testo ma solamente l’occorrenza di ogni parola.



Figura 3.2: Bag of Words

- TF-IDF: Il Term Frequency-Inverse Document Frequency è una tecnica utilizzata per valutare l’importanza di una parola all’interno di un documento rispetto a un intero corpus di documenti. Al contrario delle BoW che tengono conto solamente della presenza o meno di una parola, il tf-idf è una vera e propria funzione di peso usata per determinare l’importanza di una parola. Si basa su due concetti fondamentali: la frequenza del termine (TF), che indica quante volte una parola appare in un documento, e l’inverso della frequenza del documento (IDF), che indica quanto una parola è comune o rara in tutto il corpus. La formula del TF-IDF è data da:

$$w_{x,y} = tf_{x,y} \times \log \left( \frac{N}{df_x} \right)$$

dove:

- $tf_{x,y}$  rappresenta la frequenza di  $x$  in  $y$
- $df_x$  indica il numero di documenti contenenti  $x$
- $N$  è il numero totale di documenti

- 
- *Apprendimento e analisi:* Il corpus di dati, una volta pulito e preparato, viene utilizzato da un modello di Machine Learning per apprendere, in modo da saper discriminare testi futuri e mai visti. È necessario scegliere il modello più adeguato per il task e settare i parametri più adeguati. Esistono diverse opzioni disponibili. Una possibilità è l'utilizzo di modelli *pre-addestrati*, come *BERT*, *GPT*, *Word2Vec* e *FastText*, che offrono la possibilità di utilizzo senza la necessità di addestrare il modello da zero. Oppure, è possibile addestrare modelli *personalizzati* per adattarli alle esigenze specifiche del task, raccolgendo un corpus di dati appropriato e implementando algoritmi di machine learning specifici che si adattano alle esigenze specifiche della ricerca. Infine, un'opzione intermedia è il *fine-tuning* dei modelli pre-addestrati attraverso il settaggio e l'adattamento di questi modelli attraverso il ri-addestramento tramite nuovi dati, in genere molto più specifici rispetto al corpora sulla quale il modello è stato addestrato inizialmente.
  - *Valutazione del modello e dei risultati:* Infine, una volta che il modello scelto è stato addestrato sarà necessario valutare le capacità predittive e l'efficenza. Questo passaggio viene fatto attraverso delle metriche apposite e permette di adeguare, nel caso di risultati non soddisfacenti, l'intero processo o alcune parti, migliorando le prestazioni del modello. Va specificato che ci sono misure più adatte a modelli di *apprendimento supervisionato*, cioè in cui sono presenti etichette di classe note, utilizzate per guidare l'addestramento e quindi la ricerca dei risultati e invece misure più adatte a modelli di *apprendimento non supervisionato* in cui si cerca di trovare pattern e raggruppamenti (cluster) nei dati senza la guida di etichette di classe predefinite. Definibile anche come un'analisi nella quale sono i dati che guidano alla scoperta di risultati, pattern e informazioni interessanti. Le principali metriche di misura che si adattano meglio all'apprendimento supervisionato sono:
    - Precision: È l'abilità di un classificatore di non etichettare un'istanza positiva che è in realtà negativa. Per ogniclasse è definito come il rapporto tra veri positivi e la somma di veri e falsi positivi. Detto in un altro modo, “per tutte le istanze classificate come positive, quale percentuale era corretta?”.
    - Recall: È la capacità di un classificatore di trovare tutte le istanze positive. Per ogni classe è definito come il rapporto tra i veri positivi e la somma dei veri positivi e dei falsi negativi. Detto in altri termini, “per tutte le istanze che erano effettivamente positive, quale percentuale è stata classificata correttamente?”.
    - Accuracy: È la misura del numero di dati correttamente predetti rispetto a tutti i punti dati. Più formalmente, è definito come il numero di veri positivi e veri negativi diviso per il numero di veri positivi, veri negativi, falsi positivi e falsi negativi.
    - F-Mesaure o F1 : È una media armonica ponderata delle metriche Precision e Recall in modo tale che il punteggio migliore sia 1 e il peggiore sia 0. Come regola generale, la media ponderata di F1 dovrebbe essere utilizzata per confrontare i modelli di classificatore e non la precisione globale.

<b>Accuracy</b>	Predictions/ Classifications	$\frac{\text{Correct}}{\text{Correct} + \text{Incorrect}}$
<b>Precision</b>	Predictions/ Classifications	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$
<b>Recall</b>	Predictions/ Classifications	$\frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$
<b>F1</b>	Predictions/ Classifications	$\frac{2 * \text{True Positive}}{\text{True Positive} + 0.5 (\text{False Positive} + \text{False Negative})}$

Figura 3.3: Formule metriche di misura [29]

Le principali metriche che si adattano all'apprendimento non supervisionato sono:

- Omogeneità e coerenza: Sono misure utilizzate per valutare la qualità di un clustering. L'omogeneità misura se i cluster contengono solo elementi che appartengono alla stessa classe, mentre la coerenza misura se gli elementi appartenenti alla stessa classe sono assegnati allo stesso cluster.
- Davies-Bouldin Index: È una metrica che misura la dispersione tra i cluster, considerando sia la distanza media all'interno di ciascun cluster che la distanza tra i centroidi (punto dato centro del cluster) dei cluster.
- Silhouette score: È una metrica utilizzata per valutare la coesione e la separazione dei cluster in un dataset. Calcola quanto ogni elemento è simile agli elementi presenti nel proprio cluster rispetto agli elementi nei cluster vicini. Un punteggio silhouette uguale o vicino a +1 indica che i cluster sono ben separati. Un esempio di come viene calcolato il coefficiente silhouette sulla base di tre cluster differenti viene mostrato nella prossima pagina in Fig. 3.4

Il coefficiente silhouette è calcolato per ogni punto  $i$  come:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Dove:

- $a(i)$  è la distanza media tra il punto  $i$  e tutti gli altri punti nello stesso cluster.
- $b(i)$  è la distanza media tra il punto  $i$  e tutti i punti nel cluster più vicino, diverso dal cluster di  $i$ .

Il coefficiente per l'intero dataset è la media dei punteggi per ciascun punto.

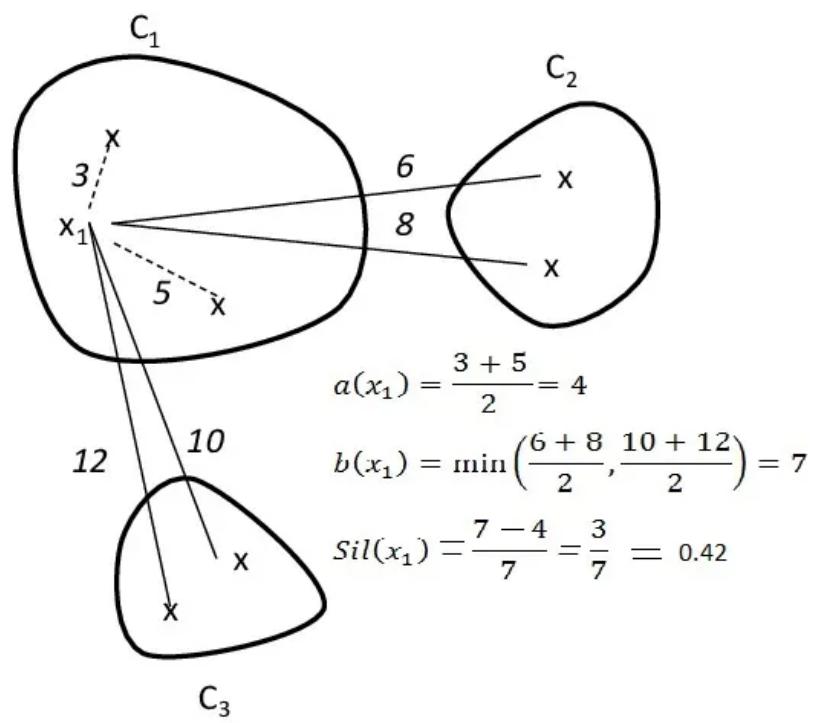


Figura 3.4: Silhouette Score [30]

### 3.1.4 Applicazioni pratiche

Grazie alla sua elasticità e adattabilità, il text mining può essere usato in una vasta gamma di settori e contesti differenti. I suoi principali utilizzi passano dal classificare: recensioni, testi brevi, post su social network, titoli di giornale e titoli di canali di comunicazione online. Per arrivare all'analisi di documenti composti da un grande numero di parole, come per esempio contratti, articoli interi, grandi corpus di testi online e molto altro. Tra le più note applicazioni di questa branca dell'NLP ci sono:

- *Sentiment analysis*: È il processo di identificazione, estrazione e analisi dei sentimenti, delle emozioni e degli atteggiamenti espressi nei testi. Viene utilizzato per determinare se un testo è positivo, neutro o negativo, e viene applicato principalmente a recensioni e commenti sui social media.
- *Classificazione dei topic*: È la classificazione automatica dei testi in base agli argomenti e ai topic trattati principalmente.
- *Hate Speech detection*: È l'identificazione automatica di linguaggio offensivo e discriminatorio all'interno dei testi.
- *Spam Filtering*: È il processo di identificazione e filtraggio automatico di messaggi indesiderati o non richiesti, come e-mail di spam, messaggi di testo non desiderati, commenti pubblicitari sui social media e testo proveniente da bot automatici.



Figura 3.5: Applicazione pratiche del text-mining



# Capitolo 4

## Metodologia

Questo capitolo è dedicato all'analisi delle tecnologie e degli strumenti impiegati nel contesto di questo studio, insieme alla metodologia adottata per raggiungere gli obiettivi di ricerca con un particolare attenzione ai passaggi pratici effettuati per ottenere i risultati.

### 4.1 Strumenti utilizzati

Per il conseguimento dell'obiettivo di ricerca e lo sviluppo di questa tesi sono stati utilizzati diversi strumenti:

#### 4.1.1 Linguaggio di programmazione e librerie

- *Python*: Python è un linguaggio di programmazione ad alto livello che nasce nei primi anni '90, interpretato e orientato agli oggetti, è fra i linguaggi più utilizzati nel campo dell'apprendimento automatico e in quello dell'elaborazione del linguaggio naturale grazie la pluralità di librerie che lo circondano.
- *NLTK*: Il Natural Language Toolkit è un insieme di librerie e programmi per l'analisi simbolica e statistica nel campo dell'elaborazione del linguaggio naturale. Sviluppata da Steven Bird ed Edward Loper, conferisce tutte le funzioni più importanti necessarie alla gestione, manipolazione e pulizia dei dati testuali [31].
- *ReGex*: Modulo standard di python, per operare con le stringhe, viene utilizzato per trovare determinati pattern di caratteri [32].
- *BERTopic*: Una libreria per l'analisi dei topic basata su BERT [33].
- *Pandas*: Una libreria per la manipolazione e l'analisi dei dati [34].
- *Scikit-learn*: Una libreria di apprendimento automatico [35].
- *Matplotlib, Wordcloud, Seaborn*: Librerie per la creazione di grafici [36, 37, 38].

---

### 4.1.2 Hardware

Il dispositivo hardware adottato per effettuare la ricerca è: AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx, 2.10 GHz, 8GB di RAM e sistema operativo Windows 10 x64.

## 4.2 Raccolta dei dati e dataset

La raccolta dei dati, come anticipato nel capitolo precedente, è il primo passaggio, ed è fondamentale una scelta il più accurata possibile per evitare bias, ridondanze e dati non coerenti con l'obiettivo di ricerca. Questa fase ha inizialmente causato delle problematiche, dopo i recenti cambiamenti all'API di Twitter [39], l'utilizzo di questo applicativo per estrarre i dati necessari al conseguimento della ricerca risultava non adeguato. Per sopperire a questa problematica è stato costruito un corpus dati utilizzando due dataset disponibili su Kaggle, entrambi sotto licenza "Attribution 4.0 International (CC BY 4.0)" [40], uno contenente 43.943 Tweet relativi al cambiamento climatico e uno contenente circa 4.5 milioni di righe fra post e commenti del social network Reddit che vanno fino a settembre 2022. Tre sono i punti importanti da sottolineare. Il primo è che tutti i dati utilizzati in questo studio sono stati sottoposti a un processo di anonimizzazione e normalizzazione al fine di garantire la protezione della privacy e dell'identità dei partecipanti. Inoltre, è importante sottolineare che tutti i dati sono stati utilizzati esclusivamente per il fine specifico della ricerca in questione e non sono stati condivisi o utilizzati per altri scopi al di fuori del contesto dello studio scientifico. Questo approccio è volto a rispettare gli standard accademici e legali relativi alla protezione dei dati e assicura l'integrità e l'eticità della ricerca. Il secondo punto è che non tutte le righe del secondo dataset sono state utilizzate per costruire il corpus, questa scelta è dovuta principalmente alla potenza di calcolo disponibile con il computer con la quale è stata effettuata la ricerca, per la precisione sono state selezionate le prime 300.000 righe e il terzo e ultimo punto che il testo di entrambi i dataset è in lingua inglese. La combinazione tra i due dataset è stata effettuata tramite python, ottenendo un terzo file csv contenente il corpo dati utilizzato per la ricerca:

```
1 import pandas as pd
2 dataset_1 = pd.read_csv('Twitter.csv')
3 dataset_2 = pd.read_csv('Reddit.csv', nrows=300000)
4 dataset_definitivo = pd.concat([dataset_1, dataset_2])
5 dataset_definitivo['dati'] = dataset_2['body'] + dataset_1['message']
6 dataset_definitivo.to_csv('corpus.csv', index=False)
```

---

## 4.3 Pre-processing dei dati

Successivamente, è stato effettuato il passaggio di preprocessing per normalizzare e pulire i dati. In particolar modo è stato effettuato tramite la libreria NLTK e tramite l'ausilio del libro *"Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, and Edward Loper"* [41, 31].

I passaggi effettuati sono:

- Import e download delle librerie e dei pacchetti necessari:

```
1 import re
2 import nltk
3 import pandas as pd
4 import os
5 from nltk.corpus import stopwords
6 from nltk.tokenize import word_tokenize
7 from nltk.stem import WordNetLemmatizer
8 import time
9 nltk.download('punkt')
10 nltk.download('stopwords')
11 nltk.download('wordnet')
```

- Inizializzazione timer per tenere conto del tempo di elaborazione:

```
1 start_time = time.time()
```

- Funzione di preprocessing contenente tutte le operazioni di normalizzazione:

```
1 def preprocess_text(text):
```

- Rimozione delle menzioni utente, degli hashtag, degli URL, dei caratteri non alfabetici e delle parole con meno di 3 caratteri, tutte utilizzando espressioni regolari (ReGex):

```
1 text = re.sub(r'@\w+', ' ', text)
2 text = re.sub(r'#\w+', ' ', text)
3 text = re.sub(r'http\S+', ' ', text)
4 text = re.sub(r'^[a-zA-Z\s]', ' ', text)
5 text = re.sub(r'\b\w{1,2}\b', ' ', text)
```

- Tokenizzazione del testo utilizzando word\_tokenize da nltk:

```
1 tokens = word_tokenize(text)
```

- Lowercasing e rimozione delle stopwords utilizzando l'insieme di stopwords predefinite da nltk:

```
1 stop_words = set(stopwords.words('english'))
2 tokens_filtered = [word for word in tokens
3 if word.lower() not in stop_words]
```

- Lemmatizzazione delle parole utilizzando WordNetLemmatizer da nltk:

```

1     lemmatizer = WordNetLemmatizer()
2     tokens_lemmatized = [lemmatizer.lemmatize(word)
3     for word in tokens_filtered]

```

- Ritorno della funzione:

```

1     return tokens_lemmatized

```

- Lettura del csv e creazione del dataframe selezionando la colonna dati dal file CSV:

```

1 try:
2     dataframe = pd.read_csv(file_csv)
3     print("Dataset trovato")
4     colonna_testo = dataframe['dati'].tolist()
5     print("Colonna raggiunta")
6 except (FileNotFoundException, KeyError) as e:
7     if isinstance(e, FileNotFoundException):
8         print("file non trovato")
9     else:
10        print("colonna non trovata")
11        exit()

```

- Applicazione della funzione a tutti i testi contenuti nella colonna del dataframe e stampa di un messaggio per ogni riga processata per monitorare l'avanzamento del processo:

```

1     preprocessed_texts = []
2 for index, row in dataframe.iterrows():
3     preprocessed_text = preprocess_text(row[colonna_testo])
4     preprocessed_texts.append(preprocessed_text)
5     print(f'Riga {index + 1} processata')

```

- Aggiunta della colonna "testo preprocessato", rimozione dei duplicati tramite il metodo drop.duplicates() della libreria Pandas e esportazione nel file .csv:

```

1     dataframe['testo_preprocessato'] = preprocessed_texts
2     dataframe.drop_duplicates(subset='testo_preprocessato',
3                               ,
4                               inplace=True)
5     output_csv = os.path.join(os.path.expanduser('~'),
6                               'OneDrive', 'Desktop', 'corpus.csv')
6     dataframe.to_csv(output_csv, index=False)

```

- Stampa del tempo di esecuzione:

```

1     end_time = time.time()
2     elapsed_time = end_time - start_time
3     print(f'Tempo di esecuzione: {elapsed_time} secondi')

```

---

Il risultato ottenuto è quindi un file csv contenente il testo normalizzato e pronto per la parte successiva di Topic Modeling.

## 4.4 Topic Modeling

Il passaggio successivo è il topic modeling, una delle tecniche di analisi testuali del text-mining. L'obiettivo di questa tecnica è quello di identificare in modo automatico o semi-automatico i temi più significativi da un corpo di documenti, estraendo pattern nascosti, senza la necessità di un'etichettatura manuale. La scelta della metodologia per questa fase è ricaduta su BERTopic, una libreria python basata sul modello *Bidirectional Encoder Representations from Transformers (BERT)*, un modello di Deep Learning preaddestrato sviluppato da Google che si basa sui Transformers. Come citato nel terzo capitolo (3.1.1), i transformer hanno rivoluzionato il campo dell'elaborazione del linguaggio naturale [25], passando da un approccio sequenziale di elaborazione, come per esempio fanno le reti neurali ricorrenti, ad un approccio parallelo. Un altro fattore chiave dei transformer è il meccanismo di attenzione, che consente al modello di pesare e dare maggiore importanza a determinate parole o token all'interno di un input. Questo meccanismo vuole simulare, come noi umani, ascoltando o leggendo un testo, selezioniamo delle parole chiave e attribuiamo delle relazioni semantiche fra le diverse parole per ottenere una comprensione complessiva e più accurata del testo [42]. L'architettura tipica di un transformer è costituita da più layer di *encoder* e *decoder*. L'encoder converte l'input testuale in una rappresentazione numerica (embedding) attraverso il meccanismo di attenzione, mentre il decoder prende in input l'embedding prodotto dall'encoder e lo utilizza per generare sequenze di output, come frasi o classificazioni. Il decoder può essere contestualizzato come un linguaggio che traduce l'embedding dell'encoder in un formato comprensibile per l'utente o per il task specifico.

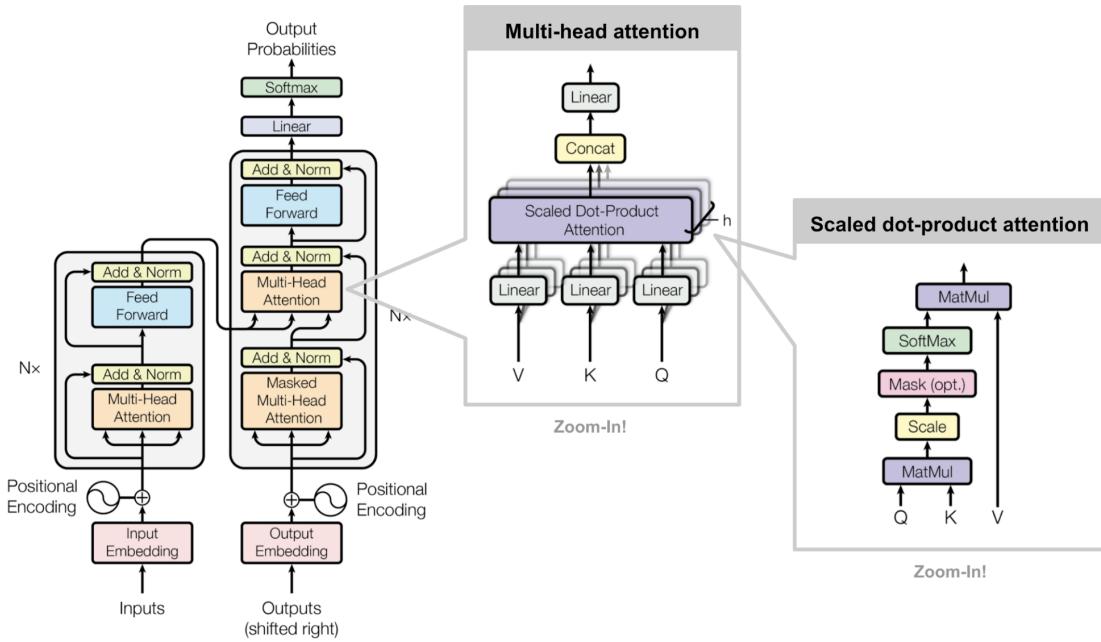


Figura 4.1: Architettura completa dei transformatori (Fig 1 e 2 in Vaswani, et al., 2017.) [25]

Nel modello BERT, l'architettura è costituita esclusivamente da encoder. Questi encoder sono formati da diversi layer, noti come *transformer blocks*, che vengono impilati l'uno sull'altro. Ogni blocco di transformer contiene più sotto-componenti, come i *multi-head self-attention mechanism* e le *feedforward neural networks*. L'encoder di BERT attua il processo di embedding attraverso un processo di auto-attenzione bidirezionale, in cui ogni parola in input viene contestualizzata rispetto a tutte le altre parole nella frase. Il modello sopperisce quindi alla mancanza del decoder attraverso un'architettura mascherata (*masked language model*). Invece di generare sequenze di output come farebbe un decoder, BERT utilizza appunto l'approccio bidirezionale. Durante questo processo, una piccola percentuale di token di input viene mascherata e il modello è addestrato per prevedere i token mascherati utilizzando i token sia a sinistra che a destra del token mascherato. Infine, l'algoritmo softmax viene utilizzato per convertire l'output dell'ultimo layer del modello in una distribuzione di probabilità su un insieme di token del vocabolario. Questo algoritmo viene quindi utilizzato per convertire gli output numerici del modello in una forma più interpretabile e utilizzabile, che rappresenta la probabilità di appartenenza a diverse categorie o classi o il token successivo in task di generazione del linguaggio naturale [43, 25].

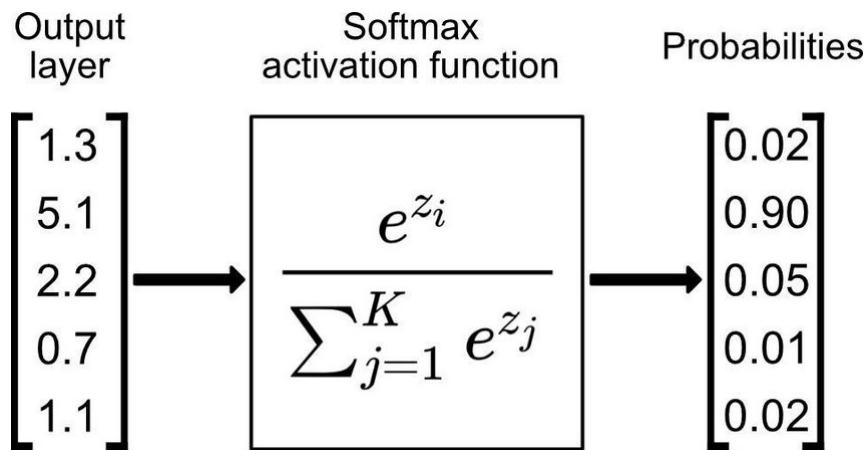


Figura 4.2: Funzione softmax [44]

## 4.5 Modello BERTopic

La libreria BERTopic, è quindi in grado di combinare la capacità del modello BERT nell'embedding con algoritmi di Clustering. La scelta di utilizzare BERTopic è stata guidata da diversi fattori chiave. Il primo è che la libreria offre un supporto attivo da parte della comunità degli sviluppatori e che la documentazione è completa e dettagliata di tutte le informazioni necessarie, inoltre, essa è in continua evoluzione e supportata dai nuovi modelli e algoritmi che vengono rilasciati. Il processo di bertopic è:

- *Embedding dei testi*: BERTopic utilizza il modello BERT per convertire i testi in rappresentazioni vettoriali numeriche. Questo passaggio cattura le relazioni semantiche e contestuali tra le parole nei testi, creando embedding che conservano informazioni sul contenuto.
- *Clustering dei vettori embedding*: Una volta ottenuti i vettori embedding per i testi, BERTopic applica algoritmi di clustering (come HDBSCAN, quello usato per questa ricerca) per raggruppare i testi simili in cluster.
- *Selezione dei topic rappresentativi*: Dopo il clustering, BERTopic identifica i topic e le keyword più rappresentative all'interno di ciascun cluster.
- *Assegnazione delle etichette ai cluster*: Per finire BERTopic assegna etichette descrittive ai cluster in base ai topic rappresentativi identificati. Queste etichette sono spesso estratte dai testi all'interno del cluster che meglio rappresentano il contenuto generale del cluster stesso.

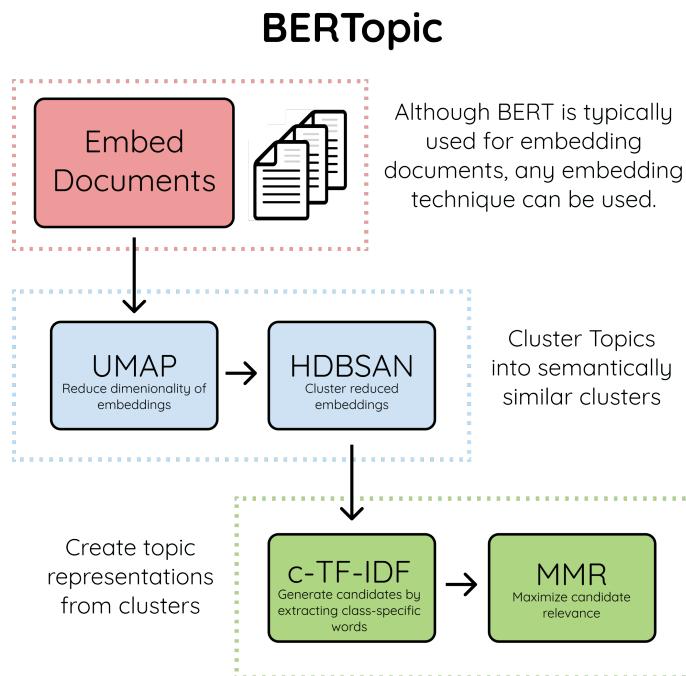


Figura 4.3: Processo di BERTopic [45]

---

## 4.6 Applicazione pratica

### 4.6.1 Creazione degli embeddings

Il primo passaggio effettuato è quello di importare le librerie necessarie e i dati sulla quale si vuole addestrare il modello, dopo di che si può procedere con la creazione degli embeddings, cioè come detto precedentemente, la conversione dei documenti in dati numerici:

```
1  import os
2  import pandas as pd
3  from bertopic import BERTopic
4  from sentence_transformers import SentenceTransformer
5  from bertopic.vectorizers import ClassTfidfTransformer
6  from umap import UMAP
7  from bertopic.representation import KeyBERTInspired
8  file_csv = os.path.join(os.path.expanduser('~'),
9  'OneDrive', 'Desktop', 'corpus.csv')
10 data_raw = pd.read_csv(file_csv)
11 data = data_raw['testo_preprocessato'].tolist()
```

Questa conversione può essere fatta con diversi modelli di trasformazione pre-addestrati, dopo qualche tentativo con alcuni modelli alternativi e vista la forte compatibilità con la lingua inglese la scelta per generare gli embedding è ricaduta sul trasformatore di frasi "*all-MiniLM-L6-v2*" e la libreria SentenceTransformer:

```
1  model = SentenceTransformer("all-MiniLM-L6-v2")
2  embeddings = model.encode(data, show_progress_bar=True)
```

### 4.6.2 Clustering

A questo punto il passaggio successivo è raggruppare gli argomenti simili insieme, (clusterizzare) in modo da estrarre i topic da questi cluster. Il clustering è una tecnica di analisi dei dati non supervisionata che raggruppa insieme i dati simili in gruppi detti cluster. Cercando di massimizzare la similarità all'interno di ciascun cluster e minimizzare la similarità tra cluster diversi. Prima di eseguire questa parte è stata diminuita la dimensionalità degli embeddings in quanto molti algoritmi di clustering possono trovare delle difficoltà nel categorizzare ampie quantità di dati. Inoltre questi passaggi vengono definiti come le miglior pratiche dalla documentazione BERTopic [45].

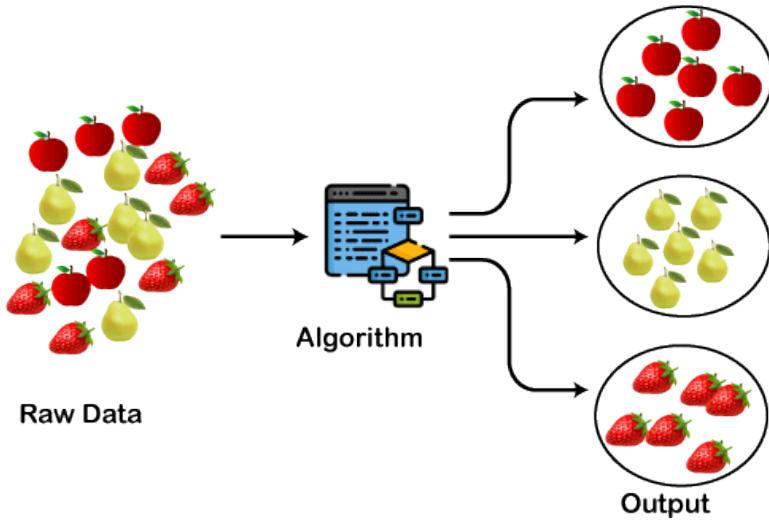


Figura 4.4: Clustering [46]

#### 4.6.3 UMAP

Per ridurre la dimensionalità degli embeddings è stato utilizzato l'algoritmo UMAP, di default BERTopic lo utilizza per ridurre la dimensionalità, ma come nel nostro caso, se si vuole utilizzare dei parametri specifici bisogna definirlo e passarlo al modello:

```

1 umap_model = UMAP(n_neighbors=20,
2 n_components=5, min_dist=1.0, metric='cosine')
3 .fit_transform(embeddings)

```

In particolare i parametri sono:

- *n\_neighbors*: Il numero di punti vicini considerati per la costruzione del grafo di vicinato.
- *n\_components*: Il numero di dimensioni nell'embedding ridotto.
- *min\_dist*: La distanza minima tra punti nell'embedding ridotto.
- *metric*: La metrica di distanza utilizzata per calcolare le distanze tra i punti.

#### 4.6.4 HDBSCAN

Dopo aver ridotto la dimensionalità dei nostri embeddings, il passaggio successivo è appunto quello di raggrupparli in cluster simili per estrarre i nostri argomenti. Con BERTopic si utilizza tipicamente il metodo HDBSCAN) poiché è perfettamente in grado di catturare strutture con densità diverse. HDBSCAN è un algoritmo *density-based* che lavora molto bene con UMAP, l'altro vantaggio di questo algoritmo è che non forza gli elementi in uno dei cluster identificati ma, piuttosto, li considera come outliers. Gli outliers sono elementi in un insieme di dati che si discostano significativamente dal resto del campione di dati. Sono quindi dei valori anomali che si trovano al di fuori del range normale di variabilità dei dati e possono influenzare

---

negativamente il processo di addestramento e i risultati dell'analisi. BERTopic fortunatamente li gestisce in modo efficiente, creando un cluster apposito alla quale vengono associati tutti gli outliers.

```
1  hdbSCAN_model = HDBSCAN(min_cluster_size=15,
2    metric='euclidean', cluster_selection_method='eom')
```

I parametri sono:

- *min\_cluster\_size*: specifica la dimensione minima del cluster. I punti che non raggiungono questa dimensione saranno etichettati come outliers.
- *metric*: indica la metrica di distanza utilizzata per calcolare le distanze tra i punti.
- *cluster\_selection\_method*: metodo utilizzato per selezionare automaticamente il numero di cluster.

#### 4.6.5 c-TF-IDF

BERTopic, al fine di ottenere una rappresentazione accurata dei topic, utilizza la vettorizzazione TF-IDF, adattata però per funzionare sui topic anziché a livello di documento. Questa rappresentazione TF-IDF modificata è chiamata c-TF-IDF e tiene conto di ciò che rende i documenti in un cluster diversi dai documenti in un altro cluster.

**c-TF-IDF**

For a term  $x$  within class  $c$ :

$$W_{x,c} = \|\mathbf{tf}_{x,c}\| \times \log\left(1 + \frac{A}{f_x}\right)$$

$\mathbf{tf}_{x,c}$  = frequency of word  $x$  in class  $c$   
 $f_x$  = frequency of word  $x$  across all classes  
 $A$  = average number of words per class

Figura 4.5: c-TF-IDF [45]

Ogni cluster viene convertito in un singolo documento anziché in una serie di elementi. Quindi, estrae la frequenza della parola  $x$  nella classe  $c$ , dove  $c$  si riferisce al cluster che è stato creato prima. Successivamente c'è il logaritmo di uno più il numero medio di parole per classe  $A$  diviso per la frequenza della parola  $x$  in tutte le classi. Si aggiunge più uno all'interno del logaritmo per forzare i valori a essere positivi. Ciò si traduce nella rappresentazione idf basata su classi di BERTopic. Come con il classico TF-IDF, si moltiplica quindi tf con idf per ottenere il punteggio di importanza di ogni parola. Questa rappresentazione TF-IDF basata su classi è abilitata per impostazione predefinita in BERTopic. Tuttavia, ho passato a BERTopic il modello c-TF-IDF per ottimizzare i parametri. Dopo aver importate

---

correttamente `bertopic.vectorized` e `ClassTfidTransformer`, assegno a una variabile il modello `ClasstFidTransformer(c-TF-IDF)` e regolo il parametro `bm25_weighting`:

```
1 ctfidf_model = ClassTfidTransformer(bm25_weighting=True)
```

- *bm25\_weighting*: Il parametro `bm25_weighting` è un valore booleano che indica se utilizzare una misura di pesatura BM-25 basata sulla classe anziché il metodo predefinito di calcolo. Su dataset di dimensioni non esageratamente grandi, questa variante può essere più efficace, come anche dimostrato nel testing dell'addestramento, mi ha portato a risultati leggermente migliori.

$$\log\left(1 + \frac{A - f_x + .5}{f_x + .5}\right)$$

$f_x$  = frequency of word  $x$  across all classes

$A$  = average number of words per class

Figura 4.6: bm25 [45]

#### 4.6.6 Addestramento del modello

Data la disponibilità di calcolo della macchina, il processo di addestramento è stato abbastanza lungo, soprattutto per provare le diverse configurazioni di parametri per cercare di ottenere i risultati più congrui a quelli attesi. Si è quindi proceduto con l'addestramento del modello, passando tutti i parametri costruiti in precedenza e aggiungendo il modello di rappresentazione *KeyBERTInspired* importato da `bertopic.representation`. Questo modello aiuta a effettuare alcune ottimizzazioni basate sulla relazione semantica tra parole chiave/frasi chiave e l'insieme di documenti in ciascun argomento. Sfrutta quindi il c-TF-IDF per creare una serie di documenti rappresentativi per argomento e utilizzarli come incorporamento degli argomenti aggiornati. Poi calcola la somiglianza tra le parole chiave candidate e l'embedding dell'argomento utilizzando lo stesso modello di embedding che ha incorporato i documenti:

```
1 representation_model = KeyBERTInspired()
2 model = BERTopic(verbose=True, nr_topics="auto",
3 ctfidf_model=ctfidf_model,
4 umap_model=umap_model,
5 representation_model=representation_model)
6 topics, probabilities = model.fit_transform(data)
7 model.save("topic_modeling_v2")
```

I passaggi sono:

- Passo a una variabile representation\_model il modello KeyBertInspired
- Creo la variabile model che conterrà il modello addestrato passandoli i parametri di finitura predisposti in precedenza.
- Salvo il modello per poterlo valutare e creare grafici di visualizzazione.

#### 4.6.7 Valutazione modello

Come spiegato al paragrafo 3.1.3, per poter proseguire con la visualizzazione e l’approfondimento dei dati è necessario valutare le capacità del modello. Questo passaggio viene fatto attraverso le metriche citate e permette di adeguare, nel caso di risultati non soddisfacenti, l’intero processo o alcune parti, per migliorare le prestazioni del modello. Per valutare il modello ho scelto di utilizzare la metrica: *Silhouette Score*, essa permette di valutare la coesione e il corretto aggregamento dei cluster. Calcola quanto ogni elemento è simile agli elementi presenti nel cluster di appartenenza rispetto ai cluster vicini al proprio. Inizialmente il punteggio non era abbastanza soddisfacente, infatti le diverse prove fatte con i dati di training riportavano un punteggio di Silhouette medio che era compreso nel range 0,177 e 0,273. Questo indicava la possibile sovrapposizione fra i cluster e/o cluster composti da pochi elementi. Con l’aggiustamento dei parametri e la fase di fine tuning i risultati sono migliorati fino a raggiungere un punteggio di: 0,598. I codici per calcolare il punteggio di silhouette medio sono stati scritti con l’ausilio della libreria *scikit-learn*, in particolar modo utilizzando le funzioni *silhouette\_score* e *pairwise\_distance*. Il procedimento è così composto:

- Importazione delle librerie e caricamento del dataset di test, del modello bertopic addestrato precedentemente e l’estrazione dei testi dal dataset:

```

1  from bertopic import BERTopic
2  from sklearn.metrics import silhouette_score,
3      pairwise_distances
4  from sklearn.feature_extraction.text
5  import CountVectorizer, TfidfTransformer
6  import pandas as pd
7  import os
8  file_csv = os.path.join(os.path.expanduser('~'),
9  'OneDrive', 'Desktop', 'corpus.csv')
10 data = pd.read_csv(file_csv)
11 modello_topic = BERTopic.load("topic_modeling_v2")
12 testo_preprocessato = data['testo_preprocessato']

```

- Ad ogni cluster viene poi assegnata un’etichetta, cioè un identificatore del raggruppamento stesso per indicare all’algoritmo a quale cluster appartiene ciascun elemento nel dataset:

```

1  labels = modello_topic.transform(testo_preprocessato)

```

- 
- Viene poi, attraverso l'oggetto CountVectorizer e la funzione TfidfTransformer, convertito il testo in una rappresentazione numerica e calcolato la rappresentazione tf-idf:

```
1     vectorizer = CountVectorizer(max_features=1000)
2     x = vectorizer.fit_transform(testo_preprocessato)
3     tfidf_transformer = TfidfTransformer()
4     rappresentazione_tfidf = tfidf_transformer.
      fit_transform(x)
```

---

- Avviene poi il calcolo delle distanze euclidee tra i punti basandosi sulla matrice calcolata nel punto precedente:

```
1     distanze = pairwise_distances(rappresentazione_tfidf,
2                                     metric='euclidean')
```

---

- Si calcola poi il coefficiente silhouette per ogni campione:

```
1     silhouette_scores = silhouette_score(distanze,
2                                         labels, metric='precomputed')
```

---

- Infine, si calcola il punteggio di silhouette medio prendendo la media dei punteggi di silhouette ottenuti e si stampa il risultato:

```
1     media_silhouette_score = silhouette_scores.mean()
2     print("Silhouette Score medio:",
       media_silhouette_score)
```

---

Un punteggio di silhouette di  $\sim 0.598$  indica quindi una discreta separazione dei cluster identificati. Il valore suggerisce che i punti all'interno di ciascun cluster sono più simili tra loro rispetto ai punti nei cluster diversi, contribuendo a una maggiore coesione all'interno dei topic rilevati.

---

## 4.7 Analisi del sentiment

L’analisi del sentiment è una tecnica di Text-mining e NLP che esplora le opinioni, gli atteggiamenti e i sentimenti delle persone nei confronti di un argomento specifico, cercando di categorizzare i testi in tre categorie: positivi, neutri e negativi (Polarità del testo). Oltre a distinguere i dati in queste diverse categorie è possibile utilizzare un approccio che mira a identificare e comprendere le emozioni espresse nei testi, consentendo di valutare la diversità emotiva delle opinioni. I due principali metodi utilizzati per effettuare questo tipo di analisi sono:

- *Analisi del sentiment basata sull’apprendimento automatico*: Il primo approccio utilizza algoritmi di machine learning per analizzare i testi e determinarne il sentimento. Come per altre task di NLP, i modelli di apprendimento automatico vengono addestrati su dataset etichettati manualmente da umani e sono in grado, una volta pronti, di classificare in modo automatico i nuovi testi. Le possibilità, come spiegato nel paragrafo 3.1.3, sono quelle di utilizzare un modello pre-addestrato, personalizzare con nuovi dati un modello esistente (fine-tuning) e infine utilizzare un modello proprio addestrato da zero. Inoltre possono essere etichettati solamente un sub-strato dei testi per poi espandere, tramite altri algoritmo di machine learning, l’etichettatura.
- *Analisi del sentiment basata su elenchi di parole*: Questo metodo si basa su elenchi predefiniti di parole associate a specifici sentimenti. Le parole nei testi vengono confrontate attraverso algoritmi informatici con questi elenchi e, in base alla presenza e alla frequenza delle parole associate a sentimenti positivi, neutri o negativi oppure alle emozioni, viene assegnato un sentimento complessivo. .

### 4.7.1 Distribuzione sentiment

La parte del corpus proveniente da Twitter è stata manualmente etichettata da 3 persone diverse ed è stato lasciato il testo solamente quando tutte e 3 gli etichettatori erano d’accordo. Questo lavoro, svolto da chi ha costruito il dataset, permette di capire, anche se solo per una parte dei dati usati per la ricerca, la distribuzione del sentimento. In particolare i testi sono stati così suddivisi:

- *2 (News)*: il testo contiene un link o del testo relativo a notizie sul cambiamento climatico.
- *1 (Pro)*: il testo supporta la convinzione del cambiamento climatico causato dall’uomo.
- *0 (Neutro)*: il testo né supporta né rigetta la convinzione del cambiamento climatico causato dall’uomo.
- *-1 (Anti)*: il testo non crede nel cambiamento climatico causato dall’uomo.

Viene quindi effettuata l’estrazione della distribuzione in percentuale per creare, successivamente, un grafico a torta usato nel capitolo successivo.

- Import delle librerie e dei dati, viene caricato un file csv ancora non presentato durante l'illustrazione del processo, preprocessato tramite la procedura spiegata nel paragrafo 4.3:

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3 import os
4
5 percorso_csv = os.path.join(os.path.expanduser('~'), 'OneDrive', 'Desktop', 'twitterv1.5_preprocessato.csv')
6 dati = pd.read_csv(percorso_csv, engine='python')
```

- Viene definito un dizionario mappatura\_etichette che mappa i valori numerici delle etichette di sentimento (-1, 0, 1, 2) a stringhe descrittive ('Negativo', 'Neutro', 'Positivo', 'Altro'). Questa mappatura viene applicata alla colonna 'sentiment' del df creato dai dati utilizzando il metodo map():

```

1 mappatura_etichette = {-1: 'Negativo', 0: 'Neutro', 1: 'Positivo',
2 2:'Altro'}
3
4 dati['sentiment']=dati['sentiment'].map(
    mappatura_etichette)
```

- Successivamente viene calcolata la distribuzione del sentimento dei tweet utilizzando il metodo value\_counts() sulla colonna 'sentiment'. Il parametro normalize=True normalizza i conteggi in modo che sommino a 1, e moltiplicando per 100 si converte la distribuzione in percentuale:

```

1 cont_sentimento = dati['sentimento'].value_counts(
    normalize=True) * 100
```

- Viene infine utilizzato matplotlib per creare un grafico a torta per rappresentare graficamente la distribuzione:

```

1 colori = ['#ff9999', '#66b3ff', '#99ff99', '#ffcc99']
2
3 plt.figure(figsize=(8, 6))
4 plt.pie(cont_sentimento, labels=cont_sentimento.index,
    autopct='%1.1f%%', colors=colori, startangle=140)
5 plt.title('Distribuzione Sentimento', fontsize=16,
    fontweight='bold', pad=20)
6 plt.axis('equal')
7
8 plt.show()
```

## 4.7.2 Parole chiave per polarità

Successivamente viene effettuata un'estrazione delle parole più significative e ricorrenti per ogni sentimento per poter contestualizzare e suddividere il lessico del

---

cambiamento in base alla polarità. Questa parte è fatta tramite la libreria scikit-learn e la classe CountVectorizer , utilizzati per convertire i testi in una matrice di conteggio delle parole. Per poi estrarre il conteggio di quante volte ogni parola appare in ciascun testo.

- Import delle librerie e caricamento dei dati, raggruppamento dei testi per sentimento:

```
7     label{lst:parolesent}
8 from sklearn.feature_extraction.text import
9     CountVectorizer
10 import os
11 import pandas as pd
12 from tqdm import tqdm
13 file_csv = os.path.join(os.path.expanduser('~'), 'OneDrive', 'Desktop', 'twitterv1.5_preprocessato.csv')
14 df = pd.read_csv(file_csv, engine='python')
15
16 testi_raggruppati = df.groupby('sentiment')[['testo_preprocessato']].apply(lambda x: ' '.join(x))
```

- Inizializzazione del CountVectorizer: e calcolo del conteggio delle parole per ciascun gruppo di testi:

```
1 vectorizer = CountVectorizer()
2 conteggi_parole=vectorizer.fit_transform(
3     testi_raggruppati)
```

- Itero su ogni sentimento e calcolo il conteggio delle parole per il sentimento corrente poi le parole più comuni vengono estratte e ordinate per conteggio in ordine decrescente. Infine, le prime 20 parole più comuni vengono memorizzate nel dizionario parole\_per\_sentimento. Il dizionario viene poi usato per creare i grafici tramite matplotlib:

```
1 parole_per_sentimento = {}
2 for sentimento, testo in tqdm(testi_raggruppati.items(), desc="Estrazione delle parole per ogni sentimento"):
3
4     conteggio_parole = conteggi_parole[testi_raggruppati.index.get_loc(sentimento), :].toarray().flatten()
5
6     parole_frequenti = [(parola, conteggio) for parola,
7         conteggio in zip(vectorizer.get_feature_names_out(), conteggio_parole)]
8
9     parole_frequenti = sorted(parole_frequenti, key=
10         lambda x: x[1], reverse=True)[:20]
11
12     parole_per_sentimento[sentimento] = parole_frequenti
```

### 4.7.3 Analisi basata su elenchi di parole

La seconda sezione di questa parte della ricerca è ricaduta sul modello di *analisi basata su elenchi di parole*, in particolare attraverso l'utilizzo del *Lexicon della NRC Word-Emotion Association (EmoLex)*, utilizzato nello studio per tracciare le emozioni dei testi e la loro proporzione. L'analisi delle emozioni, al contrario di quella della distribuzione del sentimento è stata effettuata su tutto il corpus. L'NRC Word-Emotion Association Lexicon [47, 48], è un elenco di 14182 parole inglesi codificate per otto emozioni umane: rabbia, paura, anticipazione, sorpresa, fiducia, gioia, tristezza, disgusto e due polarità (positiva e negativa). Assegna punteggi a ciascuna parola compresi tra 0 e 1, che rappresentano rispettivamente la quantità più bassa e più alta di emozione di una parola specifica. Successivamente tramite Python viene fatto un controllo in base alla presenza e alla frequenza di queste parole nei testi per trovare le proporzioni delle emozioni e la loro distribuzione. Ritornando al lessico utilizzato, lo sviluppo dello stesso da parte del National Research Council del Canada (NRC) [49], come spiegato nell'articolo *"Crowdsourcing a Word-Emotion Association Lexicon"* [47], avrebbe trovato delle difficoltà nell'essere costruito tramite annotazione di tutte le emozioni ricollegabili alla natura umana, per questo gli autori hanno deciso di costruire il lessico basandosi sulle *otto emozioni di base di Plutchik*. Robert Plutchik fu uno psicologo americano, che sviluppò il "Modello delle otto emozioni" per spiegare la complessità e la varietà delle emozioni umane. Secondo questo modello, ci sono otto emozioni primarie (quelle citate in precedenza), organizzate in un modello a forma di ruota, che può essere ulteriormente combinato per formare emozioni secondarie e terziarie.

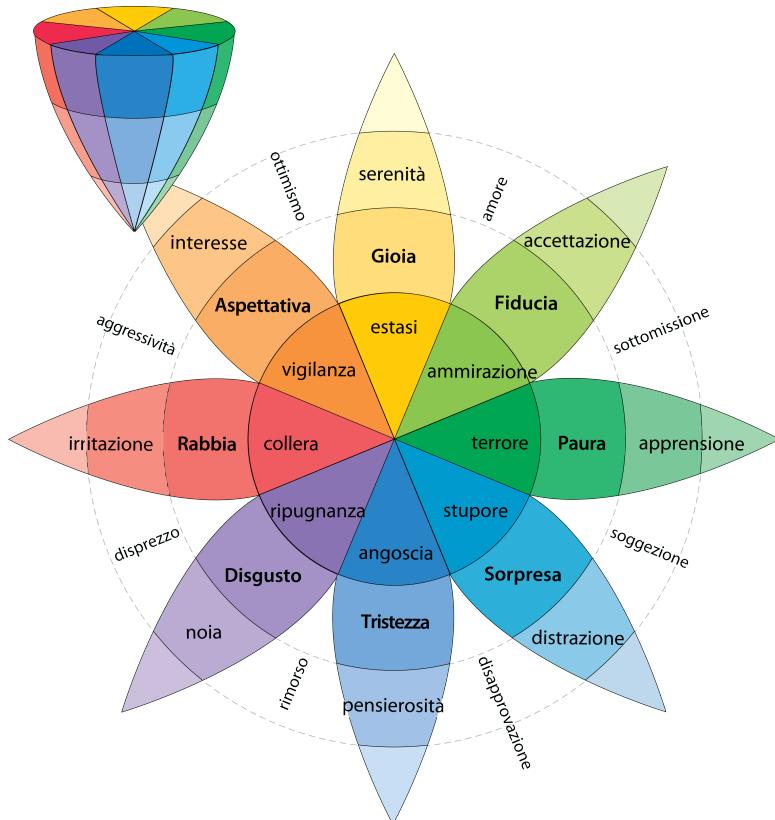


Figura 4.7: Ruota delle emozioni [50]

---

Le otto emozioni primarie sono concepite in base agli otto settori per indicare le emozioni di base: rabbia, aspettativa, gioia, fiducia, paura, sorpresa, tristezza e disgusto. Il modello si basa su una concezione bipolare delle emozioni: ogni emozione primaria ha un polo opposto, basato sulle reazioni fisiologiche che ogni emozione induce negli esseri umani. Ad esempio, la gioia è l'opposto della tristezza, la paura è l'opposto della rabbia, l'aspettativa è l'opposto della sorpresa, e il disgusto è l'opposto della fiducia. Le emozioni possono anche combinarsi per formare nuove sfumature emotive. e la loro intensità è rappresentata dalla dimensione verticale del cono, dove le emozioni si intensificano man mano che si muovono dall'esterno verso il centro della ruota, come indicato anche dal colore: una tonalità più scura indica un'intensità maggiore dell'emozione. Inoltre gioia, fiducia, anticipazione e rabbia sono considerate di valenza positiva, mentre paura, tristezza, disgusto e sorpresa denotano la polarità negativa.

#### 4.7.4 Analisi delle emozioni

Il National Research Council Canada (NRC) rende libero l'accesso al lessico per scopi di ricerca, questa ha permesso di eseguire tramite un algoritmo python l'estrazione di informazioni sulle correlazioni fra i testi e il lessico delle emozioni NRC. Il codice è così organizzato:

- Import delle librerie necessarie:

```

1
2 import pandas as pd
3 import os
4 import matplotlib.pyplot as plt

```

- Segue il caricamento del lessico, e la creazione di un dizionario (lessico\_emozioni) nella quale sono presenti le parole del lessico come chiavi e il loro valore sono le liste di emozioni associate ad esse:

```

1 file_lessico_emozioni = os.path.join(os.path.expanduser(
2     '~'),
3     'OneDrive', 'Desktop',
4     'NRC-Emotion-Lexicon-Wordlevel-v0.92.txt')
5
6 lessico_emozioni = {}
7 with open(file_lessico_emozioni, 'r', encoding='utf-8'):
8     as file:
9         for riga in file:
10             parola, emozione, valore = riga.strip().split('\t'
11             )
12             if int(valore) == 1:
13                 if parola not in lessico_emozioni:
14                     lessico_emozioni[parola] = []
15                     lessico_emozioni[parola].append(emozione)

```

- Vengono caricati i dati e estraggo la colonna 'testo\_preprocessato' dal Data-Frame 'df' e la converto in una lista di liste di parole:

---

```

1 file_csv = os.path.join(os.path.expanduser('~'),
2 'OneDrive', 'Desktop', 'corpus.csv')
3
4 df = pd.read_csv(file_csv)
5 lista_testi = df['testo_preprocessato'].apply(eval).
       tolist()

```

---

- Poi inizializzo un dizionario chiamato 'conteggi\_totali\_emozioni' con tutte le emozioni presenti nel lessico EmoLex, impostando i conteggi iniziali a zero:

```

1
2 conteggi_totali_emozioni = {emozione: 0
3 for emozioni in lessico_emozioni.values()
4 for emozione in emozioni}

```

---

- Poi si itera su ogni lista di parole nella lista estratta dal DataFrame. Per ogni parola, se la parola è presente nel lessico EmoLex, aggiorna i conteggi delle emozioni nel dizionario 'conteggio\_totali\_emozioni' incrementando i conteggi corrispondenti alle emozioni associate a quella parola:

```

1 for lista_testo in lista_testi:
2     for token in lista_testo:
3         if token in lessico_emozioni:
4             for emozione in lessico_emozioni[token]:
5                 conteggi_totali_emozioni[emozione] += 1

```

---

- Stampa il dizionario 'conteggi\_totali\_emozioni', che rappresenta i conteggi totali delle emozioni presenti nei testi:

```

1 print(conteggi_totali_emozioni)

```

---

#### 4.7.5 Analisi delle emozioni per topic

Oltre all'analisi delle emozioni su tutto il corpus è stata effettuata anche sui topic trovati dal processo di Topic Modeling:

- Import delle librerie e caricamento del lessico e la creazione di un dizionario (lessico\_emozioni) nella quale sono presenti le parole del lessico come chiavi e il loro valore sono le liste di emozioni associate ad esse:

```

1 import os
2 import matplotlib.pyplot as plt
3 from collections import defaultdict
4
5 percorso_lexicon_emozioni = os.path.join(os.path.
       expanduser('~'), 'OneDrive', 'Desktop', 'NRC-Emotion-
       Lexicon-Wordlevel-v0.92.txt')
6 lessico_emozioni = {}
7 with open(percorso_lexicon_emozioni, 'r', encoding='utf-8
       ') as file:

```

---

---

```

8     for riga in file:
9         parola, emozione, valore = riga.strip().split('\t')
10        )
11        if int(valore) == 1:
12            if parola not in lessico_emozioni :
13                lessico_emozioni[parola] = []
14                lessico_emozioni[parola].append(emozione)

```

---

- Viene definita una funzione chiamata associare\_emozioni\_a\_topic che associa le emozioni alle parole chiave di un topic. Questa funzione controlla se le parole chiave dei topic sono presenti nel lessico EmoLex e se lo sono, associa le emozioni corrispondenti:

---

```

1
2 def associare_emozioni_a_topic(parole_chiave_topic):
3     conteggio_emozioni = defaultdict(int)
4     for parola, _ in parole_chiave_topic:
5         if parola in lessico_emozioni:
6             for emozione in lessico_emozioni[parola]:
7                 conteggio_emozioni[emozione] += 1
8
9     return conteggio_emozioni

```

---

- Viene definito un dizionario chiamato topics che contiene i dati relativi ai topic. Ogni topic è rappresentato da una serie di parole chiave con il relativo peso:

---

```

1 #i dati per ogni topic sono estratti tramite la funzione
2 #BERTopic -> model.get_topic() e sono in formato "
3 #    Chiave": [('Keyword', punteggio di peso)]
4 topics = {"Topic 1": [('climate', 0.025690456210180134),
5 ('change', 0.025149237030967827), ('people',
6 0.013171093862402398), ('emission',
7 0.011635026396414245), ('fire', 0.011279541435592096),
8 ('environment', 0.01079975599399283), ('year',
9 0.010798380153162531), ('global',
10 0.010590233956477431), ('anthropogenic',
11 0.010516882821143202), ('co2', 0.010179956410201413)],
12 #altri topic }

```

---

- Vengono estratte tutte le emozioni univoche presenti nei topic e calcolate le frequenze delle emozioni per ciascun topic:

---

```

1 emozioni_univoche = set(emozione for topic in topics.
2                         values() for emozione, _ in topic)
3
4 frequenze_emozioni = {emozione: [topic.get(emozione, 0)
5                           for topic in topics.values()]} for emozione in
6                           emozioni_univoche}

```

---

- Itero attraverso tutti i topic definiti nel dizionario topics. Per ogni topic, chiama la funzione associare\_emozioni\_a\_topic passando come argomento le parole

---

chiave associate al topic. risultato viene memorizzato in un altro dizionario chiamato emozioni\_topic che viene utilizzato per stampare a schermo le emozioni dei topic:

```
1 emozioni_topic = {}
2 for nome_topic, keywords_topic in topics.items():
3     conteggio_emozioni = associare_emozioni_a_topic(
4         keywords_topic)
5     emozioni_topic[nome_topic] = conteggio_emozioni
6
7 for topic, emozioni in emozioni_topic.items():
8     print(f"Emozioni associate al topic '{topic}':")
9     for emozione, conteggio in emozioni.items():
10        print(f"{emozione}: {conteggio}")
11
12 print()
```

- Dove, per esempio, il risultato della stampa è:

```
1 Emozioni associate al topic 'Topic 1':
2 positive: 3
3 trust: 2
4 fear: 1
5 anticipation: 1
```

## 4.8 Data Visualization

Per capire, analizzare ed estrarre informazioni dai dati è necessario rappresentarli attraverso l'uso di elementi grafici comuni, facilmente interpretabili dall'uomo. Rappresentando i dati in una forma strutturata grafica è possibile incrementare l'usabilità di questi dati rendendoli più comprensibili e utilizzabili [51]. In questo progetto la maggior parte dei grafici sono stati realizzati con l'ausilio delle librerie Matplotlib e Wordcloud. Inoltre è stata creata una rete semantica grazie all'ausilio di VOSviewer e dal semantic network generator di *Clément Levallois* [52]. VOSviewer è uno strumento software utilizzato per l'analisi e la visualizzazione di reti e cluster di dati. È particolarmente popolare nell'ambito della ricerca scientifica e della bibliometria per esplorare le relazioni tra documenti, autori e parole chiave. VOSviewer consente di creare mappe di reti interattive che mostrano la connettività e la struttura dei dati in modo visuale, aiutando il processo di identificazione dei modelli, tendenze e gruppi tematici all'interno del corpus di testo [53]. La visualizzazione dei dati non solo rende più accessibili le informazioni contenute nei dataset, ma può anche rivelare relazioni e pattern altrimenti difficili da individuare mediante analisi dei dati grezzi, facilitando la presa di decisioni informate e consentendo di comunicare efficacemente risultati e conclusioni ad un pubblico più vasto. Vengono ora riportati tre codici finti per generare 3 tipologie di grafici utilizzati per illustrare i risultati nel capitolo successivo:

- 
- *Grafico a torta:*

```
17 import matplotlib.pyplot as plt
18
19 etichette = [ ]
20 valori = []
21 colori = [ ]
22
23 plt.figure(figsize=(8, 6))
24 plt.pie(valori, labels=etichette, colors=colori, autopct=
    '%1.1f%%', startangle=140)
25 plt.title('Distribuzione dei dati')
26 plt.axis('equal')
27
28 plt.show()
```

---

- *Grafico a barre:*

```
29 import matplotlib.pyplot as plt
30
31 categorie = [ ]
32 valori = []
33
34 plt.figure(figsize=(8, 6))
35 plt.bar(categorie, valori, color=' ')
36 plt.title('Distribuzione dei dati')
37 plt.xlabel('Categorie')
38 plt.ylabel('Valori')
39
40 plt.show()
```

---

- *Cloudword:*

```
43 from wordcloud import WordCloud
44 import matplotlib.pyplot as plt
45
46 # Per mostrare i risultati della ricerca sono state
        generate delle cloudword basandosi sul numero di
        documenti assegnati ad un topic e al tf-idf delle
        parole
47
48 testo = " "
49
50 wordcloud = WordCloud(width=800, height=400,
    background_color='white').generate(testo)
51
52 plt.figure(figsize=(10, 5))
53 plt.imshow(wordcloud, interpolation='bilinear')
54 plt.title('WordCloud')
55 plt.axis('off')
56
57 plt.show()
```

---



# Capitolo 5

## Risultati e Discussione

L'argomento del cambiamento climatico è intriso di complessità e interconnessioni tra i suoi diversi sotto-argomenti, rendendo la sua analisi un'impresa ricca di sfide e opportunità. La comprensione completa di questo fenomeno richiede un approccio multidisciplinare ed è evidente questa necessità soprattutto nella ricerca empirica, dove l'analisi dei dati spesso si scontra con la necessità di considerare una vasta gamma di fattori e variabili. Come illustrato nella Figura 5.1, la rete semantica generata rivela un intreccio di relazioni non solo tra argomenti strettamente correlati, ma anche tra quelli apparentemente distanti. Questa struttura di connessioni dimostra che, nonostante la diversità degli argomenti, esistono legami significativi che evidenziano la ricchezza semantica dell'argomento.

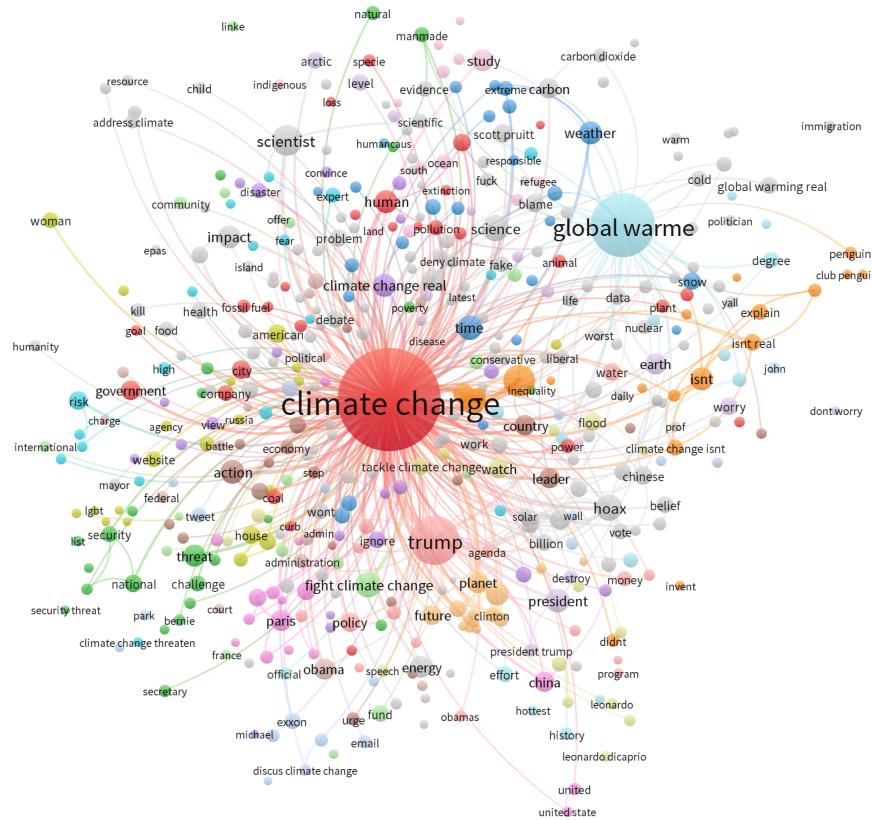


Figura 5.1: Rete semantica

---

## 5.1 Analisi dei topic

I procedimenti fino a qui elencati, hanno portato, al netto dei test e dell'ottimizzazione dei parametri, al riconoscimento di dieci significativi argomenti di diverse proporzioni. Questi topic vengono automaticamente titolati dal modello con l'attribuzione di alcune delle parole chiave più significative e sono successivamente etichettati manualmente con un titolo che descrivesse meglio l'argomento. Lo step successivo è l'estrazione delle keywords, l'analisi di alcuni documenti assegnati ai rispettivi topic e nella contestualizzazione e l'approfondimento dell'argomento. Dove possibile, i risultati vengono confrontati con i dati climatici disponibili per estrarre considerazioni, trovare correlazioni e approfondire gli esiti. Il modello BERTopic ha estratto, tramite le procedure fino ad ora spiegate, 11 argomenti, di cui il primo, il topic -1 è composto da tutti gli outliers trovati, ovvero gli elementi in un insieme di dati che si differenziano significativamente dal resto del campione di dati e quindi non vanno considerato ai fini della discussione dei risultati. La composizione degli outliers riveste però un ruolo cruciale nell'interpretazione dei risultati del modello, sebbene l'ottimizzazione dei parametri abbia contribuito a ridurre il numero di documenti assegnati al topic -1, è importante riconoscere che questa categoria continua a rappresentare una significativa percentuale del corpus dati, per un totale di 76.952 documenti rappresentativi del  $\sim 22,37\%$  del totale. Ciò può essere attribuito alla natura stessa del contesto della ricerca, dove i commenti e i post sui social network sono spesso soggetti all'influenza di troll, bot e testi contenenti bias. Pertanto, nonostante gli sforzi per ridurre l'impatto degli outliers, la presenza di tali dati rimane significante sul totale dei documenti analizzati. Si procede estraendo le informazioni sui topic e la distribuzione dei documenti assegnati a ciascun topic:

```
58 from bertopic import BERTopic  
59  
60 model = BERTopic.load("topic_modeling_v2")  
61 model.get_topic_info()
```

Non prendendo in considerazione il topic -1, cioè quello composto dagli outliers, gli altri 10 argomenti, numerati da 0 a 9, sono invece associati a specifiche tematiche chiave legate al cambiamento climatico e alle sue implicazioni. Alcuni di questi argomenti riflettono la preoccupazione globale per l'impatto del cambiamento climatico sulle persone e sull'ambiente, mentre altri affrontano questioni specifiche come la cooperazione internazionale, le azioni politiche e i dubbi sulla realtà del cambiamento climatico stesso. La suddivisione da parte del modello è così composta:

Tabella 5.1: Distribuzione documenti

Topic	Count	Name
-1	76952	-1_outliers
0	73220	0_climate_change_people_global
1	43655	1_betwe_cooperation_lecture_import
2	38644	2_ivanka_trump_merkel_melania
3	31596	3_cracks_anemone_bleached_snow
4	21380	4_disease_virus_emergency_covid
5	19254	5_manufacturing_corporate_oil_chinese
6	12138	6_warming_warmth_miami_temperature
7	10893	7_climate_change_climate_real_hoax
8	8424	8_exxon_tillerson_rex_emission
9	7787	9_bacon_watch_meat_agricolture

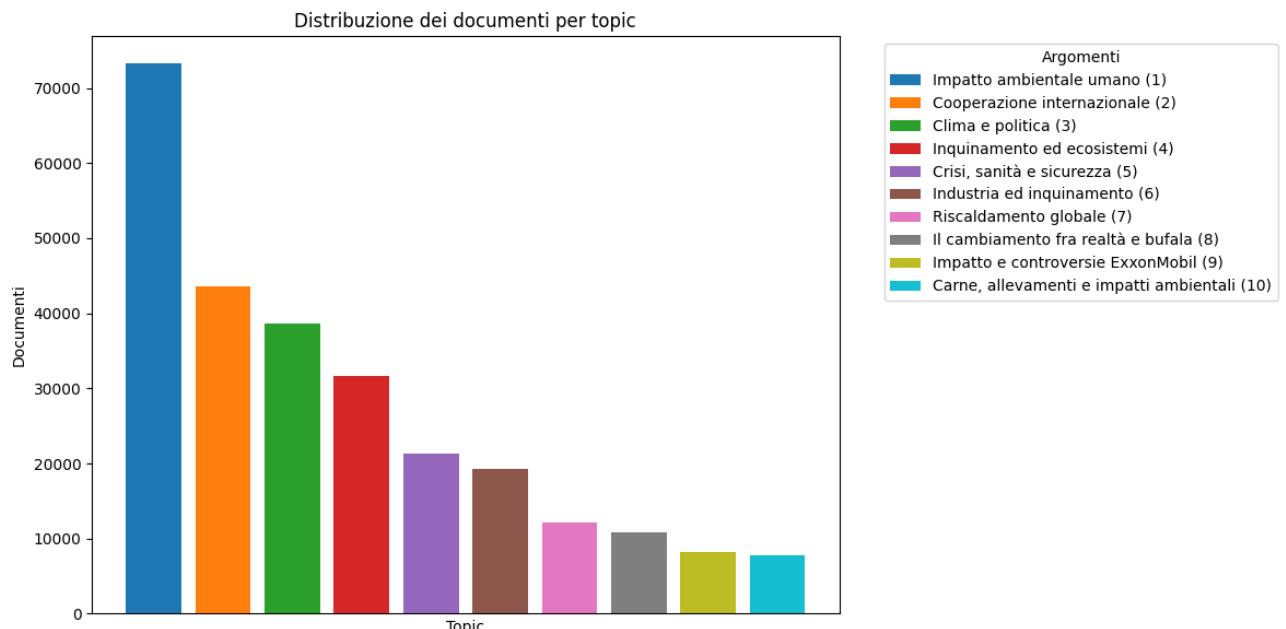


Figura 5.2: Distribuzione dei documenti per topic

---

I topic sono illustrati nella Fig. 5.1. Prima di procedere con l'analisi dettagliata di ciascun topic, ognuno di essi viene titolato e descritto brevemente nell'elenco seguente:

- Il topic 0, con 73.220 documenti, etichettato: "*Cambiamento climatico e impatto ambientale umano*", riguarda il cambiamento climatico e la sua relazione con le persone, sia in termini di impatto sulla popolazione sia come processo antropologico.
- Il topic 1, con 43.655 documenti, etichettato: "*Cooperazione internazionale e clima*", tratta della cooperazione e dell'importanza delle relazioni internazionali nel contesto del cambiamento climatico.
- Il topic 2, con 38.644 documenti, etichettato: "*Cambiamento climatico e politica*", si concentra sul rapporto fra politica e cambiamento climatico. Includendo discussioni sulle posizioni o azioni di queste figure politiche riguardo le politiche ambientali.
- Il topic 3, con 31.596 documenti, etichettato: "*Inquinamento ed ecosistemi*", riguarda la salute degli ecosistemi marini e non, in particolare i problemi legati allo sbiancamento dei coralli, del deterioramento delle barriere marine, l'inquinamento della plastica e lo scioglimento dei ghiacciai.
- Il topic 4, con 21.380 documenti, etichettato: "*Crisi, Sanità e sicurezza pubblica*", tratta sotto-argomenti come malattie, virus, emergenze e attività di spegnimento degli incendi. Si concentra in particolare su come il cambiamento climatico possa influenzare la diffusione di malattie, l'incidenza degli incendi per arrivare alle connessioni presenti fra cambiamento climatico e la pandemia di Covid-19.
- Il topic 5, con 19.254 documenti, etichettato: "*Industria, inquinamento e impatto ambientale*", si concentra sulla produzione manifatturiera, sulle società, sull'industria petrolifera cinese e non e del loro impatto sul clima.
- Il topic 6, con 12.138 documenti, etichettato: "*Riscaldamento globale*", riguarda il riscaldamento globale, il clima caldo e la temperatura crescenti. Il topic riflette la crescente preoccupazione per il riscaldamento globale e i suoi effetti sul clima, con attenzione alla minaccia di alterare l'equilibrio ecologico del pianeta.
- Il topic 7, con 10.893 documenti, etichettato: "*Il cambiamento climatico fra bufala e realtà*", tratta il dibattito sul cambiamento climatico, se è reale o un inganno, sullo scetticismo e sulle notizie legate a questo dibattito.
- Il topic 8, con 8424 documenti, etichettato: "*Impatto e controversie della società ExxonMobil*", si concentra su ExxonMobil, nota società petrolifera e Rex Tillerson l'ex CEO della società. Inoltre tratta le controversie legate a questa società e all'impatto delle imprese di questo calibro nelle emissioni di gas serra.
- Il topic 9, con 7787 documenti, etichettato: "*Consumo di carne, allevamenti e impatto ambientale*", tratta l'argomento degli allevamenti intensivi e del loro impatto ecologico, l'agricoltura e il consumo di carne.

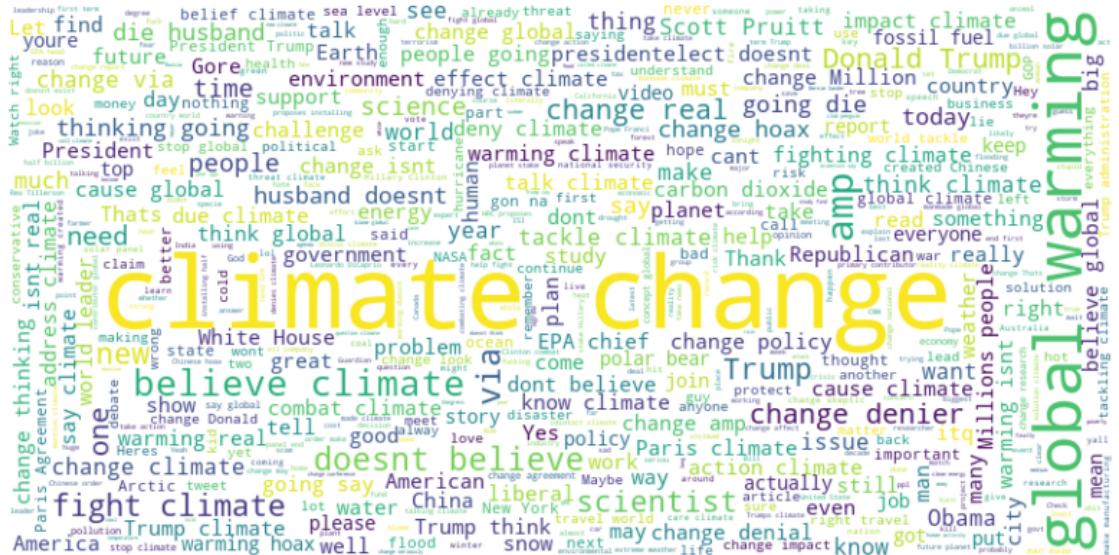


Figura 5.3: Cloudword delle 1000 parole più frequenti

Per analizzare i risultati e topic trovati, l'analisi delle keyword rappresenta un fondamentale strumento. Le keyword, o parole chiave, svolgono un ruolo essenziale nel riassumere e categorizzare il contenuto dei documenti, consentendo una rapida identificazione dei sotto-argomenti principali e delle informazioni rilevanti. Viene estratto per ogni topic un dizionario composto da topic n:[('keyword', peso della parola)..]. Il dizionario è composto dalle keyword più significative quindi con il punteggio più alto, va sottolineato che il topic è composto da più parole chiave rispetto a quelle che vengono inserite nel dizionario. L'esempio per il topic 0 è:

```
1 Topic 0: [(‘climate’, 0.025690456210180134), (‘change’,  
    0.025149237030967827), (‘people’, 0.013171093862402398), (  
    ‘emission’, 0.011635026396414245), (‘fire’,  
    0.011279541435592096), (‘environment’,  
    0.01079975599399283), , (‘global’, 0.010590233956477431),  
    (‘anthropogenic’, 0.010516882821143202), (‘co2’,  
    0.010179956410201413), (‘gas’, 0.010098380153162531), (  
    ‘fossil’, 0.008998380153162531) ]
```

Estratti attraverso:

```
1 from bertopic import BERTopic  
2 model = BERTopic.load("topic_modeling_v2")  
3 model.get_topics()
```

---

Questo ci permette di trovare le parole più significative per ogni topic la quale verrano usate nei paragrafi successivi per discutere, in modo più dettagliato, i risultati. Vengono ora presentate nel loro complesso; questo ci è utile per organizzare le informazioni in maniera strutturata e ordinata, facilitando l'identificazione dei concetti principali. La presentazione tabellare, infatti, rende evidente la relazione tra i topic e le loro parole chiave, contribuendo a una migliore comprensione e interpretazione dei risultati.

Tabella 5.2: Tabella dei Topic e delle Keyword

<b>Nome del Topic</b>	<b>Keyword</b>
Cambiamento climatico e impatto umano	Climate, Change, People, Fire, Global, Environment, Emission, Anthropogenic, CO <sub>2</sub> , Gas, Fossil
Cooperazione internazionale e clima	Between, Cooperation, Lecture, Importance, Climate, Agreement, Paris, Scientist, Act
Cambiamento climatico e politica	Trump, Ivanka, Merkel, Melania, Schwarzenegger, Donald, Policy, People, Government, Election, Dicaprio
Inquinamento ed ecosistemi	Cracks, Anemone, Bleached, Snow, Ice, Plastic, Sea, Ecosystem, Plant
Crisi, sanità e sicurezza pubblica	Disease, Virus, Emergency, Flu, Smoke, Exhausted, People, Rejected, Covid, Advice
Industria, inquinamento e impatto ambientale	Manufacturing, Chinese, Concept, Corporate, Oil, Emission, Order, Global, Greenhouse, Fossi
Riscaldamento globale	Warming, Warmth, Miami, Temperature, Mayor, Change, Surface, Rise
Il Cambiamento climatico fra bufala e realtà	Climate, Change, Real, Hoax, Warming, Fake, News, Scam, Proclimate, Lie, Media
Impatto e controversie della società ExxonMobil	Exxon, Tillerson, Rex, Emission, Exxonmobil, Change, Fossil, Know
Consumo di carne, allevamenti e impatto ambientale	Bacon, Meat, Watch, Agriculture, Vegan, Greenhouse, Mutilated, Emission

---

Per presentare gli argomenti e le loro relazioni viene creato un grafico a dispersione di similarità che ci permette di rilevare informazioni sulla struttura e sulla coerenza tematica del corpus di testo analizzato. Osservando il grafico, le celle più chiare indicano una maggiore similarità tra i cluster corrispondenti, dove la similarità è misurata in base alla somiglianza dei pesi delle parole chiave. Questo suggerisce che i documenti all'interno di cluster simili condividono argomenti o tematiche affini. Al contrario, le celle più scure denotano una minore similarità tra i cluster corrispondenti, indicando una maggiore diversità tematica tra di essi. Tale diversità potrebbe essere il risultato di sottocategorie distinte o di argomenti completamente differenti presenti nel corpus. Inoltre, la diagonale principale del grafico, dove ogni cluster è confrontato con se stesso, mostra la massima similarità possibile cioè con valore 1 e cella di colore giallo. Il grafico in Fig. 5.4 ci mostra più nello specifico come i nostri argomenti sono correlati tra di loro e quali sono le relazioni presenti fra i diversi topic. Nello specifico:

- Il Topic 0, che riguarda il cambiamento climatico e l'impatto ambientale umano, mostra un'alta similarità con il Topic 1 (cooperazione internazionale e clima) e il Topic 2 (cambiamento climatico e politica). Questa alta similarità suggerisce che questi topic condividono molti punti comuni, affrontando il cambiamento climatico da prospettive diverse ma interconnesse come la politica e la cooperazione internazionale.
- Il Topic 3, incentrato sull'inquinamento e gli ecosistemi, evidenzia una moderata similarità con i topic 0, 1 e 2, riflettendo le connessioni tra inquinamento, ecosistemi e cambiamento climatico. Inoltre, mostra una similarità con il Topic 5 (industria, inquinamento e impatto ambientale), il che indica che le discussioni sull'inquinamento industriale e i suoi effetti sugli ecosistemi sono frequentemente collegati.
- Il Topic 4, che tratta di crisi, sanità e sicurezza pubblica, presenta un'alta similarità con il Topic 0 (cambiamento climatico e impatto umano), poiché entrambi discutono dell'impatto del cambiamento climatico sulla popolazione umana, inclusi aspetti quali quelli sanitari e di emergenza.
- Il Topic 6, che si focalizza sul riscaldamento globale, mostra una similarità con i topic 0 (cambiamento climatico e impatto umano), 3 (inquinamento ed ecosistemi) e 5 (industria ed inquinamento), evidenziando che il riscaldamento globale è un elemento centrale nelle discussioni sul cambiamento climatico mettendo in risalto la correlazione fra industrie, inquinamento, depurazione degli ecosistemi e riscaldamento globale.
- Il Topic 7, che esplora il dibattito sulla realtà del cambiamento climatico, mostra una bassa similarità con molti altri topic, suggerendo che questo tema è trattato in modo relativamente isolato rispetto ad altri argomenti più concreti sugli impatti e le politiche legate al cambiamento climatico.
- Il Topic 8, che riguarda l'impatto e le controversie della società ExxonMobil, presenta una bassa similarità con molti altri topic, ad eccezione di una moderata similarità con il Topic 5. Questo riflette come le attività delle grandi industrie petrolifere siano discusse principalmente in relazione all'inquinamento industriale.

- Il Topic 9, che tratta del consumo di carne, degli allevamenti e del loro impatto ambientale, mostra una moderata similarità con i topic 0 e 3. Ciò indica che gli impatti degli allevamenti intensivi sul cambiamento climatico e sugli ecosistemi sono correlati fra di loro.

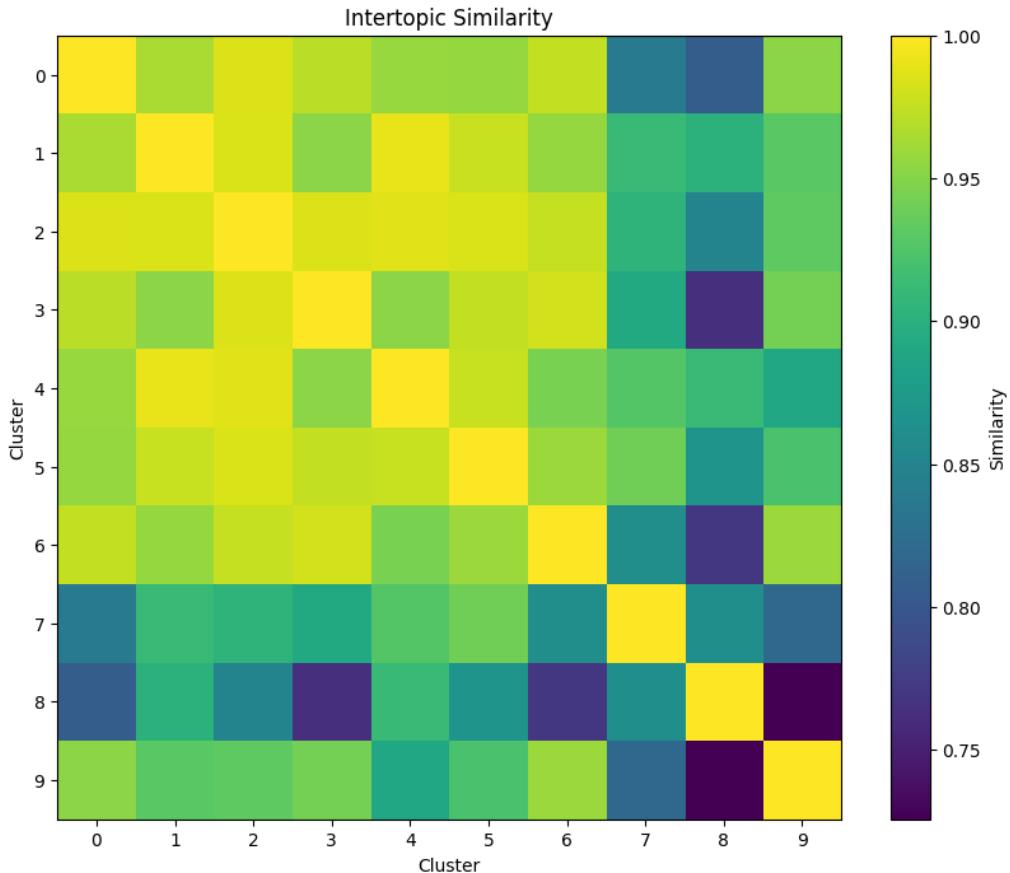


Figura 5.4: Grafico a dispersione di similarità.

Topic	Titolo
0	Cambiamento climatico e impatto umano
1	Cooperazione internazionale e clima
2	Cambiamento climatico e politica
3	Inquinamento ed ecosistemi
4	Crisi, sanità e sicurezza
5	Industria, inquinamento e impatto ambientale
6	Riscaldamento globale
7	Il Cambiamento climatico fra bufala e realtà
8	Impatto e controversie della società ExxonMobil
9	Consumo di carne, allevamenti e impatto ambientale

Per concludere, la heatmap della similarità intertopic rivela come i diversi topic estratti sul cambiamento climatico siano interconnessi. I cluster che trattano argomenti strettamente legati (ad esempio, politica, cooperazione internazionale e impatto umano) mostrano un'alta similarità tra loro, mentre argomenti più specifici o controversi (come il dibattito sulla realtà del cambiamento climatico e le controversie su ExxonMobil) tendono ad essere meno correlati con altri topic.

### 5.1.1 Cambiamento Climatico e Impatto Umano

L'argomento dibattuto più frequentemente e con la percentuale più alta, circa il 21,28%, è quello relativo al cambiamento climatico nel suo complesso, alla natura antropologica di questo cambiamento e dei fattori che lo causano. Le keywords che rappresentano questo tema sono *"Climate, Change, People, Fire, Global, Environment, Emission, Anthropogenic, CO<sub>2</sub>, Gas, Fossil"*. Come sottolineato nel Capitolo 2, è evidente come la comunità scientifica e il pubblico generale riconoscano l'impatto significativo delle attività umane sul clima globale. Diversi studi confermano che l'aumento delle emissioni di gas serra, derivanti principalmente dalla combustione di combustibili fossili, la deforestazione e l'agricoltura intensiva, siano fra i fattori più determinanti nel cambiamento globale [54]. Le emissioni di gas serra, in particolare di anidride carbonica, metano e ossido di azoto, hanno dimostrato di avere un effetto diretto sull'aumento delle temperature medie globali. Le concentrazioni di questi gas nell'atmosfera hanno raggiunto livelli senza precedenti nella storia recente, come evidenziato dai dati raccolti dal Intergovernmental Panel on Climate Change (IPCC) nelle loro valutazioni periodiche [1]. L'aumento delle temperature globali è associato a una serie di impatti climatici, tra cui lo scioglimento dei ghiacci polari, l'innalzamento del livello del mare e l'incremento della frequenza e intensità di eventi meteorologici estremi quali uragani, ondate di calore e precipitazioni intense.

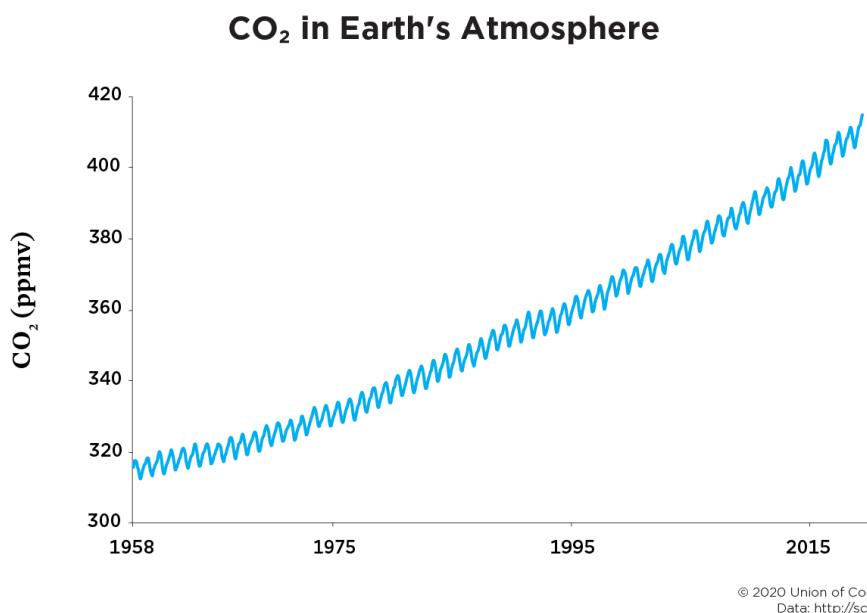


Figura 5.5: Deviazione livelli di CO<sub>2</sub> nell'atmosfera

---

## The History of Atmospheric Carbon Dioxide

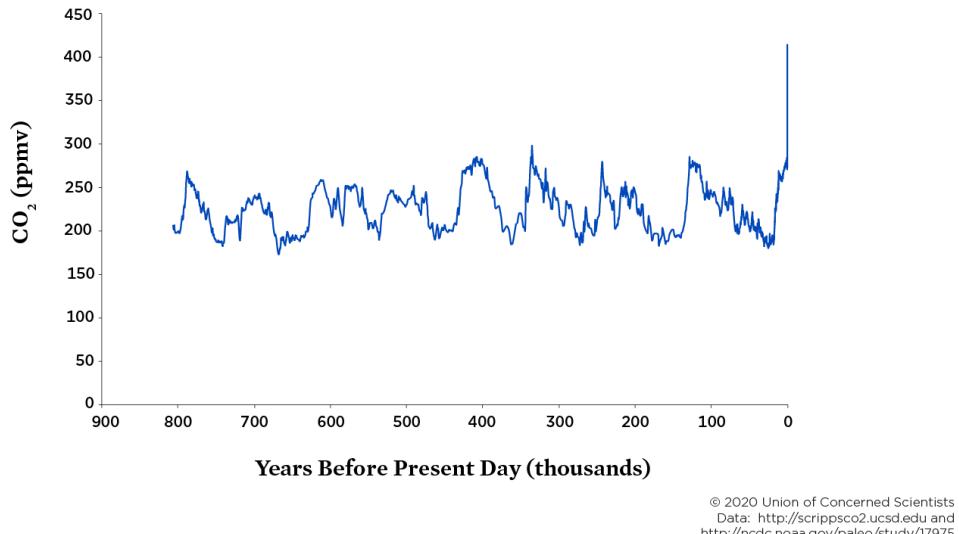


Figura 5.6: Livelli di anidride carbonica

Il grafico in Fig. 5.6 documenta i livelli di anidride carbonica da circa 800.000 anni fa. A partire dal 1760 circa, si verificò un drammatico picco al rialzo attribuito alla Rivoluzione Industriale. Tratto dai dati compilati dalla National Oceanic and Atmospheric Administration e dalla Scripps Institution of Oceanography [55, 56]. La crescente consapevolezza pubblica del ruolo dell'uomo nel cambiamento climatico è il risultato di decenni di ricerca scientifica che hanno documentato il contributo delle attività umane alle modifiche dell'atmosfera terrestre. L'analisi dei dati storici delle concentrazioni di CO<sub>2</sub>, metano e altri gas serra, insieme alla correlazione con le temperature globali, hanno fornito prove incontrovertibili dell'influenza antropica sul clima. Il progressivo aumento delle emissioni di gas serra durante la rivoluzione industriale e il loro aumento continuo nel corso del XX e del XXI secolo sono direttamente connessi all'uso diffuso di combustibili fossili, alla deforestazione su larga scala e alle pratiche agricole intensive. Ciò è supportato dai risultati del topic 0 derivanti dal modello. Se si prova ad esplorare alcuni documenti assegnati al topic, si può denotare come quando si parla di uomo e cambiamento climatico, la consapevolezza è più che tangibile. Le discussioni evidenziano un crescente interesse e una profonda preoccupazione riguardo al ruolo dell'umanità nel plasmare il futuro del nostro pianeta. Questo enunciato non solo sottolinea l'importanza della consapevolezza pubblica, ma indica anche un'inclinazione verso azioni concrete e un impegno individuale e collettivo nel perseguire soluzioni efficaci per mitigare gli impatti del cambiamento climatico. Ciò è anche in linea con i risultati illustrati nel Capitolo 2 sul Programma di sviluppo delle Nazioni Unite e il sondaggio *Peoples' Climate Vote* [11]. Questo argomento, sebbene rappresenti la quota percentuale più elevata, si riferisce in maniera generale al cambiamento climatico e al suo impatto sull'umanità. Tuttavia, saranno i temi successivi a delinearne in modo dettagliato tutte le sfaccettature.

---

Alcuni esempi di documenti concerni a queste considerazioni sono:

- *"State of the Environment report warns impact of climate change 'increasing' and 'pervasive', perhaps humans should have thought about it sooner"*
- *"[@tag]: How we know that climate change is happening—and that humans are causing it [link]. The anthropological imprint on the environment is more than evident, humanity needs to reduce its impact on the environment."*
- *"We are running out of time' on climate change, expert warns: [link] via [@tag] - #climatechange"*
- *"Im constantly on the verge of a climate change rampage, WE ARE KILLING EARTH PEOPLE WAKE UP "*
- *"[@tag]: Closing the #EmissionsGap means increased efforts from all walks of society to fight climate change: [link]"*
- *Scientists: Strong evidence that human-caused climate change intensified 2015 heat waves [link] via [@tag]*

Infine, è utile ai fini della contestualizzazione dei risultati, capire quanto determinate parole sono frequenti e importanti all'interno del topic. Il valore di peso associato dal modello alle parole più frequenti è così composto:

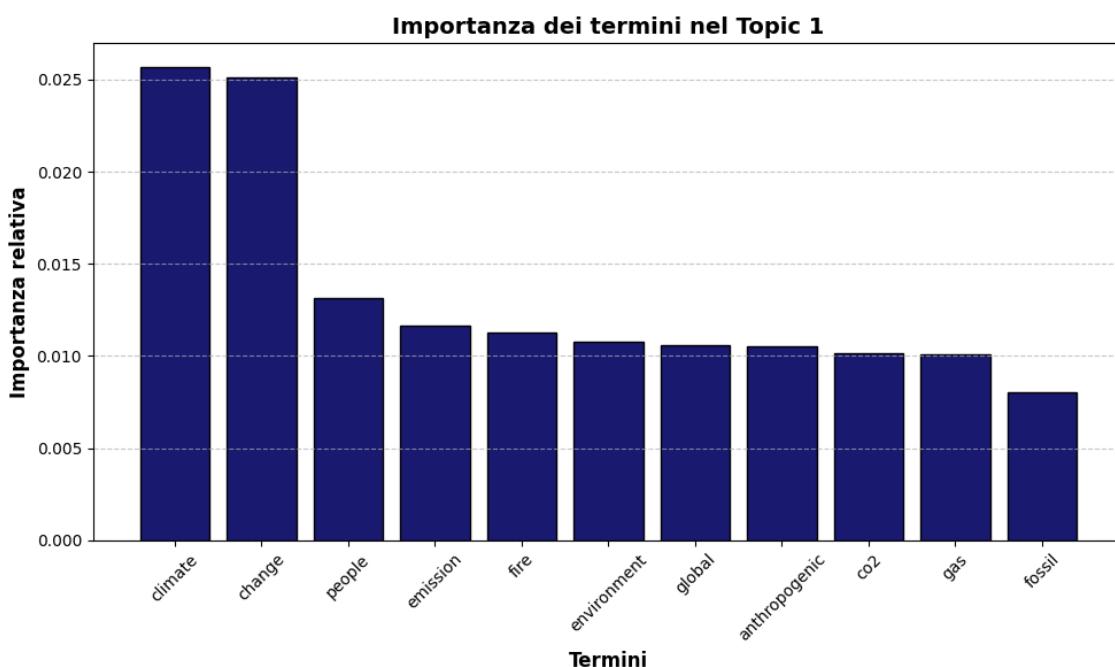


Figura 5.7: Keyword-Peso Cambiamento Climatico e impatto umano

Per esempio l'elevata ricorrenza delle parole *"emissione"*, *"CO<sub>2</sub>"* e *"Fossil"* e l'approfondimento dei testi assegnati al topic dimostrano come le emissioni di carbonio e l'effetto serra siano considerati dagli utenti come i fattori più cruciale che causano al cambiamento climatico. Oppure come la frequenza delle parole *"Anthropogenic"* e *"People"* riportino all'impatto dell'uomo sul clima e viceversa alle conseguenze del cambiamento climatico sulle persone.

---

### 5.1.2 Cooperazione Internazionale e Clima

L'argomento relativo alla cooperazione internazionale in materia di cambiamento climatico emerge come il secondo tema più centrale e ricorrente nei testi analizzati rappresentando il 12.69%. Le keyword che rappresentano questo tema sono: "*Between, Cooperation, Lecture, Importance, Climate, Agreement, Paris, Scientist, Act*". Gli accordi internazionali rappresentano uno strumento cruciale nella lotta contro questa sfida, il più significativo tra questi è l'Accordo di Parigi del 2015, un trattato globale che mira a limitare l'aumento della temperatura media globale ben al di sotto dei 2°C rispetto ai livelli preindustriali, con sforzi per limitarlo a 1.5°C. Questo accordo, firmato da 196 stati, sottolinea l'importanza della cooperazione tra le nazioni per raggiungere obiettivi comuni di mitigazione e adattamento al cambiamento climatico [57]. Oltre all'Accordo di Parigi, la cooperazione internazionale si manifesta attraverso numerose altre iniziative e piattaforme. Organizzazioni come le Nazioni Unite e il Gruppo Intergovernativo sul Cambiamento Climatico (IPCC) [58, 59], giocano un ruolo fondamentale nella facilitazione del dialogo e della collaborazione tra scienziati, politici e attivisti di diverse nazioni. Le conferenze annuali delle Parti (COP) [60], come quelle organizzate sotto l'egida della Convenzione quadro delle Nazioni Unite sui cambiamenti climatici (UNFCCC), rappresentano momenti chiave in cui vengono negoziate e rinnovate le ambizioni climatiche globali. Inoltre, la crescente rilevanza della finanza climatica e il trasferimento tecnologico tra paesi sviluppati e in via di sviluppo evidenziano ulteriormente la necessità di un approccio cooperativo. La condivisione di conoscenze scientifiche, l'implementazione di politiche comuni e il supporto reciproco nelle pratiche di adattamento e mitigazione sono essenziali per affrontare efficacemente le sfide poste dal cambiamento climatico. La cooperazione non solo rafforza la capacità globale di rispondere agli impatti climatici, ma promuove anche la giustizia climatica, assicurando che le nazioni più vulnerabili e meno responsabili delle emissioni storiche ricevano il supporto necessario per proteggere le loro popolazioni e i loro ecosistemi. Le keyword insieme all'analisi dei documenti suggeriscono con chiarezza che le istituzioni internazionali e nazionali giocano un ruolo cruciale nella lotta contro il cambiamento climatico. Sebbene l'azione individuale dei cittadini sia certamente importante, è evidente che i primi passi decisivi debbano provenire dalle istituzioni. La cooperazione internazionale tra le istituzioni rappresenta infatti l'arma più potente nella mitigazione degli impatti climatici e nell'adattamento ai cambiamenti in corso.

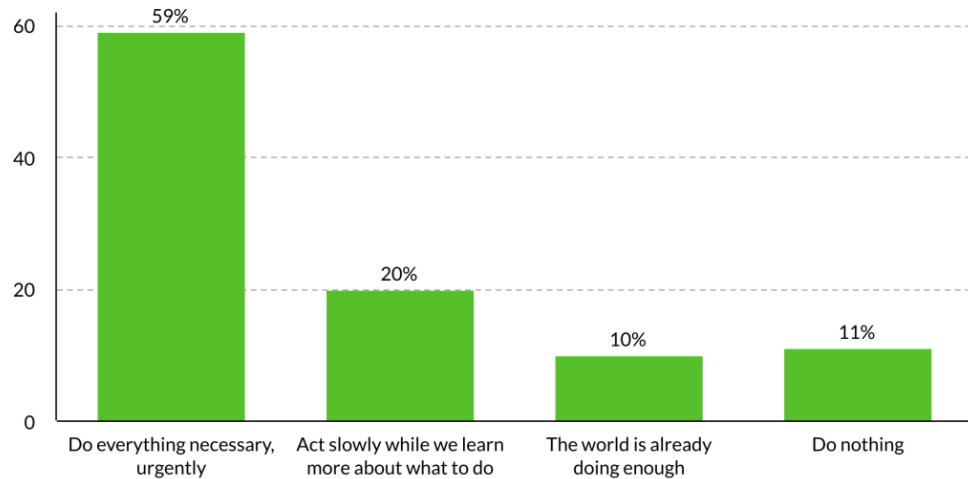


Figura 5.8: Urgenza di risposta da parte delle istituzioni tra le persone che credono nell'emergenza climatica - Fig. 4 IPCC [11]

La cooperazione internazionale tra istituzioni è identificata quindi come la strategia più efficace per affrontare il cambiamento climatico. La concertazione e l'azione collettiva permettono di mobilitare risorse, condividere conoscenze e tecnologie, e implementare politiche armonizzate su scala globale, soprattutto laddove la consapevolezza pubblica non è ampiamente sviluppata.

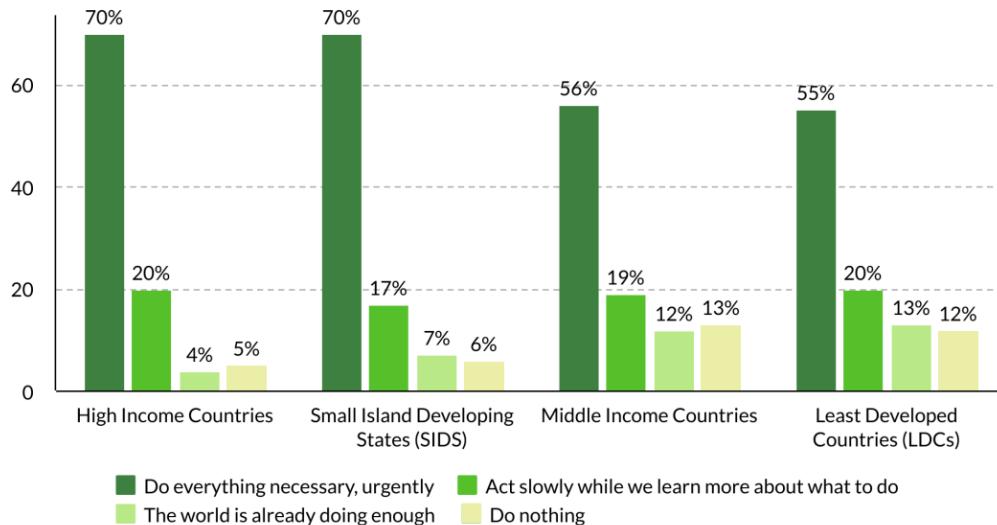


Figura 5.9: Urgenza di risposta da parte delle istituzioni tra le persone che credono nell'emergenza climatica per gruppi di paesi - Fig 5 IPCC [11]

Alcuni esempi illustrativi dei documenti assegnati a questo topic:

- ”[@tag]: All we need is. Collaboration. Communication. Cooperation. To achieve climate change action! [@tag] [@tag] [@tag]”
- ”[@tag]: In the new new world, China lectures the United States about climate change and the importance of cooperation between states”

- ”[@tag]: Last year, 190+ countries signed the #ParisAgreement promising help to those worst hit by climate change. #Promises”
- ”[@tag]: To keep global warming under 1.5C, we need to accelerate #ParisAgreement implementation; increase our ambition. [@tag]”
- ”[@tag]: We should actually all be doing more individually to fight climate change, right? Or maybe we need more international cooperation between institutions?”

Guardando il peso e la frequenza delle parole queste considerazioni appaiono in modo ancora più nitido, termini come *”Between”*, *”Cooperation”* e *”Agreement”* indicano l’interconnessione e la necessità di collaborazione tra diverse entità, siano esse nazioni, organizzazioni o individui. Queste parole chiave sottolineano che la sfida del cambiamento climatico non può essere affrontata isolatamente; piuttosto, richiede uno sforzo collettivo e coordinato. Il termine *”Paris”* si riferisce direttamente all’Accordo di Parigi, un punto di riferimento fondamentale nella politica climatica internazionale. I termini *”Lecture”* e *”Scientist”* sottolineano l’importanza dell’educazione e della sensibilizzazione sulle tematiche climatiche. Le conferenze sono strumenti fondamentali per informare e coinvolgere il pubblico, nonché per formare nuovi esperti e leader nel campo della sostenibilità, ponendo una grande fiducia, in particolar modo per chi crede nel cambiamento climatico, negli scienziati. Infine *”Importance”* indica che nelle discussioni è fortemente percepita la rilevanza delle azioni cooperative e delle politiche climatiche. Questo termine riflette la consapevolezza diffusa della necessità, come mostrato dai grafici 5.8 e 5.9, di interventi tempestivi e concertati per affrontare la crisi climatica, sottolineando l’urgenza e la serietà del problema.

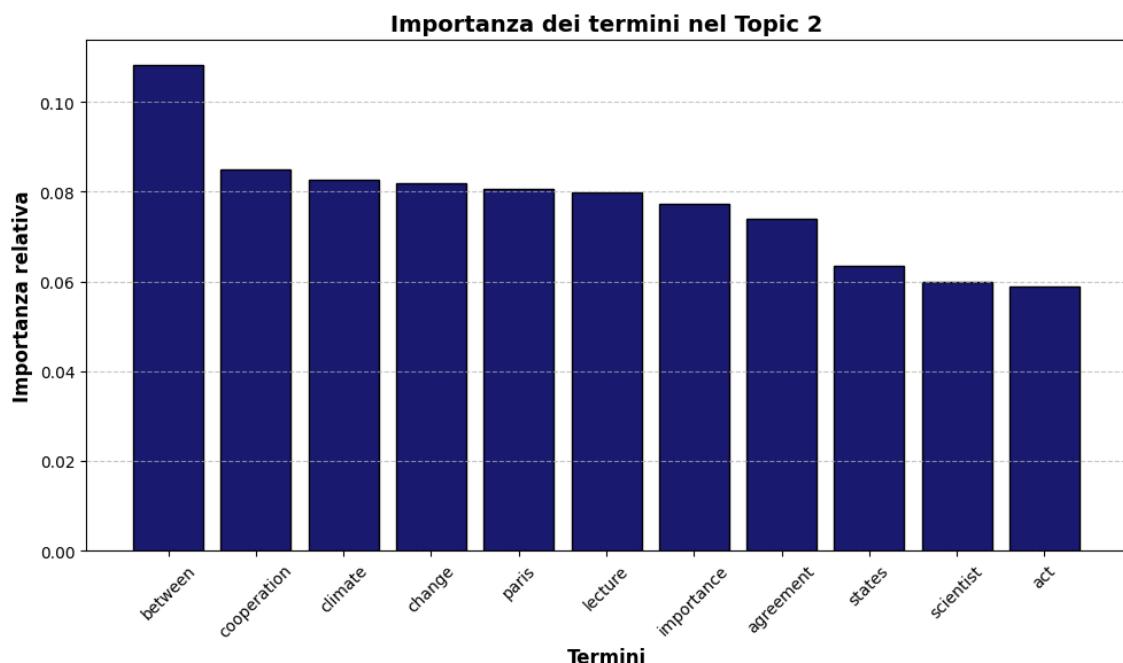


Figura 5.10: Keyword-Peso Cooperazione internazionale e Clima

### 5.1.3 Cambiamento Climatico e Politica

Dall’analisi emerge una stretta correlazione tra cambiamento climatico e politica, il topic 2, dal titolo *”Cambiamento climatico e politica”* presenta la connessione fra politica e cambiamento climatico. Le keyword che rappresentano questo tema sono: *Trump, Ivanka, Merkel, Melania, Schwarzenegger, Donald, Policy, People, Government, Election, Dicaprio*. Come presentato dalla Figura 5.4, l’argomento mostra un’ampia similitudine con il topic 1, questo perchè non si può parlare di cooperazione internazionale fra le istituzioni senza parlare di politica. Sebbene molte delle istituzioni citate nel paragrafo 5.1.2 siano di carattere non governativo, l’impronta politica nelle scelte delle nazioni sulle azioni da perseguire rimane il fattore primario. Nell’ambito delle discussioni analizzate, emergono con frequenza figure politiche di rilievo quali Ivanka Trump, Donald Trump, Angela Merkel e Melania Trump. Ivanka Trump è frequentemente menzionata in relazione alle sue posizioni su politiche ambientali e iniziative sostenibili, sottolineando il suo coinvolgimento nel dibattito pubblico sul cambiamento climatico. Donald Trump, invece, rappresenta forse il punto centrale nei dibattiti riguardanti le decisioni governative degli Stati Uniti sul cambiamento climatico e dei dibattici politici in generale, evidenziando una significativa polarizzazione delle opinioni pubbliche. L’ex cancelliere Angela Merkel si distingue come una figura di primo piano, riconosciuta per il suo ruolo di leadership nella promozione di politiche climatiche avanzate all’interno dell’Unione Europea. L’influenza di personalità politiche e pubbliche come Arnold Schwarzenegger e Leonardo DiCaprio, mette in luce come le celebrità possano fungere da catalizzatori per l’attenzione pubblica e influenzare l’agenda politica. L’intersezione tra politica e cambiamento climatico è ulteriormente accentuata dalle discussioni sulle elezioni governative [17], in cui le posizioni dei candidati sulle questioni ambientali diventano un fattore determinante per gli elettori. Queste figure non solo determinano l’agenda politica, ma contribuiscono anche a modellare le percezioni e le credenze della popolazione riguardo alla crisi climatica. Le asserzioni fatte da leader politici come Donald Trump, ad esempio, che ha spesso minimizzato l’importanza del cambiamento climatico, hanno avuto un impatto significativo sulla polarizzazione delle opinioni pubbliche. Le sue dichiarazioni e politiche hanno infatti alimentato lo scetticismo climatico tra una parte dell’elettorato, influenzando negativamente la percezione della necessità di azioni urgenti contro questa sfida. Quest’ultima considerazione è ben visibile all’interno dei documenti assegnati al topic, dove le dichiarazioni fatte da Donald Trump, sono il punto focale del discorso.

- ”[@tag]: [@tagDonaldTrump] vows to cancel ‘billions’ in climate change funds to UN, adds ‘I’m an environmentalist’ [link]”
- ”[@tag]: Trump, who once called climate change a hoax by the Chinese, tells Michigan crowd: ‘I’m an environmentalist.’ ”
- ”[@tag] [@tagLeonardoDiCaprio] why aren’t you vegan? meat industry contributes GREATLY to climate change and deforestation! ”
- ”Trumps policy’s on climate change, if we make it that far, is what’s for sure going to kill us.”
- ”politics: Merkel’s visit this week will test whether allies can persuade Trump not to blow up their efforts on global warming. . . . .”

- ”[@tag]: Angela Merkel is a former research scientist, and Trump thinks global warming is a Chinese conspiracy theory.”
- ”The most disappointing thing is that climate change should have brought government parties together, instead it has divided us’ #4corners”
- ”[@tag]: Arnold Schwarzenegger teams up with Macron to fight climate change: [link]”
- ”[@tag]: bro our next president doesn’t even believe in fucking climate change. and his VP thinks being gay is a treatable disease.”

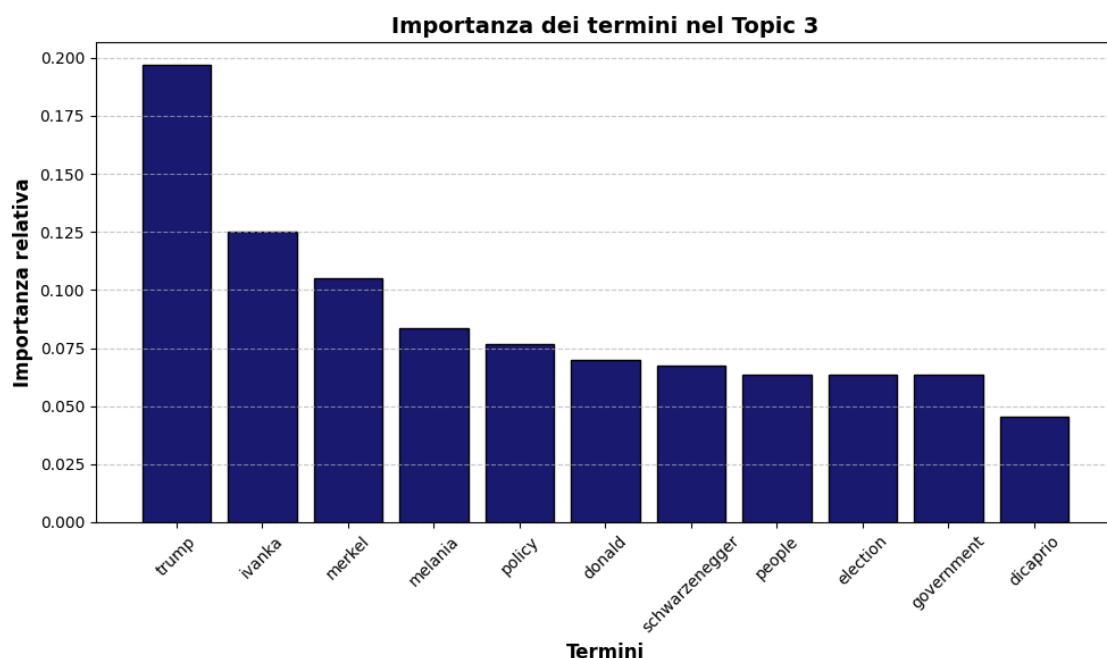


Figura 5.11: Keyword-Peso Cambiamento climatico e Politica

I testi riflettono una forte polarizzazione e divisione nelle opinioni pubbliche e politiche riguardo al cambiamento climatico. Da un lato, ci sono critiche e scetticismo verso le dichiarazioni e le politiche di Donald Trump, viste come incoerenti e pericolose. Dall’altro lato, si evidenzia l’influenza positiva delle celebrità e dei leader internazionali come Angela Merkel e Arnold Schwarzenegger nel promuovere azioni significative contro il cambiamento climatico. La divisione politica interna ed internazionale emerge chiaramente, con implicazioni significative per la cooperazione globale e l’efficacia delle politiche climatiche.

L'impatto delle politiche governative sulle azioni climatiche è inequivocabile, poiché le scelte politiche delineano le strategie nazionali ed internazionali per affrontare il cambiamento climatico. Come illustrato nella Figura 5.12, le politiche ambientali adottate durante le presidenze di Donald Trump e Joe Biden rendono evidente come l'orientamento politico di un'amministrazione possa influenzare in modo significativo la direzione delle politiche climatiche sia a livello nazionale sia internazionale.

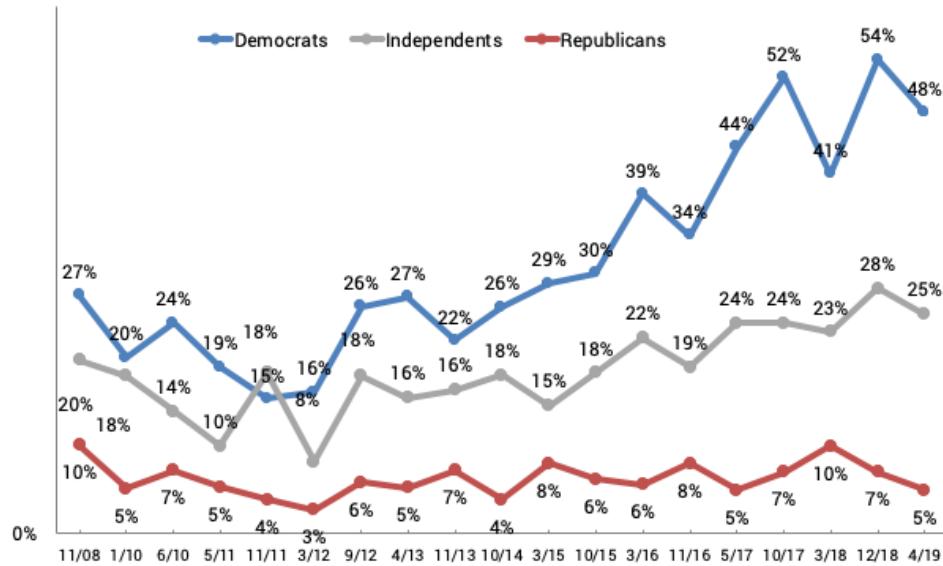


Figura 5.12: Il cambiamento climatico deve essere una priorità primaria per il Presidente e il Congresso. - Global Warming as a Priority Issue: Political Trends [61]

---

### 5.1.4 Inquinamento ed Ecosistemi

Il topic "Inquinamento ed Ecosistemi" è stato identificato come il quarto topic predominante. Tra i 343.943 documenti analizzati, 31.596 sono stati classificati sotto questo topic, rappresentando circa il 9,19% del totale. Questo dato mette in evidenza che quasi un decimo delle conversazioni si concentra sull'impatto dell'inquinamento sugli ecosistemi, suggerendo che questa è una preoccupazione di primo piano per il pubblico. Le keyword associate a questo topic sono: "*Cracks, Anemone, Bleached, Snow, Ice, Plastic, Ecosystem, Sea, Plant*". L'inquinamento degli ecosistemi rappresenta una delle principali "cracks" nella sostenibilità ambientale globale. Le sue implicazioni negative si manifestano in molteplici forme, dalle acque marine agli habitat terrestri. Un esempio emblematico è l'ecosistema corallino, dove la presenza di sostanze inquinanti, come la plastica, può avere effetti devastanti. Gli anemoni, esseri vitali all'interno di questo delicato ambiente, sono esposti ai rischi derivanti dalla contaminazione da plastica, che può causare danni irreversibili alla loro salute e alla biodiversità circostante. Inoltre, fenomeni come lo sbiancamento dei coralli, causato principalmente dall'aumento della temperatura dell'acqua e dall'acidificazione degli oceani, rappresentano un'ulteriore minaccia per gli ecosistemi corallini. Lo scioglimento dei ghiacci e la diminuzione della copertura nevosa influenzano direttamente la disponibilità di habitat per numerose specie, mettendo a rischio la sopravvivenza di piante e animali che dipendono da queste condizioni ambientali. La plastica è diventata un problema ambientale sempre più grave. Nel 1950, il mondo produceva più di 2 milioni di tonnellate di plastica all'anno [62]. Nel 2015, questa produzione annuale è aumentata fino a 400 milioni di tonnellate circa, contribuendo al proliferare dei rifiuti plastici nell'ambiente. Attualmente, circa 14 milioni di tonnellate di plastica finiscono negli oceani ogni anno, danneggiando gli habitat della fauna selvatica e gli animali che vi abitano. Senza azioni immediate, la crisi della plastica potrebbe crescere fino a rilasciare negli oceani 29 milioni di tonnellate metriche all'anno entro il 2040 [63]. Come anche evidenziato dalla curva visualizzabile nella Figura 5.13, l'aumento di produzione di plastica è in costante crescita. Sconcertante è il fatto che il 91% di tutta la plastica prodotta non venga riciclata [64], rappresentando non solo uno dei più grandi problemi ambientali del nostro tempo, ma anche un altro enorme fallimento del mercato. Considerando che la plastica impiega 400 anni per decomposizione, saranno necessarie molte generazioni prima che cessi di esistere [65]. Non è possibile prevedere quali saranno gli effetti irreversibili dell'inquinamento da plastica sull'ambiente a lungo termine.

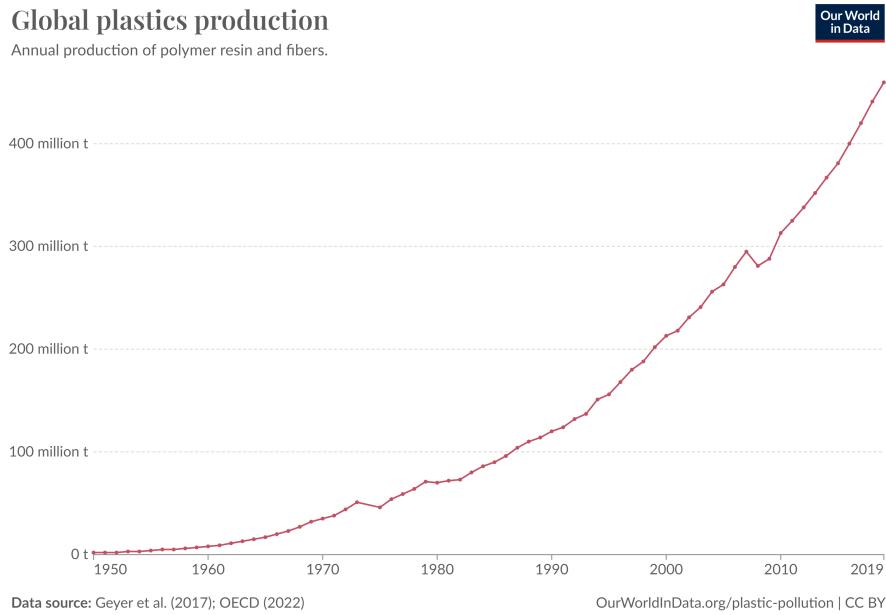


Figura 5.13: Produzione globale di plastica [65]

Dai risultati dell’analisi emerge un altro sotto-argomento di rilevanza significativa, concernente lo scioglimento dei ghiacci e l’innalzamento del livello del mare. La crisi climatica attuale manifesta un impatto accentuato sull’Artico, dove le temperature si registrano con un aumento del doppio rispetto alla media globale. Parallelamente, i livelli del mare evidenziano un tasso di aumento raddoppiato rispetto a quanto osservato nel corso della maggior parte del XX secolo. Attualmente, come osservabile anche in Figura 5.14, si osserva un incremento medio di 3,2 millimetri annui a livello mondiale, prospettando un ulteriore aumento fino a circa 0,7 metri entro la fine del presente secolo [66].

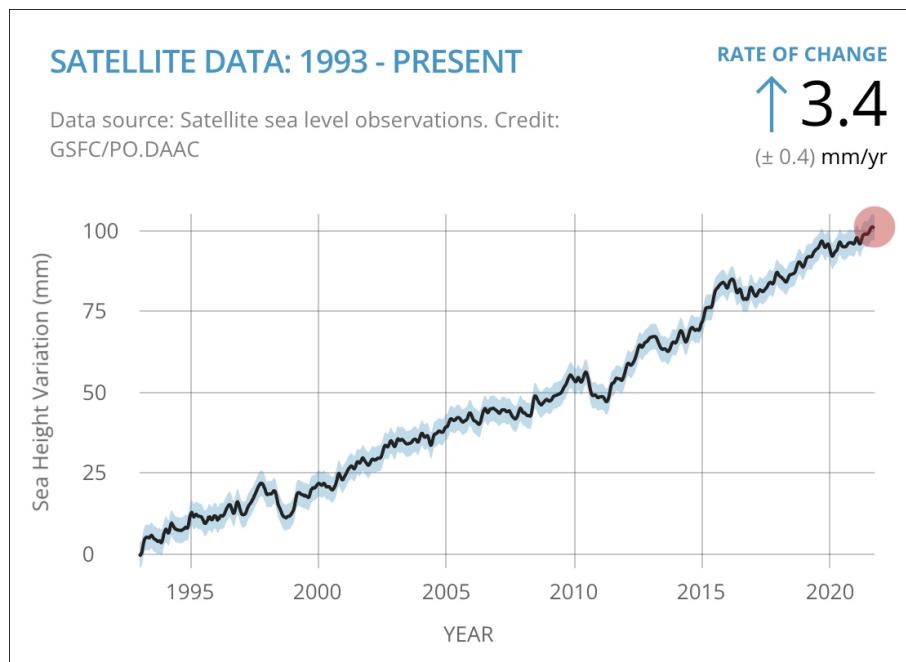


Figura 5.14: Livello del mare [66]

---

Nell'Artico, il Greenland Ice Sheet, cioè la calotta di ghiaccio in Groenlandia seconda in grandezza solo al quella Antartica, rappresenta una delle principali fonti di preoccupazione per gli esperti, poiché il suo scioglimento è identificato come il principale contribuente all'innalzamento dei livelli del mare. Questo fenomeno è strettamente correlato al disgelo del ghiaccio terrestre, il quale costituisce un fattore predominante nell'aumento complessivo dei livelli oceanici [63, 67]. Parallelamente, il continente antartico contribuisce all'innalzamento del livello del mare con una media di circa 1 millimetro all'anno, rappresentando una frazione significativa dell'aumento annuale a livello globale. I dati recenti del 2023 indicano che il continente antartico ha subito una perdita stimata di circa 7,5 trilioni di tonnellate di ghiaccio dal 1997 [68]. Ulteriori evidenze riguardano il collasso dell'ultima piattaforma di ghiaccio rimasta intatta in Canada nell'Artico, il quale ha subito una perdita del 40% della sua superficie, equivalente a circa 80 chilometri quadrati, in soli due giorni alla fine di luglio, come riportato dal Canadian Ice Service [68]. L'aumento globale della temperatura ha non solo influenzato la superficie, ma è anche la principale causa dell'acidificazione degli oceani. Anche la più piccola variazione nella scala del pH può avere un impatto significativo sull'acidità dell'oceano. L'acidificazione degli oceani ha impatti devastanti sugli ecosistemi marini e sulle specie, sulle loro reti alimentari, e provoca cambiamenti irreversibili nella qualità dell'habitat. Una volta che i livelli di pH diventano troppo bassi, organismi marini come le ostriche, le loro conchiglie e lo scheletro potrebbero persino iniziare a sciogliersi. Tuttavia, uno dei più grandi problemi ambientali derivanti dall'acidificazione degli oceani è lo sbiancamento dei coralli e la successiva perdita delle barriere coralline, rappresentando una relazione coerente con la preoccupazione dei testi analizzati. Questo è un fenomeno che si verifica quando le crescenti temperature oceaniche disturbano la relazione simbiotica tra le barriere coralline e le alghe che vi vivono all'interno, allontanando le alghe e causando la perdita dei colori naturalmente vibranti delle barriere coralline. Alcuni scienziati hanno stimato che le barriere coralline sono a rischio di essere completamente cancellate entro il 2050 [63]. Il livello del mare avrà un impatto devastante su coloro che vivono nelle regioni costiere: secondo il gruppo di ricerca e advocacy Climate Central [69], l'aumento del livello del mare in questo secolo potrebbe inondare aree costiere che ora ospitano da 340 milioni a 480 milioni di persone, costringendole a migrare in aree più sicure e contribuendo alla sovrappopolazione e allo sforzo delle risorse nelle aree in cui migrano. Bangkok (Thailandia), Ho Chi Minh City (Vietnam), Manila (Filippine) e Dubai (Emirati Arabi Uniti) sono tra le città più a rischio di aumento del livello del mare e inondazioni [67, 63]. La forte rilevanza di determinati termini quali "*Anemone*" , "*Bleached*" e "*Sea*" rivelano questa correlazione fra la preoccupazione delle persone, a volte anche inconsapevole di determinati dati scientifici e le concentrazioni di pH negli oceani e il loro innalzamento.

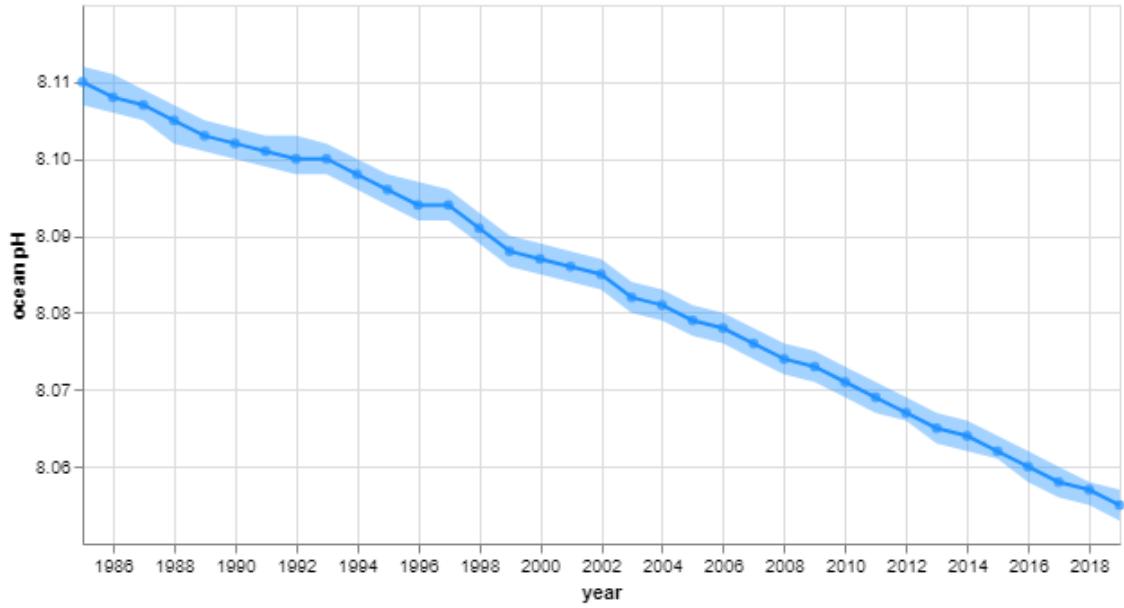


Figura 5.15: Livelli di pH nell’oceano per anno [70]

Tutte queste condiscernizioni sono ben riflesse dalle preoccupazioni espresse nei documenti assegnati nei topic e ben visibili dalle keyword:

- ”[@tag]: *No snow where there should be. Bleached anemones. Cracks in the ice. This is what climate change looks like: [link]*”
- ”[@tag]: *They push climate change down your throat so you don’t think about the massive pesticide and plastic pollution that’s go... .*”
- ”*hand to fight global warming. Plant more trees, don’t waste water. Don’t use or burn plastics. Pl don’t delete message without forwarding..2*”
- ”[@tag]: *Study finds that marine reserves are helping ecosystems cope with climate change. Read more: [link]... .*”
- ”[@tag]: *“Rising temperatures due to climate change have pushed corals beyond their tolerance levels [link]*”
- ”*The fact that people still think that sea level rise is the only thing that comes from global warming blows my mind... . ocean currents, people dying, anemones bleaching, animals that live in the snow dying: [link]*”
- ”[@tag]: *Urgent action on climate change and marine plastic pollution needed to #SaveOurOcean, help food security, #SDGs... .*”

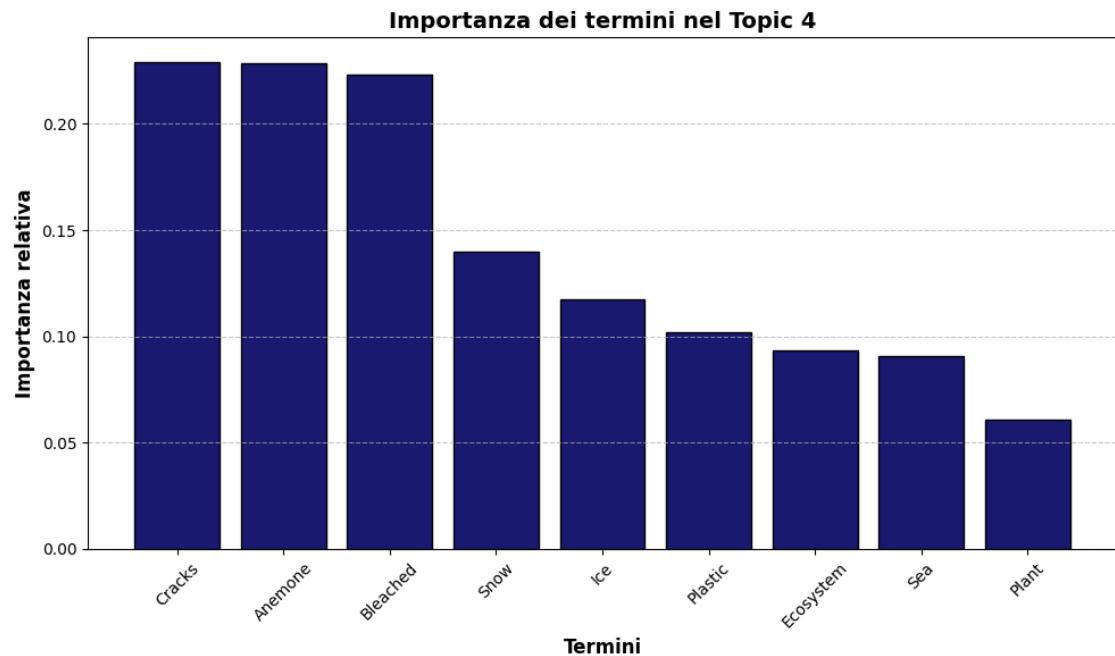


Figura 5.16: Keyword-Peso Inquinamento ed Ecosistemi

È fondamentale comprendere l'impatto sinergico di tutte queste variabili sull'ecosistema nel suo complesso e adottare strategie di mitigazione e adattamento per proteggere la biodiversità e garantire la resilienza degli ecosistemi alle pressioni antropiche. È proprio per quest'ultima considerazione, come ci mostra anche il grafico della simililità in Figura 5.4, che questo argomento è strettamente correlato al topic 1 (Cooperazione internazionale e clima), al topic 2 (Cambiamento climatico e politica) e al topic 5 (Industria, inquinamento e impatto ambientale). In primo luogo, la politica svolge un ruolo fondamentale nel plasmare le politiche ambientali nazionali e internazionali, nonché nell'adozione di misure normative volte a regolare le attività industriali e a mitigare l'inquinamento degli ecosistemi. La cooperazione internazionale diventa essenziale nel contesto della globalizzazione, poiché le questioni ambientali, come l'inquinamento degli ecosistemi, non conoscono confini nazionali e richiedono un impegno congiunto a livello globale per essere efficacemente affrontate. La comprensione e la gestione dell'impatto ambientale delle industrie richiedono politiche rigorose, incentivi per pratiche sostenibili e una cooperazione internazionale per garantire che le aziende operino nel rispetto dell'ambiente e della biodiversità.

---

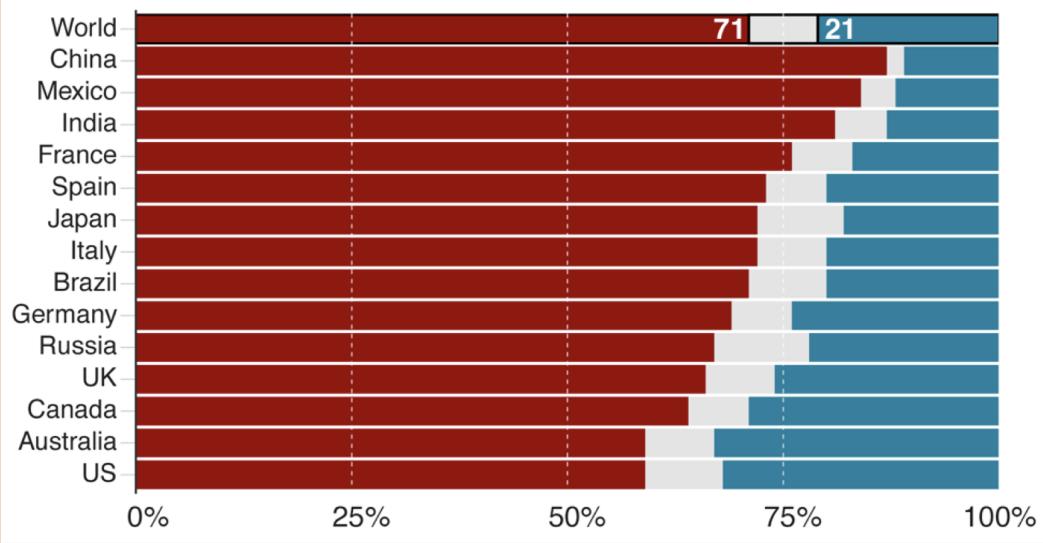
### 5.1.5 Crisi, Sanità e Sicurezza pubblica

I risultati dell'analisi rivelano che è presente una relazione intricata tra il Covid-19 e la crisi, il cambiamento climatico e le dinamiche politiche, riflettendo una complessa interazione tra questi fenomeni e le loro implicazioni per l'umanità. L'argomento legato a queste dinamiche rappresenta il 6,2% dei documenti totali e le keyword identificate, quali *"Disease"*, *"Virus"*, *"Emergency"*, *"Flu"*, *"Smoke"*, *"Exhausted"*, *"People"*, *"Rejected"*, *"Covid"* e *"Advice"*, suggeriscono un quadro articolato di questioni che includono la gestione delle emergenze sanitarie, la percezione pubblica, e le politiche governative. Una delle tendenze emerse riguarda l'impatto del Covid-19 sulle politiche ambientali e il cambiamento climatico, in particolare durante la pandemia, i lockdown e le restrizioni hanno portato a una riduzione temporanea delle emissioni di inquinanti atmosferici e dei livelli di smog in molte città. Tuttavia, questa diminuzione è stata spesso accompagnata da un aumento dell'uso di materiale plastico monouso e una maggiore produzione di rifiuti sanitari, sollevando preoccupazioni riguardo gli impatti a lungo termine sull'ambiente. L'emergenza sanitaria ha evidenziato la necessità di una maggiore resilienza e adattamento ai cambiamenti climatici. In parallelo, si osserva una dinamica politica complessa riguardo la gestione della pandemia e dei cambiamenti climatici. Le risposte dei governi alla crisi sanitaria e ambientale variano ampiamente, con alcune nazioni che adottando misure drastiche per contenere la diffusione del virus e mitigare gli effetti del cambiamento climatico, mentre altre ritardano o negano l'urgenza di agire. Queste differenze riflettono divergenze ideologiche, economiche e geopolitiche che influenzano le politiche pubbliche e la cooperazione internazionale su questioni cruciali come la salute globale e la sostenibilità ambientale. Questo si riflette altresì nelle convinzioni che gli individui nutrono al riguardo delle problematiche globali, benché di natura diversa, richiedono in tutti casi un approccio collaborativo. La Figura 5.17 ci mostra come a livello globale una significativa maggioranza del 71% delle persone consideri il cambiamento climatico una crisi altrettanto seria quanto quella del Covid-19. Tuttavia, esistono variazioni rilevanti tra i diversi paesi. La Cina e il Messico mostrano i più alti livelli di accordo, con rispettivamente circa l'87% e l'86%, indicando una crescente consapevolezza degli impatti climatici in queste nazioni. Al contrario, gli Stati Uniti e l'Australia presentano le percentuali più alte di disaccordo, rispettivamente intorno al 40% e al 32%, suggerendo una percezione pubblica differente o politiche climatiche meno stringenti. In Europa, la Spagna e l'Italia manifestano una forte preoccupazione, mentre la Germania e il Regno Unito mostrano livelli di accordo leggermente inferiori. Questi dati suggeriscono che, sebbene vi sia una consapevolezza globale della gravità del cambiamento climatico, esistono differenze significative tra le nazioni che potrebbero influenzare la capacità di raggiungere un consenso internazionale su politiche climatiche efficaci. La comparazione con il Covid-19, una crisi globale immediata e tangibile, sottolinea l'urgenza percepita della minaccia climatica, offrendo un'opportunità per promuovere azioni climatiche coordinate a livello internazionale.

## Most people think climate change is as serious as coronavirus

Percentage of people who agree that, in the long term, climate change is as serious a crisis as Covid-19

■ Agree ■ Not sure ■ Disagree



Source: Ipsos MORI

BBC

Figura 5.17: Percentuale di persone che credono, che a lungo termine, il cambiamento climatico sia una crisi seria quanto quella da Covid 19 [71]

Sebbene si possa pensare che una realtà come quella delle malattie infettive e della sanità sia ben lontana dagli effetti del cambiamento climatico, gli impatti di quest'ultimo sono in realtà evidenti in numerosi studi. Ad esempio, uno studio condotto da *Mora, Camilo and McKenzie, Tanya and Gaw, Ilyssa M. Et Al.*[72], dimostra chiaramente la correlazione tra l'aumento delle temperature globali e la diffusione di malattie infettive. oltre la metà delle malattie infettive nel mondo sono aggravate dai cambiamenti climatici. Delle 375 malattie infettive studiate, 218, il 58% sono state aggravate dal riscaldamento globale. La Fig. 5.18 ci mostra che la prima parte (a) contenente la maggior parte delle malattie, il 78% per la precisione, è esclusivamente aggravata dai pericoli climatici, mentre una porzione minore, il 19% è sia aggravata che diminuita e solo una piccola parte, il 3% è esclusivamente diminuita. Questo suggerisce che i cambiamenti climatici tendono principalmente ad aggravare le malattie piuttosto che ad alleviarle. La seconda parte (b) illustra il sottoinsieme delle malattie aggravate dai pericoli climatici rispetto a tutte le malattie infettive segnalate. Qui si osserva che il 58% delle malattie infettive segnalate è aggravato dai pericoli climatici, mentre il 42% non lo è. Inoltre, un gruppo di malattie, circa il 18% non segnalate nelle fonti GIDEON/CDC, risulta aggravato dai pericoli climatici. Questo insieme di dati evidenzia l'ampia portata con cui i cambiamenti climatici possono influenzare la diffusione e la gravità delle malattie infettive, sottolineando l'importanza di considerare i fattori climatici nella gestione e prevenzione delle malattie a livello globale.

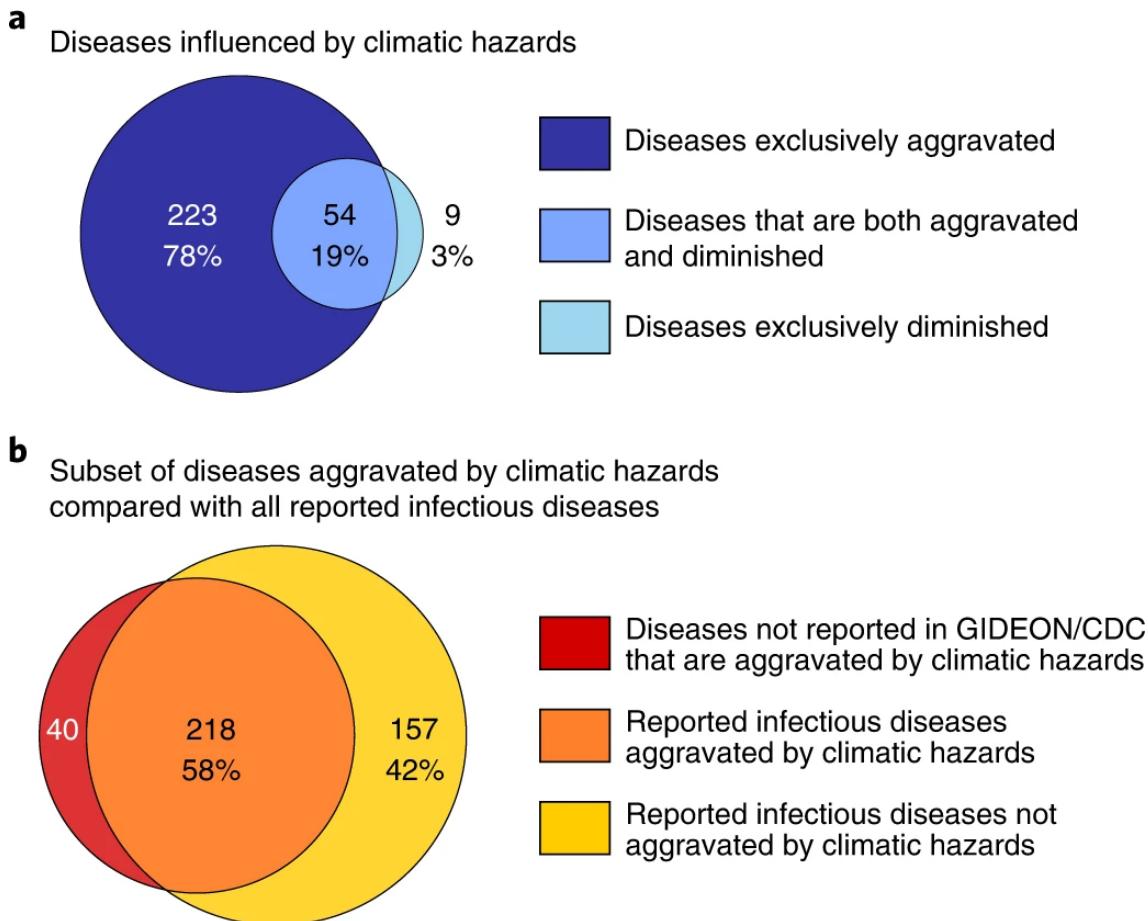


Figura 5.18: Impatto dei pericoli del cambiamento climatico sulle malattie [72]

Questi dati ci forniscono evidenze ben fondate riguardo l'influenza dei cambiamenti climatici sulla sicurezza nella vita delle persone. Proprio per questo è importante sottolineare che tali preoccupazioni e considerazioni sono chiaramente riscontrabili anche nelle opinioni delle persone, come evidenziato dalle keyword e dai documenti assegnati al topic. Questo suggerisce una consapevolezza diffusa e condivisa sia nella comunità scientifica che tra il pubblico, riflettendo una preoccupazione globale che va oltre i meri dati statistici. Alcuni esempi di documenti assegnati al topic:

- *"Most staple food crops are vulnerable to climate change and the emergence of uncontrolled diseases that backfire on us"*
- *"#flu dramatic climate changes in India favouring infectious diseases to grow. #stayprotected"*
- *"[@tag]: How climate change is pushing animal (and human diseases) to new places. [link]"*
- *"[@tag]: Ancient? viruses thawing out because of global warming bothers me more than I thought it would. We are so fucked."*
- *"[@tag]: "As climate changes, more new insects will appear, and those insects will bring new viruses and new disease"*

- 
- ”*Climate change may have helped spread Zika virus according to WHO scientists [link] #Americas*”
  - ”*[@tag]: Surely the Covid 19 pandemic has led to a decrease in pollution but right now that we were starting to use less disposable waste let's start with all these masks.. estimated production of disposable masks: [link]*”
  - ”*[@tag]: If Trump continues to treat Covid 19 as he does the devastating impacts of climate change, we are in good hands*”

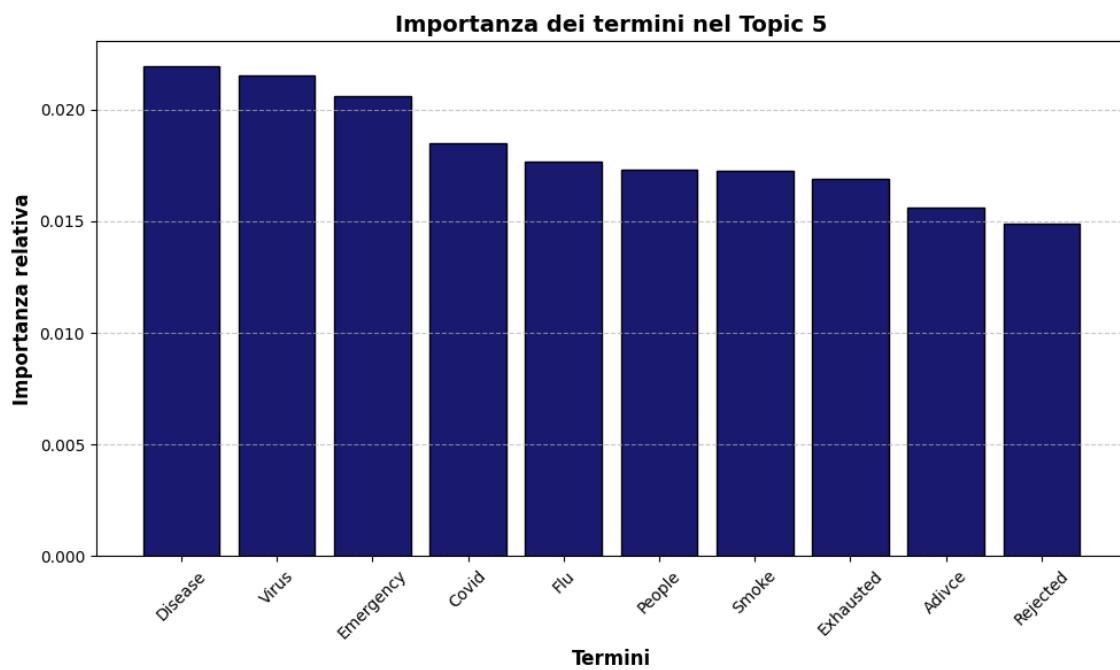


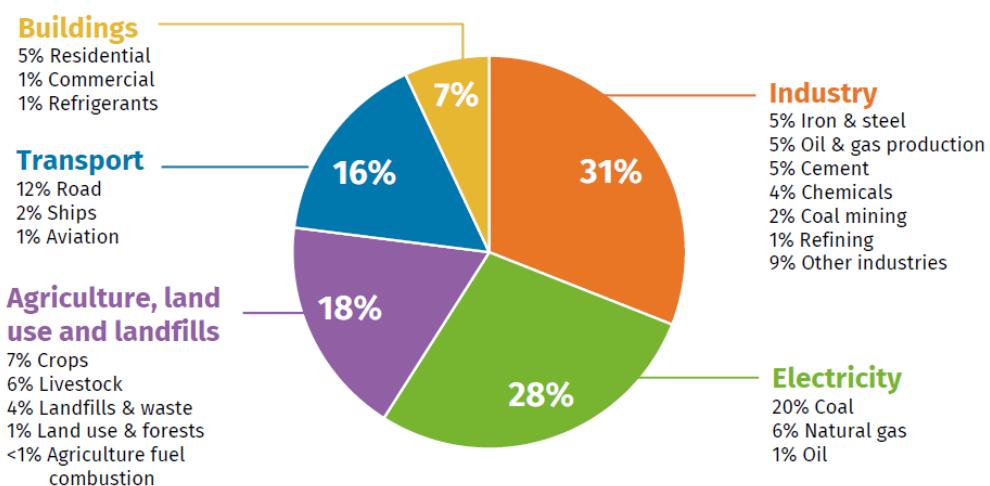
Figura 5.19: Keyword-Peso Crisi, Sanità e sicurezza pubblica

---

### 5.1.6 Industria, Inquinamento e Impatto Ambientale

I risultati presentano una relazione fra l'industria e il cambiamento climatico, il topic rappresenta circa il 5,67% dei documenti totali e le parole chiave che spiegano questo tema sono: *"Manufacturing, Chinese, Concept, Corporate, Oil, Emission, Order, Global, Greenhouse, Fossil"*. La similarità presente con il topic 3 (Inquinamento ed Ecosistemi), il topic 6 (Riscaldamento Globale) e il topic 0 (Cambiamento climatico e Impatto umano) evidenzia e mette in risalto la correlazione di questo argomento in un argomento più ampio cioè quello dell'impatto umano sulla terra e in particolar modo sulle temperature globali e sulla depurazione degli ambienti. L'industria moderna rappresenta una delle principali fonti di inquinamento e impatto ambientale a livello globale, il settore manifatturiero, che comprende una vasta gamma di attività produttive, contribuisce significativamente alle emissioni di gas serra (greenhouse) e al degrado ambientale. Dall'analisi dei risultati vi è al centro l'industria manifatturiera cinese, essendo tra le più grandi e influenti al mondo, gioca un ruolo cruciale in questo contesto. La rapidissima crescita economica della Cina ha comportato un'espansione senza precedenti delle attività industriali, con conseguenze ambientali rilevanti. Uno degli aspetti chiave da considerare è il concetto di responsabilità aziendale, le imprese globali, e in particolare quelle che operano nei settori ad alta intensità di emissioni come quello petrolifero, hanno un ruolo fondamentale nella gestione delle emissioni di inquinanti. Le aziende del settore petrolifero, ad esempio, sono tra i maggiori emettitori di anidride carbonica e altri gas serra, contribuendo in modo significativo al cambiamento climatico. Le emissioni derivanti dalla combustione di combustibili fossili rappresentano una delle principali cause dell'aumento delle concentrazioni di gas serra nell'atmosfera. Questa considerazione verrà poi approfondita per il caso specifico della ExxonMobil che è rappresentato dal topic 8 (Impatto e controversie della società ExxonMobil). Il concetto di lobbying e ordine cooperativo globale emergono come una struttura attraverso la quale le multinazionali influenzano le politiche ambientali e regolatorie nei paesi in cui operano. Questo ordine è spesso caratterizzato da una disparità tra le normative ambientali dei paesi sviluppati e quelli in via di sviluppo, con quest'ultimi che talvolta accettano standard meno rigidi per attrarre investimenti stranieri. Tuttavia, la crescente consapevolezza ambientale ha portato a una maggiore pressione sulle imprese affinché adottino pratiche più sostenibili. Le politiche aziendali green, volte a ridurre l'impronta di carbonio e promuovere l'uso di energie rinnovabili, stanno diventando sempre più comuni. Inoltre, il fenomeno della globalizzazione ha ulteriormente complicato la gestione dell'impatto ambientale dell'industria, le catene di approvvigionamento globali implicano che i processi di produzione e consumo siano distribuiti su scala internazionale, rendendo difficile l'implementazione di politiche ambientali uniformi ed efficaci. Le emissioni di gas serra, ad esempio, non rispettano i confini nazionali, rendendo necessario un approccio coordinato a livello globale per affrontare il problema. Questa considerazione va ancora una volta a sottolineare la necessità di politiche e cooperazione internazionale volte a ridurre i danni sull'ambiente di origine antropologica. La fig. 5.20 ci mostra come l'industria nel suo complesso, sia responsabile del 31% delle emissioni globali. Le sottocategorie includono ferro e acciaio (5%), gas e petrolio (5%), cemento (5%), prodotti chimici (4%), estrazione di carbone (2%), raffinazione (1%) e altri tipi d'industria (9%).

**FIGURE 6**  
**Global emissions by sector**  
Percent share of 2020 net GHG emissions



Source: Rhodium Group

Figura 5.20: Emissioni globale per settore [73]

L’impatto dell’industria sull’ambiente non si limita solo alle emissioni atmosferiche. L’inquinamento delle acque e del suolo, dovuto allo smaltimento improprio dei rifiuti industriali, rappresenta un’altra grave minaccia per gli ecosistemi e la salute umana. Le fuoriuscite di petrolio e altre sostanze tossiche possono devastare intere regioni, compromettendo la biodiversità e contaminando le risorse idriche. L’osservazione della relativa limitata frequenza di documenti associati al tema può inizialmente suggerire un basso grado di interesse e di preoccupazione nei confronti dell’argomento stesso. Tuttavia, una visione più ampia e approfondita dell’analisi rivela che il tema dell’impatto delle industrie occupa una posizione centrale, intersecando diversi ambiti tematici. L’impatto umano, intrinsecamente connesso all’attività industriale moderna, si configura come una questione che spesso richiede un approccio tecnico più dettagliato, come testimoniato anche dal contenuto dei documenti stessi. Pertanto, tale tematica non deve essere considerata solamente nell’ambito delle preoccupazioni dell’impatto delle grandi industrie, ma piuttosto come un ambito d’indagine autonomo, caratterizzato da specificità tecniche. Fra i documenti assegnati al topic evince comunque uno stato di preoccupazione nelle opinioni degli utenti, caratterizzati per lo più da variabili politiche.

- ”[@tagDonaldTrump]: *The concept of global warming was created by and for the Chinese in order to make U.S. manufacturing non-competitive.*”
- ”[@tag] *It’s the BURNING of fossil fuels that causes climate change. Not using it for other manufacturing.*”
- ”[@tag] *Human activity is the driver of climate change. Stop driving the cars that emit carbons. Stop flying, manufacturing. The data are clear.*”
- ”[@tag]: *One of world’s largest asset managers Vanguard seeks improved corporate disclosure on #climate change [link]...*”

- ”[@tag] weRE cuddling N he’s telling me how the island he’s from should incorporate green tech to mitigate climate change SCORPIO SZN POWER REAL big companies are destroying our lives”
- ”[@tag] New data reveals the stark environmental footprint of major industries worldwide. With sectors like steel, oil, and cement contributing significantly to global emissions, urgent action is needed to address industrial impact on our planet. #ClimateAction #EnvironmentalImpac”
- ”Does no one look at the emissions of big tech companies? ok oil but these servers need a huge amount of energy [@tag] [link]”
- ”White House Announces New Corporate Efforts to Fight Climate Change [link]”

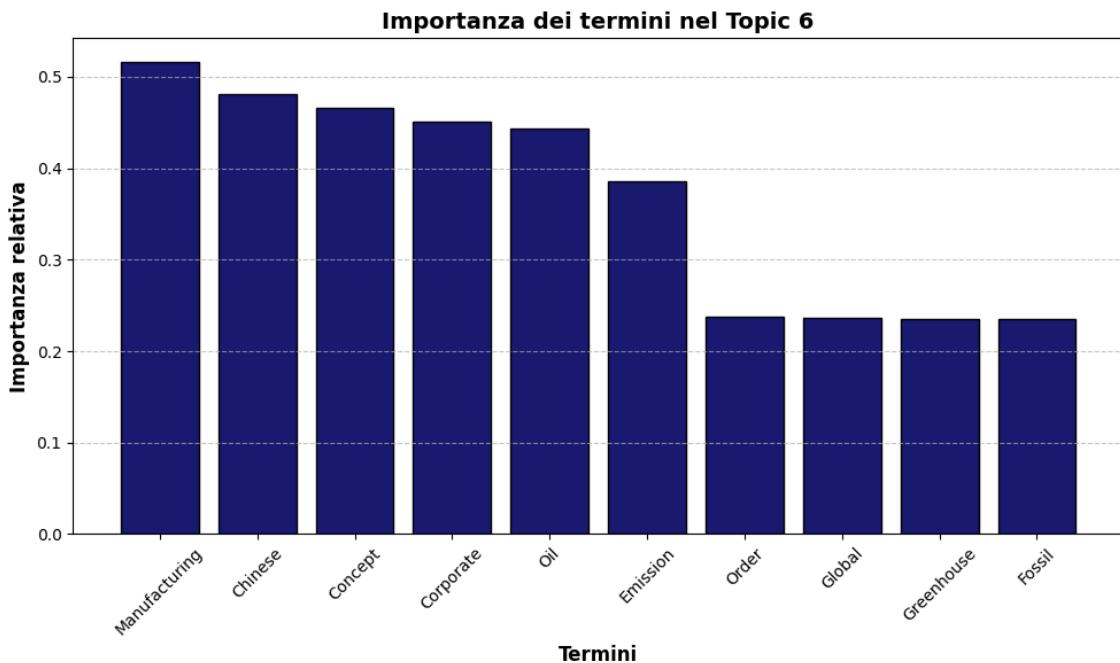
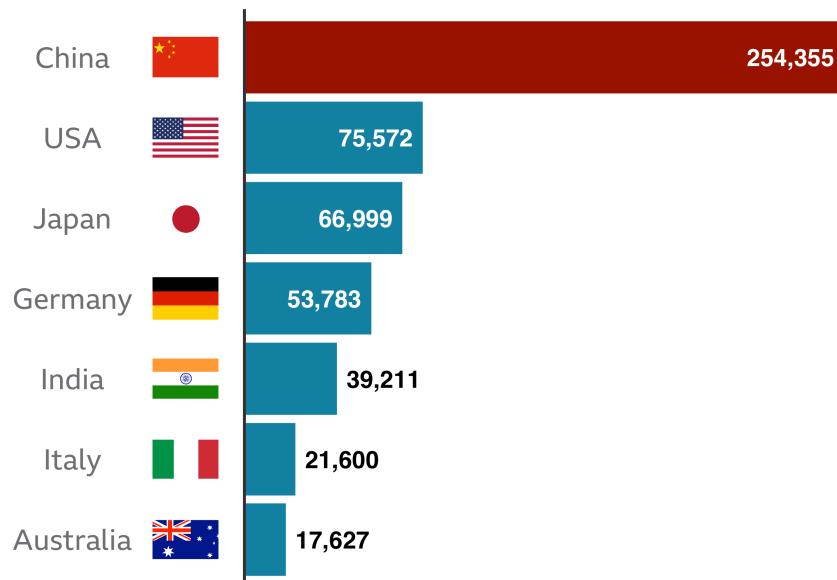


Figura 5.21: Keyword-Peso Industria inquinamento e impatto ambientale

Le parole chiave come *”Manufacturing, corporate, oil, emission, greenhouse, fossil e global”* rafforzano ulteriormente le considerazioni precedentemente espresse e sono coerenti con i dati disponibili sulle emissioni prodotte dalle aziende. Questi termini evidenziano una correlazione significativa tra l’interesse manifestato dagli utenti e l’impatto ambientale del settore industriale. Va sottolineato come la parola *”Chinese”* abbia all’interno del topic un’elevata importanza poichè effettivamente la Cina è uno dei principali attori globali sia nella produzione industriale che nelle emissioni di CO<sub>2</sub>, ma sta anche emergendo come leader nelle energie rinnovabili e nelle politiche climatiche [74]. Si contraddistingue principalmente sull’implementazione e la produzione di tecnologie a basse emissioni di carbonio e sull’utilizzo di energia pulita come quella solare come anche mostrato in Fig. 5.22 [75].

## China leads the world in solar power

Total capacity 2020 (Megawatts)



Source: International Renewable Energy Agency, 2021

BBC

Figura 5.22: Capacità energia solare [75]

In conclusione, l'interazione tra industria, inquinamento e impatto ambientale costituisce un intricato e stratificato contesto di analisi. La sua complessità intrinseca richiede una stretta cooperazione su scala globale e un vigoroso impegno da parte delle imprese, dei governi e delle istituzioni internazionali al fine di promuovere pratiche industriali sostenibili e contenere le emissioni nocive. L'analisi delle parole chiave e dei documenti pertinenti offre una buona prospettiva sulla vastità e la complessità delle sfide da affrontare in questo ambito. Solamente attraverso una concreta azione politica e da parte delle imprese attraverso l'adozione diffusa di tecnologie eco-sostenibili sarà possibile mitigare l'impatto ambientale dell'industria e assicurare la sostenibilità ambientale per le generazioni future.

---

### 5.1.7 Riscaldamento Globale

Il topic 6, etichettato: "Riscaldamento Globale" , rappresenta circa il 3,53% del totale con 12.138 documenti assegnati. Le parole chiave correlate e che spiegano questo argomento sono: "*Warming, Warmth, Miami, Temperature, Mayor, Change, Surface, Rise*". Nonostante la percentuale relativamente bassa di documenti che trattano specificamente il riscaldamento globale all'interno di questo particolare topic, l'argomento emerge come di centrale interesse nelle discussioni online. Questo è evidente dal fatto che anche negli altri topic analizzati, numerosi documenti contengono riferimenti o citazioni riguardanti il riscaldamento globale. Tale diffusione trasversale sottolinea l'importanza e la pervasività del tema, che si manifesta come probabilmente il filo conduttore fra le diverse tematiche. La correlazione presente con pressochè la maggioranza dei topic, come presentato in Fig. 5.4, evidenzia ancora di più come il riscaldamento globale sia un elemento centrale nelle discussioni sul cambiamento climatico. Più nello specifico il riscaldamento globale, rappresenta uno dei più urgenti e complessi problemi ambientali del nostro tempo. Questo fenomeno si riferisce all'aumento della temperatura media della superficie terrestre, un trend osservato a partire dalla fine del XIX secolo e attribuito principalmente all'incremento delle emissioni di gas serra derivanti dalle attività umane. Secondo l'Intergovernmental Panel on Climate Change(IPCC), la temperatura media globale è aumentata di circa 1,1°C dal periodo preindustriale ad oggi [59]. Il grafico in Fig. 5.23 offre un'analisi dettagliata delle anomalie della temperatura globale terrestre e oceanica nel periodo dal 1880 al 2022. Le anomalie di temperatura, rappresentanti le deviazioni dalla media di riferimento calcolata su un periodo base specifico, sono espresse sia in gradi Celsius che Fahrenheit sull'asse Y. Le anomalie negative, mostrate tramite barre blu, indicano anni più freddi rispetto alla media, mentre le anomalie positive, visualizzate tramite barre rosse, indicano anni più caldi. L'analisi temporale del grafico rivela alcune tendenze significative nel corso dei decenni. Tra il 1880 e il 1930, le temperature globali tendono a essere inferiori alla media di riferimento, con anomalie principalmente negative. Tra il 1930 e il 1980, si osserva una transizione verso anomalie positive, seguite da una ripresa delle anomalie negative fino agli anni '70. Tuttavia, a partire dagli anni '80, si verifica un aumento costante delle anomalie positive, indicativo di un trend di riscaldamento globale sempre più evidente.

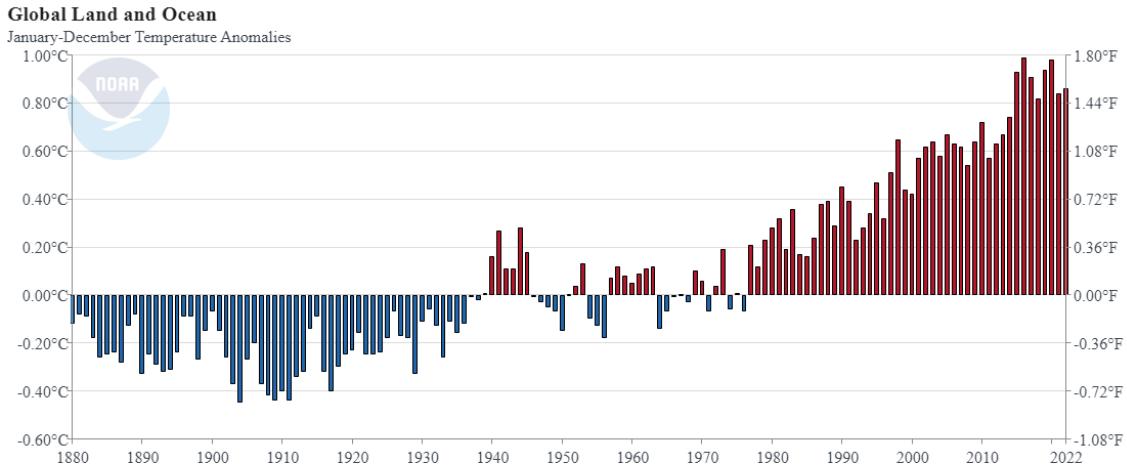


Figura 5.23: Anomalie della temperatura globale terrestre e oceanica su base annuale [76]

Il riscaldamento globale comporta pericoli significativi, come l'aumento del livello del mare, che minaccia le comunità costiere, e l'intensificazione degli eventi climatici estremi, quali uragani e ondate di calore. Inoltre, contribuisce alla perdita di biodiversità, alterando gli habitat naturali e mettendo a rischio numerose specie. Per contrastare questi pericoli, sono state intraprese diverse iniziative a livello globale come spiegato anche nel paragrafo dedicato al topic 1 (Cooperazione internazionale e clima). L'Accordo di Parigi del 2015 mira a limitare l'aumento della temperatura globale a 1,5°C sopra i livelli preindustriali, impegnando i paesi a ridurre le emissioni di gas serra e a sviluppare strategie di adattamento [57]. Questo obiettivo è cruciale poiché, secondo il rapporto dell'IPCC, mantenere l'aumento della temperatura sotto i 2°C rispetto ai livelli preindustriali ridurrebbe significativamente i rischi e gli impatti dei cambiamenti climatici. Tuttavia, limitare l'aumento a 1,5°C offrirebbe benefici ancora maggiori, proteggendo più efficacemente gli ecosistemi e le comunità umane dagli effetti più gravi del cambiamento climatico. L'Unione Europea ha inoltre siglato il Green Deal Europeo, che punta a raggiungere la neutralità climatica, cioè l'azzeramento delle emissioni, entro il 2050, attraverso una serie di misure che includono l'uso di energie rinnovabili e la promozione dell'efficienza energetica [77]. Questo perché le principali cause del riscaldamento globale sono attribuibili alle attività antropogeniche che rilasciano gas serra nell'atmosfera. I principali gas serra includono l'anidride carbonica ( $\text{CO}_2$ ), il metano ( $\text{CH}_4$ ), il protossido di azoto ( $\text{N}_2\text{O}$ ) e i gas fluorurati. La  $\text{CO}_2$ , responsabile di circa il 76% delle emissioni di gas serra globali, proviene principalmente dalla combustione di combustibili fossili per la produzione di energia, il trasporto e l'industria [78]. Il metano, che ha un potere riscaldante circa 28 volte superiore a quello della  $\text{CO}_2$  su un arco temporale di 100 anni, è emesso durante la produzione e il trasporto di carbone, petrolio e gas naturale, nonché da fonti agricole e da discariche [78]. Come appena sottolineato, le principali cause attribuibili al riscaldamento globale sono di origine antropologica, ben distinguibili dalle normali attività naturali della terra. Questa differenza viene mostrata nel grafico a Figura 5.24, che offre una rappresentazione visiva della variazione della temperatura media globale nel periodo compreso tra il 1880 e il 2020. Esso è suddiviso in diverse linee che delineano l'impatto di vari fattori sulla

temperatura globale, espressa in gradi Fahrenheit rispetto alla media di riferimento. Le linee distintive del grafico includono:

- Osservato (in nero): corrisponde alla temperatura globalmente registrata nel periodo considerato.
- Fattori Umani (in rosso): comprende le influenze derivanti dalle attività antropogeniche, quali le emissioni di gas serra e i cambiamenti nell'uso del suolo.
- Fattori Naturali (in blu): rappresenta gli impatti derivanti da fenomeni naturali, come l'attività vulcanica, l'attività solare e le variazioni orbitali.

L'analisi del grafico evidenzia il ruolo significativo delle attività umane nel determinare il riscaldamento globale, particolarmente evidente nei decenni più recenti. La stretta associazione tra la linea nera delle osservazioni e la linea rossa dei fattori umani suggerisce una correlazione sostanziale tra le emissioni antropogeniche e l'incremento delle temperature globali nel tempo. In contrasto, i fattori naturali manifestano fluttuazioni relativamente minori e non seguono lo stesso trend ascendente osservato nella linea delle osservazioni.

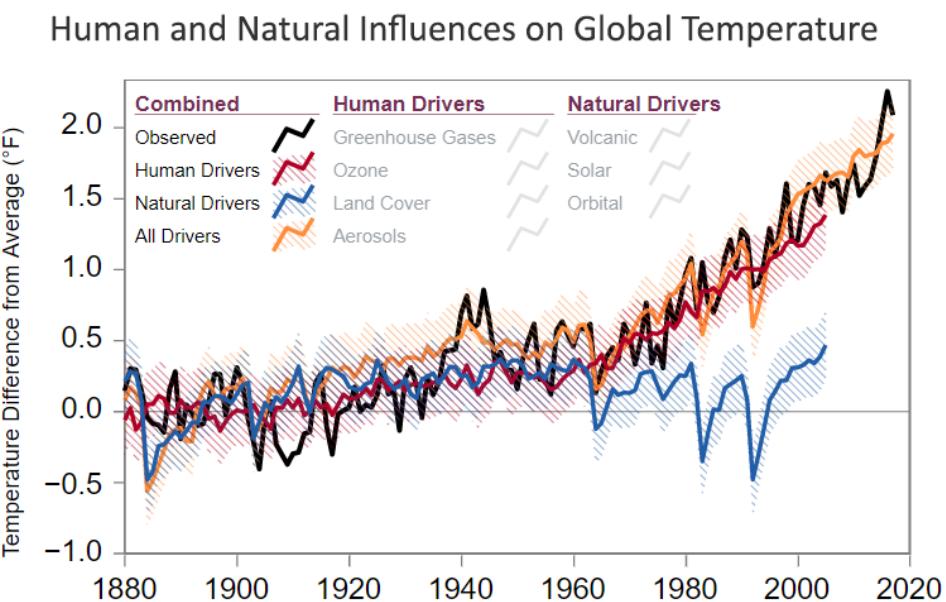


Figura 5.24: Influenze umane e naturali sulla temperatura globale [79]

Il riscaldamento globale ha inoltre implicazioni profonde per la biodiversità, come spiegato nel capitolo dedicato al topic (Inquinamento ed Ecosistemi), gli ecosistemi terrestri e marini subiscono cambiamenti a causa delle temperature crescenti e dell'acidificazione degli oceani. L'aumento della temperatura dell'acqua e l'acidificazione mettono a rischio la vita marina, comprese le barriere coralline, che sono habitat cruciali per molte specie. Miami, come sottolineato dalle keyword e dai documenti assegnati al topic, rappresenta un caso di studio emblematico per osservare gli impatti locali del riscaldamento globale. La città è particolarmente vulnerabile all'innalzamento del livello del mare a causa della sua bassa altitudine e della sua

---

posizione costiera. Secondo uno studio del National Oceanic and Atmospheric Administration (NOAA), il livello del mare a Miami è aumentato di circa 9 cm dal 1992 al 2019, con proiezioni che indicano un aumento di ulteriori 30-60 cm entro il 2060, a seconda degli scenari di emissione di gas serra [80].

Le evidenze rappresentate dai dati scientifici sono ben evidenti anche nelle preoccupazioni e nelle opinioni espresse nei documenti, ma come sottolineato in precedenza, il dibattito legato al riscaldamento globale fa da filo conduttore fra diversi sotto-argomenti, mostrando come questa tendenza è spesso associata, anche inconsciamente, al concetto stesso di cambiamento climatico. Alcuni esempi di documenti assegnati ai topic:

- ”[@tag]: *global warming real as hell. al gore told us. leo dicaprio keep tryna tell us. its 82 degrees and its halloween. summer aint here*”
- ”[@tag]: *Climate change is widely accepted, but these 9 global warming effects may still surprise you. [link]*”
- ”[@tag]: *When everyone's glad it's 80 degree weather in October but you can't stop thinking about global warming [link]*”
- ”[@tag]: *Rising temperatures put millions at risk as climate change hotspot [@tag] (world's 3rd-largest store of ice) is melting*”
- ”*Plus, with a few more years of rampant, uncontrolled climate change you might not even notice the sudden temperature change*”
- ”*While global warming has affected the entire planet in recent decades, no place has been hit harder than the Arctic. This month, temperatures have never been so high*”
- ”[@tag]: *How Miami is sinking into the sea*”
- ”[@tag]: *Miami could be underwater by 2100 due to climate change and global warming [link]*”

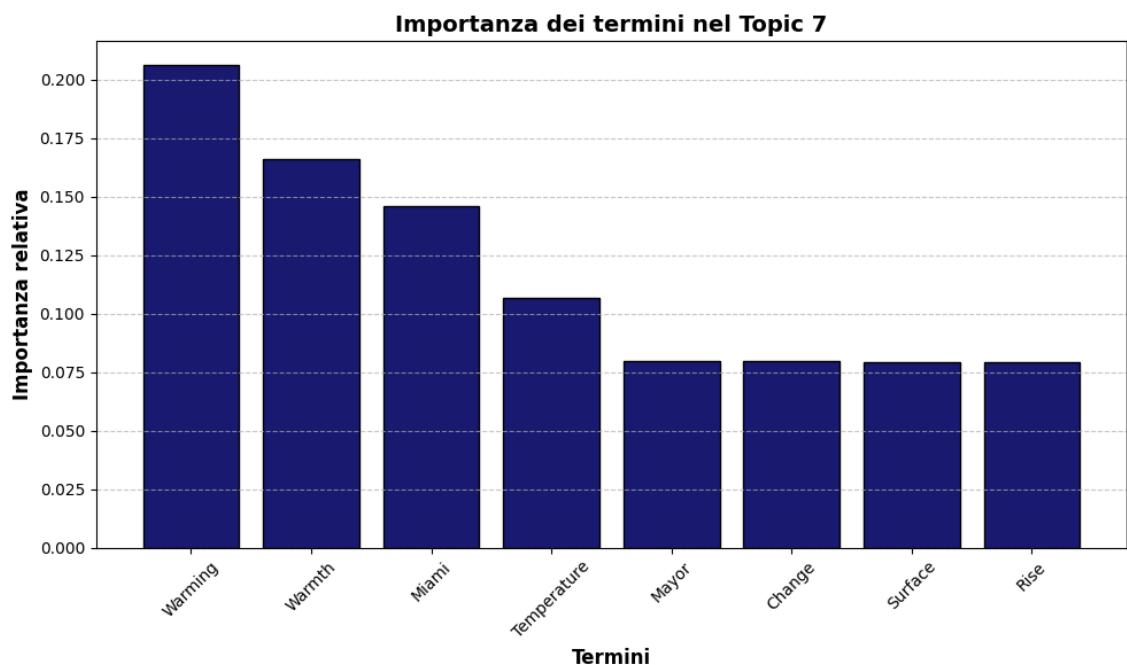


Figura 5.25: Keyword-Peso Riscaldamento Globale

---

### 5.1.8 Il Cambiamento Climatico fra Bufala e Realtà

Il topic 7, incentrato sul cambiamento climatico come bufala o come realtà, rappresenta circa il 3,16% del totale con 10.893 documenti assegnati. Le parole chiave che spiegano questo tema sono: *"Climate, Change, Real, Hoax, Warming, Fake, News, Scam, Proclimate, Lie, Media"*. Il dibattito sul cambiamento climatico è caratterizzato da una serie di opinioni contrastanti e dalla diffusa disinformazione, che spesso si manifesta attraverso la propagazione di bufale e notizie false. Tuttavia, analizzando dati scientifici e fonti affidabili, è possibile distinguere la realtà dalla menzogna e comprendere l'urgenza di affrontare questa sfida globale. Con l'avvento dei social media e delle piattaforme online, la diffusione di informazioni non verificate e, in alcuni casi, completamente false sul cambiamento climatico è stata amplificata. Le fake news sfruttano termini come *"hoax"*, *"scam"* e *"lie"* per screditare la scienza del clima e alimentare il dubbio pubblico sull'esistenza stessa del cambiamento climatico. Questo fenomeno non è solo un problema di comunicazione, ma ha anche implicazioni significative sulla formazione dell'opinione pubblica e sull'adozione di politiche ambientali efficaci. Alcuni media, spinti da interessi economici o politici, hanno contribuito a diffondere disinformazione sul cambiamento climatico, ignorando o minimizzando le prove scientifiche a favore di narrazioni che mettono in discussione il consenso scientifico. La strategia della disinformazione spesso include la selezione selettiva dei dati, l'uso di pseudo-scienza e la citazione di scienziati non qualificati, il che porta a una falsa equivalenza tra la scienza consolidata e opinioni marginali [81]. Un esempio prominente di questa dinamica, come anche evidenziato dai documenti assegnati al topic, è rappresentato dalle dichiarazioni dell'ex presidente degli Stati Uniti, Donald Trump. Durante la sua presidenza, Trump ha ripetutamente definito il cambiamento climatico una "bufala" (*"hoax"*), sostenendo che le politiche volte a contrastarlo avrebbero danneggiato l'economia americana. Le sue affermazioni sono state frequentemente citate dai media e dai social network, contribuendo a creare un ambiente di scetticismo e confusione. Nel 2012, Trump ha scritto su Twitter: *"The concept of global warming was created by and for the Chinese in order to make U.S. manufacturing non-competitive"*. Questo tipo di retorica ha avuto un impatto significativo nel modellare l'opinione pubblica e nel ritardare l'implementazione di politiche climatiche efficaci negli Stati Uniti. Tuttavia, i dati scientifici raccolti da istituzioni autorevoli come l'Intergovernmental Panel on Climate Change (IPCC) forniscono un quadro chiaro e inequivocabile del cambiamento climatico in corso. Le prove dell'aumento delle temperature medie globali, dell'innalzamento del livello del mare e degli eventi climatici estremi sono ben documentate e supportate da una vasta comunità scientifica. Ad esempio, il report del IPCC del 2021 ha confermato che molte delle trasformazioni osservate nel clima sono senza precedenti su scale temporali di centinaia di migliaia di anni, e che il ruolo umano nel riscaldamento del sistema climatico è indiscutibile [1]. Inoltre, gli impatti del cambiamento climatico sono evidenti e tangibili, con conseguenze che vanno dalla perdita di biodiversità alla minaccia per la sicurezza alimentare e idrica, fino alla devastazione di comunità costiere a causa dell'innalzamento del livello del mare. I fenomeni meteorologici estremi, come uragani, ondate di calore e siccità, sono diventati più frequenti e intensi, causando danni significativi a infrastrutture, ecosistemi e società umane. È fondamentale contrastare la disinformazione e promuovere una comprensione accurata della scienza del clima per adottare politiche e comportamenti che affrontino efficacemente questa sfida globale. In questo contesto,

l'educazione e la comunicazione scientifica giocano un ruolo cruciale. Le piattaforme social e i media devono essere utilizzati non solo per diffondere conoscenze accurate, ma anche per promuovere il pensiero critico tra il pubblico, affinché le persone possano distinguere tra informazioni veritiera e falsità. L'opinione pubblica riguardo al cambiamento climatico varia significativamente a livello globale, influenzata da fattori socio-culturali, economici e politici. Secondo un sondaggio condotto dal Pew Research Center nel 2019, una maggioranza del 67% degli intervistati in tutto il mondo riconosce che il cambiamento climatico è una minaccia seria per il proprio paese [82]. Tuttavia, esistono differenze regionali marcate: nei paesi europei e in Nord America, il consenso sulla realtà del cambiamento climatico è generalmente alto, con percentuali superiori al 70% in molte nazioni. D'altra parte, negli Stati Uniti, uno studio del Yale Program on Climate Change Communication ha rilevato che, nonostante il 69% degli americani crede che il riscaldamento globale stia avvenendo, solo il 57% attribuisce questo fenomeno principalmente alle attività umane [83]. Questo scetticismo è spesso alimentato da disinformazione e polarizzazione politica, come evidenziato dalle posizioni pubbliche di figure politiche influenti che minimizzano o negano l'esistenza del cambiamento climatico [84]. Questi dati dimostrano l'importanza di una comunicazione scientifica efficace e di iniziative educative per migliorare la comprensione pubblica del cambiamento climatico e promuovere l'adozione di comportamenti sostenibili.

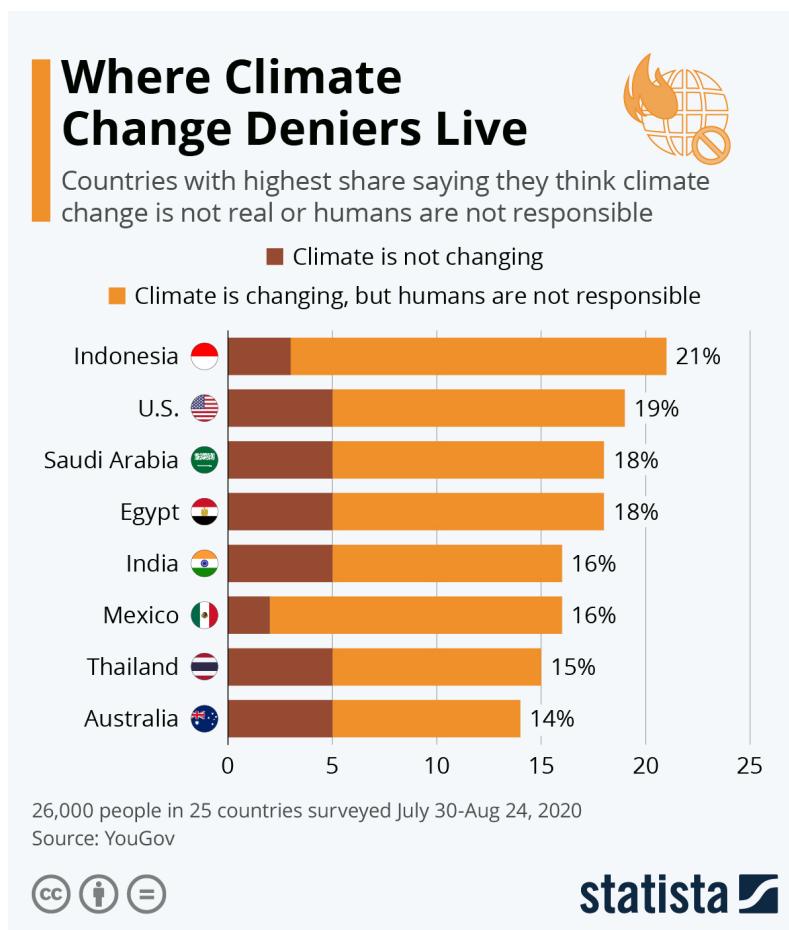
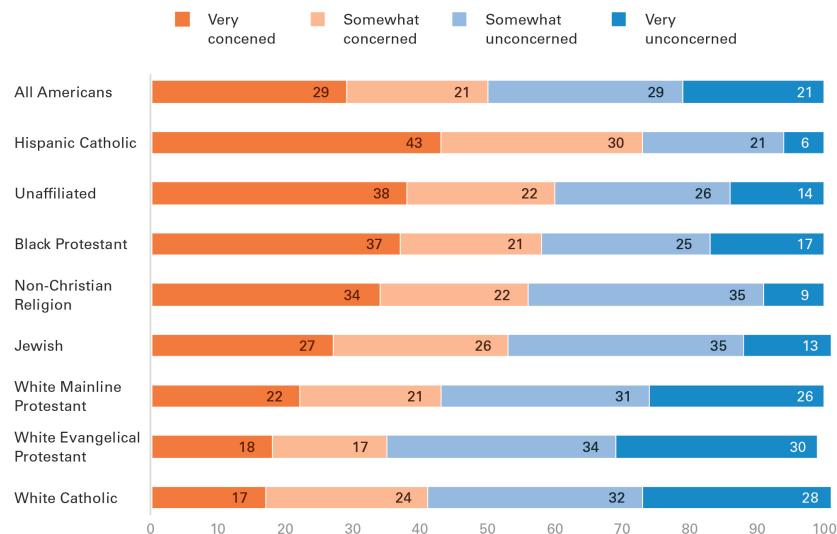


Figura 5.26: Dove vivono i negazionisti del cambiamento climatico [85]

Il grafico in Figura 5.26 illustra la percentuale di persone in vari paesi che negano

il cambiamento climatico o ritengono che esso non sia causato dall'attività umana. Il sondaggio, condotto da YouGov nel 2020 su 26.000 persone in 25 paesi, mostra che l'Indonesia ha la percentuale più alta di negazionisti climatici, con il 21% della popolazione che non crede nel cambiamento climatico. Negli Stati Uniti, il 19% delle persone è scettico riguardo alla realtà del cambiamento climatico o al ruolo umano in esso. L'Arabia Saudita ed Egitto seguono con il 18%, mentre in India e Messico il 16% della popolazione condivide queste opinioni. In Thailandia, il 15% delle persone nega il cambiamento climatico, e in Australia questa percentuale è del 14%. Questi dati evidenziano come la negazione del cambiamento climatico non sia un fenomeno isolato ma diffuso globalmente, con vari gradi di scetticismo influenzati da fattori culturali, politici ed economici. La comprensione delle differenze geografiche nella percezione del cambiamento climatico è essenziale per sviluppare strategie di comunicazione efficaci e politiche mirate che possano affrontare le preoccupazioni e le credenze delle diverse popolazioni. Un ulteriore fattore che incide sulla preoccupazione legata al cambiamento climatico è la religione [86], come mostrato in Figura 5.27 , viene mostrato che tra tutti gli americani, il 29% si dichiara "molto preoccupato" per il cambiamento climatico, mentre il 21% è "abbastanza preoccupato". Tuttavia, esistono significative variazioni tra i gruppi religiosi. I cattolici ispanici mostrano il livello più alto di preoccupazione, con il 43% "molto preoccupato" e il 30% "abbastanza preoccupato". Gli individui non affiliati a nessuna religione seguono con il 38% di "molto preoccupati" e il 22% di "abbastanza preoccupati". Al contrario, i protestanti evangelici bianchi rappresentano il gruppo con il minor livello di preoccupazione: solo il 18% si dichiara "molto preoccupato", mentre un significativo 30% è "per niente preoccupato". I protestanti della linea principale bianchi e i cattolici bianchi mostrano livelli simili di bassa preoccupazione, con il 22% e il 17% rispettivamente che si dichiarano "molto preoccupati".

**FIGURE 3. Climate Change Concern Index by Religious Affiliation**



Source: PRRI/AAR, Religion, Values, and Climate Change Survey, November 2014

Figura 5.27: Indice di preoccupazione per il cambiamento climatico per affiliazione religiosa [86]

---

Dai documenti assegnati al topic emerge una forte variabile politica, poiché le asserzioni di determinate figure politiche hanno un impatto significativo sulle opinioni e le credenze pubbliche. Come già sottolineato, le dichiarazioni dell'ex presidente degli Stati Uniti, Donald Trump, che ha spesso definito il cambiamento climatico una "bufala" e ha ritirato gli Stati Uniti dall'Accordo di Parigi, hanno contribuito a un aumento dello scetticismo tra i suoi sostenitori e in alcune regioni degli Stati Uniti. Questo fenomeno non è isolato agli Stati Uniti; simili dinamiche possono essere osservate in altri paesi dove le posizioni dei leader politici influenzano fortemente l'opinione pubblica riguardo alla scienza del clima. I documenti analizzati mostrano che la disinformazione e la polarizzazione politica giocano un ruolo cruciale nel plasmare le percezioni e le credenze delle persone sul cambiamento climatico, dimostrando come le narrazioni politiche possano distorcere la comprensione scientifica e rallentare l'azione collettiva necessaria per affrontare questa crisi globale. Alcuni esempi di documenti:

- ”[@tag]: Obamas fiddled global warming data: \*Shamelessly Manipulated\* - [link] #ClimateScam #GreenScam #TeaParty
- ”Another global warming argument bites the dust: No Increase in Global Drought Over Past 30 Yrs: [link] #ClimateScam #PJNet”
- ”[@tag]: Good. Man made global warming is the biggest hoax ever perpetrated #ClimateScam [link]”
- ”[@tag]: Friendly reminder our next president believes global warming is a Chinese hoax & hair spray can't escape your home”
- ”[@tag]: Bannon, Senior Counselor to the President on climate change: A Hoax that costs us \$4 billion per day. [link]”
- ”[@tag]: The electric car "charging stations" at the UN conference on global warming are fake. They aren't even connected to a power plant”
- ”Hey democrats let's drop the global warming bs ok even ur extremely liberal news channel is against it..fake af [link]”
- ”[@tag]: The REAL FAKE NEWS exposed: '97% of scientists agree on climate change' is an engineered hoax..[link]”
- ”[@tag]: Global warming data FAKED by gov't to fit climate change fictions #ClimateHoax #FakeNews [link] [link]”
- ”[@tag]: If a PM says he believes in climate change but approves new oil pipelines, does he actually believe in climate change?”

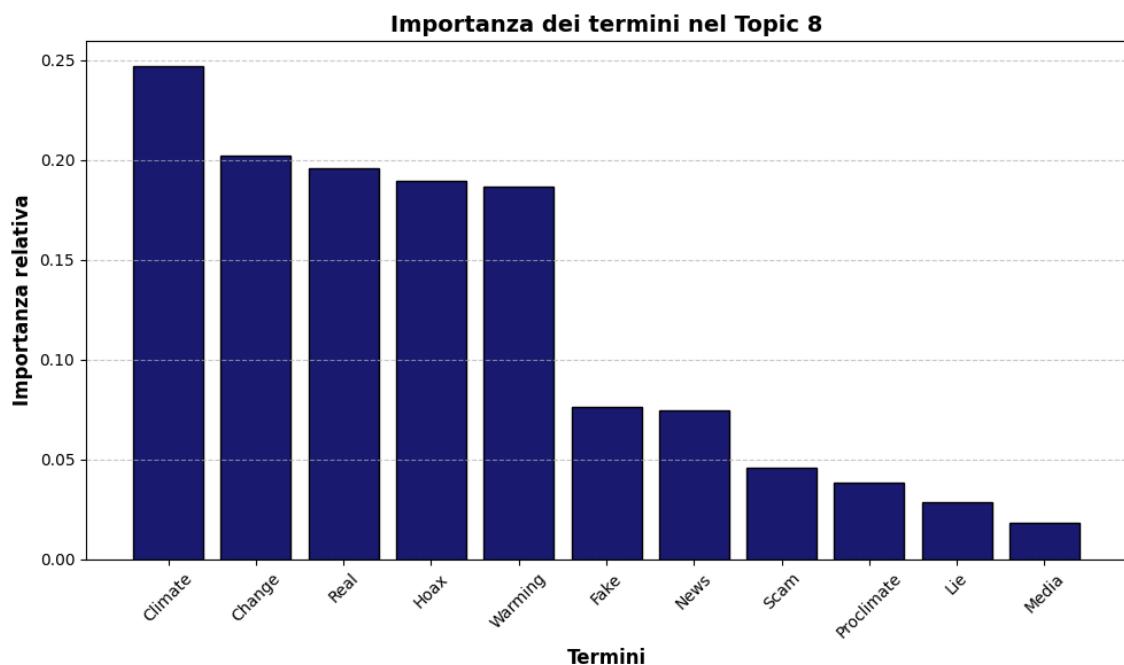


Figura 5.28: Keyword-Peso Cambiamento climatico fra bufala e realtà

Per concludere, la lotta contro la disinformazione climatica è un elemento chiave nella battaglia per la sostenibilità ambientale. Solo attraverso un impegno collettivo verso la verità scientifica e l'adozione di politiche basate su prove concrete si può sperare di mitigare gli effetti devastanti del cambiamento climatico e garantire un futuro sostenibile per le generazioni presenti e future.

### 5.1.9 Impatto e Controversie della Società Exxonmobil

Il topic 8, titolato *"Impatto e Controversie della società ExxonMobil"* rappresenta circa il 2,5% del totale con 8424 documenti assegnati, nel contesto dell'analisi è lontano dalla maggior parte degli argomenti se non per una leggera somiglianza con i topic 1 (Cooperazione internazionale e clima), 4 (Crisi, sanità e sicurezza pubblica) e 5 (Industria, inquinamento e impatto ambientale). Le parole chiave che caratterizzano questo topic sono: *"Exxon, Tillerson, Rex, Emission, Exxonmobil, Change, Fossil, Know"*. La società ExxonMobil è stata al centro di numerose controversie riguardanti il cambiamento climatico, in particolare per quanto riguarda l'impatto delle sue attività sulle emissioni di gas serra e il suo ruolo nel dibattito pubblico sull'ambiente. Fin dagli anni '70, documenti interni della Exxon (ora ExxonMobil) hanno dimostrato che l'azienda era consapevole del potenziale impatto devastante delle emissioni di CO<sub>2</sub> sul clima globale. Tuttavia, invece di agire per mitigare questi effetti, ExxonMobil è stata accusata di aver finanziato e diffuso informazioni che minimizzavano il legame tra l'uso di combustibili fossili e il riscaldamento globale, contribuendo a ritardare le politiche ambientali e il consenso scientifico su queste tematiche [87]. Evidenze riscontrabili anche nei dati mostrati in Figura 5.29, che confronta il cambiamento di temperatura osservato e le proiezioni di cambiamento di temperatura fatte da ExxonMobil.

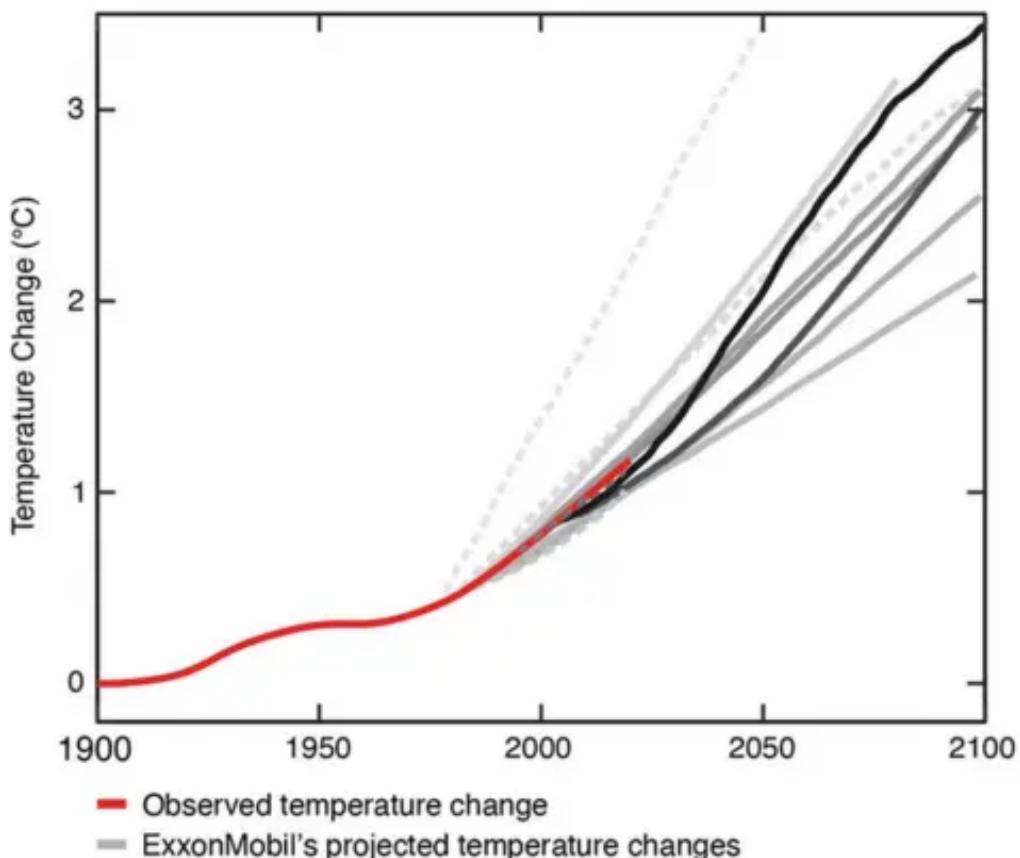


Figura 5.29: Confronto osservazioni e proiezioni di temperatura [88]

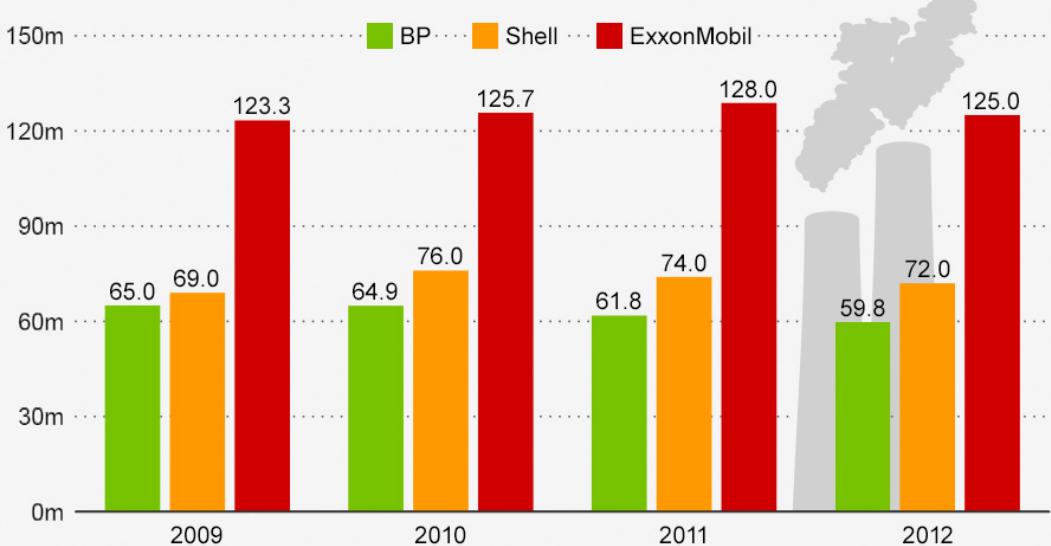
La Linea rossa: Rappresenta il cambiamento di temperatura osservato a partire

---

dal 1900 fino ai giorni nostri. Si nota un incremento graduale fino alla metà del XX secolo, seguito da un aumento più rapido e significativo a partire dagli anni '70. Le *Linee grigie*: Rappresentano le diverse proiezioni del cambiamento di temperatura fatte da ExxonMobil per il periodo dal 1950 al 2100. Queste proiezioni mostrano un aumento delle temperature che segue una tendenza simile a quella osservata, ma con un leggero disallineamento iniziale rispetto ai dati osservati. La *Linea nera*: Presenta una media delle proiezioni di ExxonMobil, che si allinea abbastanza bene con i dati osservati fino al presente, indicando che le previsioni di ExxonMobil erano in gran parte accurate riguardo al trend generale di aumento delle temperature. Viene mostrato quindi che ExxonMobil aveva una comprensione piuttosto accurata del potenziale aumento della temperatura globale dovuto alle emissioni di gas serra, ma nonostante questa consapevolezza, l'azienda è stata criticata per aver minimizzato pubblicamente il rischio del cambiamento climatico e per aver ostacolato le politiche di mitigazione. Una figura centrale in queste controversie è Rex Tillerson, che ha servito come CEO di ExxonMobil dal 2006 al 2016. Durante il suo mandato, ExxonMobil ha continuato a investire pesantemente nell'estrazione di combustibili fossili, nonostante le crescenti preoccupazioni globali riguardo ai cambiamenti climatici. Tillerson ha spesso sostenuto la necessità di bilanciare la domanda energetica mondiale con le preoccupazioni ambientali, una posizione che è stata vista come un tentativo di giustificare la continua espansione delle attività di ExxonMobil nel settore dei combustibili fossili [89]. Le azioni di ExxonMobil hanno avuto un impatto significativo sia a livello ambientale che politico. Le emissioni di gas serra derivanti dalle operazioni della compagnia hanno contribuito all'aumento delle temperature globali, mentre le sue attività di lobbying, cioè modalità d'azione con cui una azienda si inserisce nel panorama politico, esercitando la propria pressione sul sistema stesso, hanno influenzato le politiche climatiche di numerosi governi, spesso ostacolando l'adozione di misure più stringenti per la riduzione delle emissioni. Questo ha sollevato interrogativi etici e legali sulla responsabilità delle grandi aziende nel contrastare il cambiamento climatico e sulla necessità di una maggiore trasparenza nelle loro operazioni e strategie di comunicazione [90].

## ExxonMobil Exceeds Rivals in Emissions

Comparison of greenhouse gas emissions 2009-2012 (in million metric tons)



statista The Statistics Portal @StatistaCharts

Source: ExxonMobil, Shell, BP

Figura 5.30: Confronto emissioni di gas serra 2009-2012

Il grafico in Figura 5.30 mostra un confronto delle emissioni di gas serra in milioni di tonnellate metriche tra tre grandi compagnie petrolifere: BP, Shell, ed ExxonMobil, nel periodo dal 2009 al 2012. I dati evidenziano che ExxonMobil ha costantemente superato i suoi rivali in termini di emissioni. Nel 2009, ExxonMobil ha emesso 123,3 milioni di tonnellate, mentre BP e Shell hanno emesso rispettivamente 65,0 e 69,0 milioni di tonnellate. Questa tendenza è continuata negli anni successivi, con ExxonMobil che ha raggiunto il picco di 128,0 milioni di tonnellate nel 2011, rispetto ai 61,8 milioni di BP e 74,0 milioni di Shell nello stesso anno. Anche nel 2012, ExxonMobil ha mantenuto emissioni superiori (125,0 milioni di tonnellate) rispetto a BP (59,8 milioni di tonnellate) e Shell (72,0 milioni di tonnellate). Il grafico vuole quindi sottolineare come ExxonMobil abbia avuto un impatto ambientale significativamente maggiore rispetto ai suoi concorrenti durante questo periodo. L'influenza dei privati e delle industrie, come anche spiegato al paragrafo 5.1.6, sulla politica e sull'inquinamento, è un tema di grande rilevanza. Le loro attività di lobbying possono avere un impatto diretto sulle decisioni politiche, spesso a discapito dell'interesse pubblico e della salute ambientale. La capacità di tali aziende di finanziare campagne e sostenere politicamente candidati favorevoli ai loro interessi economici rende difficile l'adozione di regolamenti più severi sulle emissioni di gas serra. Questo scenario non solo compromette gli sforzi per combattere il cambiamento climatico, ma solleva anche questioni etiche sulla responsabilità e l'integrità delle politiche aziendali. La pressione esercitata dalle industrie può ritardare o indebolire le iniziative globali volte alla sostenibilità ambientale, dimostrando l'importanza di una governance più rigorosa e di una maggiore responsabilità corporativa nel contesto della crisi climatica. Questa discrepanza con l'interesse pubblico e la preservazione della salute ambientale è evidente anche nei documenti assegnati

al topic, inoltre è evidente un rapporto di questo argomento con la politica visto il ruolo ricoperto da Rex Tillerson di segretario degli Stati Uniti sotto la presidenza Trump:

- ”*Exxon Mobil being sued over climate change cover-up [@tag] [link] via [@tag] Back #JILL; support renewables*”
- ”*[@tag]: #Trump’s Cabinet: Exxon shills, climate change deniers, and criminals. Way to #draintheswamp, Donnie.*”
- ”*[@tag], [@tag], [@tagExxonMobil] Didn’t Exxon scientists do some of the original research on climate change?*”
- ”*[@tag]: ’ExxonMobil has a long history of peddling misinformation on climate change.’ [@tag] in @NewYorker #ExxonKnew [link]*”
- ”*[@tag]: Asked if charges that #ExxonMobil withheld information on climate change, #Tillerson declines to answer twice. #SecretaryOfSt... .*”
- ”*[@tag]: [@tag]: Rex you’re not the CEO of Exxonmobil anymore. U can Acknowledge climate change. 6th graders can explain the climate change better than you*”
- ”*[@tag]: Rex Tillerson’s ExxonMobil is planning for global warming but he will do everything humanly possible to stop his country from climate change*”
- ”*[@tag]: Little is known about how climate change regulation and green tech would impact exxonmobil. That could be [link]*”

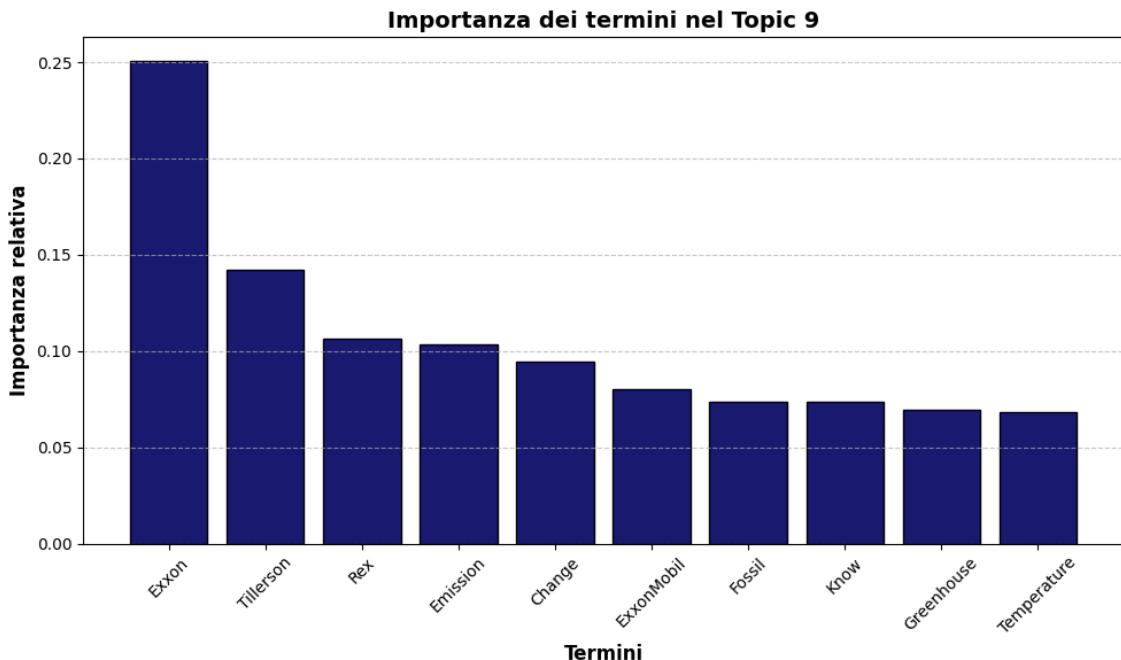


Figura 5.31: Keyword-peso Impatto e controversie Exxonmobil

---

Il termine "Exxon" domina nettamente, suggerendo un forte interesse e una frequenza elevata di menzioni relative alla compagnia petrolifera ExxonMobil nelle conversazioni sul cambiamento climatico. Seguito da "Tillerson" e "Rex", il grafico indica anche un'attenzione significativa verso l'ex amministratore delegato della compagnia, Rex Tillerson. Altri termini rilevanti come "Emission", "Change", "Fossil", "Greenhouse" e "Temperature" confermano che il topic comprende discussioni sulle emissioni di gas serra e le problematiche ambientali. Il grafico in Figura 5.31 riflette come le tematiche legate a ExxonMobil siano centrali nelle discussioni ambientali, sottolineando la concordanza con le opinioni degli utenti, che tendono a focalizzarsi sulle responsabilità delle grandi compagnie petrolifere come ExxonMobil nei confronti del cambiamento climatico e dell'inquinamento, denotando un interesse pubblico rilevante e una percezione critica dell'opinione pubblica nei confronti delle attività della compagnia. Per concludere, la controversia attorno a ExxonMobil evidenzia le sfide che la società globale deve affrontare nel mitigare il cambiamento climatico e la necessità di una governance più rigorosa delle attività delle grandi imprese energetiche. Il caso di ExxonMobil fa da esempio centrale sull'importanza di una politica ambientale basata su dati scientifici e su una gestione responsabile delle risorse naturali.

---

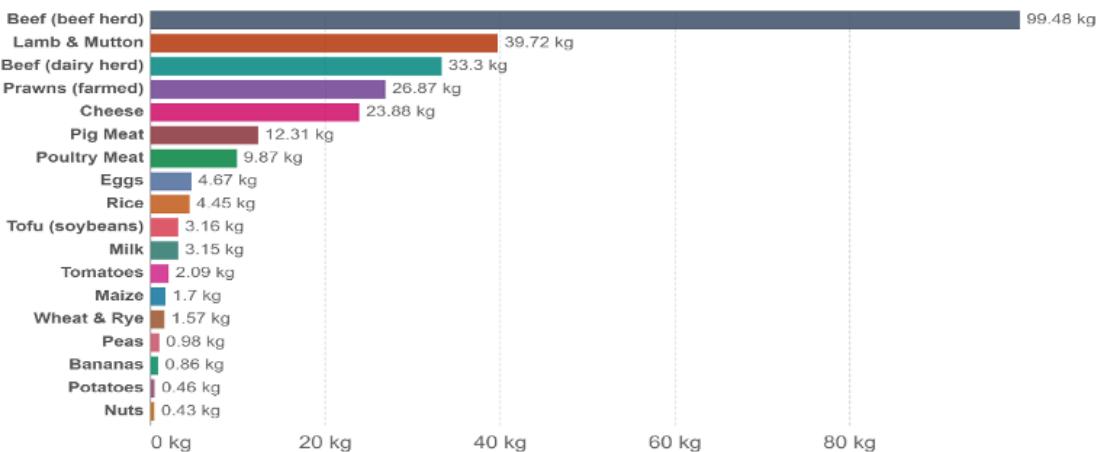
### 5.1.10 Consumo di Carne, Allevamenti e Impatto

L'ultimo topic, etichettato "*Consumo di carne, Allevamenti e Impatto Ambientale*" è composto da 7787 documenti rappresentando il 2,26% del totale. Le parole chiave estratte e che rappresentano il topic sono: "*Bacon, Meat, Watch, Agriculture, Vegan, Greenhouse, Mutilated, Emission*". Il consumo di carne e gli allevamenti intensivi sono temi centrali nei dibattiti contemporanei sull'ambiente e sulla sostenibilità. L'impatto degli allevamenti intensivi e delle pratiche agricole sull'ambiente è significativo e multidimensionale, riguardando principalmente le emissioni di gas serra, l'uso del suolo e delle risorse idriche, la deforestazione e la perdita di biodiversità. Gli allevamenti intensivi sono responsabili di una porzione considerevole delle emissioni globali di gas serra. Secondo la FAO (Food and Agriculture Organization), il settore zootecnico contribuisce al 14,5% delle emissioni totali di gas serra antropogeniche, principalmente attraverso il metano prodotto dalla fermentazione enterica nei ruminanti, il protossido di azoto derivato dalla gestione del letame e l'anidride carbonica risultante dalla deforestazione per la creazione di pascoli [91]. Inoltre, gli allevamenti richiedono ampie superfici di terreno per il pascolo e la produzione di mangimi. Circa il 70% delle terre agricole mondiali è destinato all'allevamento di bestiame e alla produzione di foraggi, contribuendo significativamente alla deforestazione, in particolare nelle regioni tropicali [92]. La deforestazione, a sua volta, è una delle principali cause della perdita di biodiversità e dell'aumento delle emissioni di carbonio atmosferico, poiché gli alberi abbattuti liberano il carbonio immagazzinato. L'uso delle risorse idriche è un altro aspetto critico. La produzione di carne richiede volumi d'acqua significativamente maggiori rispetto alla produzione di alimenti vegetali. Secondo il Water Footprint Network, per produrre un chilogrammo di carne bovina sono necessari circa 15.400 litri d'acqua, una quantità molto superiore rispetto a quella richiesta per produrre un chilogrammo di cereali [93]. Questo elevato consumo idrico contribuisce alla scarsità d'acqua in molte regioni, aggravando le crisi idriche globali. Le pratiche agricole intensive associate agli allevamenti possono anche provocare la degradazione del suolo e l'inquinamento delle risorse idriche. L'uso eccessivo di fertilizzanti e pesticidi per coltivare i mangimi può portare alla contaminazione delle acque superficiali e sotterranee, mentre la gestione inadeguata del letame può causare l'eutrofizzazione dei corsi d'acqua, con conseguenze dannose per gli ecosistemi acquatici [92]. Il grafico in Figura 5.32 mostra le emissioni di gas serra per chilogrammo di vari prodotti alimentari, misurate in equivalenti di anidride carbonica. I dati evidenziano che la carne bovina (allevamento per carne) è di gran lunga la più impattante, con 99,48 kg di CO<sub>2</sub>eq per chilogrammo di prodotto, seguita da agnello e montone con 39,72 kg e carne bovina da latticini con 33,3 kg. Altri prodotti di origine animale come gamberi allevati, formaggio, carne di maiale e pollame hanno emissioni significativamente inferiori, ma ancora sostanziali rispetto ai prodotti vegetali. Tra i prodotti vegetali, il tofu (3,16 kg) e il latte (3,15 kg) presentano emissioni relativamente basse, mentre prodotti come i piselli (0,98 kg) e le patate (0,46 kg) risultano tra i meno impattanti. Il grafico illustra chiaramente come la produzione di carne, in particolare quella bovina, abbia un impatto ambientale molto più elevato rispetto alla produzione di alimenti vegetali, enfatizzando l'importanza di considerare le scelte alimentari per ridurre le emissioni globali di gas serra.

## Greenhouse gas emissions per kilogram of food product

Our World  
in Data

Emissions are measured in carbon dioxide equivalents (CO<sub>2</sub>eq). This means non-CO<sub>2</sub> gases are weighted by the amount of warming they cause over a 100-year timescale.



Source: Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers.

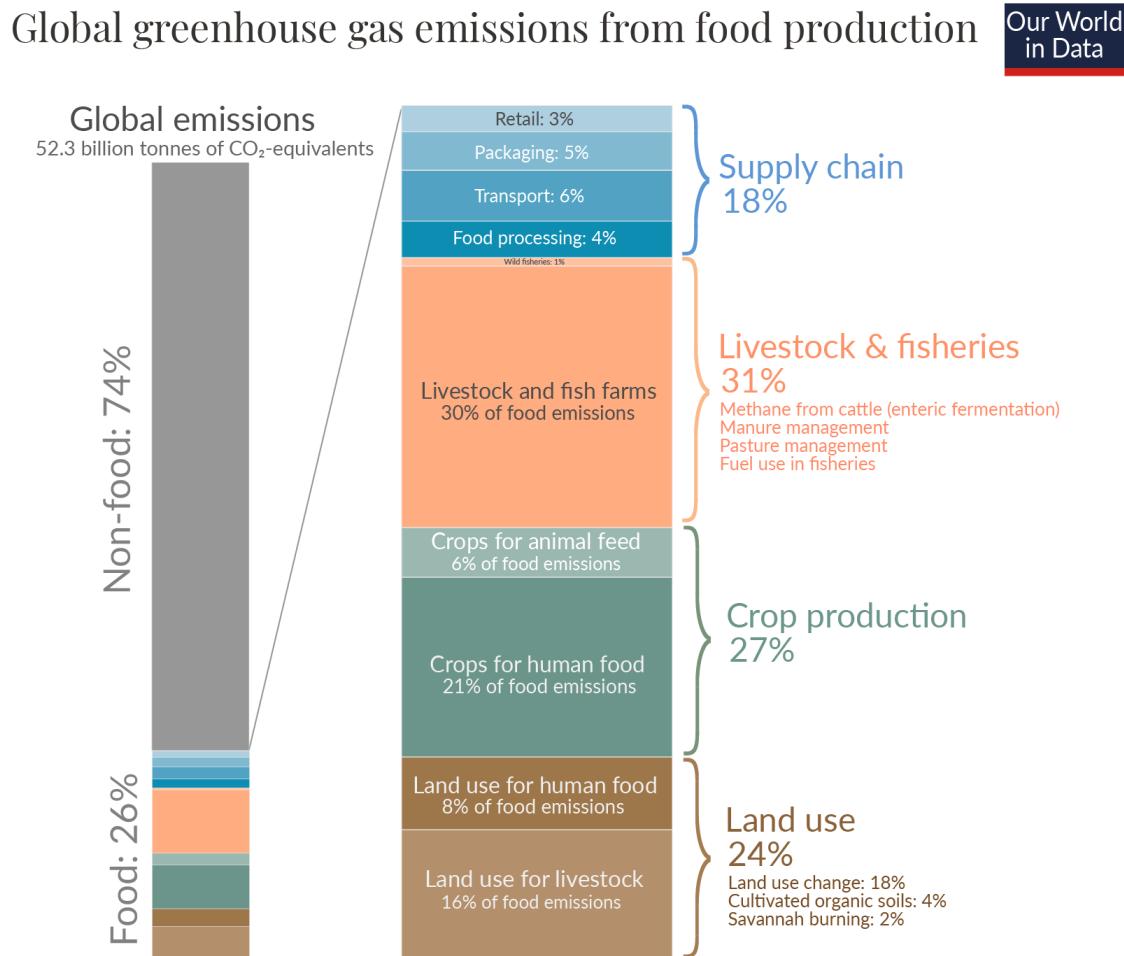
Note: Greenhouse gases are weighted by their global warming potential value (GWP100). GWP100 measures the relative warming impact of one molecule of a greenhouse gas, relative to carbon dioxide, over 100 years.  
OurWorldInData.org/environmental-impacts-of-food • CC BY

Figura 5.32: Emissione di gas serra per chilogrammo di cibo prodotto [94]

Il grafico in Figura 5.33 illustra le emissioni globali di gas serra derivanti dalla produzione alimentare, suddivise in diverse categorie. La produzione di cibo rappresenta il 26% delle emissioni globali, pari a 52,3 miliardi di tonnellate di CO<sub>2</sub> equivalenti. All'interno di questo 26%, le emissioni sono ulteriormente suddivise in quattro principali settori: la catena di approvvigionamento, gli allevamenti e la pesca, la produzione di colture e l'uso del suolo.

- *Catena di approvvigionamento* (18%): Include le emissioni derivanti dalla vendita al dettaglio (3%), imballaggio (5%), trasporto (6%) e trasformazione degli alimenti (4%). Questo settore rappresenta una parte significativa delle emissioni, ma è meno impattante rispetto agli allevamenti e alla produzione agricola.
- *Alllevamenti e pesca* (31%): Comprende le emissioni di metano provenienti dalla fermentazione enterica nei bovini, la gestione del letame, la gestione dei pascoli e l'uso di carburante nella pesca. Questo settore è il maggior contributore alle emissioni della produzione alimentare, sottolineando l'impatto significativo degli allevamenti intensivi sul clima.
- *Produzione di colture* (27%): Le colture destinate all'alimentazione umana costituiscono il 21% delle emissioni alimentari, mentre le colture per l'alimentazione animale rappresentano il 6%. Questo dato evidenzia che una parte sostanziale delle risorse agricole è destinata alla produzione di foraggi per il bestiame.
- *Uso del suolo* (24%): Include il cambiamento di uso del suolo (18%), la coltivazione di suoli organici (4%) e la combustione delle savane (2%). L'uso del suolo per l'allevamento rappresenta il 16% delle emissioni alimentari, mentre l'uso del suolo per l'alimentazione umana è responsabile dell'8%. Questo settore mette in luce l'impatto della deforestazione e della conversione delle terre per scopi agricoli.

Il grafico evidenzia quindi che la produzione di carne e i relativi usi del suolo rappresentano una porzione significativa delle emissioni globali di gas serra.



Data source: Joseph Poore & Thomas Nemecek (2018). Reducing food's environmental impacts through producers and consumers. Published in *Science*. Licensed under CC-BY by the author Hannah Ritchie (Nov 2022).

Figura 5.33: Emissioni di gas serra nella produzione di cibo [95]

Dai documenti la preoccupazione per l'impatto del consumo di carne sull'ambiente è ben evidente ma evince una forte polarizzazione tra difensori e critici del veganesimo. Chi sostiene il veganesimo evidenzia dati sulle emissioni di gas serra dell'allevamento animale, mentre i critici sollevano questioni legate alla complessità dei sistemi alimentari e alle sfide pratiche della transizione verso una dieta vegana:

- ”[@tag] [@tagDiCaprio] why aren't you vegan? meat industry contributes GREATLY to climate change and deforestation!”
- ”[@tag]: Meat production is a leading cause of climate change, water waste, deforestation, extinction. #WorldVeganDay”
- ”[@tag]: Eating meat is a leading cause of habitat destruction, species extinction and climate change: [link]”
- ”[@tag]: Meat industry's devastating role in climate change highlighted in new Leonardo DiCaprio film [link]”

- ”[@tag]: Did you know the avg American eats more than 210 lbs of meat every yr?! Here’s how that’s affecting climate change:[link]”
- ”[@tag] [@tag] How inconvenient. Let’s have a climate change fundraiser where we serve bacon cheeseburger... [link]”
- ”LMAO Damn, all vegans are slow. The animals we eat are those affected by climate change, they do not cause or afflict...”
- ”[@tag]: seriously just be vegan at least once a week. animal agriculture is the biggest contributor to global warming, try the fuck...”
- ”[@tag]: vegans don’t fkn care that lions eat meat bc they’re biologically designed to and their eating habits aren’t causing climate change!!!”
- ”[@tag]: ‘Republicans are to blame for global warming,’ said the 17 y/o vegan trans feminist who drives a 1970s chevy pickup..”

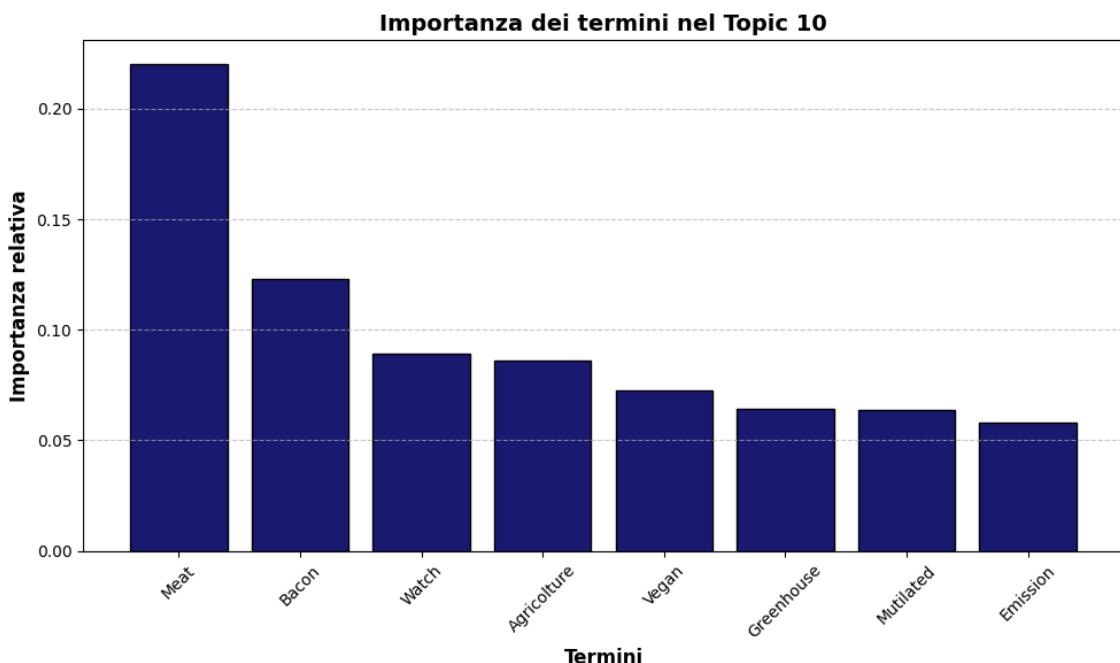


Figura 5.34: Keyword-Peso Consumo di carne, allevamenti e impatto

I termini ”*Bacon*” e ”*Meat*” indicano un forte interesse e una frequenza elevata di discussioni sui prodotti derivati dalla carne, spesso al centro dei dibattiti sui benefici e sui costi del consumo di carne, con una polarizzazione tra difensori del consumo tradizionale e critici per ragioni di salute e ambientali. ”*Agriculture*” evidenzia il legame tra produzione animale e pratiche agricole, con discussioni su come gli allevamenti intensivi influenzino l’uso del suolo, la deforestazione e la perdita di biodiversità. Il termine ”*Mutilated*” indica una critica verso le pratiche di trattamento degli animali, riflettendo preoccupazioni etiche. ”*Greenhouse*” e ”*Emission*” sono centrali nelle discussioni sull’impatto ambientale degli allevamenti, con una consapevolezza crescente dell’opinione pubblica riguardo alle emissioni di gas serra e al loro

---

contributo al cambiamento climatico, evidenziando la necessità di ridurre le emissioni attraverso pratiche agricole sostenibili e una riduzione del consumo di carne. "Vegan" rappresenta le alternative dietetiche al consumo di carne, con discussioni sui benefici ambientali e salutari delle diete vegane e sugli aspetti etici del trattamento degli animali, riflettendo un cambiamento nelle preferenze dei consumatori verso stili di vita più sostenibili. "Watch" suggerisce un crescente interesse pubblico verso la trasparenza e la responsabilità delle pratiche agricole, nonché un impegno verso un consumo informato e consapevole. Le keyword analizzate evidenziano come il consumo di carne e le pratiche agricole associate siano centrali nei dibattiti sull'ambiente, con una particolare attenzione agli impatti negativi sul clima e alla ricerca di alternative più sostenibili, evidenziando la necessità di politiche e pratiche più sostenibili nel settore agroalimentare. Per concludere gli allevamenti intensivi e le pratiche agricole contribuiscono significativamente al cambiamento climatico, alla perdita di biodiversità, alla deforestazione, alla scarsità d'acqua e alla degradazione degli ecosistemi. Questi impatti sottolineano la necessità urgente di adottare pratiche agricole più sostenibili e di promuovere diete a minore impatto ambientale per mitigare i danni ambientali associati alla produzione di carne.

---

## 5.2 Analisi del sentiment

Lo studio dei sentimenti e delle emozioni espresse sui social media in relazione al cambiamento climatico riveste un’importanza cruciale. Le emozioni influenzano significativamente la percezione pubblica e possono modellare il dibattito sociale, guidare le politiche ambientali e determinare il coinvolgimento delle comunità nelle iniziative di mitigazione e adattamento. Comprendere quali emozioni predominano in queste discussioni fornisce un quadro chiaro dei sentimenti collettivi e delle preoccupazioni principali della popolazione. L’analisi del sentiment, come spiegato in precedenza si è suddivisa in due parti:

- Un analisi della distribuzione e delle parole chiave associate ad ogni polarità per la sola parte del corpus dati proveniente da Twitter.
- Analisi delle emozioni sull’intero corpus basandosi sul lessico: *”Lexicon NRC Word-Emotion Association (EmoLex)“*.

I risultati che seguono forniscono una visione approfondita delle emozioni prevalenti e della loro distribuzione all’interno del corpus. Questa sezione illustra le principali emozioni identificate e le loro frequenze relative, offrendo una comprensione più chiara del sentimento generale riguardante il cambiamento climatico sui dati analizzati.

### 5.2.1 Distribuzione e parole chiave

L’etichettatura manuale effettuata sul dataset contenente i testi provenienti da Twitter, come spiegato al paragrafo 4.7.1, ci permette di analizzare in che modo questa parte dei dati viene classificata e quali sono le distribuzioni e le parole chiave associate ad ogni polarità. Il grafico a torta in Figura 5.35 mostra la distribuzione del sentimento suddivisa in quattro categorie: “Negativo” (9,1%), “Positivo” (52,3%), “Neutro” (17,6%) e “Altro” (21,1%).

Polarità	Etichettato	Significato
-1	Negativo	Il testo non crede nel cambiamento climatico causato dall’uomo.
0	Neutro	Il testo nè supporta nè rigetta la convinzione del cambiamento climatico causato dall’uomo.
1	Positivo	Il testo supporta la convinzione del cambiamento climatico causato dall’uomo.
2	Altro	Il testo contiene un link o del testo relativo a notizie sul cambiamento climatico.

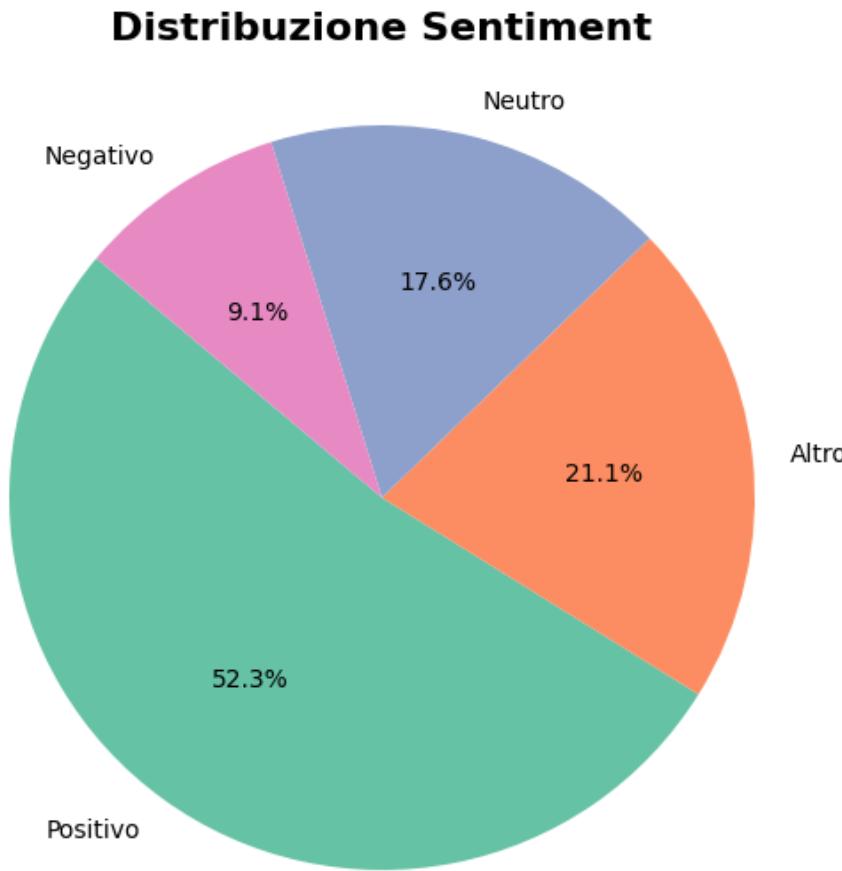


Figura 5.35: Grafico a Torta distribuzione sentiment

- La maggior parte dei testi rientra nella categoria ‘Positivo’, che rappresenta più della metà della distribuzione. Questo suggerisce una risposta prevalentemente positiva all’interno del set di dati esaminato.
- La categoria ‘Altro’ detiene poco più di un quinto del totale, indicando una parte significativa di sentimenti che non si adattano alle classiche classificazioni positivo-negativo-neutro. Considerazione comprensibile anche esplorando il dataset, composto da una moltitudine di righe relative a notizie e/o contenenti link diretti a notizie scientifiche, politiche e istituzionali.
- I sentimenti ‘Neutri’ costituiscono una frazione minore ma comunque maggiore al negativo, il che potrebbe implicare che le risposte tendono a polarizzarsi verso il positivo o il neutro piuttosto che verso il negativo.
- Infine, i sentimenti ‘Negativi’ sono i meno rappresentati in questa distribuzione rappresentando poco meno di un decimo del totale.

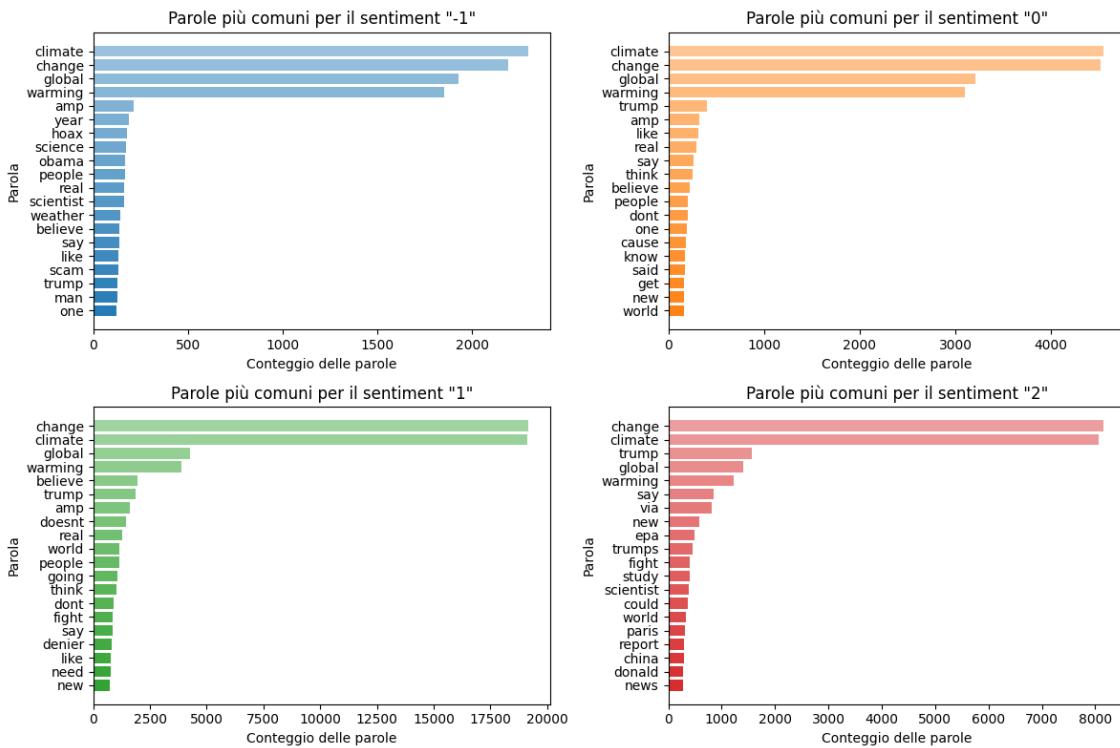


Figura 5.36: 20 Parole più comuni per ogni polarità

I grafici in Figura 5.36 presentano le parole più comuni associate ai diversi sentimenti: *"Negativo"*, *"Neutro"*, *"Positivo"*, *"Altro"*. Per il sentimento negativo (-1), le parole principali includono *"climate"*, *"change"*, *"global"*, *"warming"*, *"hoax"*, *"science"*, *"real"*, e *"scam"*, suggerendo un marcato scetticismo verso il cambiamento climatico. La presenza di termini come *"hoax"* e *"scam"* evidenzia la percezione del cambiamento climatico come una costruzione ingannevole. La frequenza elevata di termini scientifici come *"science"* e *"scientist"* indica che queste discussioni coinvolgono argomentazioni che sfidano le evidenze scientifiche consolidate, vedendole come una cospirazione. Nel sentimento neutro (0), le parole chiave come *"climate"*, *"change"*, *"global"*, *"warming"*, *"trump"*, *"like"*, e *"real"* suggeriscono discussioni generali sull'argomento, senza una chiara posizione emotiva. La presenza del termine *"trump"* indica che queste discussioni possono avere un contesto politico, mantenendo però un tono neutrale. Per il sentimento positivo (1), le parole principali sono *"change"*, *"climate"*, *"global"*, *"warming"*, *"believe"*, e *"real"*, suggerendo un forte riconoscimento e accettazione del cambiamento climatico come fenomeno reale. La presenza del termine *"believe"* denota fiducia nelle evidenze scientifiche riguardanti il cambiamento climatico. Termini come *"real"* e *"world"* indicano un approccio pragmatico e globale alla questione. Nel sentimento *"Altro"* (2), le parole chiave includono *"change"*, *"climate"*, *"trump"*, *"global"*, *"warming"*, *"via"*, *"new"*, e *"epa"*, indicando che queste discussioni possono riguardare notizie, rapporti o testi relativi al cambiamento climatico. La frequenza di termini come *"news"* e *"report"* suggerisce che molti dei testi in questa categoria trattano di informazioni e aggiornamenti riguardanti il cambiamento climatico, con un'enfasi sulle notizie politiche e istituzionali, come indicato dalla presenza dei termini *"trump"* e *"epa"*. Osservando i grafici, emerge chiaramente che in tutti i sentimenti, le parole *"climate"* e *"change"* sono predominanti, il che è coerente con il tema centrale del cambiamento climatico.

La presenza del termine *"trump"* in tutti i sentimenti sottolinea la politicizzazione delle discussioni sul cambiamento climatico. Il contrasto tra le parole nei sentimenti negativi e positivi evidenzia la divisione tra chi riconosce il cambiamento climatico e chi ne è scettico. In seguito in Figura 5.37 viene mostrata la Word Cloud delle parole più frequenti per ogni polarità.

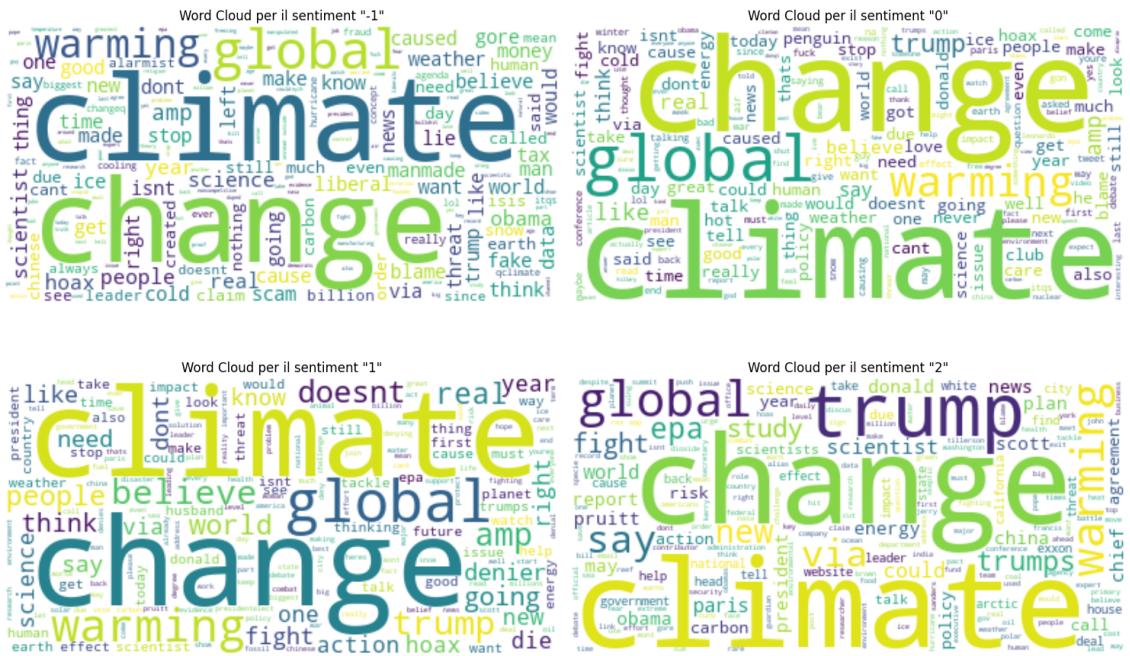


Figura 5.37: Word Cloud delle parole più frequenti per ogni polarità

### 5.2.2 Analisi delle emozioni complessiva

L'analisi delle emozioni sui dati dei social media relativi al cambiamento climatico ha rivelato una distribuzione eterogenea delle emozioni, come illustrato nel grafico a torta in Figura 5.38. Il grafico evidenzia che la paura è l'emozione dominante con il 23.5%, seguita dalla fiducia al 19.4%. Questi dati suggeriscono una forte preoccupazione tra gli utenti riguardo agli effetti del cambiamento climatico, accompagnata tuttavia da un significativo livello di fiducia, probabilmente nei confronti delle informazioni scientifiche o delle azioni intraprese per mitigare tali effetti.

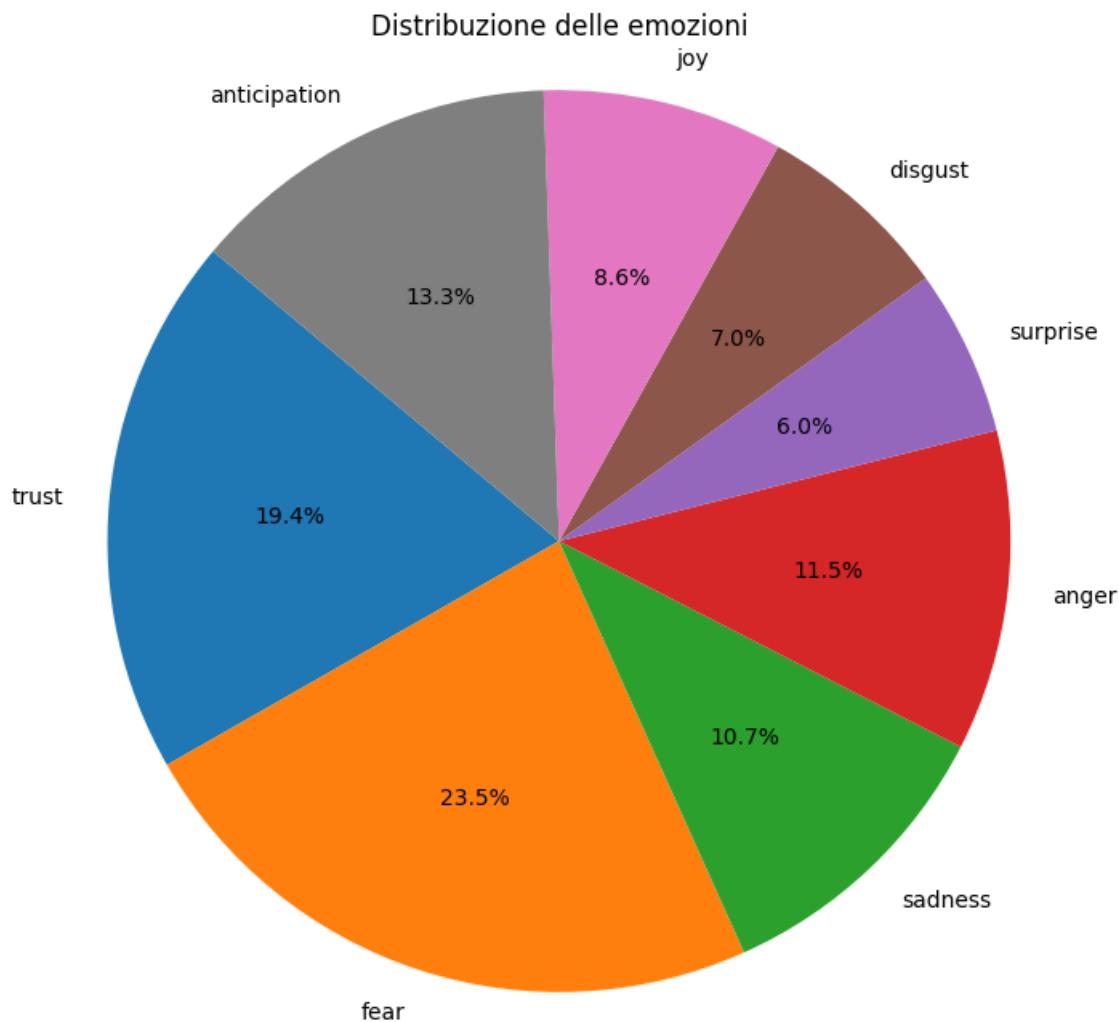


Figura 5.38: Grafico a torta della distribuzione delle emozioni nel dataset

L'anticipazione rappresenta il 13.3% del sentimento complessivo, indicando una sensazione di aspettativa e forse di speranza riguardo agli sviluppi futuri e alle soluzioni potenziali. La tristezza (10.7%) e la rabbia (11.5%) evidenziano una risposta emotiva negativa consistente, suggerendo un senso di perdita e frustrazione diffusi tra gli utenti. Emozioni meno prevalenti come la gioia (8.6%), il disgusto (7.0%) e la sorpresa (6.0%) forniscono ulteriori spunti di riflessione: la gioia potrebbe essere legata a notizie positive o successi nelle iniziative ambientali, mentre il disgusto e la sorpresa potrebbero riflettere reazioni a eventi climatici estremi o comportamenti politici-istituzionali percepiti come irresponsabili. La predominanza di emozioni

negative, in particolare la paura, potrebbe indicare la necessità di campagne comunicative più efficaci per informare e rassicurare il pubblico. La significativa presenza di fiducia e anticipazione suggerisce comunque una base solida su cui costruire narrative positive e proattive.

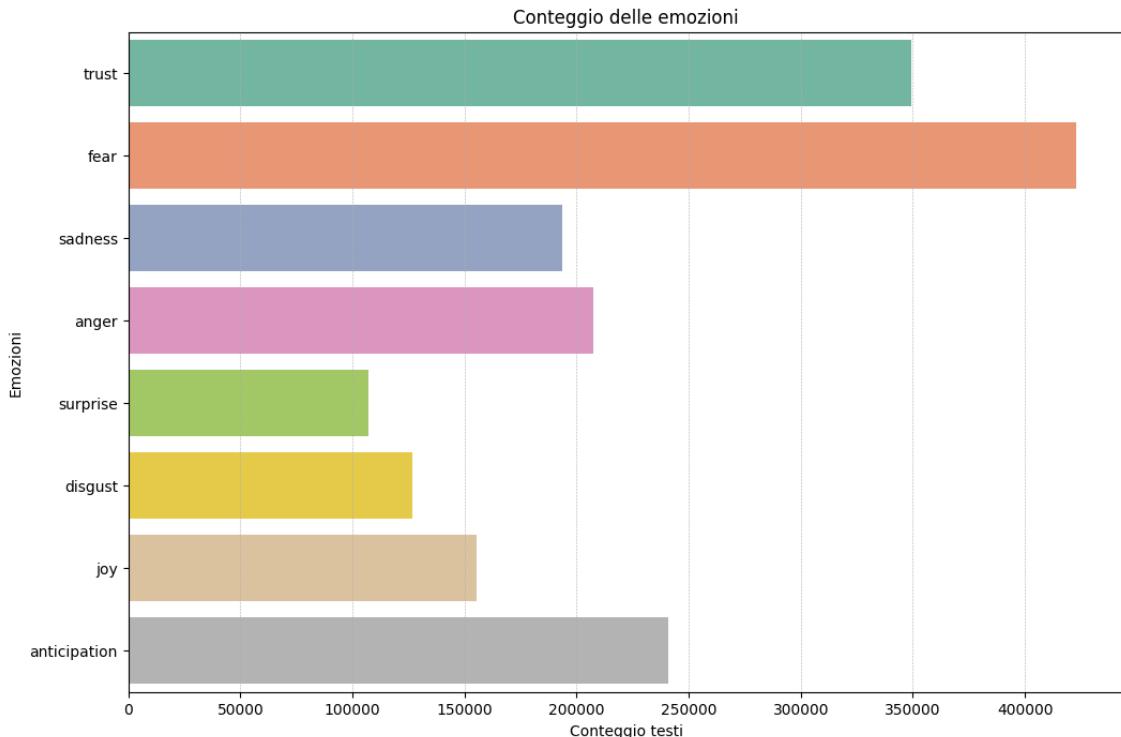


Figura 5.39: Conteggio testi per emozione

Il grafico a barre fornito (Fig 5.39) illustra la distribuzione delle emozioni nel dataset analizzato. I conteggi possono superare la totalità dei dati poiché ogni documento può essere assegnato a più emozioni contemporaneamente. La "paura" emerge come l'emozione predominante, con oltre 400.000 occorrenze, seguita da "fiducia" e "anticipazione" con conteggi superiori a 300.000 e 250.000 rispettivamente. Le emozioni come "rabbia" e "tristezza" mostrano conteggi significativi, intorno ai 200.000, mentre "gioia" e "disgusto" hanno una presenza minore, con conteggi attorno ai 150.000. "Sorpresa" risulta essere l'emozione meno rappresentata con meno di 100.000 occorrenze.

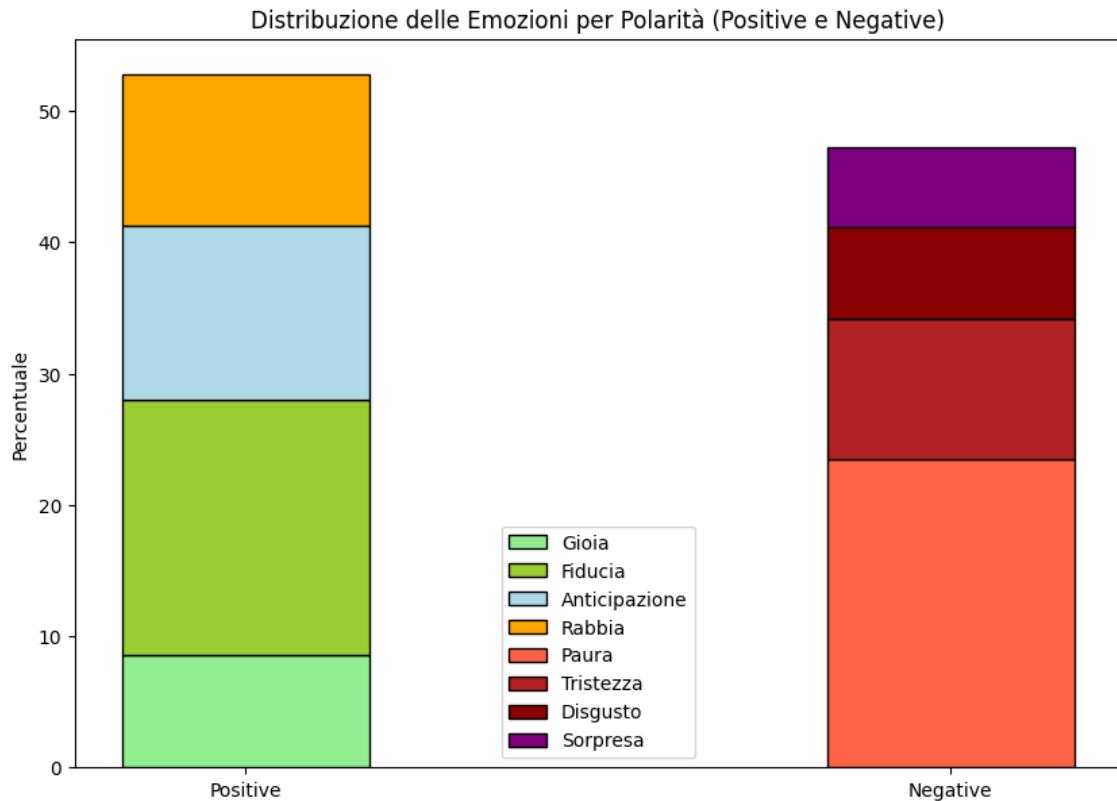


Figura 5.40: Distribuzione delle emozioni per polarità

Il grafico sopra (Fig 5.40) illustra la distribuzione complessiva delle emozioni categorizzate in base alla loro valenza nella teoria di Plutchik, con le emozioni positive (gioia, fiducia, anticipazione e rabbia) confrontate con quelle negative (paura, tristezza, disgusto e sorpresa). Dall'analisi, emerge che il 52.8% delle emozioni analizzate nei dati è di valenza positiva, mentre il 47.2% è di valenza negativa. Questo equilibrio indica che le discussioni sul cambiamento climatico evocano una gamma diversificata di emozioni, con una leggera preponderanza di sentimenti positivi. In particolare, la fiducia (19.4%) e l'anticipazione (13.3%) sono le emozioni positive più prevalenti, suggerendo che una parte significativa degli utenti manifesta una visione ottimistica e proattiva riguardo al cambiamento climatico. D'altro canto, la paura (23.5%) è l'emozione negativa più rappresentata, riflettendo preoccupazioni diffuse e timori associati alle conseguenze del cambiamento climatico. Questa distribuzione delle emozioni è significativa per comprendere come la narrativa sul cambiamento climatico viene percepita e discussa. La presenza sostanziale di emozioni positive potrebbe essere indicativa di un'adesione crescente alle misure proattive per affrontare il cambiamento climatico. Tuttavia, la concomitante presenza di emozioni negative sottolinea l'urgenza e la serietà percepite del problema.

### 5.2.3 Analisi delle emozioni per topic

Distribuzione delle emozioni per i diversi Topic

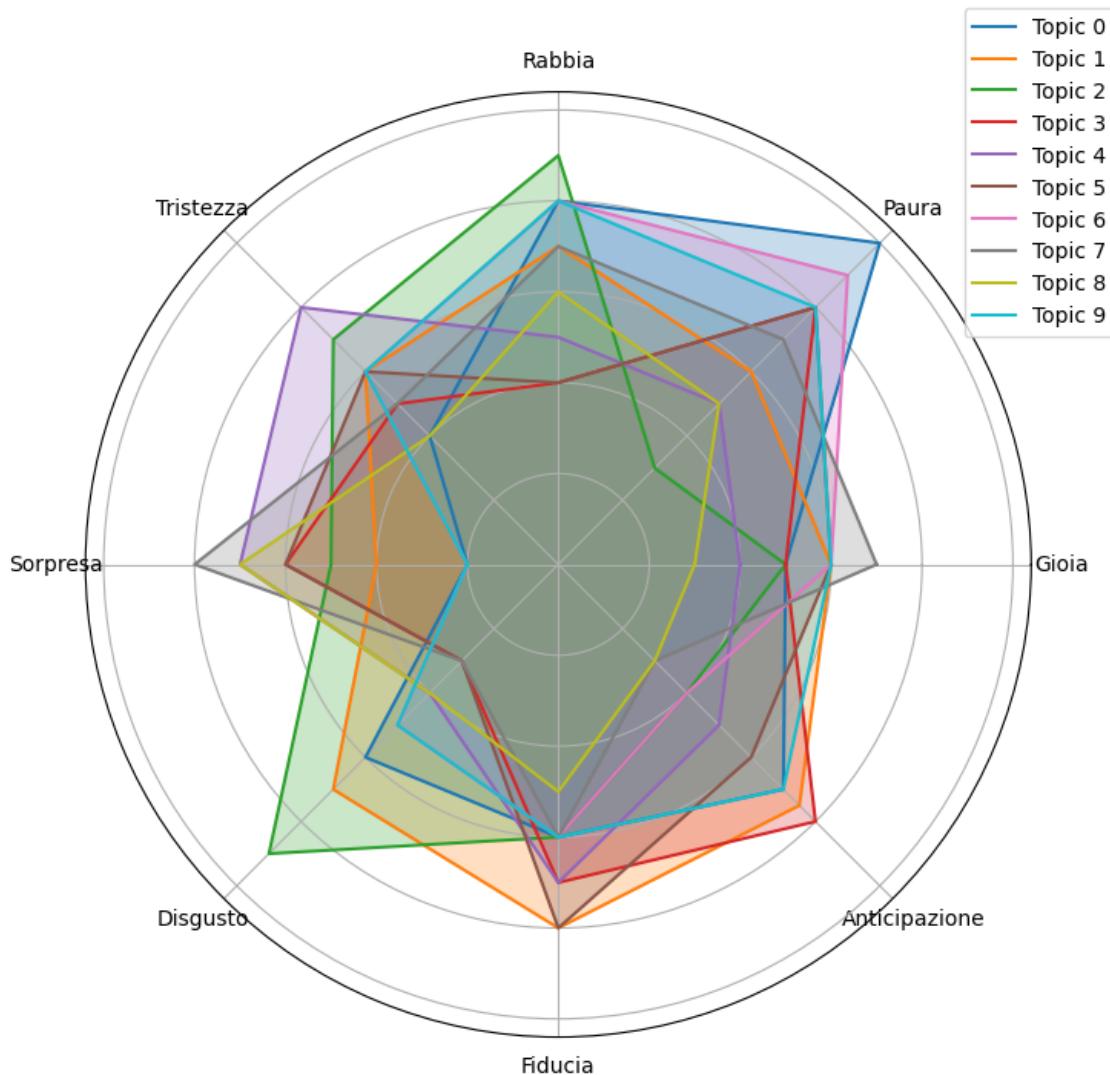


Figura 5.41: Distribuzione delle emozioni per topic

Il grafico radar presentato in Figura 5.41 illustra la distribuzione delle emozioni associate a diversi topic riguardanti il cambiamento climatico. L'analisi evidenzia una predominanza di emozioni di valenza positiva, come anche visto in Figura 5.40 , con il 52.8% delle emozioni analizzate che rientra in questa categoria. In particolare, emozioni come fiducia, gioia, anticipazione e rabbia sono frequentemente associate a vari temi. Ad esempio, il topic 1 (Cooperazione internazionale e clima) mostra elevati livelli di fiducia e anticipazione, indicando un ottimismo e una speranza significativi nelle soluzioni collaborative internazionali per affrontare il cambiamento climatico. La presenza di emozioni positive è evidente anche nei topic 0 (Cambiamento climatico e impatto umano) e 6 (Riscaldamento globale) , dove queste emozioni si potrebbero riflettere un desiderio di azione e miglioramento. Nonostante la prevalenza di emozioni positive, emozioni di valenza negativa come paura, tristezza, disgusto e sorpresa sono più che presenti, rappresentando il 47.2% delle

---

emozioni analizzate, per esempio nel topic 2 (Cambiamento climatico e politica) e 5 (Industria, inquinamento e impatto ambientale) evincono livelli elevati di disgusto e paura, suggerendo una forte preoccupazione e frustrazione del pubblico riguardo alle risposte politiche e agli effetti negativi delle attività industriali sull'ambiente. Invece per esempio il topic 7 (Il cambiamento climatico fra bufala e realtà) è associato alla sorpresa, evidenziando lo stupore e la perplessità del pubblico di fronte alle controversie e alle cospirazioni. Per concludere, il grafico radar fornisce una visione più dettagliata delle diverse reazioni emotive associate ai vari topic del cambiamento climatico. La predominanza di emozioni di valenza positiva sottolinea un certo grado di ottimismo e fiducia del pubblico nelle soluzioni proposte per affrontare il cambiamento climatico. Tuttavia, la presenza significativa di emozioni negative riflette un senso di urgenza e preoccupazione. Questa distribuzione emotiva complessa evidenzia come ogni argomento venga rappresentato di per sé con emozioni e sentimenti diversi, e di come tali emozioni possano influenzare le percezioni e le azioni della società.



# Capitolo 6

## Conclusioni

### 6.1 Approccio innovativo

L’obiettivo principale di questo studio è esaminare le correlazioni esistenti tra gli argomenti e le emozioni che compongono il vasto dibattito sul cambiamento climatico. Attraverso un’analisi approfondita, si mira a individuare i principali fuochi d’attenzione all’interno di questa vasta e complessa questione, grazie a tale approccio è possibile ottenere una visione più completa e dettagliata dell’atteggiamento della società nei confronti di sfide di portata globale. L’incrocio di queste informazioni con i dati climatici disponibili, potrebbe, con ulteriori sviluppi, aprire la strada a nuove direzioni di ricerca e progresso. Questi risultati potrebbero quindi stimolare ulteriori approfondimenti, contribuendo così alla crescita e all’avanzamento della conoscenza nel campo del cambiamento climatico e delle sue interazioni con la società. Le nuove ricerche potrebbero abbracciare una vasta gamma di campi disciplinari, spaziando dal marketing per le campagne di sensibilizzazione, allo studio sui rapporti clima-società fino agli studi computazionali e alle ricerche climatiche stesse. Lo studio vuole sottolineare l’importanza dell’integrazione di approcci multidisciplinari per fornire nuovi strumenti per affrontare efficacemente queste sfide, migliorando la comprensione delle dinamiche sociali, psicologiche e comportamentali coinvolte. L’utilizzo di un approccio olistico per ricerche di questa tipologia è in grado di promuovere un’impronta positiva nell’affrontare una delle problematiche più urgenti del nostro tempo.

### 6.2 Limitazioni e aree di miglioramento

Nonostante l’approccio e i risultati promettenti e coerenti con le evidenze scientifiche, esistono diverse limitazioni che devono essere considerate per migliorare ulteriormente la ricerca. La raccolta dei dati relativi al dibattito sul cambiamento climatico è stata effettuata attraverso fonti online specifiche, il che potrebbe introdurre bias legati alla rappresentatività delle opinioni espresse sui social media. Inoltre, la natura dinamica e in continua evoluzione del dibattito online rende difficile catturare una visione completa e statica delle tendenze e delle emozioni predominanti. Un’altra limitazione riguarda l’accuratezza dei modelli di topic modeling e sentiment analysis utilizzati, che, nonostante siano potenti strumenti analitici, possono essere influenzati da vari fattori quali la qualità e la quantità dei dati, nonché le specifiche impostazioni algoritmiche adottate. Il campione di dati utilizzato, composto da 343.934

---

righe di dati, consente di ottenere una panoramica dettagliata e rappresentativa delle opinioni espresse online, permettendo di identificare tendenze, temi ricorrenti e variazioni di sentimenti in maniera robusta. Tuttavia, nonostante la quantità di dati raccolti, esistono alcune limitazioni intrinseche che devono essere riconosciute. In primo luogo, il campione potrebbe non essere completamente rappresentativo dell'intera popolazione globale che partecipa al dibattito sul cambiamento climatico. Le opinioni espresse online possono essere influenzate da vari fattori, come la demografia degli utenti di social media, le piattaforme utilizzate, e le dinamiche di partecipazione online, che potrebbero non riflettere perfettamente la diversità di opinioni esistente nella società. Inoltre, la qualità dei dati è un altro aspetto cruciale da considerare, nonostante il volume, la presenza di dati duplicati, rumorosi o irrilevanti può influenzare negativamente la precisione delle analisi. La natura eterogenea delle fonti online e la variabilità nel modo in cui le opinioni vengono espresse rendono complesso garantire un'accurata rappresentazione dei sentimenti e dei temi dominanti. Per migliorare future ricerche, sarebbe vantaggioso implementare modelli di machine learning più avanzati e personalizzati, aumentando la precisione delle analisi e permettendo una comprensione più dettagliata delle interazioni tra variabili climatiche e sociali. Per esempio per la parte di analisi legata ai sentimenti potrebbe portare a risultati migliori un approccio basato su apprendimento automatico invece che su elenchi di parole. Inoltre, un'area cruciale di miglioramento riguarda l'analisi dei dati su scala temporale. Integrare una dimensione temporale nelle analisi potrebbe rivelare come le emozioni e gli argomenti del dibattito evolvono nel tempo, fornendo una comprensione più profonda delle dinamiche e delle reazioni della società agli eventi climatici e alle politiche ambientali. Questo approccio temporale potrebbe evidenziare pattern emergenti e cambiamenti significativi nelle percezioni pubbliche, contribuendo a identificare momenti chiave per interventi di sensibilizzazione. In parallelo, un approccio su scala temporale potrebbe consentire di analizzare i risultati in correlazione con i dati climatici per evidenziare eventuali cambiamenti di opinioni sulla base degli eventi e dei cambiamenti climatici in atto. In conclusione, sebbene lo studio presenti dei risultati significativi nella comprensione delle dinamiche del dibattito sul cambiamento climatico, riconoscere e affrontare le limitazioni esistenti è essenziale per migliorare ulteriormente la qualità e l'impatto delle future ricerche in questo ambito.

### 6.3 Conclusioni e risultati chiave

Il presente studio ha analizzato 343.943 testi riguardanti il cambiamento climatico nell'arco di 34 mesi su Twitter e dalla creazione della piattaforma fino a settembre 2020 su Reddit. Attraverso l'applicazione del modello BERTopic, l'analisi delle distribuzioni delle polarità e delle parole chiave ad esse associate, e l'analisi del sentimento basata sulla teoria delle emozioni di Plutchik, è stata esaminata l'attenzione e l'interconnessione tra gli argomenti legati al cambiamento climatico. Dall'analisi del modello sono emersi dieci argomenti significativi: *Cambiamento climatico e impatto umano*: il topic più frequente, in cui vengono menzionate le cause antropiche del cambiamento climatico, le preoccupazioni e l'impatto che questo fenomeno ha sull'umanità. *Cooperazione internazionale e clima*: include riferimenti alle istituzioni che collaborano per mitigare gli effetti del cambiamento climatico e alla necessità espressa dagli utenti di una maggiore collaborazione internazionale. *Cambiamen-*

---

*to climatico e politica*: comprende i documenti riguardanti il ruolo cruciale della politica nella lotta contro il cambiamento climatico. *Inquinamento ed ecosistemi*: contiene preoccupazioni e riferimenti alla deturpazione degli ecosistemi e dell'ambiente. *Crisi, sanità e sicurezza pubblica*: evidenzia parallelismi tra il Covid-19 e il cambiamento climatico come crisi, e il rapporto tra la propagazione delle malattie e il cambiamento climatico, con attenzione alla sicurezza pubblica. *Industria, inquinamento e impatto ambientale*: fa riferimento all'impatto delle attività produttive industriali sull'ambiente, sugli ecosistemi e sulle emissioni. *Riscaldamento globale*: include preoccupazioni e riferimenti relativi al riscaldamento globale, alle anomalie delle temperature e al contributo antropologico all'innalzamento delle temperature. *Il cambiamento climatico tra bufala e realtà*: raccoglie espressioni di scetticismo e negazione del cambiamento climatico, con la rabbia come emozione prominente e tentativi di persuasione verso le bufale. *Impatto e controversie della società ExxonMobil*: contiene riferimenti all'impatto ambientale di grandi società petrolifere e alle controversie legate alle loro responsabilità. *Consumo di carne, allevamenti e impatto*: include riferimenti all'impatto ambientale degli allevamenti intensivi legati al consumo di carne e alle conseguenze che ne derivano. L'analisi del sentiment ha mostrato che la discussione positiva riguardo al cambiamento climatico è leggermente più diffusa di quella negativa. L'emozione più evocata nelle discussioni su Reddit e Twitter è la paura, seguita dalla fiducia. Le emozioni meno suscite sono la sorpresa e il disgusto. La paura espressa online è probabilmente legata alle gravi conseguenze dei disastri ambientali già avvenuti. Tuttavia, molte persone hanno mostrato fiducia verso le istituzioni e la scienza nell'affrontare la situazione attuale. Per quanto concerne le emozioni associate ai topic, i topic 1 (cooperazione internazionale) e 6 (riscaldamento globale) mostrano alti livelli di fiducia e anticipazione, indicando ottimismo e speranza nelle soluzioni proposte. Il topic 0 (cambiamento climatico e impatto umano) presenta un equilibrio tra emozioni positive come gioia e fiducia, e negative come paura, suggerendo un desiderio di azione mitigato da preoccupazioni. Al contrario, i topic 2 (cambiamento climatico e politica) e 5 (industria, inquinamento e impatto ambientale) evidenziano alti livelli di paura e disgusto, riflettendo preoccupazioni significative riguardo alle risposte politiche e agli effetti negativi delle attività industriali. Il topic 7 (cambiamento climatico fra bufala e realtà) è associato alla sorpresa, sottolineando la perplessità del pubblico di fronte alle controversie. I topic 3 (inquinamento ed ecosistemi) e 4 (crisi, sanità e sicurezza pubblica) mostrano emozioni miste, con una presenza significativa di fiducia e anticipazione ma anche di rabbia e tristezza, indicando un mix di speranza per la gestione della crisi ambientale e frustrazione per le sfide esistenti. Il cambiamento climatico, in quanto problema globale, è da anni ampiamente dibattuto su diverse piattaforme. Comprendere la percezione pubblica è fondamentale per lo sviluppo economico, l'elaborazione delle politiche, le decisioni sullo stile di vita e le politiche di sensibilizzazione. Le piattaforme social offrono una finestra unica sulle preoccupazioni, speranze e opinioni delle persone, rendendo possibile un approccio più informato e partecipativo nella lotta contro il cambiamento climatico. Future ricerche potrebbero approfondire ulteriormente l'interazione tra le diverse emozioni e gli argomenti trattati, contribuendo a una comprensione più completa della dinamica discorsiva sul cambiamento climatico. Solo unendo scienza, analisi, sensibilizzazione e partecipazione collettiva, possiamo sperare di affrontare con successo una delle sfide più grandi del nostro tempo.



# Bibliografia

- [1] Intergovernmental Panel on Climate Change's (IPCC). *AR6 Synthesis Report: Climate Change 2023*. 2023. URL: <https://www.ipcc.ch/report/sixth-assessment-report-cycle/>.
- [2] E.M. et. al. Cody. "Climate change sentiment on Twitter: An unsolicited public opinion poll". In: *PLOS ONE* 10.8 (2015). DOI: [10.1371/journal.pone.0136092](https://doi.org/10.1371/journal.pone.0136092).
- [3] S. K. Yeo et al. "The influence of temperature on #ClimateChange and #GlobalWarming discourses on Twitter". In: *JCOM* 16.05 (2017). DOI: [10.22323/2.16050201](https://doi.org/10.22323/2.16050201).
- [4] World Bank. *World Bank Climate Change*. 2024. URL: <https://www.worldbank.org/en/topic/climatechange>.
- [5] CultureDigitali. *La politica nell'era digitale, come i nuovi media plasmano l'opinione pubblica*. 2022. URL: <https://www.culturedigitali.org/la-politica-nellera-digitale-come-i-nuovi-media-plasmano-lopinione-pubblica/>.
- [6] Shikha K. Ratan, Tarun Anand e Jyoti Ratan. "Formulation of Research Question—Stepwise Approach". In: *Journal of Indian Association of Pediatric Surgeons* 24.1 (2019), pp. 15–20. DOI: [10.4103/jiaps.JIAPS\\_76\\_18](https://doi.org/10.4103/jiaps.JIAPS_76_18).
- [7] Rachael L. Shwom et al. "269Public Opinion on Climate Change". In: *Climate Change and Society: Sociological Perspectives*. Oxford University Press, ott. 2015. ISBN: 9780199356102. DOI: [10.1093/acprof:oso/9780199356102.003.0009](https://doi.org/10.1093/acprof:oso/9780199356102.003.0009). eprint: [https://academic.oup.com/book/0/chapter/157167035/chapter-ag-pdf/44944687/book\\\_9871\\\_section\\\_157167035.ag.pdf](https://academic.oup.com/book/0/chapter/157167035/chapter-ag-pdf/44944687/book\_9871\_section\_157167035.ag.pdf). URL: <https://doi.org/10.1093/acprof:oso/9780199356102.003.0009>.
- [8] European Investment Bank. *The EIB Climate Survey 2020-2021*. 112 Pages, PDF/EN. 2021. ISBN: 978-92-861-5021-0. DOI: [10.2867/5219](https://doi.org/10.2867/5219).
- [9] Pew Research Center. *Pew Research Center Pew Global Attitudes Project: Spring 2006 Survey*. Report. 15 Nation Survey. 2006. URL: <http://pewrsr.ch/10FbavU>.
- [10] Daniel Read et al. "What Do People Know About Global Climate Change? Survey Studies of Educated Laypeople". In: *Risk Analysis* 14.6 (1994), pp. 971–982. DOI: [10.1111/j.1539-6924.1994.tb00066.x](https://doi.org/10.1111/j.1539-6924.1994.tb00066.x). URL: <https://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.1994.tb00066.x>.
- [11] United Nations Development Programme. *People's climate vote*. 2021. URL: <https://www.undp.org/publications/peoples-climate-vote>.

- [12] Stuart Capstick et al. “International trends in public perceptions of climate change over the past quarter century: Erratum”. In: *Wiley Interdisciplinary Reviews: Climate Change* 6 (nov. 2014). DOI: 10.1002/wcc.321.
- [13] Stefan Wagner et al. “Challenges in Survey Research”. In: ago. 2020, pp. 93–125. ISBN: 978-3-030-32488-9. DOI: 10.1007/978-3-030-32489-6\_4.
- [14] Samson Ebenezar Uthirapathy e Dominic Sandanam. “Topic Modelling and Opinion Analysis On Climate Change Twitter Data Using LDA And BERT Model.” In: *Procedia Computer Science* 218 (2023). International Conference on Machine Learning and Data Engineering, pp. 908–917. ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2023.01.071>. URL: <https://www.sciencedirect.com/science/article/pii/S1877050923000716>.
- [15] Fang Qiao e Jago Williams. “Topic Modelling and sentiment analysis of global warming tweets”. In: *Journal of organizational and end user computing* 34.3 (dic. 2021), pp. 1–18. DOI: 10.4018/joeuc.294901. URL: <https://doi.org/10.4018/joeuc.294901>.
- [16] Emelie Rosenberg et al. “Sentiment analysis on Twitter data towards climate action”. In: *Results in Engineering* 19 (2023), p. 101287. ISSN: 2590-1230. DOI: <https://doi.org/10.1016/j.rineng.2023.101287>. URL: <https://www.sciencedirect.com/science/article/pii/S2590123023004140>.
- [17] Matthew Burgess et al. “Climate change opinion and recent presidential elections”. In: (gen. 2024). DOI: 10.5281/zenodo.10494414.
- [18] Anastasia Nikolaidou et al. “Assessing the EU Climate and Energy Policy Priorities for Transport and Mobility through the Analysis of User-Generated Social Media Content Based on Text-Mining Techniques”. In: *Sustainability* 16 (mag. 2024), p. 3932. DOI: 10.3390/su16103932.
- [19] Sylaios Georgios Effrosynidis Dimitrios e Avi Arampatzis. “Exploring climate change on Twitter using seven aspects: Stance, sentiment, aggressiveness, temperature, gender, topics, and disasters”. In: *PLoS one* 17.9 (set. 2022), e0274213. DOI: 10.1371/journal.pone.0274213. URL: <https://doi.org/10.1371/journal.pone.0274213>.
- [20] I. (n.d.) Chiari. *Introduzione alla linguistica computazionale*. EditoriLaterza, 2007. ISBN: 9788842082095.
- [21] A. M. TURING. “I.—COMPUTING MACHINERY AND INTELLIGENCE”. In: *Mind* LIX.236 (ott. 1950), pp. 433–460. ISSN: 0026-4423. DOI: 10.1093/mind/LIX.236.433. eprint: <https://academic.oup.com/mind/article-pdf/LIX/236/433/30123314/lix-236-433.pdf>. URL: <https://doi.org/10.1093/mind/LIX.236.433>.
- [22] Responsa. *Il linguaggio naturale nell'intelligenza artificiale: il Natural Language Processing (NLP)*. 2020. URL: <https://responsa.ai/il-linguaggio-naturale-nellintelligenza-artificiale-il-natural-language-processing-nlp/#:~:text=Il%20NLP%20nasce%20tradizionalmente%20negli,date%20in%20pasto%E2%80%9D%20ai%20computer..>
- [23] John Hutchins. “The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954”. In: (gen. 2004).

- [24] Sinteticamente. *Elaborazione del linguaggio naturale definizione e storia*. 2023. URL: <https://sinteticamente.com/elaborazione-del-linguaggio-naturale-definizione-e-storia/>.
- [25] Vaswani. et. al. “Attention is all you need”. In: *arXiv.org* (). URL: <https://arxiv.org/abs/1706.03762>.
- [26] Palwinder Kaur Mangat e Dr. Kamaljit Singh Saini. “Chapter 26 - Relevance of data mining techniques in real life”. In: *System Assurances*. A cura di Prashant Johri et al. Emerging Methodologies and Applications in Modelling. Academic Press, 2022, pp. 477–502. ISBN: 978-0-323-90240-3. DOI: <https://doi.org/10.1016/B978-0-323-90240-3.00026-6>. URL: <https://www.sciencedirect.com/science/article/pii/B9780323902403000266>.
- [27] Analyticsinsight. *The future of data revolution will be unstructured data*. 2020. URL: <https://www.analyticsinsight.net/the-future-of-data-revolution-will-be-unstructured-data/>.
- [28] C. Zhai e S. Massung. “Text Mining and its Applications to Intelligence, CRM and Knowledge Management”. In: *Text Data Management and Analysis* 1 (2007), pp. 1–37. URL: [https://link.springer.com/chapter/10.1007/978-3-642-20813-1\\_1](https://link.springer.com/chapter/10.1007/978-3-642-20813-1_1).
- [29] Morten Goodwin et al. “Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook”. In: *ICES Journal of Marine Science* 79 (gen. 2022), Figure.3. DOI: [10.1093/icesjms/fsab255](https://doi.org/10.1093/icesjms/fsab255).
- [30] Shubham Koli. *How to Evaluate the Performance of Clustering Algorithms Using Silhouette Coefficient*. 2023. URL: <https://medium.com/@MrBam44/how-to-evaluate-the-performance-of-clustering-algorithms-3ba29cad8c03>.
- [31] nltk. *Natural Language Toolkit*. 2024. URL: <https://www.nltk.org>.
- [32] Python Software Foundation. *re — Regular expression operations*. 2024.
- [33] Maarten Grootendorst. “BERTopic: Neural topic modeling with a class-based TF-IDF procedure”. In: *arXiv preprint arXiv:2203.05794* (2022).
- [34] The pandas development team. *pandas-dev/pandas: Pandas*. Ver. latest. Feb. 2020. DOI: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134). URL: <https://doi.org/10.5281/zenodo.3509134>.
- [35] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [36] J. D. Hunter. “Matplotlib: A 2D graphics environment”. In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [37] Andreas Mueller. *WordCloud for Python documentation*. Accessed: 2024-05-27. 2023. URL: [https://amueller.github.io/word\\_cloud/](https://amueller.github.io/word_cloud/).
- [38] Michael L. Waskom. “seaborn: statistical data visualization”. In: *Journal of Open Source Software* 6.60 (2021), p. 3021. DOI: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021). URL: <https://doi.org/10.21105/joss.03021>.
- [39] Twitter. *Twitter API Documentation*. 2024. URL: <https://developer.twitter.com/en/docs/twitter-api>.

- [40] Creative Commons. *Creative commons by 4.0*. 2024. URL: <https://creativecommons.org/licenses/by/4.0/deed.en>.
- [41] Ewan Klein Steven Bird e Edward Loper. *Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit*. 2009. URL: [http://nltk.org/book\\_1ed/](http://nltk.org/book_1ed/).
- [42] HumanAi. *Transformer, la rete neurale che presta attenzione*. 2023. URL: <https://www.humai.it/il-transformer-la-rete-neurale-che-presta-attenzione/#:~:text=Il%20Transformer%20utilizza%20un%20meccanismo,1%27elaborazione%20del%20linguaggio%20naturale..>
- [43] neptune.ai. *Bert and the transformer architecture*. 2024. URL: <https://neptune.ai/blog/bert-and-the-transformer-architecture>.
- [44] Nabi Nabiyev e Saber Malekzadeh. *Anomalous Sound Localization Estimation - Scientific Figure on ResearchGate*. Available from: [https://www.researchgate.net/figure/Work-principles-of-softmax-functionfig3\\_349662206](https://www.researchgate.net/figure/Work-principles-of-softmax-functionfig3_349662206). Feb. 2021. DOI: 10.13140/RG.2.2.25949.95201.
- [45] Maartengr. *BERTopic*. 2024. URL: <https://maartengr.github.io/BERTopic/index.html>.
- [46] Javatpoint ml. *Clustering*. 2024. URL: <https://static.javatpoint.com/tutorial/machine-learning/images/clustering-in-machine-learning.png>.
- [47] Saif M. Mohammad e Peter D. Turney. “Crowdsourcing a Word-Emotion Association Lexicon”. In: *Computational Intelligence* 29.3 (2013), pp. 436–465.
- [48] Saif Mohammad e Peter Turney. “Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon”. In: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Los Angeles, CA: Association for Computational Linguistics, 2010, pp. 26–34. URL: <https://aclanthology.org/W10-0204>.
- [49] National Research Council del Canada. *National Research Council del Canada*. 2024. URL: <https://nrc.canada.ca/en>.
- [50] Italian Six Seconds. *La ruota delle emozioni di plutchik*. 2022. URL: <https://italia.6seconds.org/2020/10/la-ruota-delle-emozioni-di-plutchik/>.
- [51] Manuela Aparicio e Carlos J. Costa. “Data visualization”. In: *Commun. Des. Q. Rev* 3.1 (gen. 2015), pp. 7–11. DOI: 10.1145/2721882.2721883. URL: <https://doi.org/10.1145/2721882.2721883>.
- [52] Clement Levallois et al. “Translating upwards: Linking the neural and social sciences via neuroeconomics”. In: *Nature reviews. Neuroscience* 13 (ott. 2012), pp. 789–97. DOI: 10.1038/nrn3354.
- [53] Centre for Science e The Netherlands Technology Studies Leiden University. *Vosviewer*. 2024. URL: <https://www.vosviewer.com>.
- [54] Ucsusa et. al. *Are humans major cause global warming?* 2021. URL: <https://www.ucsusa.org/resources/are-humans-major-cause-global-warming>.

- [55] Scripps. 2024. URL: <https://scripps.ucsd.edu>.
- [56] NOAA partners with Scripps Institution of Oceanography on unmanned systems. 2024. URL: <https://www.noaa.gov/media-release/noaa-partners-with-scripps-institution-of-oceanography-on-unmanned-systems>.
- [57] United Nations. 2024. URL: <https://unfccc.int/process-and-meetings/the-paris-agreement>.
- [58] United Nations. 2024. URL: <https://unric.org/it/>.
- [59] Intergovernmental Panel on Climate Change. 2024. URL: <https://www.ipcc.ch>.
- [60] United Nations. 2024. URL: <https://unfccc.int/process/bodies/supreme-bodies/conference-of-the-parties-cop>.
- [61] Anthony Leiserowitz et al. *Climate Change in the American Mind, April 2022*. New Haven, CT: Yale University e George Mason University, 2022.
- [62] ourworldindata. *global plastic production*. 2024. URL: <https://ourworldindata.org/plastic-pollution#global-plastic-production>.
- [63] Earth.org. *The biggest environmental problems of our lifetime*. 2024. URL: <https://earth.org/the-biggest-environmental-problems-of-our-lifetime/>.
- [64] ourworldindata. *plastic pollution*. 2024. URL: <https://ourworldindata.org/plastic-pollution>.
- [65] ourworldindata. *how much plastic and waste do we produce*. 2024. URL: <https://ourworldindata.org/faq-on-plastics#how-much-plastic-and-waste-do-we-produce>.
- [66] Nasa. *Nasa sea level*. 2024. URL: <https://sealevel.nasa.gov/faq/8/is-the-rate-of-sea-level-rise-increasing/>.
- [67] Our world in data. *Our World in data website*. 2024. URL: <https://ourworldindata.org>.
- [68] Canadian Ice Service. 2024. URL: <https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/about-ice-service.html>.
- [69] Climate central. *Climate research*. 2024. URL: <https://www.climatecentral.org/what-we-do>.
- [70] Anthroponumbers. *The Human Impacts Database*. 2024. URL: <https://www.anthroponumbers.org/catalog/>.
- [71] We forum. *Climate action top global agenda covid 19*. 2024. URL: <https://www.weforum.org/agenda/2020/05/climate-action-top-global-agenda-covid-19/>.
- [72] Camilo Mora, Tanya McKenzie, Ilyssa M. Gaw et al. “Over half of known human pathogenic diseases can be aggravated by climate change”. In: *Nature Climate Change* 12 (2022), pp. 869–875. DOI: 10.1038/s41558-022-01426-1. URL: <https://doi.org/10.1038/s41558-022-01426-1>.

- [73] Alfredo Rivera et al. “Global Greenhouse Gas Emissions: 1990-2020 and Preliminary 2021 Estimates”. In: (2021). Rhodium Group provides the most up-to-date global and country-level GHG emissions estimates each year. In 2021, as countries and economies began to recover from the worst of the pandemic, global emissions bounced back—rising 4.6% from 2020 levels.
- [74] Carbon Brief. *The Carbon Brief Profile: China*. Accessed: 2024-05-30. 2023. URL: <https://interactive.carbonbrief.org/the-carbon-brief-profile-china/>.
- [75] BBC News. *China: Insert Article Title Here*. Accessed: 2024-05-30. 2023. URL: <https://www.bbc.com/news/world-asia-china-57483492>.
- [76] National Centers for Environmental Information (NCEI). *Global Climate Report - Annual 2022*. National Oceanic e Atmospheric Administration (NOAA). Gen. 2022. URL: <https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213> (visitato il 03/06/2024).
- [77] European Commission, Directorate-General for Research e Innovation. *European Green Deal – Research innovation call*. Publications Office of the European Union, 2021. DOI: doi/10.2777/33415.
- [78] World Meteorological Organization. *State of global climate*. 2024. URL: <https://wmo.int/publication-series/state-of-global-climate>.
- [79] U.S. Environmental Protection Agency. *Causes of Climate Change*. <https://www.epa.gov/climatechange-science/causes-climate-change>.
- [80] William Sweet et al. *Global and regional sea level rise scenarios for the United States*. Rapp. tecn. Center for Operational Oceanographic Products e Services (U.S.), 2017. DOI: <https://doi.org/10.7289/v5/tr-nos-coops-083>.
- [81] Naomi Oreskes e Erik M. Conway. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York: Bloomsbury Press, 2010. ISBN: 978-1-59691-610-4.
- [82] Pew Research Center. *A Changing Climate: Global Attitudes and Government Policies*. 2019. URL: <https://www.pewresearch.org/global/2019/04/18/a-changing-climate-global-attitudes-and-government-policies/>.
- [83] Anthony Leiserowitz et al. “Climate Change in the American Mind: April 2020”. In: (2020). URL: <https://climatecommunication.yale.edu/publications/climate-change-in-the-american-mind-april-2020/>.
- [84] Riley E. Dunlap e Aaron M. McCright. “Challenging climate change: The denial countermovement”. In: *Climate change and society: Sociological perspectives*. Oxford, UK: Oxford University Press, 2015, pp. 300–332.
- [85] YouGov. *Where Climate Change Deniers Live*. Accessed: 2024-06-05. 2020. URL: <https://www.statista.com/statistics/19449/climate-change-skeptics-worldwide/>.
- [86] PRRI/AAR. *Climate Change Concern Index by Religious Affiliation*. <https://www.prri.org/research/climate-change-religion-survey>. Accessed: 2024-06-05. 2014.
- [87] Naomi Oreskes. *The Collapse of Western Civilization: A View from the Future*. Columbia University Press, 2015.

- [88] G. Supran, S. Rahmstorf e N. Oreskes. “Assessing ExxonMobil’s global warming projections”. In: *Science* 379.6628 (2023), eabk0063. DOI: 10.1126/science.abk0063. eprint: <https://www.science.org/doi/pdf/10.1126/science.abk0063>. URL: <https://www.science.org/doi/abs/10.1126/science.abk0063>.
- [89] Steve Coll. *Private Empire: ExxonMobil and American Power*. Penguin Press, 2012.
- [90] Geoffrey Supran e Naomi Oreskes. “Assessing ExxonMobil’s climate change communications (1977–2014)”. In: *Environmental Research Letters* 12.8 (2017), p. 084019.
- [91] Food and Agriculture Organization of the United Nations. *Tackling climate change through livestock: A global assessment of emissions and mitigation opportunities*. Rome, Italy: FAO, 2013. URL: <http://www.fao.org/3/i3437e/i3437e.pdf>.
- [92] H. Steinfeld et al. *Livestock’s long shadow: Environmental issues and options*. Rome, Italy: Food e Agriculture Organization of the United Nations, 2006. URL: <http://www.fao.org/3/a0701e/a0701e.pdf>.
- [93] M. M. Mekonnen e A. Y. Hoekstra. “A Global Assessment of the Water Footprint of Farm Animal Products”. In: *Ecosystems* 15.3 (2012), pp. 401–415. DOI: 10.1007/s10021-011-9517-8. URL: <https://link.springer.com/article/10.1007/s10021-011-9517-8>.
- [94] Hannah Ritchie, Pablo Rosado e Max Roser. “Environmental Impacts of Food Production”. In: *Our World in Data* (2022). <https://ourworldindata.org/environmental-impacts-of-food>.
- [95] Hannah Ritchie. “Food production is responsible for one-quarter of the world’s greenhouse gas emissions”. In: *Our World in Data* (2019). <https://ourworldindata.org/food-ghg-emissions>.



## **Ringraziamenti**

*A conclusione di questo lavoro, mi prendo qualche riga di spazio per menzionare chi mi ha accompagnato durante questo percorso.*

*In primis un sincero ringraziamento al professore Enrico Mensa e alla Dott.ssa Sabrina Villata per il prezioso supporto nel condurre la ricerca e nella stesura dell'elaborato.*

*Grazie ai miei amici Marco e Gabbo per essere stati sempre presenti anche durante questa ultima fase del mio percorso di studi. Grazie per aver condiviso con me tante giornate di studio e tanti momenti di spensieratezza.*

*Grazie al mio conquilino Fabrizio con la quale ho trascorso 3 anni indimenticabili che hanno reso questo percorso più leggero.*

*Un immenso grazie a Controllato e ACC per essermi vicino da sempre.*

*Grazie alla mia collega Federica che ha condiviso con me gioie e fatiche di questi anni trascorsi insieme.*

*Infine, per concludere, un grazie di cuore a mia mamma, mio papà e mio fratello per avermi sempre sostenuto e per avermi permesso di portare a termine gli studi universitari.*