



Università
Ca' Foscari
Venezia

Corso di Laurea magistrale
in Management e sostenibilità

Tesi di Laurea

Greenwashing nella rendicontazione di sostenibilità e intelligenza artificiale

Una revisione sistematica della letteratura

Relatore

Prof. Antonio Costantini

Correlatore

Prof. Salvatore Russo

Laureanda

Alessia Cosentino
Matricola 867424

Anno Accademico

2025/2026

*A te, nonna Amelia, la mia roccia,
la donna più forte che conosca,
dedico questo traguardo con tutto il mio cuore.
Spero di averti resa fiero di me
e che la tua forza e il tuo coraggio
possano ispirarmi e guidarmi
durante tutto il mio cammino.*

Dichiarazione del ricorso a strumenti di Intelligenza Artificiale (IA)

La sottoscritta Alessia Cosentino, matricola numero 867424, iscritta presso Ca' Foscari al corso di Laurea magistrale in Management e sostenibilità, autrice della tesi di laurea "Greenwashing nella rendicontazione di sostenibilità e intelligenza artificiale: una revisione sistematica della letteratura", relatore Antonio Costantini, consapevole delle sanzioni previste dall'art. 76 del Testo unico, D.P.R. 28/12/2000 n.445, e della decadenza dei benefici prevista dall'art. 75 del medesimo Testo unico in caso di dichiarazioni false o mendaci, sotto la propria personale responsabilità ai sensi degli artt. 46 e 47 del D.P.R. 28/12/2000 n. 445,

DICHIARA

di aver fatto ricorso a tecnologie di intelligenza artificiale generativa per supportare vari aspetti della ricerca e della scrittura della sua tesi di laurea, in particolare:

- *Claude*, tramite l'applicazione *claude.ai*, con il modello *Sonnet 5*;
- nel capitolo 3 lo strumento è stato impiegato per rielaborare alcune parti del testo da me redatto, al fine di migliorarne la chiarezza e la coerenza espositiva. Nello stesso capitolo mi è servito anche per elaborare tabelle e immagini esplicative dei risultati della revisione sistematica della letteratura;
- nel capitolo 4 l'ho utilizzato per rielaborare le spiegazioni dei singoli studi analizzati, così da renderne più chiara l'esposizione di obiettivi, metodologie, dati e risultati, e per elaborare le tabelle di sintesi presenti nel capitolo.

La correttezza dei contenuti è stata verificata confrontando ogni passaggio rielaborato con le fonti originali, per controllarne la fedeltà e la correttezza. I dati riportati nelle tabelle e nelle immagini sono stati verificati uno per uno.

DICHIARA INOLTRE

che tutto il materiale generato dall'IA è stato attentamente verificato, rivisto e rielaborato dall'autrice, che si assume la piena responsabilità per la qualità, l'accuratezza e l'integrità scientifica del lavoro.

INDICE

Indice delle abbreviazioni.....	5
Indice delle figure	7
Indice delle tabelle	9
Introduzione	11
Capitolo 1 La sostenibilità aziendale e la rendicontazione ESG.....	15
1.1. La sostenibilità aziendale	15
1.2. Il report di sostenibilità.....	19
1.3. L'evoluzione normativa europea dalla NFRD alla CSRD	24
1.3.1 Il principio della doppia materialità	27
1.4. Gli European Sustainability Reporting Standards – ESRS.....	29
1.5. La normativa digitale e tecnologica	33
1.5.1 La Green Claims Directive (GCD)	34
1.5.2 L'EU AI Act	35
1.5.3 L'Explainable AI (XAI) nell'auditing.....	37
Capitolo 2 Greenwashing nella rendicontazione ESG.....	41
2.1 Greenwashing: definizioni, tipologie ed evoluzione.....	41
2.1.1 Definizioni.....	41
2.1.2 Tipologie e tassonomie	43
2.1.3 Evoluzione digitale.....	46
2.2 Il disallineamento tra ESG Talk e ESG Walk: misurazione del gap prestazionale	47
2.3 Dalla comunicazione all'auditing: il fenomeno dell'assurance-washing.....	50
2.4 Le determinanti interne ed esterne del comportamento opportunistico.....	52
2.5 Le teorie organizzative.....	53
2.5.1 Institutional Theory.....	53
2.5.2 Legitimacy Theory.....	56
2.6 Le teorie economiche e relazionali	58
2.6.1 Signaling Theory	58
2.6.2 Resource Based View Theory.....	60
2.6.3 Stakeholder Theory	61
2.7 Le teorie comportamentali.....	63

2.7.1	La razionalità limitata.....	63
2.7.2	Il disimpegno morale	64
2.7.3	La dissonanza cognitiva	65
2.7.4	La teoria dei prospetti	65
Capitolo 3 Systematic literature review (SLR).....		69
3.1	Pianificazione della SLR.....	69
3.2	Conduzione della SLR	70
3.3	Rendicontazione e divulgazione dei risultati della SLR.....	73
Capitolo 4 Evidenze dalla letteratura sul ruolo dell'intelligenza artificiale nella rendicontazione ESG		87
4.1	Premessa	87
4.2	L'IA come strumento di rilevazione e contrasto del greenwashing	89
4.2.1	L'NLP tradizionale.....	90
4.2.2	I modelli pre-addestrati.....	96
4.2.3	I modelli generativi e il deep learning.....	102
4.2.4	I sistemi agentici	106
4.2.5	Che cosa misurano davvero le prestazioni dichiarate.....	108
4.2.6	Dall'individuazione al contrasto	111
4.2.7	Sintesi	118
4.3	L'IA generativa come agevolatore.....	120
4.3.1	Il rischio teorizzato	121
4.3.2	L'evidenza sperimentale	122
4.3.3	Il greenwashing algoritmico osservato.....	125
4.3.4	Complementarità istituzionale	126
4.3.5	Sintesi	129
4.4	Il controllo umano e istituzionale.....	131
4.4.1	Responsabilità algoritmica e supervisione umana	132
4.4.2	La supervisione interna	135
4.4.3	Il controllo esterno del mercato.....	138
4.4.4	Il presidio dell'assurance	140
4.4.5	Il contesto istituzionale.....	142

4.4.6 Sintesi	146
4.5 Considerazioni finali	147
4.5.1 I fattori che distinguono il ruolo deterrente da quello agevolatore.....	148
4.5.2 Un confronto con altre revisioni della letteratura	149
4.5.3 I limiti della letteratura.....	150
 Conclusioni.....	 153
Bibliografia.....	157
Fonti normative	171
Sitografia.....	173

Indice delle abbreviazioni

AI-GW: AI-based greenwashing detection

BERT: Bidirectional Encoder Representations from Transformers

CSR: Corporate Social Responsibility

CSRD: Corporate Sustainability Reporting Directive

CTI: Cheap Talk Index

DID: Staggered Difference-in-Differences

EFRAG: European Financial Reporting Advisory Group

EPA: Environmental Protection Agency

ESAP: European Single Access Point

ESG: Environmental, Social and Governance

ESGSI: ESG-washing Severity Index

ESRS: European Sustainability Reporting Standards

ETR: Environmental Target Responsibility system

GCD: Green Claims Directive

GPAI: General Purpose AI

GRI: Global Reporting Initiative

GSI: Greenwashing Severity Index

GWL: Corporate Greenwashing Index

IA: Intelligenza Artificiale

IFRS: International Financial Reporting Standards

ISSB: International Sustainability Standards Board

LDA: Latent Dirichlet Allocation

LLM: Large Language Model

MDGs: Millennium Development Goals

NFRD: Non-Financial Reporting Directive

NLP: Natural Language Processing

ONG: Organizzazioni Non Governative

PCCS: Packaging Circularity Commitment Score

PMI: Piccole e Medie Imprese

PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses

PSM: Propensity Score Matching

RAG: Retrieval-Augmented Generation
RBV: Resource-Based View
ROE: Return On Equity
RQ: Research Questions
SASB: Sustainability Accounting Standards Board
SDGs: Sustainable Development Goals
SEC: Securities and Exchange Commission
SEDI: Specific Environmental Disclosure Index
SLR: Systematic Literature Review
SRI: Socially Responsible Investing
TBL: Triple Bottom Line
TCFD: Task Force on Climate-related Financial Disclosures
TF-IDF: Term Frequency - Inverse Document Frequency
WCED: World Commission on Environment and Development
XAI: Explainable Artificial Intelligence

Indice delle figure

Figura 1 - Le tre dimensioni della sostenibilità	16
Figura 2 - Triple Bottom Line	17
Figura 3 - GRI standards	20
Figura 4 - I 17 Obiettivi di Sviluppo Sostenibile (SDGs).....	22
Figura 5 - Principi GRI.....	23
Figura 6 - Il principio della doppia materialità	28
Figura 7 - La struttura degli ESRS.....	30
Figura 8 - Cronoprogramma di entrata in vigore dei vincoli dell'AI Act (previsioni ad aprile 2024)	35
Figura 9 - Flusso logico e addestramento di un modello surrogato globale	38
Figura 10 - Mappa di co-occorrenza delle parole chiave nella letteratura sul greenwashing	41
Figura 11 - Tipologia d'impresa basata su performance ambientale e comunicazione....	42
Figura 12 - Classificazione multilivello del greenwashing.....	44
Figura 13 - Orientamento e tipologie delle asserzioni ecologiche di tipo "claim"	45
Figura 14 - le conseguenze del greenwashing.....	47
Figura 15 - Tassonomia greenwashing-assurance	51
Figura 16 - Il divario tra teoria e pratica (policy-practice decoupling) a confronto con l'allineamento di un'organizzazione ideale	55
Figura 17 - Il divario tra mezzi e fini (means-ends decoupling) a confronto con l'allineamento di un'organizzazione ideale	55
Figura 18 - Classificazione delle dinamiche di legittimazione (pragmatica, morale e cognitiva).....	57
Figura 19 - processo di segnalazione tra insider e outsider	59
Figura 20 - la stakeholder theory e la rete di relazioni tra l'impresa e i suoi portatori di interesse.....	62
Figura 21 - Il ruolo del disimpegno morale nei processi di inizio, facilitazione e perpetuazione dell'opportunismo aziendale.....	64
Figura 22 - diagramma di flusso PRISMA del processo di selezione degli articoli.....	73
Figura 23 - Andamento temporale delle pubblicazioni incluse (2022-2026)	81
Figura 24 - Distribuzione dell'ESGSI: l'indice (in blu) come differenza tra punteggio di tono (in verde) e punteggio di sostanza (in rosso)	91

Figura 25 - Distribuzione dei valori del GSI	91
Figura 26 - Framework AI- GW	94
Figura 27 - Distribuzione dell'informativa climatica fra le quattro categorie TCFD	97
Figura 28 - Quadro metodologico per la costruzione dell'indice di greenwashing (GWL)	98
Figura 29 - Architettura del modello per l'analisi delle informative di sostenibilità	99
Figura 30 - Schema del processo di analisi e classificazione del testo dei report.....	100
Figura 31 - Classificazione per i pilastri ambientale (A), sociale (B) e di governance (C)	101
Figura 32 - L'architettura a due stadi del sistema DeepGreen.....	102
Figura 33 - Il funzionamento della RAG nella valutazione degli indicatori dei piani di transizione	103
Figura 34 - Lo schema del sistema ibrido a due passaggi	105
Figura 35 - Composizione del punteggio ESG per pilastro	109
Figura 36 - Andamento greenwashing nelle imprese cinesi per livello di adozione dell'IA, 2011-2019.....	112
Figura 37 - I tre percorsi attraverso cui l'adozione dell'IA agisce sul comportamento dichiarativo.....	114
Figura 38 - Relazioni stimate fra adozione dell'IA, greenwashing e brownwashing	115
Figura 39 - Quadro concettuale per la valutazione delle implicazioni dell'IA sulla rendicontazione di sostenibilità	121
Figura 40 - Quadro concettuale dell'adozione dell'IA, del greenwashing e delle forze disciplinanti.....	127
Figura 41 - Framework sull'uso responsabile dell'IA e della supervisione degli amministratori.....	136
Figura 42 - Quadro dei meccanismi attraverso cui i tribunali ambientali incidono sul greenwashing	144

Indice delle tabelle

Tabella 1 - NFRD e CSRD somiglianze e differenze	27
Tabella 2 - Riduzione dei requisiti obbligatori "shall" e dei requisiti volontari "may"	32
Tabella 3 - Elenco dei 51 studi inclusi nella SLR	73
Tabella 4 - Tipologia di autore e tipologia di risorsa degli studi inclusi	81
Tabella 5 - Distribuzione degli articoli inclusi per rivista di pubblicazione.....	82
Tabella 6 - Natura metodologica dei contributi inclusi.....	84
Tabella 7 - le tre variabili chiave	93
Tabella 8 - Punteggi assegnati da tre modelli multimodali agli stessi report.....	107
Tabella 9 - Quadro degli studi di rilevazione	118
Tabella 10 - Studi sull'effetto dell'adozione dell'IA.....	119
Tabella 11 - I quattro contributi sulla RQ2	129
Tabella 12 - Confronto fra gli approcci alla revisione del greenwashing.....	134

Introduzione

Negli ultimi anni, la rendicontazione di sostenibilità ha conosciuto una progressiva istituzionalizzazione, passando da una pratica prevalentemente volontaria a un ambito soggetto a obblighi normativi (Todaro et al., 2024). In tale contesto, la Corporate Sustainability Reporting Directive (CSRD) e gli European Sustainability Reporting Standards (ESRS) hanno introdotto un quadro comune per la comunicazione delle informazioni di sostenibilità da parte delle imprese europee rientranti nel relativo ambito di applicazione, prevedendo obblighi informativi e forme di attestazione esterna (Fornasari e Traversi, 2024). Tale quadro mira a ridurre le asimmetrie informative tra imprese e stakeholder e ad accrescere la comparabilità e l'attendibilità di informazioni la cui selezione e rappresentazione sono state a lungo caratterizzate da un'ampia discrezionalità aziendale. L'ampliamento degli obblighi informativi, tuttavia, non si traduce necessariamente in una maggiore fedeltà della rappresentazione delle pratiche e delle performance aziendali. Il greenwashing, inteso come disallineamento tra la comunicazione di sostenibilità e le condotte effettive dell'impresa, può persistere anche in presenza di obblighi di rendicontazione, assumendo forme differenti. Tra queste rientrano la formulazione di impegni generici e difficilmente verificabili (cheap talk), la divulgazione selettiva delle informazioni favorevoli (cherry-picking) e un'adesione meramente formale agli standard, volta ad accrescere la credibilità percepita dell'impresa senza un corrispondente cambiamento delle pratiche aziendali. In questo scenario si inserisce la crescente diffusione dell'intelligenza artificiale (IA), le cui applicazioni interessano sia la predisposizione sia l'analisi dei report di sostenibilità. Da un lato, l'IA può supportare i soggetti incaricati del controllo nell'esame di ingenti volumi di documenti e nell'individuazione di elementi potenzialmente riconducibili a una comunicazione ingannevole. Dall'altro, può essere impiegata dalle imprese per migliorare la chiarezza e l'efficacia comunicativa dei report, ma anche per attenuare la visibilità delle criticità e adeguare i contenuti ai criteri utilizzati dagli strumenti di controllo. L'impiego dell'IA nella rendicontazione di sostenibilità presenta, pertanto, una potenziale ambivalenza (Liu et al., 2026): può rafforzare la capacità di individuare il greenwashing, ma anche favorire l'elaborazione di strategie comunicative più sofisticate, capaci di eludere i controlli. Da tale ambivalenza muove l'interrogativo centrale della presente tesi, che si propone di indagare se, e a quali condizioni, l'IA applicata ai report di sostenibilità contribuisca a contrastare oppure ad agevolare il greenwashing. Per approfondire le

diverse dimensioni del problema, tale interrogativo è stato articolato in tre domande di ricerca (RQ):

- RQ1: In che modo l'IA può essere impiegata per individuare anomalie documentali e contrastare il greenwashing nei report di sostenibilità?
- RQ2: In che modo l'IA può facilitare la manipolazione semantica, agevolando il greenwashing nei report di sostenibilità?
- RQ3: In che modo l'IA, unita al controllo umano e istituzionale, può limitare i rischi algoritmici e preservare la credibilità dei report di sostenibilità?

Sul piano metodologico, la ricerca adotta una systematic literature review (SLR), condotta secondo il protocollo in tre fasi proposto da Tranfield et al. (2003) e in conformità con le linee guida PRISMA 2020 (Page et al., 2021). La scelta di tale metodo risponde alla natura del tema indagato, che costituisce un ambito di studio recente, in rapida espansione e distribuito fra discipline differenti. In simili condizioni, un approccio sistematico e replicabile consente di contenere il rischio che la selezione delle fonti risulti condizionata dalle ipotesi iniziali del ricercatore. La ricerca bibliografica è stata condotta sul database Scopus e limitata agli studi in lingua inglese pubblicati fra il 2020 e il 2026. Al termine delle fasi di selezione, il corpus finale risulta composto da 51 contributi, di cui 44 studi primari e 7 revisioni della letteratura. Il lavoro si articola in quattro capitoli.

Il *primo capitolo* delinea il contesto di riferimento della ricerca. Ricostruisce il processo attraverso cui la sostenibilità è stata integrata nella gestione aziendale e l'evoluzione del quadro normativo europeo che ha reso obbligatoria la rendicontazione di sostenibilità, fino ai più recenti interventi regolatori in materia di intelligenza artificiale.

Il *secondo capitolo* è dedicato al greenwashing nella rendicontazione di sostenibilità. Ne esamina la definizione, le principali forme di manifestazione e le determinanti, anche alla luce dei quadri teorici che ne spiegano l'adozione da parte delle imprese.

Il *terzo capitolo* illustra l'impianto metodologico della revisione sistematica e descrive le fasi attraverso cui gli studi del corpus sono stati individuati, selezionati e analizzati.

Il *quarto capitolo* presenta e discute i risultati della revisione in relazione alle tre RQ.

Le conclusioni, infine, riprendono le risposte alle RQ e indicano alcune possibili direzioni per la ricerca futura.

Capitolo 1 La sostenibilità aziendale e la rendicontazione ESG

1.1 La sostenibilità aziendale

Il ruolo dell'impresa all'interno della società è cambiato nel tempo insieme al contesto economico e culturale. Nel secondo dopoguerra, in particolare durante il tardo capitalismo, si è affermata una visione dell'azienda orientata quasi esclusivamente alla massimizzazione del valore per gli azionisti, nota come *shareholder theory* (Zhang, 2011). Milton Friedman ne è diventato il principale portavoce nel 1970, con il saggio "*The Social Responsibility of Business Is to Increase Its Profits*", in cui sosteneva che «esiste una e solo una responsabilità del business – usare le sue risorse e impegnarsi in attività volte ad aumentare i suoi profitti [...]»¹ (Friedman, 1970). Per l'economista, il dirigente aziendale è un dipendente dei proprietari, e il suo compito primario è «condurre gli affari in accordo con i loro desideri, che in genere consisteranno nel fare più soldi possibile [...]»² (Friedman, 1970). L'organizzazione, di conseguenza, doveva rispondere in primo luogo all'interesse economico di breve periodo di proprietari e azionisti.

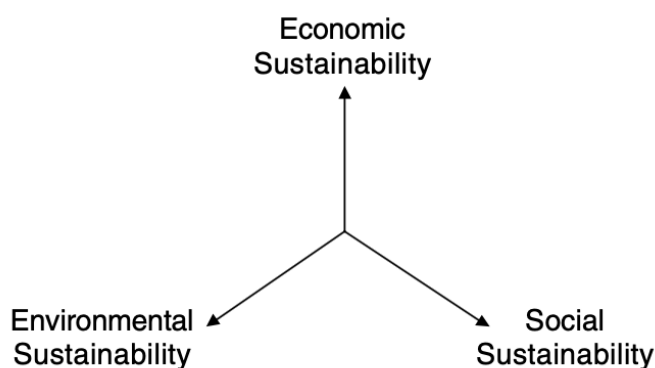
Questa impostazione ha iniziato a incrinarsi quando la società ha preso coscienza dei danni provocati da oltre un secolo di sviluppo industriale incontrollato. I disastri di Seveso, Chernobyl e Bhopal, insieme all'attuale emergenza climatica, hanno reso visibili le conseguenze di una gestione poco attenta sia alla tutela ambientale sia ai diritti dei lavoratori. Diventava sempre più difficile continuare a considerare l'ambiente naturale come un bacino inesauribile di risorse e, al tempo stesso, un deposito illimitato per gli scarti. Le imprese hanno dovuto quindi ampliare il proprio raggio di responsabilità, passando dal solo interesse degli azionisti alla *stakeholder theory*, teorizzata da R. Edward Freeman negli anni Ottanta. Secondo questa teoria, l'impresa deve rispondere anche alle aspettative di tutti i soggetti che partecipano, direttamente o indirettamente, alla sua attività: gli *stakeholder* (Parmar et al., 2010).

¹ Testo originale: «there is one and only one social responsibility of business--to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud» (Friedman, 1970).

² Testo originale: «That responsibility is to conduct the business in accordance with their desires, which generally will be to make as much money as possible while conforming to their basic rules of the society, both those embodied in law and those embodied in ethical custom» (Friedman, 1970).

Questo cambiamento di prospettiva ha trovato una consacrazione istituzionale nel 1987, quando la World Commission on Environment and Development (WCED)³ ha pubblicato il Rapporto Brundtland⁴, portando all'attenzione mondiale il concetto di *sviluppo sostenibile* (figura 1), definito come lo sviluppo che «consente alla generazione presente di soddisfare i propri bisogni senza compromettere la possibilità delle generazioni future di soddisfare i propri»⁵ (WCED, 1987). Con questa definizione la sostenibilità viene inquadrata come una priorità globale, che pone l'accento sulla giustizia intergenerazionale e sulla necessità di ragionare nel lungo periodo (Rogers e Hudson, 2011). Restava però un problema pratico: tradurre principi macroeconomici così ampi in strumenti utilizzabili dalle singole aziende.

Figura 1 - Le tre dimensioni della sostenibilità



Fonte: Dyllick e Hockerts, 2002

Uno di questi strumenti è la *Triple Bottom Line (TBL)* in figura 2, elaborata nel 1997 da John Elkington come applicazione pratica della stakeholder theory. La TBL affianca ai tradizionali indicatori economici anche i risultati sociali e ambientali, valutando l'andamento dell'impresa attraverso tre linee interconnesse. Secondo Elkington (1997, come citato in Alhaddi, 2015), la dimensione economica comprende, oltre alla redditività interna, gli effetti delle attività aziendali sullo sviluppo e sulla stabilità del sistema economico circostante; la dimensione sociale riguarda l'adozione di pratiche manageriali

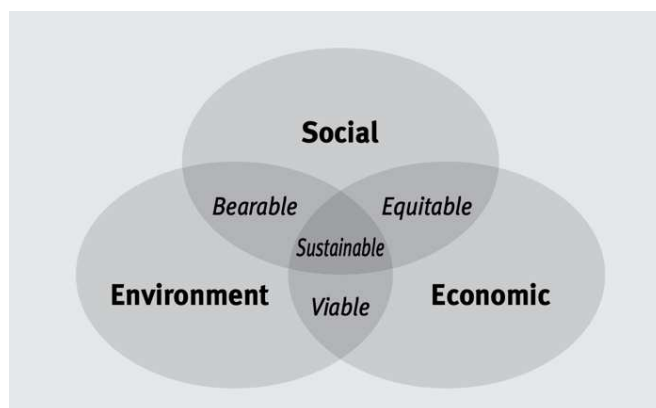
³ Commissione indipendente dell'ONU creata nel 1983 sotto la presidenza di Gro Harlem Brundtland, incaricata di riesaminare i problemi ambientali globali e proporre soluzioni strategiche per lo sviluppo.

⁴ Noto anche come "Our Common Future".

⁵ Testo originale: «Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs» (World Commission on Environment and Development, 1987).

eque, orientate alla tutela dei lavoratori e al sostegno delle comunità locali; la dimensione ambientale richiede infine processi capaci di preservare il capitale naturale per le generazioni future. In quest'ottica gli investimenti in sostenibilità smettono di essere un costo improduttivo e diventano una leva per innovazione, resilienza e vantaggio competitivo.

Figura 2-Triple Bottom Line



Fonte: Rogers e Hudson, 2011

Nella pratica, però, accanto alle sinergie la TBL genera anche tensioni difficili da conciliare. Il rispetto di standard ecologici più severi spinge spesso le aziende a ottimizzare i processi e a ridurre gli sprechi, con un guadagno di efficienza economica quasi immediato (Rogers e Hudson, 2011). Quando la tutela socio-ambientale entra in conflitto con il profitto a breve termine, però, mantenere in equilibrio le tre dimensioni richiede una leadership diffusa a tutti i livelli dell'organizzazione. Non è un risultato che si ottiene automaticamente.

A guidare questo mutamento strategico concorrono due tipi di forze complementari: i *fattori push* (di spinta), di natura coercitiva, legati alle pressioni della società civile, alle richieste degli investitori e soprattutto all'evoluzione normativa⁶; e i *fattori pull* (di attrazione), che rendono le pratiche sostenibili delle vere opportunità di business: incentivi economici pubblici, certificazioni ambientali volontarie, mutamenti nelle preferenze dei consumatori. Si stima che oggi circa due miliardi di persone nel mondo orientino i propri acquisti premiando i marchi percepiti come socialmente responsabili (Todaro et al., 2024).

⁶ Vedi paragrafo 1.3.

La sostenibilità, per rispondere davvero alle sfide dei nuovi scenari competitivi, deve quindi permeare tutti i processi decisionali dell'organizzazione. È questo il senso della definizione di *sostenibilità aziendale* proposta da Dyllick e Hockerts (2002), che la descrivono come la capacità di «soddisfare i bisogni degli stakeholder diretti e indiretti di un'impresa (come azionisti, dipendenti, clienti, comunità, ecc.), senza compromettere la sua capacità di soddisfare anche i bisogni degli stakeholder futuri»⁷.

Questa definizione segna un cambiamento netto rispetto alla concezione storica della *Corporate Social Responsibility (CSR)*, a lungo intesa come un'attività filantropica, isolata o accessoria rispetto alle reali operazioni dell'azienda (Todaro et al., 2024). Un approccio così concepito mostra però limiti evidenti: subordina troppo facilmente gli interessi sociali ed ecologici al calcolo finanziario, e si rivela inefficace proprio nei casi in cui profitti privati e interessi pubblici entrano in conflitto. Per superare questi limiti, mondo accademico e imprenditoriale hanno iniziato a promuovere una strategia proattiva, che integra le considerazioni ambientali e sociali direttamente nella pianificazione del core business, trasformandole da elemento accessorio a parte integrante dell'obiettivo aziendale, orientato all'innovazione sistemica per la creazione di valore condiviso (*shared value*).

L'azienda punta così a costruire un'offerta di valore che generi contemporaneamente benefici economici, sociali e ambientali nel lungo periodo, a vantaggio di tutti gli stakeholder. Tradurre questa strategia in risultati oggettivi e misurabili ha portato all'affermazione dei *parametri ESG (Environmental, Social, Governance)*⁸, oggi l'approccio più accreditato per valutare la sostenibilità d'impresa. I parametri ESG fungono sia da criterio di misurazione sia da tattica di investimento, incorporando variabili extra-finanziarie nelle scelte strategiche aziendali (Li et al., 2021): la *sfera ambientale (E)* osserva le ricadute ecologiche delle operazioni (emissioni, efficienza energetica, trattamento degli scarti); la *sfera sociale (S)* guarda alle dinamiche legate alle risorse umane, all'inclusione e alla tutela della salute; la *sfera governance (G)* si concentra sui meccanismi di amministrazione societaria, sull'integrità morale e sulle logiche retributive.

⁷ Testo originale: «meeting the needs of a firm's direct and indirect stakeholders (such as shareholders, employees, clients, pressure groups, communities etc), without compromising its ability to meet the needs of future stakeholders as well» (Dyllick e Hockerts, 2002)

⁸ La genesi formale di questi parametri risale al documento "Who Cares Wins", promosso nel 2004 dall'ONU in collaborazione con diverse entità del settore finanziario (Li et al., 2021).

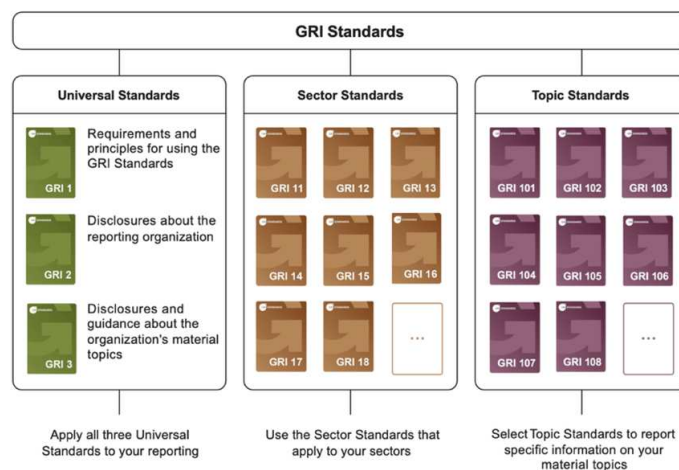
Un buon posizionamento ESG produce inoltre un "effetto assicurativo", che protegge la stabilità finanziaria e il valore azionario dell'impresa anche nelle fasi di maggiore incertezza del mercato (Li et al., 2021). Perché questo effetto assicurativo non resti solo teorico, l'adozione dei criteri ESG richiede una rigorosa accountability, elemento fondante della stessa governance (Li et al., 2021): l'impegno sostenibile deve tradursi in parametri oggettivi, secondo il principio per cui "ciò che viene misurato viene fatto" (Rogers e Hudson, 2011). Altrimenti il rischio è che resti una dichiarazione di intenti priva di riscontro. È per garantire questa trasparenza agli stakeholder che si è affermata la necessità di strumenti di comunicazione formalizzati, aprendo la strada alla rendicontazione non finanziaria e alla stesura del report di sostenibilità (Todaro et al., 2024).

1.2 Il report di sostenibilità

Le radici della rendicontazione di sostenibilità affondano negli anni Sessanta, quando negli Stati Uniti e nel Regno Unito la responsabilità sociale d'impresa comincia a suscitare un interesse più ampio; nei primi anni Settanta il tema si estende all'Europa continentale (Dyllick e Hockerts, 2002). Tra gli anni Settanta e Ottanta, sotto la spinta dei movimenti ambientalisti e sociali che denunciano gli effetti negativi delle attività produttive, alcune imprese pioniere iniziano a pubblicare su base volontaria i primi report ambientali e sociali. Si tratta però di documenti sporadici e privi di uno schema comune (Fornasari e Traversi, 2024). Negli anni successivi si afferma, sul versante aziendale, il concetto di eco-efficienza, che mette in rapporto il valore economico creato dall'impresa con il suo impatto ecologico complessivo e che le imprese adottano sempre più spesso come principio guida del proprio impegno verso lo sviluppo sostenibile (Dyllick e Hockerts, 2002). Alla fine degli anni Novanta compaiono i primi principi e le prime linee guida internazionali, che migliorano in modo significativo la qualità e la trasparenza delle informazioni divulgate (Fornasari e Traversi, 2024). Si fa strada così una visione più integrata, che ha come riferimento la TBL: i framework di rendicontazione nascono per aiutare le imprese a comunicare agli stakeholder i propri impatti economici, ambientali e sociali in modo trasparente, affidabile e comparabile (Kücükgül et al., 2022). Oggi le imprese redigono questi documenti per motivi molto concreti. Tra il management e gli altri stakeholder, come dipendenti, clienti, fornitori e azionisti, esiste un'asimmetria informativa sulle prestazioni di sostenibilità. Pubblicando il report, l'impresa cerca di

ridurla e di segnalare all'esterno i propri risultati positivi, così da consolidare la propria legittimità. A volte la spinta è più elementare: si pubblica il report perché lo fanno anche gli altri, adeguandosi a pratiche ormai comunemente accettate (Kücükgül et al., 2022). Attraverso il reporting, oltre a comunicare, l'azienda testimonia la propria capacità di creare valore condiviso e mantiene un confronto costante con gli stakeholder. Per ridurre il divario informativo tra l'azienda e l'ambiente esterno, diversi enti internazionali hanno elaborato linee guida dedicate. Due dei principali e più diffusi framework di riferimento a livello globale sono quelli emanati dal *Sustainability Accounting Standards Board (SASB)* e dalla *Global Reporting Initiative (GRI)*. Il SASB, nato negli Stati Uniti nel 2011, ha elaborato standard che guardano ai temi di sostenibilità capaci di incidere sui risultati economico-finanziari e, di conseguenza, sulla capacità dell'impresa di creare valore nel lungo periodo. I destinatari dell'informativa sono gli investitori e i mercati finanziari (de Villiers et al., 2024a). La particolarità del SASB sta nel fatto che, per ciascun comparto industriale, vengono isolati i pochi fattori ESG che hanno un impatto concreto e misurabile in quel contesto, con i relativi indicatori quantitativi (de Villiers et al., 2024a). Il SASB ha progressivamente perso la propria autonomia: nel 2021 è confluito nella Value Reporting Foundation e nel 2022 nell'*International Sustainability Standards Board (ISSB)*, all'interno della IFRS Foundation. Su un piano diverso si colloca la GRI (figura 3), costituita nel 1997. È da quel momento che si consolida il rapporto tra le imprese e i framework di rendicontazione, nel frattempo moltiplicatisi. Ancora oggi la GRI è considerata uno dei riferimenti più autorevoli e più utilizzati a livello mondiale per la redazione dei report di sostenibilità (Kücükgül et al., 2022).

Figura 3 - GRI standards



Fonte: GRI 1 Foundation 2021

La diffusione del framework trova conferma nei dati dell'indagine KPMG del 2022: tra le 250 maggiori imprese al mondo per fatturato (G250) il 78% adotta i GRI Standards, contro il 73% del 2020, mentre nel campione più ampio delle N100, composto da 5.800 aziende, la quota si attesta al 68% (Fornasari e Traversi, 2024).

L'impianto del GRI ha conosciuto diverse evoluzioni, fino alla pubblicazione dei GRI Standards nel 2021⁹, che ancorano la rendicontazione a principi di *verificabilità*, *completezza* e *comparabilità* (Todaro et al., 2024), articolandosi in tre serie: gli *standard universali*, gli *standard settoriali* e gli *standard tematici*.

- gli *standard universali* (GRI 1, 2 e 3) sono obbligatori per chiunque utilizzi il *framework*, poiché fissano i principi fondamentali e guidano all'identificazione dei temi materiali;
- gli *standard settoriali* che aiutano a individuare le tematiche più rilevanti per lo specifico settore di appartenenza;
- gli *standard tematici*, che contengono le direttive puntuali per rendicontare gli impatti relativi ai singoli temi materiali, precedentemente identificati dall'azienda.

Il framework si fonda inoltre su quattro concetti chiave (GRI, 2021):

- il concetto di *impatto*, riferito agli effetti positivi o negativi generati sull'economia, sull'ambiente e sulle persone;
- i *temi materiali*, ovvero gli ambiti in cui tali impatti risultano più significativi e prioritari;
- la *due diligence* ossia il processo attraverso cui l'impresa identifica, previene, mitiga e rendiconta i propri impatti negativi;
- gli *stakeholder*, tutti i soggetti i cui interessi sono influenzati dalle attività aziendali.

Per garantire la massima trasparenza, il sistema distingue due livelli di applicazione: un'impresa può dichiarare di aver redatto il report "*in conformità*" (in accordance) con i GRI Standards solo se soddisfa integralmente nove requisiti procedurali; in caso di adozione parziale, può affermare di averlo redatto solo "*con riferimento*" (with reference) a essi (GRI, 2021).

⁹ Entrati ufficialmente in vigore per tutti i report pubblicati a partire dal 1° gennaio 2023

Il perimetro d'azione del GRI si è poi ampliato per rispondere alle sfide istituzionali dell'Agenda 2030¹⁰ delle Nazioni Unite, adottata nel settembre 2015 e sottoscritta da 193 Paesi, con il proprio fulcro nei Sustainable Development Goals (SDGs)¹¹ riportati in figura 4.

Figura 4 - I 17 Obiettivi di Sviluppo Sostenibile (SDGs)



Fonte: Randers et al., 2018

Rispetto ai precedenti *Millennium Development Goals (MDGs)*, introdotti nel 2000 e concentrati principalmente sui paesi in via di sviluppo, gli SDGs chiamano all'azione tutti i Paesi senza distinzioni e collegano in modo esplicito prosperità economica, inclusione sociale e tutela ambientale, coerentemente con la logica della TBL. Il World Business Council for Sustainable Development (WBCSD)¹² li definisce un mezzo efficace per permettere alle imprese di strutturare e comunicare in modo trasparente il proprio contributo allo sviluppo sostenibile (Kücükgül et al., 2022).

¹⁰ L'Agenda è stata approvata dall'Assemblea Generale dell'ONU con la risoluzione "Trasformare il nostro mondo". Tiene insieme le tre dimensioni della sostenibilità: quella economica, quella sociale e quella ambientale. I suoi principi sono riassunti nelle cosiddette "5 P": Persone, Pianeta, Prosperità, Pace e Partnership. L'attuazione chiama in causa gli Stati, la società civile e le imprese.

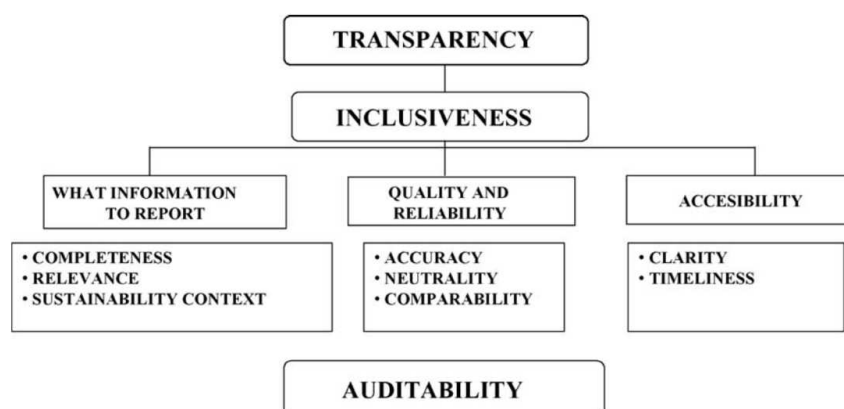
¹¹ Gli obiettivi sono 17. Al loro interno si contano 169 target, la cui attuazione è misurata da 330 indicatori.

¹² È stata fondata nel 1992 in vista del Summit di Rio da un gruppo di leader aziendali, per favorire il dialogo politico sulla sostenibilità. Conta circa 150 membri e ha avuto un ruolo chiave nella diffusione dell'eco-efficienza nel settore privato.

È per rispondere a questa esigenza di misurazione oggettiva che dopo l'adozione degli SDGs nel settembre 2015, la GRI ha pubblicato una serie di guide operative su cosa rendicontare e su come farlo. Tra queste, l'SDG Compass accompagna l'impresa in un percorso in cinque fasi, che parte dalla comprensione degli SDGs e si chiude con la rendicontazione e la comunicazione dei risultati (Kücükgül et al., 2022). Il bisogno di strumenti di questo tipo è reale: secondo un'indagine PwC del 2019, 820 imprese su 1.141 si erano impegnate sugli SDGs, ma solo 157 ne citavano i target e, tra queste, appena l'8% aveva fissato obiettivi quantificabili con un monitoraggio periodico dei progressi (Kücükgül et al., 2022).

Nonostante la diffusione globale, l'impianto del GRI non è privo di limiti, ed è qui che si comincia a intravedere il problema che occupa il resto di questa tesi. Parte della letteratura accademica evidenzia come l'adozione del framework non garantisca automaticamente un reale mutamento organizzativo verso la sostenibilità; il rischio è che il GRI venga usato per fare "vetrina" (*window-dressing*), diventando un mezzo per legittimare le decisioni del management senza un impegno effettivo (Moneva et al., 2006). Una delle criticità principali sta nella flessibilità delle linee guida, che permette alle aziende di selezionare e divulgare solo gli indicatori che migliorano la propria immagine, omettendo o mascherando le aree più problematiche: una pratica nota come cherry-picking. A livello metodologico, inoltre, il GRI viene criticato per l'assenza di indicatori realmente integrati: mantenendo separate le tre dimensioni della TBL, invece di calcolarne il rapporto complessivo, rischia di ridurre la sostenibilità a una semplice fornitura di dati isolati, una frammentazione che si presta facilmente a fenomeni di greenwashing (Moneva et al., 2006).

Figura 5 - Principi GRI



Fonte: Moneva et al., 2006

Per contenere queste derive opportunistiche, le linee guida GRI hanno introdotto un'architettura di principi fondata su trasparenza, inclusività e verificabilità, insieme a criteri complementari orientati alla completezza e all'affidabilità del dato. La pratica, però, dimostra i limiti anche di questa architettura teorica. L'aderenza rigorosa al principio di completezza si traduce spesso in un sovraccarico informativo che disorienta il lettore più che aiutarlo a individuare le reali criticità aziendali. L'inclusività, a sua volta, viene interpretata in modo soggettivo, permettendo alle organizzazioni di coinvolgere soprattutto gli stakeholder che influenzano le performance di business, a scapito di una visione più ampia. E l'assenza di direttive restrittive sui "confini" organizzativi lascia alle imprese un margine sufficiente per escludere dalla rendicontazione le attività indirette più inquinanti della propria catena del valore, contraddicendo di fatto l'obiettivo di trasparenza totale che le linee guida si erano proposte (Moneva et al., 2006).

Queste criticità dipendono tutte da un limite di fondo: l'approccio volontario (soft law) su cui si basa il GRI. Senza vincoli obbligatori, a partire dall'assenza di una verifica indipendente di terza parte, l'adozione di questi framework è rimasta spesso una semplice riforma amministrativa, incapace di impedire lo scostamento tra la narrazione aziendale e la performance materiale effettiva (Todaro et al., 2024). Colmare questo vuoto ha richiesto il passaggio a un regime normativo vincolante (hard law): il legislatore europeo è intervenuto con misure coercitive precise, trasformando la divulgazione non finanziaria da esercizio facoltativo a obbligo legale (Todaro et al., 2024). Un percorso avviato dalla Non-Financial Reporting Directive (NFRD) e portato a termine dalla Corporate Sustainability Reporting Directive (CSRD), di cui il paragrafo successivo ricostruisce le tappe.

1.3 L'evoluzione normativa europea dalla NFRD alla CSRD

Il primo passo verso l'obbligatorietà legislativa si concretizza con la *Direttiva 2014/95/UE*, entrata in vigore a dicembre 2014 e comunemente nota come *Non-Financial Reporting Directive (NFRD)* (Tylec, 2025). Il provvedimento si rivolgeva a grandi enti di interesse pubblico, società quotate, istituti di credito, imprese di assicurazione, con una media di oltre 500 dipendenti durante l'esercizio (Todaro et al., 2024), imponendo loro di pubblicare, nella relazione sulla gestione o in un rapporto autonomo, una dichiarazione non finanziaria su ambiente, questioni sociali e relative al personale, diritti umani e lotta alla corruzione (Tylec, 2025).

L'impianto della NFRD si è rivelato presto lacunoso, per due ordini di motivi.

Il primo è quantitativo: il suo perimetro applicativo coinvolgeva appena 11.700 imprese in tutta l'Unione Europea (Tylec, 2025).

Il secondo è metodologico: l'assenza di standardizzazione obbligatoria lasciava alle aziende un'eccessiva libertà nella scelta delle linee guida di riferimento, compromettendo la comparabilità dei report e ostacolando il confronto tra le performance delle diverse organizzazioni (Moneva et al., 2006).

A ciò si aggiungeva l'assenza di un obbligo di verifica indipendente sui contenuti: il ruolo dei revisori era circoscritto al controllo formale della semplice presentazione del documento (Tylec, 2025).

Questa carenza di controlli ha finito per intaccare la credibilità delle dichiarazioni societarie, esponendo le informazioni a dubbi di affidabilità e favorendo, come già visto nel paragrafo precedente a proposito del GRI, logiche di cherry-picking (Moneva et al., 2006). Il risultato è stato, in molti casi, una conformità di facciata più vicina al greenwashing che a una rendicontazione trasparente. Di fronte all'impossibilità per investitori e stakeholder di valutare con chiarezza le performance aziendali, il legislatore europeo ha compreso la necessità di cambiare rotta, abbandonando la flessibilità iniziale a favore di regole vincolanti e standardizzate.

Questo cambiamento si inserisce in un mutamento più ampio del contesto politico europeo, orientato verso obiettivi ecologici e sociali decisamente più ambiziosi, inaugurato con il Green Deal europeo¹³ nel 2019 (Todaro et al., 2024).

Il Green Deal punta a una transizione ecologica ed economica radicale, con l'obiettivo di raggiungere la neutralità climatica entro il 2050 e di rendere l'Europa il primo continente a emissioni nette di gas serra azzerate (Tylec, 2025).

Per guidare questa transizione l'Europa ha introdotto anche il Piano d'azione per finanziare la crescita sostenibile¹⁴ e la Tassonomia europea¹⁵ (Todaro et al., 2024), con

¹³ Introdotto dalla Commissione Europea, il Green Deal è il pacchetto strategico volto a trasformare l'UE in un'economia sostenibile, dissociando la crescita economica dall'uso delle risorse e garantendo una transizione equa verso la neutralità climatica.

¹⁴ Il Piano d'azione per finanziare la crescita sostenibile (Commissione Europea, 8 marzo 2018) fissa tre obiettivi strategici: riorientare i capitali verso investimenti ecosostenibili, gestire i rischi climatico-sociali e promuovere la trasparenza finanziaria a lungo termine. Il conseguimento di tali traguardi pone come preconditione indispensabile la divulgazione di dati societari di sostenibilità comparabili e affidabili.

¹⁵ Il Reg. (UE) 2020/852 (Tassonomia europea) definisce i criteri per classificare le attività economiche come ecosostenibili rispetto a sei obiettivi prioritari, mirando a orientare gli investimenti verso la transizione e a contrastare il greenwashing.

l'obiettivo di convogliare i capitali verso investimenti responsabili. Questo obiettivo richiede, a monte, la massima trasparenza sui dati socio-ambientali, allo stesso livello di rigore già garantito dalla rendicontazione finanziaria (Tylec, 2025).

Da questa esigenza nasce la *Corporate Sustainability Reporting Directive (CSRD)*, sancita dalla *Direttiva (UE) 2022/2464* ed entrata in vigore a gennaio 2023. Con la CSRD il legislatore europeo abbandona la terminologia di informativa "non finanziaria", riconoscendone la piena rilevanza economica, e impone l'integrazione strutturale delle tematiche di sostenibilità nella pianificazione e nella rendicontazione societaria obbligatoria, insieme a requisiti rigorosi di standardizzazione e controllo (Todaro et al., 2024).

Rispetto alla NFRD, la CSRD introduce almeno quattro elementi di novità.

Il *primo* riguarda il perimetro di applicazione, ampliato da circa 11.700 a circa 49.000 imprese operanti nell'Unione Europea (Tylec, 2025): alle grandi società si aggiungono progressivamente le piccole e medie imprese (PMI) quotate nei mercati regolamentati, con la sola esclusione delle microimprese.

Il *secondo* è l'equiparazione tra rendicontazione finanziaria e rendicontazione di sostenibilità: i dati socio-ambientali devono essere collocati in un'apposita sezione della relazione sulla gestione, e non più pubblicati come report separato, proprio per evitare che vengano percepiti come un elemento accessorio.

Il *terzo* è il vincolo alla digitalizzazione del reporting. Le organizzazioni devono adottare un formato elettronico che permetta l'analisi computazionale dei dati e l'etichettatura (*tagging*) delle singole voci secondo la tassonomia digitale europea (Tylec, 2025). Ne beneficia sia chi redige il report sia chi lo consulta, perché i dati così marcati possono confluire in archivi centralizzati. L'intera infrastruttura è pensata per convergere nel futuro European Single Access Point (ESAP).

Il *quarto* elemento, probabilmente il più determinante per contrastare le pratiche opportunistiche analizzate finora, è l'introduzione dell'obbligo di verifica indipendente (*assurance*) sui contenuti divulgati.

Il legislatore ha previsto un percorso di controllo graduale: dapprima un'attestazione basata su un livello di sicurezza limitato (*limited assurance*), fino al futuro rilascio di un'attestazione a sicurezza ragionevole (*reasonable assurance*), la cui fattibilità sarà valutata entro il 1° ottobre 2028.

È un tassello normativo decisivo, perché definisce i presupposti operativi e legali necessari per dare credibilità e consistenza ai dati (tabella 1).

Tabella 1 - NFRD e CSRD somiglianze e differenze

Comparison criteria	NFRD	CSRD
Purpose	Information to the extent necessary to understand the company's development situation, results, position and impact of its operation	Information necessary to understand the company's impact on sustainability issues, and details needed to comprehend how sustainability issues affect the company's development, performance and position
Minimum content	Five sustainability issues (environmental, social and labour issues, respect for human rights, anti-corruption and anti-bribery)	Three sustainability issues (environmental, social and governance); sector-specific, sector-neutral information identified by EFRAG
Perspective	From the outside towards inside	From the inside towards outside
Time horizons	Not specified	Short, medium and long term
Target group	Financial resources providers	Stakeholders (in a broad sense)
Materiality principle	Materiality principle, "comply or explain" principle	Double materiality principle (obligation to report the environmental and social impact of the company, as well as the impact (risk) of ESG factors on the company)
Links to financial reporting	Optional reporting, as separate reports or as part of the management board report (report on operation). Mandatory references and additional explanations of other information included in the annual report	Mandatory reporting as part of the report on operation. Mandatory references to additional explanations related to other information included in the report on operation and the annual report
External assurance	Optional	Mandatory (emphasising the materiality analysis process)
Scope	Large undertakings which are public-interest entities with more than 500 employees and public-interest entities that are parent companies in a large group of companies with more than 500 employees	All listed companies, including small and medium-sized listed companies with fewer than 500 employees
Reporting standards	No uniform standards - voluntary guidelines	Uniform sustainability reporting system - mandatory use of a set of EU sustainability reporting standards - some standards under development

Fonte: Tylec, 2025

1.3.1 Il principio della doppia materialità

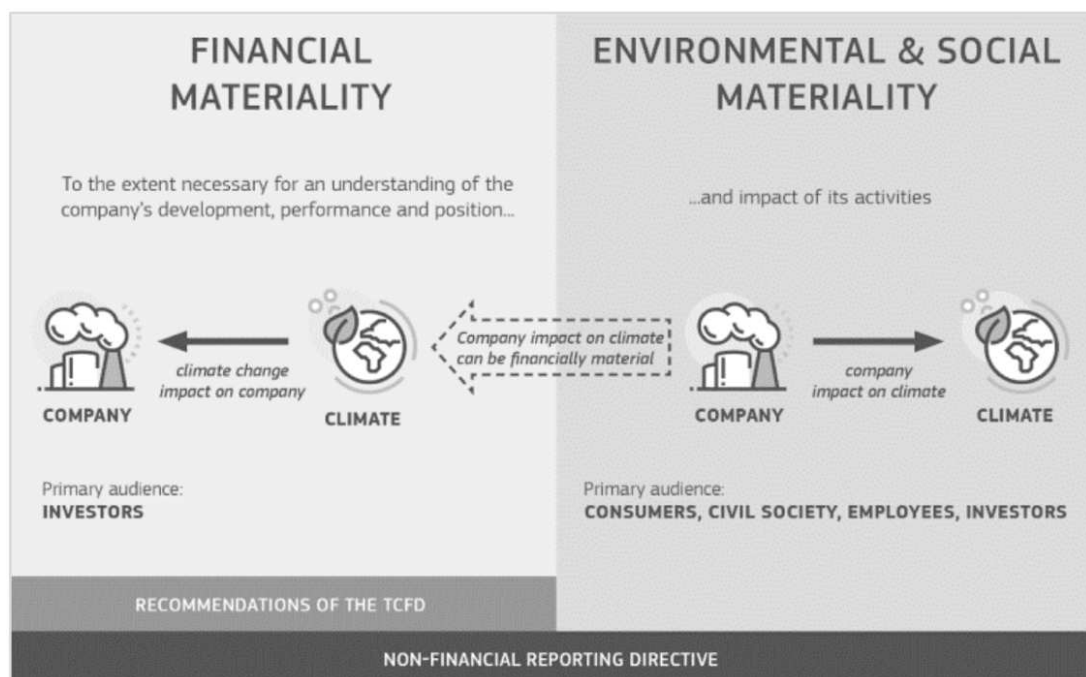
L'innovazione concettuale più profonda della nuova CSRD è l'obbligo di applicare il *principio della doppia materialità*.

Questo meccanismo obbliga le aziende ad abbandonare le misurazioni a senso unico, per abbracciare una duplice e speculare prospettiva di indagine (Todaro et al., 2024):

- la materialità d'impatto (prospettiva *inside-out*), impone di mappare come le attività dell'azienda e la sua intera catena del valore incidano sul mondo circostante, generando effetti positivi o negativi sull'ambiente, sui diritti umani e sul tessuto sociale. L'obiettivo è tradurre l'impronta reale lasciata dall'impresa sulla collettività e sulla natura;
- la materialità finanziaria (prospettiva *outside-in*), valuta invece il fenomeno in senso inverso, esaminando come i cambiamenti ecologici possano ripercuotersi sulla stabilità e sullo sviluppo economico dell'organizzazione.

È una prospettiva utile sia al management, per prevenire i rischi legati a una scarsa cura delle dinamiche socio-ambientali, sia agli investitori, per valutare la resilienza del modello di business.

Figura 6 - Il principio della doppia materialità



Fonte: Todaro et al., 2024

La vera discontinuità, rispetto al regime precedente, sta nell'obbligo di valutare le due dimensioni in modo del tutto indipendente. Sotto la NFRD, il principio di "*comply or explain*"¹⁶ lasciava troppo margine, in quanto un'azienda poteva giustificare l'omissione

¹⁶ Principio normativo (traducibile come "conformati o spiega") che consente alle imprese di scegliere se aderire a una determinata disposizione oppure, in caso di mancata applicazione, di giustificarne pubblicamente i motivi.

di informazioni critiche con l'assenza di ricadute finanziarie immediate, finendo per agevolare, più o meno consapevolmente, pratiche di greenwashing (Moneva et al., 2006). Il nuovo impianto europeo chiude questo spiraglio: se un tema è rilevante per il suo impatto esterno, non può essere de-prioritizzato solo perché l'azienda non ne percepisce un rischio economico diretto. La rendicontazione trasparente degli impatti, in questa logica, diventa una questione di interesse pubblico (GRI, 2021), che riduce le storiche asimmetrie informative e restringe il margine per narrazioni di comodo (Tylec, 2025). Il valore di questa scelta normativa emerge con più chiarezza se confrontata con gli standard internazionali dell'*International Sustainability Standards Board (ISSB)*, che opera nella fondazione *International Financial Reporting Standards (IFRS)*¹⁷ e ha costruito una linea di base globale pensata quasi esclusivamente per i mercati di capitali, fondata sulla single materiality, la sola rilevanza finanziaria.

In questa prospettiva outside-in, alle aziende viene richiesto di comunicare solo i rischi e le opportunità socio-ambientali capaci di incidere sulle decisioni degli investitori. Questo modello ha suscitato perplessità tra studiosi e società civile a causa del timore di una "cattura finanziaria" del concetto di sostenibilità, un sistema che finirebbe per dipendere dagli interessi dei grandi investitori, trascurando le responsabilità che le imprese hanno verso la collettività e l'ambiente (de Villiers et al., 2024a).

Anche per questo la scelta dell'Unione Europea di imporre la doppia materialità appare, tra le due impostazioni, quella più efficace per contrastare il greenwashing alla radice (Tylec, 2025).

1.4 Gli European Sustainability Reporting Standards – ESRS

Un impianto normativo così ambizioso richiedeva un sistema di metriche standardizzate, capace di garantire una rendicontazione uniforme e comparabile tra le imprese. L'Unione Europea ha affidato questo compito tecnico all'*European Financial Reporting Advisory Group (EFRAG)*¹⁸, che il 31 luglio 2023 ha adottato, tramite *Regolamento Delegato (UE) 2023/2772*, un primo nucleo di dodici standard, gli *European Sustainability Reporting*

¹⁷ La Fondazione IFRS è l'organizzazione indipendente responsabile della definizione dei principi contabili globali, adottati oggi in circa 160 giurisdizioni. Nel novembre del 2021 ha istituito l'ISSB per affiancare lo storico comitato contabile (IASB) e promulgare i primi standard internazionali di sostenibilità, entrati ufficialmente in vigore nel 2024 con le sigle IFRS S1 e IFRS S2.

¹⁸ L'European Financial Reporting Advisory Group (EFRAG) è un organismo tecnico indipendente fondato nel 2001. Storicamente incaricato di fornire consulenza alla Commissione Europea sull'adozione dei principi contabili internazionali (IFRS), è stato ufficialmente designato come organo preposto all'elaborazione dei nuovi standard europei di rendicontazione della sostenibilità (ESRS).

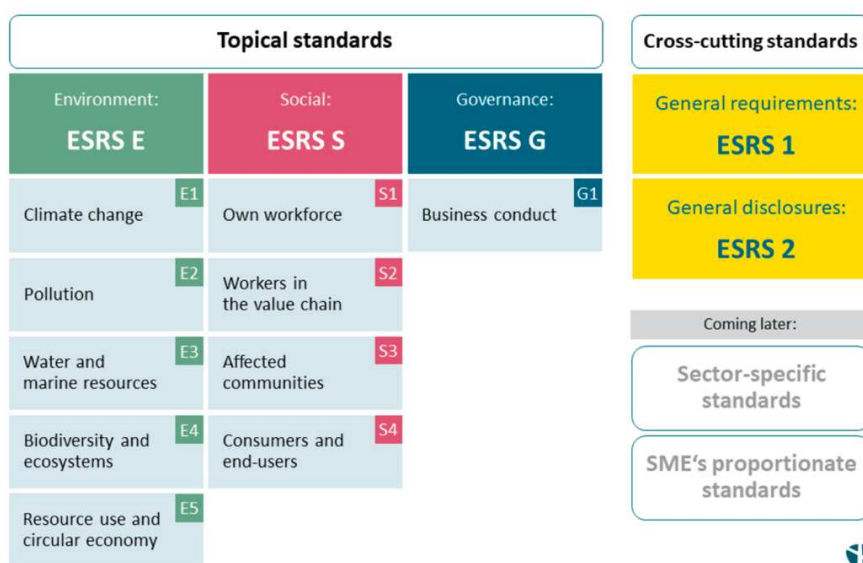
Standards (ESRS) (Tylec, 2025). Questi documenti fissano in modo vincolante le informazioni che le organizzazioni devono fornire su impatti, rischi e opportunità relativi alle tre dimensioni ESG.

L'EFRAG ha puntato fin dall'inizio a creare un linguaggio condiviso, capace di limitare la frammentazione delle regole a livello internazionale, cercando l'interoperabilità con i framework già riconosciuti a livello globale, in primo luogo i GRI e gli standard dell'ISSB. Questa ricerca di armonizzazione si arresta però di fronte a un pilastro che l'Europa non è disposta a cedere: il principio della doppia materialità. Il dialogo con gli altri standard facilita la raccolta dei dati, ma gli ESRS mantengono ferma la prospettiva d'impatto (inside-out), garantendo così che le performance socio-ambientali abbiano lo stesso valore legale dei tradizionali indicatori finanziari.

L'architettura di questo linguaggio è stata disegnata dall'EFRAG seguendo un rigoroso assetto gerarchico, definito "a ombrello". Il sistema si articola in due macrocategorie fondamentali:

- gli standard trasversali (*cross-cutting standards*) dettano i principi metodologici e le informazioni di base applicabili a ogni organizzazione, a prescindere dal settore economico di appartenenza;
- gli standard tematici (*topical standards*), ideati per indagare in modo verticale singole questioni riconducibili alle dimensioni ambientale, sociale e di governance.

Figura 7 - La struttura degli ESRS



Fonte: Fornasari e Traversi, 2024

Il nucleo fondante è costituito dai due principi trasversali, l'ESRS 1 (*General Requirements*) e l'ESRS 2 (*General Disclosures*):

- l'ESRS 1 fissa le direttive generali necessarie per la stesura della dichiarazione di sostenibilità e orienta le imprese nell'applicazione pratica della doppia materialità; per questo richiede un'analisi della catena del valore (*value chain*) che spinga l'indagine oltre il perimetro delle attività dirette, includendo impatti, rischi e opportunità generati sia a monte (*upstream*) che a valle (*downstream*) (EFRAG, 2025b);
- l'ESRS 2, invece, riguarda le informative di carattere generale che tutte le realtà soggette alla direttiva sono obbligate a comunicare, in tema di governance, strategia e pianificazione a breve, medio e lungo periodo (EFRAG, 2025c).

Gli *standard tematici* sono concepiti per mappare in modo specifico e dettagliato le performance aziendali nei tre pilastri ESG (Fornasari e Traversi, 2024):

- *E - Environmental* (Pilastro Ambientale da ESRS E1 a ESRS E5) copre cambiamenti climatici (ESRS E1), inquinamento (ESRS E2), acque e risorse marine (ESRS E3), biodiversità ed ecosistemi (ESRS E4) e uso delle risorse ed economia circolare (ESRS E5);
- *S - Social* (Pilastro Sociale da ESRS S1 a ESRS S4) riguarda la forza lavoro interna (ESRS S1), i lavoratori nella catena del valore (ESRS S2), le comunità colpite dalle attività aziendali (ESRS S3) e i consumatori (ESRS S4);
- *G - Governance* (Pilastro di Governance ESRS G1) disciplina la trasparenza su cultura d'impresa, contrasto alla corruzione, whistleblowing e, dove applicabile, benessere animale.

La rendicontazione dei requisiti tematici dipende dall'esito della valutazione della *doppia materialità* svolta dall'impresa: se un tema non genera impatti significativi né comporta rischi finanziari rilevanti, l'azienda può omettere le relative divulgazioni. Resta però un vincolo importante, che a nostro avviso è tra i più efficaci dell'intero impianto ESRS nel prevenire nuove forme di cherry-picking: l'impresa deve comunque possedere e produrre elementi informativi solidi e verificabili a supporto della conclusione di non-rilevanza, così da impedire l'omissione ingiustificata di temi scomodi (de Villiers et al., 2024a).

L'architettura degli ESRS ha subito un adeguamento importante nel 2025. Dopo aver raccolto le difficoltà pratiche segnalate dalle imprese nei primi approcci alla rendicontazione, la Commissione Europea ha lanciato nel febbraio 2025 il pacchetto di semplificazione "*Omnibus I*", con l'obiettivo di alleggerire il carico burocratico senza però ridurre le ambizioni ambientali fissate dal Green Deal. Il percorso si è concluso a fine 2025 con l'approvazione del *Draft Amended ESRS*¹⁹. Sul piano operativo, i requisiti obbligatori (*shall datapoints*) sono stati ridotti di circa il 61%, passando da 803 a 314, mentre quelli ad applicazione volontaria (*may datapoints*) sono stati eliminati del tutto (tabella 2).

Tabella 2 - Riduzione dei requisiti obbligatori "shall" e dei requisiti volontari "may"

Datapoints (excluding MDR/GDR)	ESRS as enacted in 2023 ¹¹	Amended ESRS Revision 2025	Overall reduction
Total 'shall'	803	314	-60.9%
Total 'may'	270	0	-100.0%
Total	1073	314	-70.7%

Fonte: EFRAG, 2025a

La riforma introduce criteri più flessibili, come l'esenzione per costi o sforzi sproporzionati e la possibilità di usare stime medie di settore per la catena del valore, e rende più rapida anche la valutazione della doppia materialità, spostando il focus sulle informazioni davvero rilevanti per gli stakeholder. Resta da vedere, e sarà uno dei nodi da tenere presenti nei prossimi capitoli, se questa semplificazione riduca soltanto la burocrazia superflua o rischi anche di restringere il perimetro di trasparenza appena conquistato.

Un ulteriore sviluppo è avvenuto il 3 luglio 2026, quando la Commissione Europea ha adottato due regolamenti delegati per rendere operativo l'impianto di semplificazione. Il primo provvedimento ha modificato il *Regolamento Delegato (UE) 2023/2772*, riducendo di oltre il 60% i dati obbligatori degli ESRS per le grandi imprese e introducendo criteri applicativi più flessibili (Commissione Europea, 2026c).

¹⁹ Il documento è stato approvato dal Sustainability Reporting Board (SRB) dell'EFRAG il 28 novembre 2025 e pubblicato ufficialmente il 3 dicembre 2025. La revisione sancisce la definitiva transizione degli standard europei verso un framework orientato alla fair presentation (corretta rappresentazione).

Il secondo provvedimento integra la *Direttiva 2013/34/UE*²⁰ introducendo uno standard di rendicontazione di sostenibilità ad uso volontario (voluntary standard) per le aziende fino a 1.000 dipendenti non soggette direttamente alla CSRD. Questo strumento definisce il cosiddetto «tetto alla catena del valore» (value chain cap), impedendo alle grandi imprese di richiedere ai propri fornitori o partner informazioni superiori a quelle previste dallo standard volontario. Entrambi gli atti, una volta superata la fase di controllo di Parlamento e Consiglio ed effettuata la pubblicazione in Gazzetta Ufficiale, diventeranno vincolanti a partire dagli esercizi finanziari che iniziano il 1° gennaio 2027, con la possibilità per le imprese di applicarli su base volontaria già dal 2026.

Nel complesso, il percorso ricostruito in questo capitolo, dai primi modelli teorici di sostenibilità aziendale, attraverso l'evoluzione del reporting volontario, fino all'obbligo normativo della CSRD e alla sua traduzione tecnica negli ESRS, mostra un fenomeno che si ripete a ogni passaggio: ogni volta che la rendicontazione ESG si allontana da un controllo verificabile, si apre uno spazio per pratiche opportunistiche. È in questo spazio che si colloca il greenwashing, il tema al centro del capitolo che segue.

1.5 La normativa digitale e tecnologica

Per contrastare le dinamiche opportunistiche analizzate nel secondo capitolo, l'intervento del legislatore europeo si è sviluppato su due fronti complementari, con l'intento di allineare la transizione ecologica a quella digitale.

Da una parte, la *Green Claims Directive* interviene per ripristinare la trasparenza e contrastare apertamente il greenwashing, riducendo gli squilibri informativi del mercato. Dall'altra, attraverso l'*EU AI Act* e l'applicazione di modelli di intelligenza artificiale spiegabile (*Explainable AI - XAI*), le istituzioni mirano a scongiurare il rischio che la gestione digitale dei parametri ESG e la revisione automatizzata, si convertano in nuovi mezzi di opacità.

²⁰ La *Direttiva 2013/34/UE del Parlamento europeo e del Consiglio*, del 26 giugno 2013, è relativa ai bilanci d'esercizio, ai bilanci consolidati e alle relative relazioni di talune tipologie di imprese. La direttiva è stata modificata dalla CSRD (Direttiva (UE) 2022/2464) per integrare gli obblighi di rendicontazione di sostenibilità.

1.5.1 La Green Claims Directive (GCD)

Per contrastare queste pratiche ingannevoli, a marzo 2023 la Commissione Europea ha proposto la *Green Claims Directive (GCD)* (Commissione Europea, 2023).

La proposta chiedeva alle imprese di fondare ogni dichiarazione ambientale rivolta ai consumatori su prove scientifiche e su un'analisi dell'intero ciclo di vita del prodotto, così da evitare che il miglioramento di un aspetto nascondesse danni ad altri indicatori ecologici. Regolava inoltre in modo rigoroso i claim sulla neutralità climatica: le aziende avrebbero dovuto distinguere in modo trasparente la reale riduzione delle emissioni dalle semplici compensazioni ottenute acquistando crediti di carbonio (Carreño, 2023).

Il punto più innovativo riguardava però il superamento dell'auto-certificazione: la documentazione a supporto dei claim avrebbe dovuto essere verificata ex ante da un ente terzo indipendente (Commissione Europea, 2023).

Proprio questo obbligo è stato al centro delle critiche sugli oneri amministrativi per le imprese. Nel giugno 2025 la Commissione ha annunciato l'intenzione di ritirare la proposta.

Pur in assenza di un ritiro formale, i negoziati tra Parlamento e Consiglio si sono fermati. Sul piano normativo, un argine più immediato al greenwashing arriva invece dalla *Direttiva (UE) 2024/825* sulla responsabilizzazione dei consumatori per la transizione verde, applicabile dal 27 settembre 2026.

È questa direttiva a vietare le asserzioni ambientali generiche, come "eco-compatibile" o "amico della natura", quando l'impresa non può dimostrare prestazioni ambientali eccellenti riconosciute (Direttiva UE 2024/825).

Se approvata, la GCD avrebbe completato la disciplina sul reporting extra-finanziario analizzata nei capitoli precedenti.

La *Corporate Sustainability Reporting Directive (CSRD)* impone la rendicontazione dei dati ESG secondo il principio della doppia materialità. La GCD avrebbe impedito che quelle stesse informazioni venissero piegate a scopi di marketing, assicurando che alle dichiarazioni di sostenibilità corrispondessero azioni reali e misurabili. La vicenda mostra quanto la verifica indipendente delle dichiarazioni ambientali resti un nodo politico ancora aperto.

1.5.2 L'EU AI Act

L'Unione Europea ha inquadrato la transizione tecnologica con l'EU AI Act (*Regolamento UE 2024/1689*), entrato in vigore ad agosto 2024 (figura 8) (Cors e Thiébaut, 2025).

Figura 8 - Cronoprogramma di entrata in vigore dei vincoli dell'AI Act (previsioni ad aprile 2024)



Fonte: Wilson Sonsini, 2024

L'AI Act è la prima normativa orizzontale e globale in materia, mirata a garantire lo sviluppo di sistemi algoritmici sicuri, trasparenti e antropocentrici. Il regolamento punta a tutelare la salute, la sicurezza e i diritti fondamentali, e fra i beni protetti include esplicitamente l'ambiente (Regolamento UE 2024/1689). La sua portata è extraterritoriale: vale per chiunque immetta sul mercato europeo sistemi di intelligenza artificiale o modelli per scopi generali (*General Purpose AI - GPAI*), ovunque abbia la sede operativa. Il punto di contatto con la disciplina ESG nasce dalla CSRD, che trasforma la rendicontazione extra-finanziaria in un insieme di dati standardizzati ed etichettati, elaborabili in automatico dai software. Misurazione, aggregazione e validazione delle informazioni socio-ambientali lungo le catene del valore passeranno quindi sempre più per processi informatici e formati digitali strutturati, il che facilita l'auditing e il confronto tra report (Fornasari e Traversi, 2024).

Tuttavia, questa pervasività tecnologica comporta rischi strutturali emergenti. Senza controlli rigorosi e presidi di governance algoritmica, i software potrebbero trasformarsi in nuovi e sofisticati strumenti di manipolazione informativa, vanificando di fatto le tutele di trasparenza e i diritti fondamentali, inclusi quelli ecologici, introdotti dalle direttive europee.

L'impiego di modelli algoritmici opachi o privi di supervisione per processare o redigere le dichiarazioni non finanziarie favorisce infatti il *digital greenwashing* o *machine-washing*. Questa pratica sfrutta l'automazione per generare informazioni ambientali distorte o allucinate, occultando le reali violazioni operative ed eludendo i requisiti di conformità attraverso una reportistica puramente cosmetica e disaccoppiata dalla realtà materiale dell'impresa (Forliano et al., 2025).

L'EU AI Act classifica l'intelligenza artificiale in quattro livelli di rischio, dal più alto al più basso, applicando regole proporzionate per ciascuna fascia (Regolamento UE 2024/1689):

- *rischio inaccettabile*: applicazioni che violano i diritti fondamentali e la dignità umana, come i sistemi di manipolazione subliminale del comportamento o il riconoscimento delle emozioni sul luogo di lavoro;
- *rischio alto*: sistemi impiegati in settori aziendali critici, quali la selezione del personale, la determinazione dei prezzi assicurativi o la gestione delle infrastrutture, ammessi sul mercato previa idonea valutazione della conformità;
- *rischio limitato*: tecnologie che interagiscono direttamente con gli utenti o generano contenuti artificiali, come i chatbot, soggette a specifici obblighi di informazione e trasparenza;
- *rischio minimo*: applicazioni di uso comune e ampiamente diffuse, le quali sono del tutto esenti da vincoli o obblighi regolatori aggiuntivi.

Norme specifiche disciplinano inoltre i modelli per scopi generali (GPAI), compresa l'intelligenza artificiale generativa (Kim et al., 2025). I fornitori di tali modelli devono garantire la trasparenza tecnica, rispettare la direttiva europea sul diritto d'autore e pubblicare una sintesi dettagliata dei dati utilizzati per l'addestramento. Qualora un modello GPAI presenti rischi sistemici significativi, si applicano obblighi aggiuntivi di cybersicurezza e test avversari (Regolamento UE 2024/1689).

Alle imprese che integrano nei propri processi tecnologie ad alto rischio o modelli GPAI, l'EU AI Act chiede il rispetto di alcuni requisiti di conformità, pensati per rendere verificabili le decisioni automatizzate:

- *gestione dei rischi (art. 9)*: un processo continuo e documentato, che accompagna il sistema lungo tutto il ciclo di vita e serve a individuare e ridurre le criticità;
- *governo dei dati (art. 10)*: dataset di addestramento, convalida e prova appropriati, accurati e privi di distorsioni sistematiche. La norma chiede in particolare di evitare i circuiti di feedback, in cui dati storici distorti si auto-alimentano e amplificano le anomalie nel tempo;
- *tracciabilità (art. 12)*: la registrazione automatica degli eventi, che permette di ricostruire a posteriori il percorso decisionale dell'algoritmo;

- *supervisione umana (art. 14)*: un presidio umano effettivo sul funzionamento del sistema;
- *sicurezza informatica (art. 15)*: il mantenimento di adeguati livelli di robustezza tecnica e resilienza strutturale contro attacchi informatici volti ad alterare dati e output.

L'efficacia di questi controlli è garantita dall'obbligo di monitoraggio e segnalazione dei malfunzionamenti. I fornitori devono infatti notificare tempestivamente alle autorità competenti qualsiasi "incidente grave", nozione che include esplicitamente non solo le violazioni dei diritti fondamentali, ma anche eventuali gravi danni all'ambiente (Regolamento UE 2024/1689). Queste tutele contano molto contro le nuove forme di greenwashing tecnologico. Senza verifiche rigorose, affidare ai software il calcolo dei dati ESG o la stesura del bilancio di sostenibilità può replicare le distorsioni già presenti nei dati storici, oppure produrre narrazioni scollegate da ciò che l'impresa fa davvero (Forlano et al., 2025). L'articolo 50 dell'AI Act interviene proprio su questo punto: i contenuti generati artificialmente vanno contrassegnati con una marcatura permanente ed elaborabile automaticamente, così che la loro origine resti riconoscibile.

1.5.3 L'Explainable AI (XAI) nell'auditing

L'evoluzione tecnologica della rendicontazione di sostenibilità si scontra con il rischio crescente del cosiddetto *assurance-washing*, un'estensione del greenwashing in cui le imprese scelgono appositamente sistemi di revisione deboli o compiacenti per ottenere una legittimazione formale, eludendo reali controlli sulle performance. L'impiego di algoritmi complessi, come le reti neurali profonde, rischia di aggravare questo fenomeno. Le loro logiche di calcolo sono incomprensibili per i revisori umani, poiché operano come "scatole nere" (Adadi e Berrada, 2018). Questa impenetrabilità tecnica rende l'attività di controllo estremamente superficiale, esponendo l'intero processo di auditing al rischio di un fallimento metodologico sistemico.

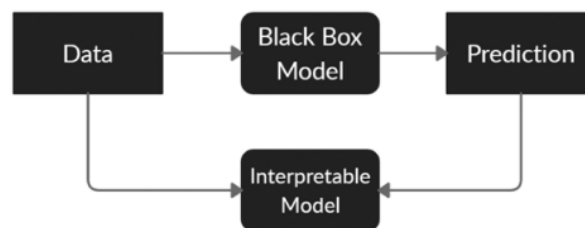
Affidare la validazione dei dati ESG a modelli opachi rischia di nascondere le manipolazioni, trasformando l'audit in una procedura di facciata e alimentando l'*assurance-washing* algoritmico. Per risolvere il problema dell'opacità, l'Intelligenza Artificiale Spiegabile (XAI) diventa un requisito essenziale. La XAI comprende un insieme di metodi e tecniche volti a rendere i modelli algoritmici comprensibili e interpretabili

dagli esseri umani, preservando al contempo l'accuratezza predittiva del sistema (Das e Rad, 2020).

Nei processi di auditing, le tecniche di spiegabilità a posteriori vengono implementate attraverso due diverse metodologie (Chamola et al., 2023):

- spiegazioni globali (*global explanations*): utilizzano modelli semplici e trasparenti (come gli alberi decisionali) per mostrare ai revisori la logica generale con cui ragiona l'algoritmo (come rappresentato in figura 9);

Figura 9 - Flusso logico e addestramento di un modello surrogato globale



Fonte: Dwivedi et al., 2023

- spiegazioni locali (*local explanations*): analizzano una singola decisione presa dal sistema per capire esattamente quanto ogni dato iniziale abbia influenzato quel risultato specifico.

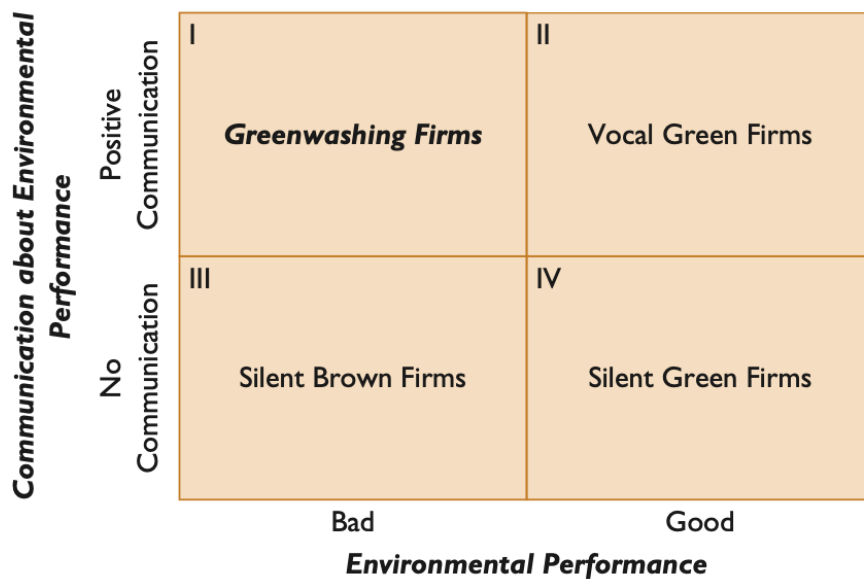
Grazie a questi strumenti, i revisori possono tracciare il percorso decisionale dell'algoritmo, valutare la rilevanza attribuita ai singoli fattori ESG e intercettare distorsioni nei dati. La XAI assicura, così, che l'audit rimanga un processo di verifica concreto e confutabile (Chamola et al., 2023).

Dal punto di vista storico, la paternità dell'espressione è attribuita all'attivista Jay Westervelt (de Freitas Netto et al., 2020), che nel 1986 la usò per criticare alcune strutture alberghiere che giustificavano il riutilizzo degli asciugamani con presunte motivazioni di risparmio idrico, mascherando così l'assenza di impegni ambientali concreti. Da questa prima denuncia, la letteratura accademica ha sviluppato nel tempo diversi inquadramenti teorici del fenomeno (Yang et al., 2020).

Una prima definizione descrive il *greenwashing* come «la disseminazione deliberata, a fini commerciali (e, pertanto, verso clienti e consumatori), di informazioni inaccurate, infondate, fuorvianti o non veritiere riguardo alle prestazioni o gli attributi ambientali di un'organizzazione, o dei suoi prodotti e servizi» (Todaro et al., 2024).

Delmas e Burbano ne propongono una definizione altrettanto influente, secondo cui il *greenwashing* è «l'intersezione di due comportamenti aziendali: scarse prestazioni ambientali e una comunicazione positiva sulle prestazioni ambientali»²² (Delmas e Burbano, 2011).

Figura 11 - Tipologia d'impresa basata su performance ambientale e comunicazione



Fonte: Delmas e Burbano, 2011

²² Testo originale: «...the intersection of two firm behaviors: poor environmental performance and positive communication about environmental performance» (Delmas e Burbano, 2011).

La matrice presente in figura 11 incrocia il dato oggettivo della performance ecologica con la propensione dell'impresa a comunicare tali tematiche all'esterno, delineando quattro distinti profili aziendali (Delmas e Burbano, 2011):

- *greenwashing firms* che, pur avendo prestazioni ambientali inadeguate ("brown"), sfruttano una narrazione positiva per ingannare gli stakeholder;
- *vocal green firms* in cui un impegno ecologico reale corrisponde a una divulgazione trasparente;
- *silent brown firms* che evitano di esporsi, per non incorrere in sanzioni o accuse di falsità;
- *silent green firms*, oggettivamente sostenibili ma reticenti per timore dello scetticismo del mercato.

Il valore di questo modello sta soprattutto nell'aver mostrato che la deriva verso l'opportunismo ambientale è una traiettoria lungo cui l'impresa si sposta nel tempo: può arrivarci per deterioramento delle proprie prestazioni reali, mantenendo però intatta la narrazione positiva (da *vocal green* a *greenwasher*); oppure può arrivarci decidendo di abbandonare il silenzio per lanciare campagne di marketing ecologico puramente di facciata (da *silent brown* a *greenwasher*).

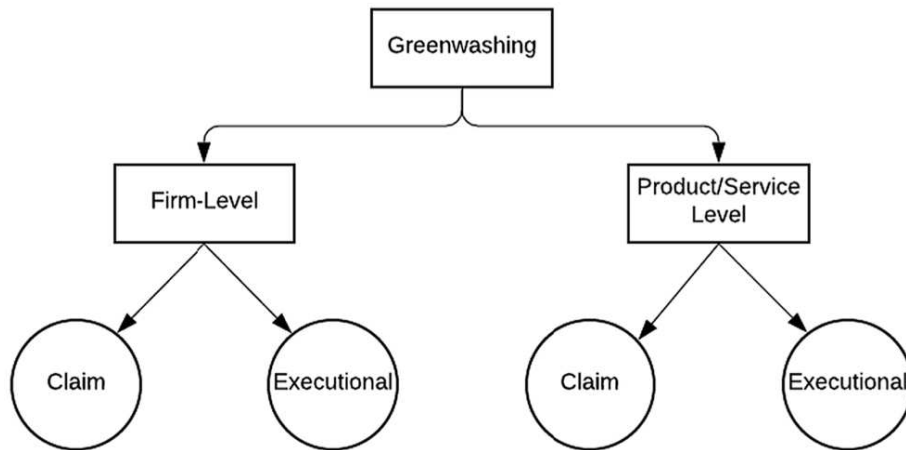
2.1.2 Tipologie e tassonomie

Nel corso degli anni le manifestazioni del greenwashing sono passate da semplici forme di pubblicità ingannevole legate alle caratteristiche ecologiche di singoli prodotti (*product-level greenwashing*) a strategie più complesse di manipolazione narrativa a livello aziendale (*firm-level greenwashing*) (de Freitas Netto et al., 2020).

Nei report ESG contemporanei il greenwashing si manifesta raramente come menzogna grossolana o fabbricazione totale dei dati; più spesso assume la forma di un ambientalismo aziendale simbolico (*symbolic corporate environmentalism*), costruito attraverso il ricorso alla narrazione: non potendo dimostrare un impatto reale sui bilanci, molte aziende preferiscono "segnalare" la propria conformità morale con retoriche suggestive (green talk), formule vaghe e allineamento formale a macro-obiettivi globali come gli SDGs dell'ONU, senza integrare realmente questi valori nei processi operativi (Todaro et al., 2024). Per mappare questo insieme di comportamenti manipolativi, sono state sviluppate dettagliate tassonomie, distinguendo tra diverse tipologie e livelli di

gravità. Una delle distinzioni fondamentali è quella tra il greenwashing testuale (*claim greenwashing*) e il greenwashing esecutivo (*executional greenwashing*), come riportato in figura 10 (de Freitas Netto et al., 2020).

Figura 12 - Classificazione multilivello del greenwashing.



Fonte: de Freitas Netto et al., 2020

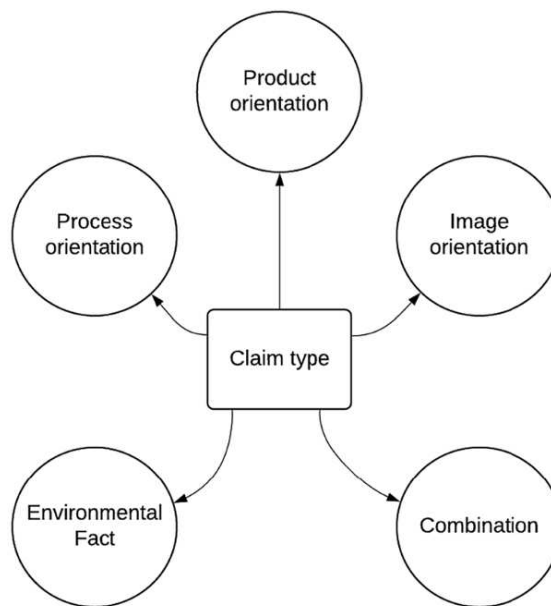
Il primo si serve di argomenti testuali espliciti o impliciti per creare una falsa credenza di sostenibilità, e si manifesta principalmente attraverso tre dinamiche individuate da Kangun, Carlson e Grove (1991, citato in de Freitas Netto et al., 2020): le affermazioni false (*false claims*), l'omissione di informazioni rilevanti (*omission*) e l'impiego di termini vaghi o ambigui (*vagueness*).

Sul piano pratico, questo tipo di *claim greenwashing* si esprime attraverso comportamenti riconoscibili (de Freitas Netto et al., 2020):

- omissione selettiva (*selective disclosure*), che divulga in modo sproporzionato indicatori ambientali benefici, occultando deliberatamente informazioni negative o prestazioni scarse su altre dimensioni critiche di sostenibilità;
- disaccoppiamento (*decoupling*), che si realizza quando gli annunci formali dell'impresa non si traducono in alcuna modifica sostanziale delle pratiche economiche effettive;
- *fuzzy reporting*, una condotta in cui l'impresa sfrutta la rendicontazione non finanziaria come un canale di comunicazione unidirezionale e asimmetrico per distorcere la verità o proiettare una rappresentazione eccessivamente ottimistica delle proprie performance di CSR;

- deviazione dell'attenzione (*attention deflection*): l'uso di espressioni prive di un significato scientifico o giuridico rigoroso (come “eco-friendly”, “naturale” o “puro”) per confondere il destinatario e sviare l'attenzione dalle scarse performance ecologiche complessive dell'impresa;
- certificazioni fittizie e marchi dubbi (*dubious authorizations and labels*), che imitano loghi di certificazioni ufficiali senza averne davvero i requisiti.

Figura 13 - Orientamento e tipologie delle asserzioni ecologiche di tipo "claim"



Fonte: de Freitas Netto et al., 2020

Come illustrato in figura 13, la comunicazione ecologica può essere classificata secondo la tassonomia originaria di Carlson, Grove e Kangun (1993, citato in de Freitas Netto et al., 2020).

Questo modello scompone l'orientamento dei *green claim* in cinque categorie (de Freitas Netto et al., 2020):

- Orientamento al prodotto (*product orientation*), incentrato sugli specifici attributi ecologici del bene offerto
- Orientamento al processo (*process orientation*), che riguarda l'efficienza ecologica dei metodi di produzione
- Orientamento all'immagine (*image orientation*), che rafforza il profilo "verde" dell'organizzazione senza legarlo a un prodotto specifico

- Dichiarazioni fattuali (*environmental fact*), affermazioni apparentemente oggettive sullo stato dell'ambiente usate per accreditarsi come voce autorevole
- Combinazione (*combination*), di due o più di questi approcci nello stesso messaggio

Accanto al *claim greenwashing* si colloca l'*executional greenwashing*, che al posto dei testi fuorvianti ricorre a stimoli visivi e uditivi che evocano la natura: sfondi di paesaggi incontaminati, colori associati all'ecologia, immagini di specie minacciate, tutti elementi pensati per generare una percezione automatica di ecocompatibilità, condizionando inconsciamente le decisioni dei lettori meno esperti (de Freitas Netto et al., 2020).

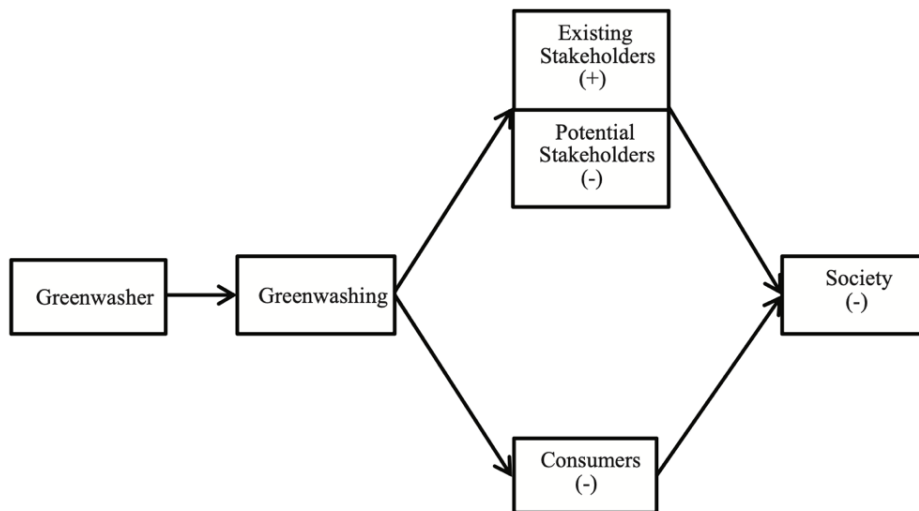
2.1.3 Evoluzione digitale

Quando queste strategie manipolatorie si spostano e si amplificano nell'ambiente online, nasce un fenomeno sempre più pervasivo: il *digital greenwashing* (Eze, 2025), che sfrutta sistematicamente canali social, siti web aziendali e piattaforme di e-commerce per diffondere messaggi ambientali distorti o privi di fondamento.

La sua insidia principale sta nella fluidità del web: i contenuti online sono spesso effimeri e transitori, il che rende molto più difficile per gli stakeholder tracciarne nel tempo la veridicità. Come nell'*executional greenwashing*, la manipolazione digitale punta soprattutto sull'immagine: narrazioni visive, grafiche interattive, loghi ecologici che l'impresa si attribuisce da sé, fotografie ad alta definizione. L'effetto cercato è una sensazione immediata di responsabilità socio-ambientale, che aggira la richiesta di prove documentali. Il meccanismo funziona perché online si legge in fretta e in superficie, e distinguere un impegno autentico da una semplice operazione di cosmesi digitale diventa difficile (Eze, 2025).

Le conseguenze di questa pratica cambiano a seconda del tipo di stakeholder coinvolto. Quando l'impegno nella CSR è autentico, l'impresa risulta più attraente e le intenzioni di acquisto aumentano. Il greenwashing rovescia l'effetto, ma non allo stesso modo per tutti. Gli stakeholder attuali possono ricavarne un vantaggio temporaneo, legato agli extra-profitti, mentre la società nel suo complesso registra comunque una perdita netta. Per gli stakeholder potenziali, in particolare per gli investitori orientati a progetti realmente ecologici, l'inganno funziona invece come una barriera, perché impedisce la corretta allocazione dei capitali.

Figura 14 - le conseguenze del greenwashing



Fonte: Yang et al., 2020

Oltre a erodere la fiducia verso il brand, se non viene contenuto il fenomeno genera danni sistemici a lungo termine: allontana i capitali dal mercato verde e incentiva le imprese a mantenere condotte dannose, producendo esternalità negative a scapito del benessere collettivo (Yang et al., 2020).

2.2 Il disallineamento tra ESG Talk e ESG Walk: misurazione del gap prestazionale

Il *greenwashing* può essere spiegato attraverso il concetto di disallineamento tra ciò che le aziende dichiarano in tema di sostenibilità (*ESG Talk*) e ciò che mettono in pratica (*ESG Walk*). Le conseguenze vanno oltre la comunicazione, perché narrazioni distorte ingannano gli investitori nell'allocazione delle risorse e compromettono il rapporto di fiducia con gli stakeholder (Zioło et al., 2024), permettendo alle imprese di incassare subito i vantaggi reputazionali di un'immagine "green" mentre nascondono performance socio-ambientali gravemente insufficienti.

Questo fenomeno è l'espressione pratica del *decoupling* (disaccoppiamento), attraverso cui le imprese costruiscono una conformità di facciata per garantirsi l'approvazione sociale. Nel contesto ESG, il *decoupling* si traduce in una sorta di adesione cerimoniale, secondo cui si firmano policy formali, si istituiscono comitati per la sostenibilità, ci si associa agli SDGs dell'ONU, ma tutto ciò resta sulla carta, senza mai entrare nei reali processi decisionali del management. È una strategia con cui l'azienda accontenta le

richieste esterne, tenendo però il proprio *core business* al riparo dai costi e dalle complessità di una vera transizione ecologica (Moneva et al., 2006).

Il risultato è che la comunicazione viaggia su un binario parallelo e scollegato dalle logiche gestionali, riducendo la sostenibilità a una pura operazione estetica.

Per tradurre questo divario in grandezze oggettive, letteratura empirica, investitori e agenzie di rating analizzano i grandi database ESG.

Nel database Thomson Reuters ASSET4²³, ad esempio, le metriche vengono sdoppiate per far emergere il disallineamento tra i "*drivers*", impegni formali, policy, annunci programmatici, cioè il Talk; e gli "*outcomes*", risultati fisici e misurabili, come le tonnellate di CO2 emesse o il tasso di infortuni sul lavoro, cioè il Walk.

Un punteggio elevato nei drivers ma deficitario negli outcomes fornisce, di fatto, la dimostrazione matematica del decoupling.

Parallelamente all'analisi dei database quantitativi, per comprendere il disallineamento serve anche un'analisi sistematica delle narrative d'impresa contenute nella rendicontazione extra-finanziaria. La letteratura ha storicamente quantificato la retorica dell'*ESG Talk* ricorrendo alla *content analysis* dei report di sostenibilità, un processo strutturato per raccogliere ed esaminare i dati, classificando le informazioni qualitative e quantitative in categorie concettuali predefinite (Küçükgül, 2020). Attraverso questa tecnica i ricercatori catalogano i messaggi aziendali in tassonomie specifiche e ne valutano il grado di ingannevolezza, mappando la presenza di espressioni vaghe, omissioni o palesi falsità (de Freitas Netto et al., 2020).

L'interpretazione di questo divario, però, è in continua evoluzione.

In passato lo scostamento tra Talk e Walk veniva letto quasi esclusivamente come prova di opportunismo. Luisi e Cucaro (2026) propongono una lettura diversa, secondo cui la comunicazione aziendale non si limita a raccontare passivamente ciò che è già stato fatto. In un contesto normativo sempre più esigente come quello della CSRD, le parole assumono un ruolo attivo: prendere un impegno pubblico e comunicarlo diventa il primo passo per innescare il cambiamento, spingendo l'organizzazione a modificare concretamente le proprie prassi (*walking*).

²³Thomson Reuters ASSET4 è un database internazionale specializzato in dati ESG, che monitora le performance aziendali raccogliendo circa 900 indicatori qualitativi e quantitativi da fonti pubbliche.

Secondo questa lettura, le organizzazioni sviluppano una capacità dinamica discorsiva: scelgono i temi e i canali con cui rispondere alle pressioni istituzionali, e su questi modellano la propria strategia di comunicazione.

Questa competenza si fonda su tre pilastri (Luisi e Cucaro, 2026):

- percezione discorsiva (*discursive sensing*): l'impresa intercetta le nuove istanze normative o sociali in tema di sostenibilità e le fa entrare nei propri flussi comunicativi;
- posizionamento discorsivo (*discursive seizing*): la narrazione viene collocata rispetto al destinatario e al canale, con l'obiettivo di acquisire legittimità;
- trasformazione discorsiva (*discursive transforming*): la retorica resta ancorata alla strategia di lungo periodo anche quando gli scenari operativi cambiano.

Resta comunque un confine netto da tracciare tra questa flessibilità strategica e la manipolazione. Se la narrazione si disallinea in modo definitivo dalla realtà gestionale, l'impresa finisce nel greenwashing patologico, esponendosi a gravi rischi reputazionali; se invece la comunicazione spinge verso traguardi futuri (*l'aspirational talk*), le dichiarazioni pubbliche assumono la valenza di un vincolo morale che obbliga l'intera struttura a un miglioramento progressivo delle proprie performance (Luisi e Cucaro, 2026).

Per misurare concretamente questo divario, secondo Luisi e Cucaro (2026) conta soprattutto la coerenza strategica delle informazioni pubblicate, quello che gli autori chiamano "allineamento configurazionale", illustrato attraverso la matrice Piattaforma-Materialità.

La comunicazione aziendale risulta credibile, e il rischio di opportunismo si riduce, solo quando due fattori coincidono:

- *la materialità settoriale*, per cui la narrazione deve riflettere oggettivamente i rischi reali del business (per un'azienda carbon-intensive, l'assenza di risposte concrete sui rischi fisici è un segnale d'allarme di possibile decoupling, mentre per un'impresa tecnologica la credibilità si misura più sulla trasparenza etica e di governance);

- *la scelta coerente dei canali digitali*, poiché l'uso di piattaforme ad alto coinvolgimento emotivo per veicolare asserzioni vaghe può essere sintomo di campagne di facciata, mentre canali più istituzionali risultano più adatti a comunicazioni tecniche per investitori e specialisti.

Allineare rischi materiali, tematiche e canali comunicativi trasforma quindi il report da semplice vetrina a strumento strategico (Luisi e Cucaro, 2026). Resta però un problema pratico non da poco: per uno stakeholder esterno è quasi impossibile verificare da solo questa coerenza. Serve quindi l'intervento di revisori indipendenti che certifichino la corrispondenza tra ESG Talk ed ESG Walk.

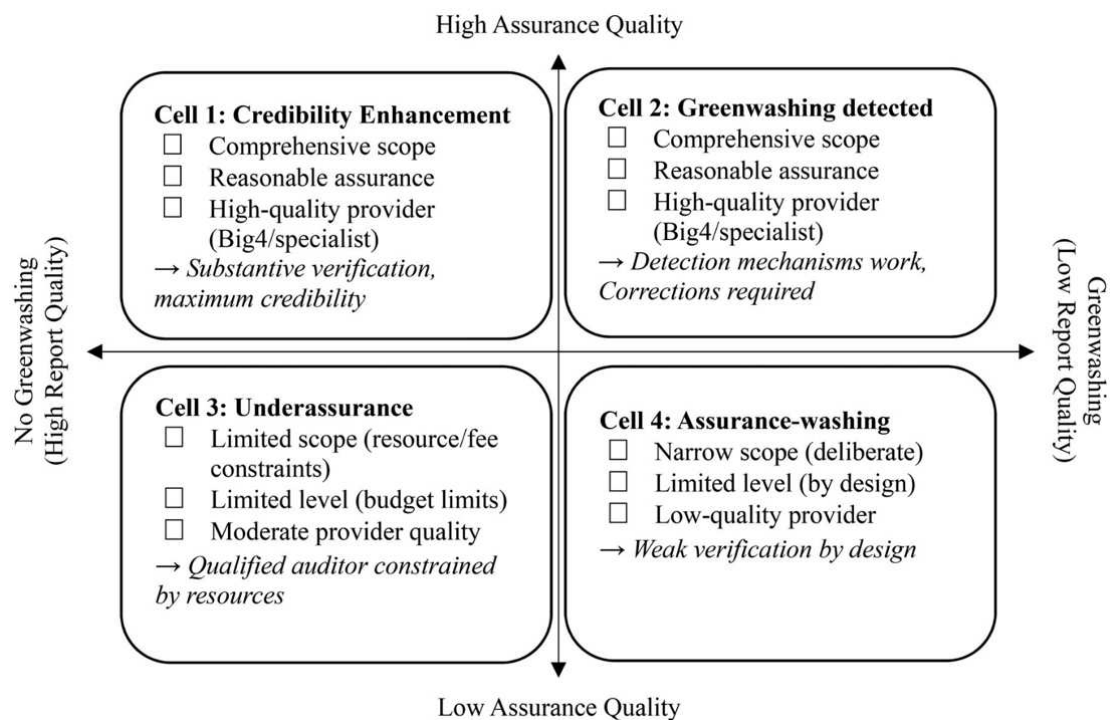
2.3 Dalla comunicazione all'auditing: il fenomeno dell'assurance-washing

La richiesta di verifiche indipendenti sui report di sostenibilità è cresciuta molto negli ultimi anni. In teoria l'audit dovrebbe dare credibilità ai dati, ridurre le asimmetrie informative e rendere le imprese più responsabili verso la collettività. Nei fatti la certificazione non basta a blindare l'integrità aziendale: il greenwashing ha raggiunto anche i processi di revisione, dando origine all'*assurance-washing*. Il termine indica i controlli esterni che nascono già svuotati e si riducono a una formalità, invece di presidiare la sostanza.

Il meccanismo si sviluppa attraverso tre leve principali (Sari et al., 2026):

- limitazione dell'ambito di verifica (*narrow scope*), che sottopone a controllo solo un sottoinsieme ristretto di indicatori, lasciando scoperto il resto del report;
- livello di affidabilità ridotto (*limited level*), che punta sistematicamente sulla *limited assurance*, un esame superficiale basato su procedure analitiche blande, evitando la *reasonable assurance*, che richiederebbe campionamenti rigorosi;
- scelta del revisore (*low-quality provider*), affidandosi a soggetti non attrezzati o sfruttando una dipendenza economica per condizionare persino i grandi network di revisione.

Figura 15 - Tassonomia greenwashing-assurance



Fonte: Sari et al., 2026

L'assurance non finanziaria smette così di essere uno strumento di trasparenza e diventa un "rituale cerimoniale", usato per rassicurare fittiziamente gli investitori mentre si neutralizzano i controlli e si costruisce una narrazione aziendale ingannevole (Sari et al., 2026).

È importante non confondere questo comportamento con la cosiddetta *underassurance*, che costituisce un limite involontario, causato da vincoli di budget o carenze strutturali di un'azienda che vorrebbe comunque fare le cose per bene. L'assurance-washing, invece, è una scelta strategica e calcolata (Sari et al., 2026).

I manager scelgono deliberatamente standard di controllo leggeri per evitare un'indagine approfondita, sfruttando la natura volontaria del reporting o la debolezza dei sistemi sanzionatori.

Il paradosso è evidente: se la revisione si riduce a una formalità, l'audit non finanziario non funziona più da scudo contro l'opportunismo, anzi finisce per legittimarlo, rendendolo ancora più difficile da individuare (Sari et al., 2026).

2.4 Le determinanti interne ed esterne del comportamento opportunistico

Quando le organizzazioni decidono di sviare verso pratiche di rendicontazione ingannevoli, alimentando il divario tra promesse e azioni reali, non lo fanno quasi mai per caso o per un semplice errore. Si tratta, al contrario, di una scelta strategica ponderata, spesso basata su un calcolo di convenienza economica (Lyon e Maxwell, 2006).

Avviare una vera transizione verso la sostenibilità richiede capitali immediati e costi di riconversione industriale notevoli, mentre i ritorni economici di questi sforzi sono spesso immateriali e visibili solo a distanza di anni (Tylec, 2025). Di fronte a questo squilibrio temporale, manipolare i dati ESG diventa un'alternativa razionale ogni volta che il costo della menzogna, o della conformità di facciata, viene percepito come inferiore allo sforzo necessario per un cambiamento sistemico reale.

Questo tipo di comportamento opportunistico nasce dall'intersezione di due macro-forze contrapposte (Delmas e Burbano, 2011):

- le *determinanti esterne* includono tutte quelle pressioni provenienti dall'ambiente circostante, come l'evoluzione della sensibilità dei consumatori, le richieste dei mercati dei capitali e lo scrutinio delle organizzazioni non governative (ONG), che impongono all'organizzazione l'urgenza di proiettare un'immagine di assoluta responsabilità ecologica per mantenere la propria "licenza di operare";
- le *determinanti interne* riguardano i conflitti d'interesse e le debolezze strutturali annidate nell'architettura aziendale. Tra questi spiccano i sistemi di remunerazione del management legati a risultati finanziari di breve termine e una cultura d'impresa che concepisce la sostenibilità come uno strumento di marketing.

Elencare questi fattori di rischio, però, non basta a spiegare come si traducano concretamente in pratiche di inganno e decoupling. Per capirlo davvero occorre guardare ai meccanismi organizzativi e cognitivi che li mettono in moto e, nei paragrafi seguenti, proveremo a farlo attraverso l'analisi di alcune teorie.

2.5 Le teorie organizzative

Le determinanti interne ed esterne appena descritte mostrano che fenomeni opportunistici come il greenwashing non si spiegano guardando all'impresa come a un soggetto isolato, mosso dal solo reddito. Per capire il disallineamento tra le promesse dichiarate (*talk*) e le azioni concrete (*walk*) è possibile fare affidamento alla prospettiva delle teorie organizzative, che spiegano perché le aziende preferiscano spesso soluzioni simboliche a cambiamenti sostanziali. I due approcci che seguono, la *institutional theory* e la *legitimacy theory*, condividono un assunto di fondo: l'azienda è un'entità sociale che, per garantirsi risorse e sopravvivenza nel lungo periodo, deve conformarsi in modo credibile alle aspettative, ai valori e alle norme del contesto in cui opera. La teoria neo-istituzionale, oggetto della prossima sezione, in particolare parla di disaccoppiamento e conformità rituale: sottoposte a forti pressioni sociali e reputazionali, le imprese tendono a fornire risposte di facciata per preservare la propria legittimità.

2.5.1 Institutional Theory

La *teoria neo-istituzionale* aiuta a comprendere come le pressioni del contesto esterno spingano le organizzazioni ad allinearsi a modelli di comportamento formalizzati, determinando spesso una netta separazione tra le dichiarazioni pubbliche e le pratiche reali (Meyer e Rowan, 1977, citato in Bromley e Powell, 2012). Alla base di questo approccio, c'è la consapevolezza che le organizzazioni non operano in un "vuoto sociale", limitandosi a inseguire l'efficienza tecnica e l'ottimizzazione produttiva. Come evidenziato dalla letteratura classica, infatti, gli ambienti istituzionali impongono regole, strutture e cerimoniali a cui le imprese devono necessariamente conformarsi per ottenere supporto e legittimità sociale, a prescindere dall'utilità tecnica immediata di tali adempimenti (Scott, 1987).

Per comprendere la dinamica di queste pressioni, W. Richard Scott identifica tre pilastri²⁴ fondamentali che sostengono e stabilizzano l'ordine istituzionale (Scott, 2014):

²⁴ Il modello dei tre pilastri (regolativo, normativo e culturale-cognitivo) è stato originariamente teorizzato dall'autore nella prima edizione del suo volume *Institutions and Organizations* (Sage, 1995). Nel presente elaborato si fa riferimento all'articolo del 2014, in cui lo stesso Scott propone una riflessione retrospettiva ripercorrendo la genesi e l'evoluzione del proprio framework concettuale

- il pilastro regolativo (*regulative*), che riguarda le norme formali, le leggi, le sanzioni e i sistemi di incentivi istituiti dalle autorità per regolare il comportamento aziendale;
- il pilastro normativo (*normative*), che definisce i valori, i modelli di comportamento appropriati e le aspettative morali condivise all'interno di una comunità professionale, delineando ciò che "*dovrebbe essere fatto*";
- il pilastro culturale-cognitivo (*cultural-cognitive*), che fa riferimento agli assunti profondi e alle convinzioni consolidate nella società.

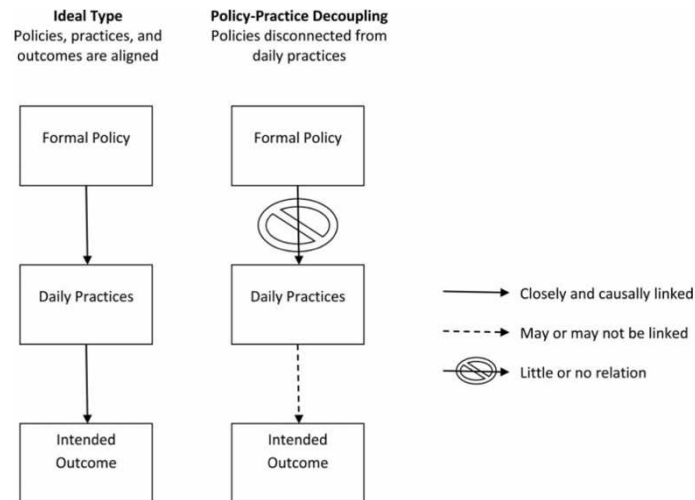
Per garantirsi risorse e sopravvivenza, le imprese tendono ad allinearsi a questi pilastri, innescando un processo di omologazione strutturale noto come *isomorfismo istituzionale*²⁵ (DiMaggio e Powell, 1983, citato in Scott, 1987), guidato da tre tipi di pressioni: *coercitiva* (regolamenti e mandati legali), *mimetica* (l'imitazione dei concorrenti percepiti come leader in contesti di elevata incertezza) e *normativa* (la diffusione di standard professionali e accademici). È proprio questa ricerca di legittimità esterna a poter generare fenomeni patologici di disaccoppiamento (*decoupling*). Quando le imprese subiscono forti pressioni reputazionali senza avere le risorse, o la reale volontà, di modificare le proprie routine operative, tendono ad adottare soluzioni di facciata per compiacere gli stakeholder esterni (Meyer e Rowan, 1977, citato in Bromley e Powell, 2012). Sotto questo profilo, il greenwashing viene concettualizzato una *strategia di disaccoppiamento simbolico*, attuata per proiettare un'immagine di conformità ecologica formale lasciando del tutto inalterate le routine industriali e le condotte operative dannose.

Nel tempo, il concetto di disaccoppiamento si è arricchito distinguendo due manifestazioni (Bromley e Powell, 2012):

- il *policy-practice decoupling* (figura 16) rappresenta la frattura classica tra *walk* e *talk*: le politiche e le promesse ambientali contenute nei report vengono adottate su un piano puramente formale, restando completamente disconnesse dalle attività quotidiane e sistematicamente ignorate nel nucleo operativo;

²⁵ I concetti di isomorfismo istituzionale e la relativa tripartizione in pressioni di tipo coercitivo, mimetico e normativo sono stati formalizzati originariamente in: P. J. DiMaggio e W. W. Powell, "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields", in *American Sociological Review*, vol. 48, n. 2, 1983, pp. 147-160.

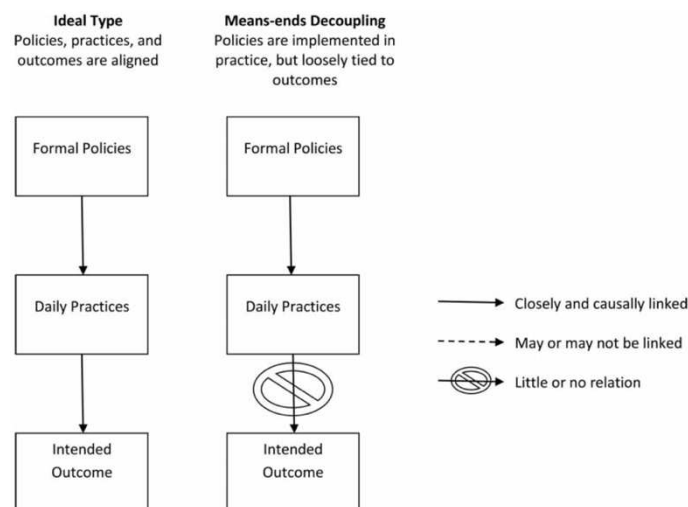
Figura 16 - Il divario tra teoria e pratica (*policy-practice decoupling*) a confronto con l'allineamento di un'organizzazione ideale



Fonte: Bromley e Powell, 2012

- il *means-ends decoupling* (figura 17), invece, si verifica in contesti di elevata razionalizzazione e forte spinta all'audit, in cui le politiche di sostenibilità vengono effettivamente implementate e le routine lavorative modificate, ma il legame causale tra queste attività procedurali (i mezzi) e l'effettivo raggiungimento dell'obiettivo dichiarato (i fini) resta opaco o inefficace. L'adempimento procedurale e la misurazione diventano fini a sé stessi, drenando risorse e attenzione senza alcuna prova tangibile di un beneficio ecologico reale.

Figura 17 - Il divario tra mezzi e fini (*means-ends decoupling*) a confronto con l'allineamento di un'organizzazione ideale



Fonte: Bromley e Powell, 2012

2.5.2 Legitimacy Theory

Per capire cosa spinga i vertici aziendali a pubblicare informazioni extra-finanziarie, uno dei modelli più diffusi è la *teoria della legittimità*. In questa lettura il reporting ESG è una mossa strategica attraverso cui il management allinea l'immagine dell'impresa alle aspettative del pubblico, disinnescando minacce reputazionali o, nei casi peggiori, maschera pratiche di greenwashing.

Il punto di partenza di questo approccio è la definizione di Mark C. Suchman (1995), che descrive la *legittimità* come «*la percezione o l'assunto generalizzato che le azioni di un'entità siano desiderabili, corrette o appropriate all'interno di un sistema socialmente costruito di norme, valori, credenze e definizioni*»²⁶ (Suchman, 1995).

Alla base c'è l'idea di un contratto sociale implicito tra l'azienda e la società in cui opera: il diritto di un'impresa a operare va guadagnato continuamente, producendo risultati che la collettività consideri desiderabili e garantendo vantaggi ai gruppi da cui dipende il proprio consenso.

I valori e le priorità della società, tuttavia, non sono fissi nel tempo.

Basti pensare a quanto accaduto tra gli anni '60 e '70, quando la crescente sensibilità verso i danni ambientali provocati dall'industria ha fatto crollare la fiducia dell'opinione pubblica verso le imprese. Quando un'azienda non riesce a mettersi al passo con questi nuovi valori sociali, si crea un vero e proprio divario, il cosiddetto *legitimacy gap*. Se questo strappo nel contratto sociale non viene ricucito, l'impresa si espone a boicottaggi, interventi legislativi punitivi e danni d'immagine, arrivando persino a perdere la sua "licenza di operare".

A seconda del tipo di valutazione fatta dagli stakeholder, si identificano tre forme principali di legittimità (Suchman, 1995):

- *legittimità pragmatica*: si fonda su una logica utilitaristica e di interesse personale. Gli stakeholder supportano l'azienda perché ne ricavano un vantaggio pratico o economico diretto. Non a caso, è proprio in questo ambito che prolifera il greenwashing, poiché la comunicazione di facciata viene usata per assecondare i calcoli di convenienza del pubblico;

²⁶ Testo originale: «*a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions*» (Suchman, 1995).

- *legittimità morale*: si basa su principi etici indipendenti dai vantaggi economici immediati. L'azienda guadagna legittimità perché le sue azioni sono percepite come intrinsecamente "giuste" e capaci di generare un reale benessere per la società e l'ambiente;
- *legittimità cognitiva*: è il livello più radicato e profondo. L'esistenza e il modo di operare dell'impresa sono considerati talmente naturali da essere dati per scontati.

Figura 18 - Classificazione delle dinamiche di legittimazione (pragmatica, morale e cognitiva)

	Actions	Essences	
Episodic	Exchange	Disposition Interest	Pragmatic Legitimacy
Continual	Influence	Character	
Episodic	Consequential	Personal	Moral Legitimacy
Continual	Procedural	Structural	
Episodic	Comprehensibility Predictability Plausibility		Cognitive Legitimacy
Continual	Taken-for-Grantedness Inevitability Permanence		

Fonte: Suchman, 1995

Quando si verificano crisi o eventi che mettono in pericolo la reputazione aziendale, i report ESG si trasformano in veri e propri scudi difensivi per il management (Patten, 1992).

Lindblom (1994, citato in Cho e Patten, 2007) individua quattro strategie comunicative con cui le imprese arginano le minacce alla propria legittimità:

1. *educare e informare* gli stakeholder sui reali miglioramenti apportati;
2. *alterare le percezioni* degli stakeholder rispetto a eventi critici, senza modificare le prassi operative;
3. *distogliere l'attenzione* dalle questioni più delicate, dirottando comunicazione e reporting su attività collaterali ad alto impatto emotivo positivo;
4. *ridefinire le aspettative esterne*, cercando di ricalibrare gli standard sociali di idoneità per renderli in linea con le attuali capacità dell'organizzazione.

L'uso concreto di queste strategie difensive è documentato dallo studio di Dennis M. Patten (1992), che analizzò le reazioni al disastro della Exxon Valdez²⁷ del 1989 concentrandosi sulle compagnie petrolifere non coinvolte nell'incidente. Patten scoprì che, proprio nel 1989, i report annuali di queste compagnie registrarono un aumento drastico delle informazioni ambientali, passando da una media di 0,61 a 1,90 pagine. L'aumento serviva a calmare l'opinione pubblica e a scongiurare nuove normative stringenti sul settore, senza alcun miglioramento ecologico alle spalle.

Il dato più sorprendente è che chi inquina di più tende a parlare di più di sostenibilità.

Lo studio di Cho e Patten (2007) ha confermato che, a parità di dimensioni e settore, le aziende con i peggiori record ambientali redigono report più corposi e ricchi di dichiarazioni "verdi" rispetto ai concorrenti più virtuosi, ed evitano sistematicamente di fornire dati quantitativi precisi, come i costi per sanzioni o contenziosi, perché i numeri metterebbero in luce le reali inefficienze gestionali. Preferiscono invece saturare i documenti di contenuti qualitativi e narrativi: dichiarazioni di intenti, promesse sul riciclo, politiche che restano confinate alla teoria.

Alla luce di queste evidenze, il greenwashing appare come la manifestazione più estrema della teoria della legittimità: una facciata costruita per salvare le apparenze e proteggere il business.

2.6 Le teorie economiche e relazionali

Le teorie organizzative viste finora considerano la rendicontazione ESG quasi come un obbligo di facciata. Le *teorie economico-relazionali*, invece, la leggono come una leva strategica che il management usa per dialogare con gli stakeholder, valorizzare l'azienda e ridurre i gap informativi sul mercato.

Questo cambio di prospettiva si articola in tre framework interconnessi: la *signaling theory*, la *resource-based view* e la *stakeholder theory*.

2.6.1 Signaling Theory

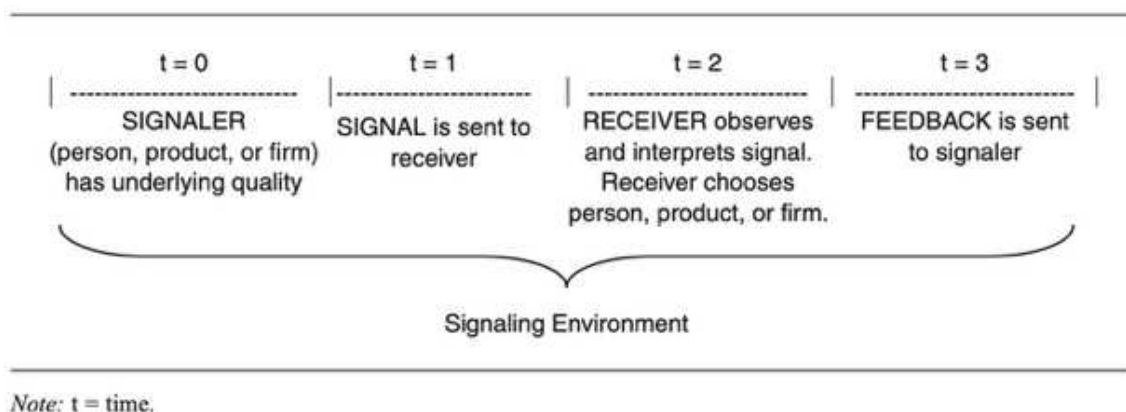
Per capire cosa spinga le aziende a rendicontare le proprie performance extra-finanziarie, e a scivolare, talvolta, in pratiche di greenwashing, è utile la *teoria dei segnali* (Connelly et al., 2011).

²⁷ Il 24 marzo 1989 la Exxon Valdez si incagliò su una scogliera nel Golfo dell'Alaska e versò in mare oltre 11 milioni di galloni (circa 41,6 milioni di litri) di petrolio.

Il punto di partenza è l'*asimmetria informativa*, teorizzata da Michael Spence nel 1973, secondo cui in qualsiasi mercato esistono due soggetti che non condividono le stesse informazioni. I vertici aziendali (*insider*) conoscono la reale qualità e i difetti della propria organizzazione; il mercato (*outsider*), cioè investitori, consumatori e stakeholder, non ne dispone, pur avendone bisogno per decidere in modo consapevole.

Chi possiede l'informazione deve quindi mandare un "segnale" per dimostrare il proprio valore, mentre chi lo riceve deve valutarne la credibilità (Connelly et al., 2011).

Figura 19 - processo di segnalazione tra insider e outsider



Fonte: Connelly et al., 2011

Applicato alla sostenibilità aziendale, il report ESG diventa esattamente questo: un segnale strategico e volontario. Poiché gli stakeholder non possono ispezionare direttamente le routine produttive né calcolare le esternalità ecologiche, l'impresa usa la rendicontazione per comunicare il proprio impegno. Le aziende realmente virtuose (le *vocal green*) hanno tutto l'interesse a mandare un segnale forte per distinguersi dai concorrenti meno attenti (le *brown firms*).

Perché questo meccanismo funzioni davvero e riesca a separare le aziende virtuose da quelle inattive, creando un *separating equilibrium*, Spence sottolinea che il segnale deve essere costoso (Karasek e Bryant, 2012).

Riorganizzare la filiera produttiva, abbattere le emissioni, sottoporsi ad audit rigorosi comporta investimenti strutturali enormi (Connelly et al., 2011; 2025).

Le imprese sostenibili sostengono questi costi, e per questo il loro report ESG è un segnale difficile da imitare. Le imprese con scarse performance ambientali (*brown firms*), al contrario, sono fortemente tentate da condotte opportunistiche basate su "chiacchiere a basso costo" (*cheap talk*), che generano proprio quel disallineamento tra ESG Talk ed ESG Walk²⁸.

La teoria dei segnali spiega quindi perché il greenwashing prolifera: le aziende sfruttano la natura volontaria del reporting per emettere segnali a basso costo, slogan vaghi, impegni formali non vincolanti, simulando un impegno che nella realtà operativa non esiste (Connelly et al., 2011; 2025). Quando questi segnali simbolici si moltiplicano senza controllo regolatorio, si rischia di arrivare a un equilibrio di coalizione (*pooling equilibrium*), in cui i riceventi non distinguono più le aziende sostenibili da quelle opportunistiche, con l'effetto paradossale di penalizzare anche le imprese autenticamente sostenibili, distorcendo l'allocazione dei capitali. Questo scenario, peraltro, sta cambiando rapidamente con le nuove architetture tecnologiche.

Da una parte, l'integrazione di strumenti come l'intelligenza artificiale o l'analisi avanzata dei dati offre l'opportunità di blindare l'oggettività del segnale ESG, riducendo lo spazio per narrazioni cerimoniali.

Dall'altra, però, la stessa evoluzione digitale fornisce alle aziende strumenti per manipolare le percezioni in modo più sofisticato: nell'ecosistema online, fatto di comunicazioni veloci ed effimere, il segnale di sostenibilità si sposta sempre di più sull'estetica e sull'emotività, e l'IA generativa rischia di facilitare narrazioni "green" perfette nella forma ma prive di sostanza misurabile²⁹.

2.6.2 Resource Based View Theory

La *resource-based view* (RBV) trae origine dal contributo di Birger Wernerfelt del 1984, che sposta l'attenzione dal posizionamento dell'impresa nel mercato, alle risorse di cui essa dispone (Lockett et al., 2008). Nel 1991 Barney ne precisa le condizioni, sostenendo che il successo competitivo deriva dalla capacità di controllare e sfruttare un insieme unico di risorse e competenze (Lockett et al., 2009). Per garantire un vantaggio duraturo, queste risorse devono soddisfare i quattro criteri del modello VRIN³⁰: creare valore

²⁸ Vedi paragrafo 2.2.

²⁹ Questo è uno dei punti centrali della tesi: vedi capitolo 3 e capitolo 5.

³⁰ Formulato da Barney nel 1991 nell'articolo «Firm Resources and Sustained Competitive Advantage».

(*valuable*), essere rare (*rare*), risultare difficilmente imitabili (*inimitable*) e non avere validi sostituti (*non-substitutable*).

Il vantaggio si mantiene solo quando questa combinazione è protetta e impossibile da replicare per i concorrenti (Barney, 2001). La sostenibilità aziendale, in questa luce, cessa di essere un obbligo normativo e diventa una risorsa strategica, in quanto un impegno ecologico sostanziale ottimizza i processi produttivi, riduce i costi e aumenta il valore percepito dagli stakeholder; una corretta gestione ambientale e sociale abbassa i rischi reputazionali e legali, aumentando la resilienza dell'impresa (Fornasari e Traversi, 2024). Avviare una transizione ecologica vera, però, richiede capitali ingenti e sforzi organizzativi rilevanti, che molte imprese, soprattutto in condizioni di stress finanziario, non hanno (Tylec, 2025). Il greenwashing emerge, in questi casi, come un espediente estetico per camuffare la mancanza di innovazioni reali, simulando un vantaggio competitivo senza sostenerne i costi industriali effettivi.

A distanza di trent'anni dalla formulazione originaria, Barney et al. (2021) hanno riletto la teoria alla luce dei meccanismi di generazione e ripartizione del valore economico. Secondo questa impostazione, l'impresa genera valore quando il prezzo che i clienti sono disposti a riconoscere eccede il costo sostenuto per realizzare il bene o il servizio. Le risorse necessarie a tale scopo, tuttavia, solo di rado sono nella piena disponibilità dell'impresa, poiché provengono in misura rilevante da dipendenti, fornitori e partner. Tali soggetti sono spesso chiamati a effettuare investimenti specifici, il cui rendimento dipende dalla combinazione con le risorse di quella determinata impresa e si ridurrebbe sensibilmente in un impiego alternativo. La loro disponibilità a sostenerli è subordinata all'aspettativa di partecipare al valore che contribuiscono a generare. Di conseguenza, un'impresa orientata a destinare il risultato economico ai soli azionisti incontra maggiori difficoltà nell'attrarre le risorse di cui necessita. Gli autori non ritengono, peraltro, che ogni stakeholder vanti un titolo sul valore residuo e riconoscono la complessità di conciliare interessi spesso divergenti. In questa lettura, la RBV presenta significativi punti di contatto con la stakeholder theory, esaminata nel paragrafo successivo.

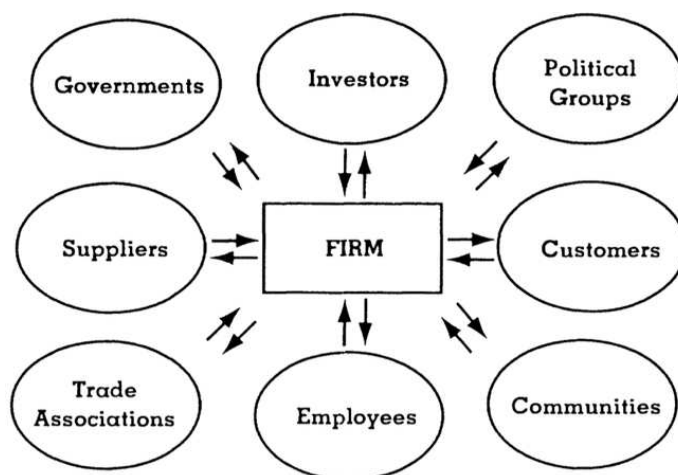
2.6.3 Stakeholder Theory

La *stakeholder theory* assume una prospettiva prettamente strategico-relazionale.

In quest'ottica, l'organizzazione va oltre la funzione di strumento per massimizzare la ricchezza degli azionisti (*shareholder wealth maximization*), e si configura come un

ecosistema, o *value network* (Freeman et al., 2020). È una rete in cui dipendenti, fornitori, clienti, investitori e comunità locali cooperano per scambiarsi valore (Donaldson e Preston, 1995). Diventa quindi compito primario del management bilanciare attivamente questi interessi contrastanti, garantendo la tenuta del patto fiduciario ed evitando conflitti che potrebbero minare la sopravvivenza a lungo termine dell'azienda (Freeman et al., 2004).

Figura 20 - la stakeholder theory e la rete di relazioni tra l'impresa e i suoi portatori di interesse



Fonte: Donaldson e Preston, 1995

Il report ESG diventa, in questa logica, il principale canale di mediazione tra l'azienda e il suo network. Per rispondere alle crescenti pressioni dei consumatori, delle istituzioni e degli investitori *SRI* (*Socially Responsible Investing*), le organizzazioni usano la pubblicazione dei propri indicatori non finanziari per negoziare costantemente i termini del contratto sociale: un reporting trasparente permette di superare la semplice esposizione di dati storici e di promuovere un vero *stakeholder engagement*, trasformando le istanze esterne in strategie aziendali condivise (Fornasari e Traversi, 2024). Il greenwashing, analizzato attraverso la *stakeholder theory*, rappresenta un profondo e gravissimo fallimento relazionale. Anche se nel brevissimo termine può fruttare extra-profitti e ritorni d'immagine per gli azionisti, logora irreparabilmente il legame di fiducia con la collettività. Le reazioni sanzionatorie del mercato di fronte alle manipolazioni (*deceptive manipulation*) sono infatti severe: i consumatori maturano un profondo scetticismo (*green skepticism*) che spesso sfocia in boicottaggi o in azioni di *shaming* digitale, gli investitori sostenibili ritirano i propri capitali e l'intera società,

infine, subisce un danno dovuto all'inefficienza del "mercato verde", premiando operatori opportunistici a scapito di chi investe in vera sostenibilità (Yang et al., 2020).

L'introduzione della CSRD e del principio della doppia materialità rappresenta, in questo senso, la definitiva istituzionalizzazione giuridica dei principi della stakeholder theory nell'ordinamento europeo³¹. Obbligando il management a considerare con pari rigore sia gli impatti esterni della propria catena del valore (prospettiva *inside-out*) sia i rischi finanziari che tali dinamiche riversano sull'azienda stessa (prospettiva *outside-in*), il legislatore ha di fatto vietato di subordinare le legittime aspettative della collettività al mero interesse finanziario.

2.7 Le teorie comportamentali

2.7.1 La razionalità limitata

Per analizzare a fondo il greenwashing, è utile abbandonare i modelli macro-sistemici analizzati in precedenza, in favore di un approccio micro-cognitivo e comportamentale. Le teorie tradizionali spiegano, infatti, il motivo per cui le aziende facciano rendicontazione di sostenibilità, spronate dalla reputazione e dalle pressioni degli stakeholder. L'approccio comportamentale, invece, chiarisce come chi lavora nell'organizzazione percepisca e metta in pratica queste scelte. La letteratura evidenzia che la manipolazione dei dati è spesso il risultato di bias cognitivi e vulnerabilità psicologiche che condizionano i manager all'interno di contesti complessi e incerti. Esplorare le basi psicologiche della CSR permette, quindi, di mappare quegli ostacoli mentali che frenano la reale transizione ecologica, favorendo derive opportunistiche anche in soggetti partiti originariamente con buone intenzioni (Moore, 2008).

Alla base di questa dinamica c'è il concetto di *razionalità limitata* di Herbert Simon, molto usato oggi nell'etica aziendale (Moore, 2008). Il compito di questa teoria è quello di «sostituire la razionalità globale dell'uomo economico con un tipo di comportamento razionale che sia compatibile con l'accesso alle informazioni e le capacità computazionali effettivamente possedute dagli organismi»³² (Simon, 1955).

Le decisioni dei manager sono condizionate da limiti di tempo, informazioni parziali e risorse cognitive finite che, spesso, creano dei veri e propri vincoli biologici e psicologici.

³¹ Vedi paragrafo 1.3.1.

³² Testo originale: «...replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms...» (Simon, 1955).

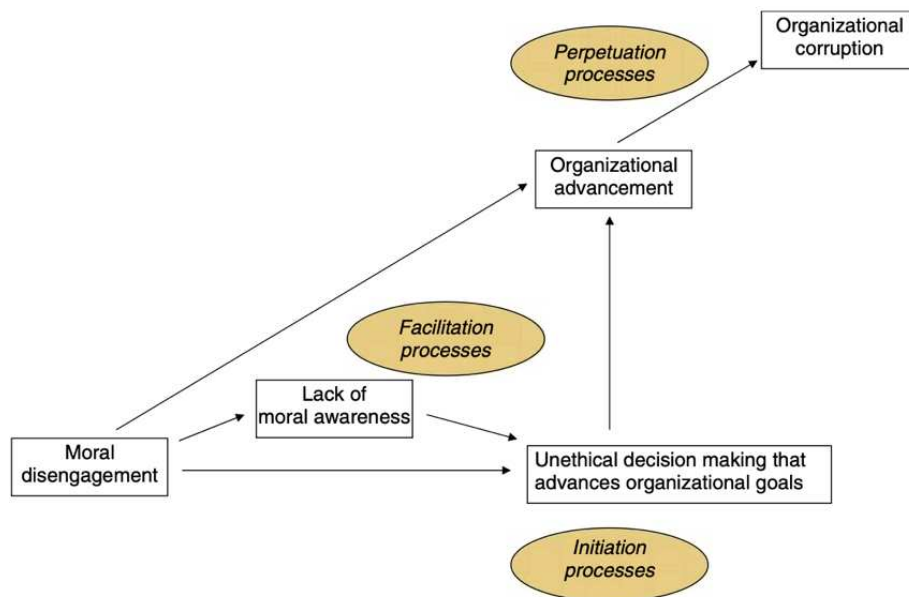
Strategicamente, i decisori finiscono per accontentarsi di soluzioni *soddisfacenti*, invece di cercarne di *ottimali*.

Come sottolinea Simon, in simili contesti decisionali complessi «*il giocatore, invece di cercare la mossa "migliore", deve solo cercare una mossa "buona"*»³³ (Simon, 1955).

2.7.2 Il disimpegno morale

La *teoria del disimpegno morale* di Albert Bandura del 1999 spiega, poi, come sia possibile che manager apparentemente onesti firmino report di sostenibilità fuorvianti, senza provare un profondo senso di colpa. Secondo questo approccio, gli individui sono in grado di "spegnere" temporaneamente la propria autocensura etica, riuscendo così a giustificare e normalizzare condotte dannose o ingannevoli, a vantaggio dell'organizzazione (Bandura, 1999).

Figura 21 - Il ruolo del disimpegno morale nei processi di inizio, facilitazione e perpetuazione dell'opportunismo aziendale



Fonte: Moore, 2008

I manager neutralizzano il proprio giudizio morale affidandosi principalmente a due scappatoie (Moore, 2008):

³³ Testo originale: «The player, instead of seeking for a "best" move, needs only to look for a "good" move.» (Simon, 1955).

- la *diffusione della responsabilità*, per cui ci si assolve pensando "lo fanno tutti i nostri concorrenti", imitando le pratiche di facciata diffuse nel settore, il comportamento perde la sua connotazione di colpa individuale;
- il *confronto vantaggioso* che minimizza il danno ambientale esaltando altri meriti dell'azienda. Attraverso queste scorciatoie cognitive, il greenwashing smette di essere percepito come una truffa e diventa, agli occhi di chi lo pratica, una necessità strategica.

2.7.3 La dissonanza cognitiva

Questo processo psicologico si lega strettamente alla *teoria della dissonanza cognitiva* di Leon Festinger del 1957, secondo cui il management di un'azienda inquinante vive una costante contraddizione. Nonostante conosca i reali impatti ecologici interni, esso subisce forti pressioni esterne per proiettare un'immagine "green". Per alleviare questo disagio mentale scatta l'autoinganno. Le riforme strutturali sono lunghe e costose, perciò il manager si convince che i piccoli interventi simbolici già adottati siano sufficienti. Il disimpegno morale interviene ancora prima, bloccando la dissonanza sul nascere, e porta a credere alle proprie narrazioni edulcorate e a slogan vaghi come "eco-friendly". Il greenwashing diventa così un processo quasi involontario: chi diffonde il report ingannevole è spesso il primo a crederci (Moore, 2008).

2.7.4 La teoria dei prospetti

La *teoria dei prospetti* di Kahneman e Tversky del 1979 offre un'ottima chiave di lettura per queste dinamiche. Il concetto di base è che le nostre scelte dipendono fortemente da come ci viene presentato un problema e dalla nostra innata avversione alle perdite. In pratica, cambiamo atteggiamento verso il rischio, rimanendo molto prudenti quando si tratta di accumulare guadagni, ma diventando azzardati pur di evitare una perdita certa. Applicata alla sostenibilità, la teoria descrive bene la posizione di chi gestisce un'azienda altamente inquinante (*brown firm*), che si trova spesso a operare nel "dominio delle perdite". Il crollo del rating ESG, la fuga degli investitori, una sanzione imminente: sono minacce percepite come danni certi e ravvicinati. Sostenere i costi di una vera transizione ecologica appare allora meno conveniente di una scommessa, e l'avversione alle perdite spinge verso scelte azzardate (Kahneman e Tversky, 1979): si accettano rischi reputazionali enormi, truccando i dati ambientali od omettendo quelli negativi, nella

speranza di non essere scoperti da revisori e ONG. Il greenwashing diventa così una scommessa disperata per evitare una perdita immediata, che finisce però per compromettere la reputazione dell'azienda nel lungo periodo. Mettendo insieme questi quattro meccanismi, razionalità limitata, disimpegno morale, dissonanza cognitiva e avversione alle perdite, si vede come l'opportunismo aziendale nasca sia da un calcolo economico freddo sia dai limiti cognitivi che chiunque, in una posizione di responsabilità, si porta dietro.

Capitolo 3 Systematic literature review (SLR)

Dopo aver inquadrato il contesto teorico, normativo e tecnologico nei capitoli precedenti, il focus della tesi si sposta ora sull'analisi empirica, con una revisione sistematica della letteratura (*systematic literature review - SLR*), per indagare se l'intelligenza artificiale (IA) agisca da *deterrente* o da *agevolatore* del greenwashing nei report di sostenibilità.

Una SLR si articola tipicamente in tre fasi sequenziali (Tranfield et al., 2003):

1. *Pianificazione della revisione*: definizione degli obiettivi e sviluppo del protocollo di ricerca;
2. *Conduzione della revisione*: identificazione, selezione, valutazione e sintesi dei lavori pertinenti;
3. *Rendicontazione e divulgazione*: formulazione di un resoconto descrittivo e tematico degli articoli selezionati.

3.1 Pianificazione della SLR

La prima fase della SLR consiste nella *pianificazione della revisione*, un momento preliminare in cui si traccia il percorso metodologico dell'indagine.

L'obiettivo è costruire un vero e proprio protocollo formale, che guidi la selezione dei testi e limiti i pregiudizi personali (Tranfield et al., 2003).

In questa tesi è stata utilizzata la revisione sistematica della letteratura poiché, spesso, gli studi manageriali tradizionali rischiano di risultare poco oggettivi, selezionando solo le fonti che confermano le idee di partenza.

La SLR costituisce l'approccio più adatto per evitare questo bias, poiché si fonda su criteri rigidi di trasparenza, esaustività e tracciabilità, garantendo un processo di selezione degli articoli del tutto verificabile e replicabile (Tranfield et al., 2003).

Inoltre, la SLR si rivela lo strumento più adatto in quanto il legame tra IA e sostenibilità aziendale è un tema recente e con una mole di pubblicazioni in forte crescita.

Permette, quindi, di trattare gli studi già pubblicati come veri e propri dati primari, aiutando a isolare le evidenze più solide e a individuare eventuali lacune teoriche (Kunisch et al., 2023).

Per garantire la trasparenza e l'affidabilità del lavoro, questo studio segue le linee guida internazionali *PRISMA - Preferred Reporting Items for Systematic reviews and Meta-Analyses* (Page et al., 2021; Maraolo, 2021).

Partendo da queste basi metodologiche, sono state formulate tre domande di ricerca (*research questions - RQ*) per guidare l'indagine:

- *RQ1*: In che modo l'IA viene impiegata per individuare anomalie documentali e contrastare il greenwashing nei report di sostenibilità?
- *RQ2*: In che modo l'IA può facilitare la manipolazione semantica, agevolando il greenwashing nei report di sostenibilità?
- *RQ3*: In che modo l'IA, unita al controllo umano e istituzionale, può limitare i rischi algoritmici e preservare la credibilità dei report di sostenibilità?

3.2 Conduzione della SLR

La seconda fase della SLR riguarda la *conduzione* vera e propria della revisione, ovvero l'identificazione, la selezione, la valutazione e la sintesi dei lavori pertinenti (Tranfield et al., 2003). Per garantire trasparenza e replicabilità, il processo è stato guidato da un insieme di criteri di inclusione ed esclusione definiti a priori, in linea con le indicazioni dello standard PRISMA 2020 (Page et al., 2021):

Identificazione delle fonti: la ricerca è stata condotta esclusivamente sul database Scopus, scelto per l'ampiezza della copertura disciplinare nell'area economico-aziendale e per l'elevato standard qualitativo delle riviste indicizzate (Tranfield et al., 2003).

La stringa di ricerca booleana è stata applicata ai campi titolo, abstract e parole chiave (TITLE-ABS-KEY) ed è stata costruita per mantenere il focus sui due elementi centrali della tesi: l'*intelligenza artificiale*, intesa in senso ampio, e il *greenwashing*, comprensivo delle sue principali varianti terminologiche (cheap talk, cherry-picking, ESG-washing).

Le parole chiave inserite sono: "artificial intelligence", "machine learning", "deep learning", "natural language processing", "large language model*", "LLM*" AND "greenwash*", "ESG-washing", "ESG washing", "cheap talk", "cherry-pick*".

Attraverso queste parole chiave, sono stati individuati 387 articoli.

Selezione e screening: ai 387 risultati sono stati applicati una serie di *filtri* per delimitare il campo di indagine.

- la *tipologia di documento* è stata limitata agli articoli (*article*), sia empirici che concettuali/teorici, in linea con l'approccio di Tranfield et al. (2003), ai contributi di rassegna (*review*) e ai capitoli di libro (*book chapter*);
- la ricerca è stata limitata agli studi pubblicati nell'*arco temporale* 2020–2026³⁴;
- sono stati considerati solo gli studi in *lingua* inglese;
- le *aree disciplinari* incluse sono: *Economics, Econometrics and Finance - Business, Management and Accounting - Social Sciences - Environmental Science - Computer Science - Energy - Decision Sciences - Engineering - Arts and Humanities - Mathematics - Multidisciplinary - Psychology - Earth and Planetary Sciences - Medicine - Chemical Engineering - Physics and Astronomy - Materials Science - Chemistry - Biochemistry, Genetics and Molecular Biology - Agricultural and Biological Sciences.*

L'applicazione congiunta di questi quattro filtri ha ridotto il campione a 301 risultati, di cui 188 articoli, 97 capitoli di libri e 16 rassegne.

È necessario chiarire perché sono state incluse anche le rassegne (*review*).

Il database Scopus non le classifica in modo omogeneo, infatti alcune sono etichettate come "review", mentre la maggior parte è registrata come "article".

Se la ricerca fosse stata limitata ai soli articoli, alcune rassegne sarebbero state escluse in modo del tutto casuale, dipendendo dalle etichette del database e non dai criteri fissati.

Sono stati inclusi anche i capitoli di libro (*book chapter*) poiché, trattandosi di un tema molto recente, limitarsi agli articoli di rivista avrebbe significato perdere i primi studi teorici, che spesso vengono pubblicati all'interno di monografie o volumi collettivi. Inoltre, non è stato applicato alcun filtro all'area disciplinare, includendole tutte. Selezionare solo l'area "Business, Management and Accounting" avrebbe escluso studi importanti pubblicati su riviste di informatica, scienze ambientali o analisi dei dati, fondamentali per la ricerca metodologica.

³⁴ Il 2026 è da intendersi fino alla data di conduzione della ricerca (settembre 2026).

Valutazione della rilevanza: la valutazione della pertinenza degli studi è stata condotta attraverso una scrematura progressiva su tre livelli di lettura, in linea con l'approccio PRISMA (Page et al., 2021).

Nella *prima fase* sono stati letti i titoli dei 301 risultati, escludendo soltanto i contributi palesemente estranei, ossia quelli che trattano il greenwashing senza alcuna componente di intelligenza artificiale e quelli che trattano l'intelligenza artificiale senza alcun riferimento alla rendicontazione o alla comunicazione ambientale d'impresa.

Il campione si è ridotto a 149.

Nella *seconda fase*, sono stati letti gli abstract dei 149 residui applicando un criterio di pertinenza più stringente, riducendo ulteriormente il campione a 89.

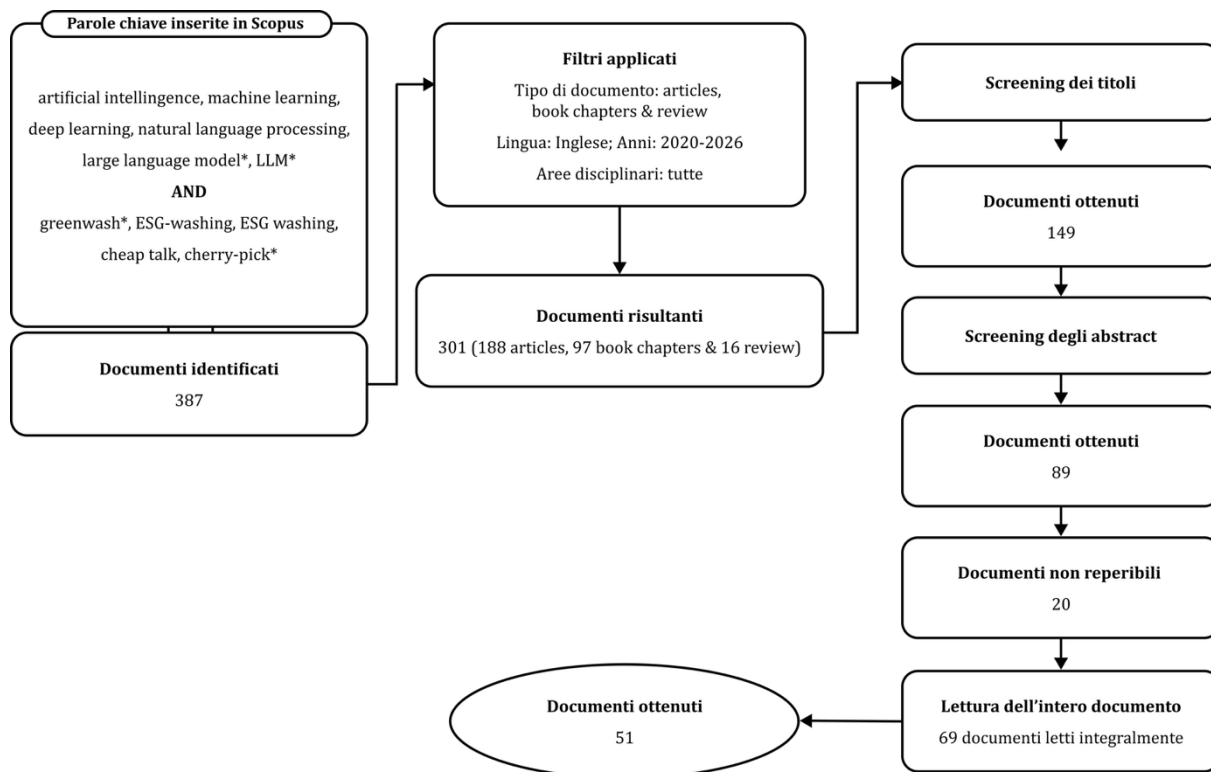
Nella *terza fase*, 20 degli 89 studi rimasti, non sono risultati reperibili in versione integrale e sono stati classificati come “esclusi per irreperibilità”, lasciando 69 testi disponibili alla lettura. Sono stati quindi letti integralmente i 69 studi, valutando principalmente due elementi: research question coperta; risultato rispetto al ruolo di deterrente o agevolatore dell'IA.

La lettura integrale ha condotto alla selezione finale di 51 risultati.

Estrazione dei dati: parallelamente alla lettura, per ciascun articolo è stata compilata una scheda su foglio Excel, con i dati principali: autore/i; anno di pubblicazione; titolo dell'articolo; rivista di pubblicazione; research question coperta; tipologia di IA; IA come deterrente o agevolatore; decisione finale di inclusione o esclusione.

Diagramma di flusso PRISMA: l'intero processo di selezione è sintetizzato nel diagramma di flusso riportato in figura 22 che riproduce, secondo lo standard PRISMA, le fasi di identificazione, screening e valutazione della rilevanza, con l'indicazione del numero di articoli risultante da ciascuna fase.

Figura 22 - diagramma di flusso PRISMA del processo di selezione degli articoli



Fonte: elaborazione personale

3.3 Rendicontazione e divulgazione dei risultati della SLR

La terza fase della SLR riguarda la *rendicontazione e la divulgazione* dei risultati, ovvero la formulazione di un resoconto descrittivo e tematico degli articoli selezionati (Tranfield et al., 2003). La tabella 3 riporta l'elenco completo dei 51 studi inclusi nella revisione, con l'indicazione di autori, anno, sede di pubblicazione, titolo e research question (RQ) a cui ciascun contributo risponde.

Tabella 3 - Elenco dei 51 studi inclusi nella SLR

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
1	Bai C., Yao D. & Xue Q.	2025	<i>Energy Economics</i>	Does artificial intelligence suppress firms' greenwashing behavior? Evidence from robot adoption in China	RQ1/ RQ2

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
2	Bennaceur F., Bendob A. & Othman A.H.A.	2026	<i>Journal of Financial Reporting and Accounting</i>	Leveraging AI for ESG disclosures: a deep dive into UAE's Islamic and conventional banking sectors	RQ1
3	Bergougui B., Sulimany H.G.H. & Alharbi A.A.	2026	<i>Corporate Social Responsibility and Environmental Management</i>	Combating ESG Greenwashing Through AI Models: Evidence From Disaggregated AI Technologies, Mechanisms, and Thresholds	RQ1/ RQ2
4	Bingler J.A. et al.	2022	<i>Finance Research Letters</i>	Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures	RQ1
5	Bingler J.A. et al.	2024	<i>Journal of Banking and Finance</i>	How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk	RQ1
6	Cheng Z. & Yang G.	2026	<i>Socio-Economic Planning Sciences</i>	The impact of artificial intelligence on corporate ESG greenwashing	RQ1/ RQ2
7	Colesanti Senni C. et al.	2025	<i>Environmental Research Communications</i>	Using AI to assess corporate climate transition disclosures	RQ1
8	Davidescu A.A. et al.	2026	<i>Sustainability</i>	Detecting Greenwashing in ESG Disclosure: An NLP-Based Analysis of Central and Eastern European Firms	RQ1
9	de Villiers C., Dimes R. & Molinari M.	2024	<i>Sustainability Accounting, Management and Policy Journal</i>	How will AI text generation and processing impact sustainability reporting? Critical analysis, a conceptual framework and avenues for future research	RQ2

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
10	Dimes R. et al.	2026	<i>Business Strategy and the Environment</i>	The Impact of Generative AI on the Perpetuation and Detection of Greenwashing in Sustainability Reports: Insights From an Experiment	RQ2
11	dos Santos I.A.A. et al.	2025	<i>Corporate Social Responsibility and Environmental Management</i>	Greenwashing in Brazilian Corporations: A Machine Learning Approach to Unmask the Discourse-Practice Gap	RQ1
12	Ferro A., Marazzina D. & Stocco D.	2025	<i>Corporate Social Responsibility and Environmental Management</i>	Uncovering ESG Ratings: The (Im)Balance of Aspirational and Performance Features	RQ1
13	Gorovaia N. & Makrominas M.	2025	<i>European Financial Management</i>	Identifying greenwashing in corporate-social responsibility reports using natural-language processing	RQ1
14	Gu Y. et al.	2026	<i>International Journal of Accounting Information Systems</i>	An LLM-based agentic system for greenwashing detection	RQ1
15	Harke G.	2026	<i>Impact of Market Sentiment on Green Valuations (volume collettaneo)</i>	Algorithmic Accountability in Greenwashing Detection: Hybrid Governance in ESG Auditing	RQ3
16	He L.-Y. & Wang L.	2025	<i>Energy Economics</i>	Can artificial intelligence curb greenwashing? Firm-level evidence based on large language model	RQ1/ RQ2

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
17	Jakubczak J.K., Chmielewska-Muciek D. & Iwanicka K.	2026	<i>Sustainability</i>	Evaluating Multimodal AI for Greenwashing Detection: A Comparative Analysis of ChatGPT, Claude, and Gemini in ESG Reports	RQ1
18	Janik A. & Ryszko A. (<i> rassegna</i>)	2026	<i>Sustainability</i>	Greenwashing in Sustainability Reporting: A Systematic Literature Review of Strategic Typologies and Content-Analysis-Based Measurement Approaches	RQ1
19	Jiang N. et al.	2026	<i>Socio-Economic Planning Sciences</i>	Responsible use of artificial intelligence and the mitigation of corporate greenwashing: The role of executive oversight	RQ3
20	Ke Z. et al.	2026	<i>International Review of Economics and Finance</i>	Can environmental courts inhibit corporate greenwash? Evidence from heterogeneous LLMs	RQ1
21	Khan M.K. et al.	2024	<i>Business Strategy and the Environment</i>	The automated sustainability auditor: Does artificial intelligence curtail greenwashing behavior in Chinese firms?	RQ1/ RQ2
22	Kılınç Y., İnce M.R. & Badem A.C.	2026	<i>Discover Sustainability</i>	A multi-dimensional textual framework for detecting greenwashing in sustainability reporting	RQ1
23	Lagasio V.	2024	<i>International Review of Financial Analysis</i>	ESG-washing detection in corporate sustainability reports	RQ1
24	Lei X., Kocoglu M. & Nikbakht E.	2026	<i>Finance Research Letters</i>	The announcement effect: How AI adoption reduces corporate ESG greenwashing	RQ1/ RQ2

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
25	Liu J., Yuan Y. & Zhu Z. <i>(rassegna)</i>	2026	<i>Journal of Risk and Financial Management</i>	The Role of Artificial Intelligence in Enhancing ESG Disclosure Quality in Accounting	RQ1/ RQ3
26	Lorkiewicz W. & Wasilewski A.	2026	<i>Applied Soft Computing</i>	A hybrid intelligence approach for assessing packaging sustainability claims in e-commerce	RQ1
27	Lublóy Á., Keresztúri J.L. & Berlinger E. <i>(rassegna)</i>	2025	<i>Journal of Environmental Management</i>	Quantifying firm-level greenwashing: A systematic literature review	RQ1
28	Łudzińska K. <i>(rassegna)</i>	2025	<i>Economics and Environment</i>	The firm-level AI-ESG performance nexus: a rapid literature review and future research agenda	RQ1/ RQ2
29	Moodaley W. & Telukdarie A. <i>(rassegna)</i>	2023	<i>Sustainability</i>	Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review	RQ1
30	Mousavian Anaraki S.A., Croce D. & Basili R. <i>(rassegna)</i>	2025	<i>Sustainable Futures</i>	Large language models for sustainability reporting: A systematic review and research agenda	RQ1
31	Pant P. et al.	2026	<i>Business Strategy and the Environment</i>	Deceptive Tone in Sustainable Supply Chain Disclosure and Its Impact on Firm Performance: The Role of Organizational Structure	RQ1
32	Ren X. et al.	2025	<i>Journal of Accounting Literature</i>	The impact of artificial intelligence on corporate greenwashing: evidence from the Chinese listed firms	RQ1/ RQ2
33	Sari R.C. et al.	2025	<i>Business Strategy & Development</i>	Artificial Intelligence-Based ESG Greenwashing Detection: Road to Net Zero Carbon and Its Impact on Corporate Performance	RQ1

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
34	Song X. et al.	2026	<i>Journal of Business Research</i>	Enhancing global value chain transparency: A multi-agent AI framework for detecting ESG greenwashing with open-source innovations	RQ1
35	Suárez-Giri F., Mayoral Blaya S. & Rodríguez López R.	2026	<i>Global Finance Journal</i>	Greenwashing in biodiversity reporting: The role of financial risk exposure and conservation policy stringency	RQ1
36	Telukdarie A., Nyathi M.H.L. & Fabchi R.J.	2026	<i>Frontiers in Sustainability</i>	Environmental, social, and governance sustainability: an AI-centric approach driving data standardization and automation	RQ1
37	Tian X. & Shi D.	2025	<i>Sustainability</i>	Facilitating or Inhibiting: A Study on the Impact of Artificial Intelligence on Corporate Greenwashing	RQ1/ RQ2
38	Vilanova M.E.M. & Bazanini R.	2026	<i>Cleaner Waste Systems</i>	Digital mirage of sustainability: Algorithmic greenwashing and informational asymmetries in the Brazilian beef supply chain	RQ2
39	Vilchez Olivares P.A. & Astorga De La Cruz B.J. (<i>rassegna</i>)	2026	<i>Sustainability</i>	AI and Data Analytics in Sustainable Financial Reporting and ESG Disclosure: A Systematic Literature Review	RQ1/ RQ3
40	Wang X. et al.	2025	<i>Environmental Sciences Europe</i>	The impact of artificial intelligence on corporate greenwashing: evidence from the New Generation of AI Innovation and Development Pilot Zones Policy	RQ1/ RQ2
41	Wang X., Gao X. & Sun M.	2025	<i>EPJ Data Science</i>	Construction and analysis of corporate greenwashing index: a deep learning approach	RQ1
42	Xi M., Xia R. & Pan W.	2026	<i>Journal of Digital Economy</i>	AI adoption and corporate green behavior: Institutional isomorphism in the algorithmic age	RQ1/ RQ2

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
43	Xing C. et al.	2026	<i>British Accounting Review</i>	The effect of CSR assurance on subsequent corporate greenwashing: Suggestion acquisition or opinion shopping?	RQ3
44	Xu C. et al.	2026	<i>Computational Economics</i>	DeepGreen: Effective LLM-Driven Greenwashing Monitoring System Designed for Empirical Testing — Evidence from China	RQ1
45	Yadav A. & Asongu S.A.	2026	<i>Corporate Social Responsibility and Environmental Management</i>	The AI Sustainability Paradox: How Verification and Regulation Synergize to Curb Greenwashing in Emerging Markets	RQ2/ RQ3
46	Yu X. et al.	2026	<i>Business Strategy and the Environment</i>	"Good Citizen" or "Good Actor": The Asymmetric Impact of Artificial Intelligence on Firms' Tactical Environmental Information Disclosure	RQ1/ RQ2
47	Zhang D.	2024	<i>Energy Economics</i>	The pathway to curb greenwashing in sustainable growth: The role of artificial intelligence	RQ1/ RQ2
48	Zhao Q. et al.	2026	<i>Journal of Cleaner Production</i>	Green or Greenwashing? A novel deep learning method integrating multi-source data for ESG greenwashing detection	RQ1
49	Zheng M. et al.	2026	<i>International Review of Financial Analysis</i>	Can artificial intelligence mitigate greenwashed green credit? Evidence from loan contracts of Chinese listed firms	RQ3
50	Zhou C., Liu L. & Wang H.	2026	<i>Corporate Social Responsibility and Environmental Management</i>	How Government Environmental Targets Shape Firms' Environmental Disclosure Strategy: Evidence From LLM-Based Textual Analysis	RQ1

N	Autori	Anno	Sede di pubblicazione	Titolo	RQ
51	Zhou Z., Gao S. & Wang Z.	2026	<i>Information Sciences</i>	AuraESG: A cognitive architecture for auditable ESG analysis through knowledge-infused causal reasoning and hybrid deliberation	RQ1

Fonte: elaborazione personale

Come mostra la tabella 3, la distribuzione per domanda di ricerca è fortemente sbilanciata: 28 contributi rispondono alla sola RQ1 e altri 13 la affiancano alla RQ2, per un totale di 41 studi che interrogano, in un modo o nell'altro, la capacità dell'intelligenza artificiale di rilevare il greenwashing o di ridurne l'incidenza.

Alla RQ2 rispondono 3 contributi, mentre alla RQ3 altri 4, ai quali si aggiungono 2 lavori che coprono insieme la prima e la terza e 1 che copre la seconda e la terza.

Questo squilibrio indica che la maggior parte della letteratura si è concentrata su come rilevare il greenwashing.

Al contrario, pochissimi studi (quasi tutti degli ultimi due anni) hanno analizzato l'intelligenza artificiale generativa come strumento che facilita il greenwashing, rendendolo più difficile da scoprire.

Infine, la terza area di indagine, relativa alla governance umana e istituzionale dei rischi dell'IA nella rendicontazione, risulta ancora meno esplorata.

Va inoltre segnalato che sei dei contributi sopra elencati sono revisioni sistematiche o analisi bibliometriche. La loro funzione, all'interno della tesi, è di inquadramento e di confronto.

La tabella 4 riporta la tipologia degli autori e delle risorse bibliografiche incluse.

Il campione è quasi interamente accademico e composto da articoli pubblicati su riviste sottoposte a revisione paritaria. L'unica eccezione riguarda il capitolo di Harke (2026), firmato da un ricercatore indipendente.

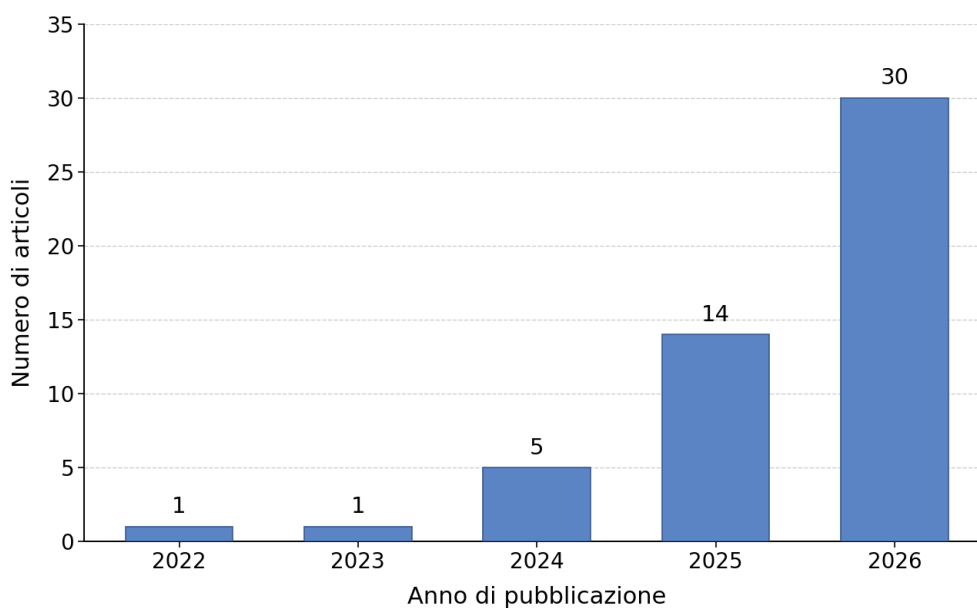
Tabella 4 - Tipologia di autore e tipologia di risorsa degli studi inclusi

Categoria	Variabile	N	%
Tipologia di autore		51	
	Accademico	50	98,0%
	Non accademico	1	2,0%
Tipologia di risorsa		51	
	Articolo su rivista	50	98,0%
	Capitolo di libro	1	2,0%

Fonte: elaborazione personale

La figura 23 mostra che la produzione si concentra in modo netto nel periodo più recente: 1 solo contributo nel 2022 e 1 nel 2023, 5 nel 2024, 14 nel 2025 e 30 nel 2026, mentre il biennio 2020–2021 non compare affatto. L'andamento è coerente con la rapida evoluzione dei modelli di IA applicati alla sostenibilità, e conferma che il tema è in una fase di espansione iniziale.

Figura 23 - Andamento temporale delle pubblicazioni incluse (2022-2026)



Fonte: elaborazione personale

In tabella 5 sono riportati i 51 contributi del campione, distribuiti su 34 sedi diverse (33 riviste e un volume). Solo otto di queste ospitano più di un lavoro, mentre ben 26 sedi compaiono una sola volta. Questa forte dispersione è tipica dei temi di ricerca molto recenti, che non hanno ancora riviste di riferimento consolidate. Sebbene 35 lavori formino un solido nucleo economico-aziendale (management, sostenibilità, finanza e accounting), i restanti 16 provengono da altri ambiti: economia, scienze ambientali e informatica.

Tabella 5 - Distribuzione degli articoli inclusi per rivista di pubblicazione

Sede di pubblicazione	Area disciplinare	N
<i>Applied Soft Computing</i>	Informatica e data science	1
<i>British Accounting Review</i>	Accounting e finanza	1
<i>Business Strategy & Development</i>	Management e sostenibilità d'impresa	1
<i>Business Strategy and the Environment</i>	Management e sostenibilità d'impresa	4
<i>Cleaner Waste Systems</i>	Scienze ambientali	1
<i>Computational Economics</i>	Economia	1
<i>Corporate Social Responsibility and Environmental Management</i>	Management e sostenibilità d'impresa	5
<i>Discover Sustainability</i>	Management e sostenibilità d'impresa	1
<i>Economics and Environment</i>	Economia	1
<i>Energy Economics</i>	Economia	3
<i>Environmental Research Communications</i>	Scienze ambientali	1
<i>Environmental Sciences Europe</i>	Scienze ambientali	1

Sede di pubblicazione	Area disciplinare	N
<i>EPJ Data Science</i>	Informatica e data science	1
<i>European Financial Management</i>	Accounting e finanza	1
<i>Finance Research Letters</i>	Accounting e finanza	2
<i>Frontiers in Sustainability</i>	Management e sostenibilità d'impresa	1
<i>Global Finance Journal</i>	Accounting e finanza	1
<i>Impact of Market Sentiment on Green Valuations</i> (volume collettaneo)	Accounting e finanza	1
<i>Information Sciences</i>	Informatica e data science	1
<i>International Journal of Accounting Information Systems</i>	Accounting e finanza	1
<i>International Review of Economics and Finance</i>	Accounting e finanza	1
<i>International Review of Financial Analysis</i>	Accounting e finanza	2
<i>Journal of Accounting Literature</i>	Accounting e finanza	1
<i>Journal of Banking and Finance</i>	Accounting e finanza	1
<i>Journal of Business Research</i>	Management e sostenibilità d'impresa	1
<i>Journal of Cleaner Production</i>	Scienze ambientali	1
<i>Journal of Digital Economy</i>	Economia	1
<i>Journal of Environmental Management</i>	Scienze ambientali	1
<i>Journal of Financial Reporting and Accounting</i>	Accounting e finanza	1

Sede di pubblicazione	Area disciplinare	N
<i>Journal of Risk and Financial Management</i>	Accounting e finanza	1
<i>Socio-Economic Planning Sciences</i>	Economia	2
<i>Sustainability</i>	Management e sostenibilità d'impresa	6
<i>Sustainability Accounting, Management and Policy Journal</i>	Accounting e finanza	1
<i>Sustainable Futures</i>	Management e sostenibilità d'impresa	1
Totale		51

Fonte: elaborazione personale

La tabella 6 evidenzia una distinzione fondamentale: l'insieme degli studi si divide in due grandi filoni. Quasi la metà degli studi usa l'IA come *strumento* per analizzare i report e scoprire il greenwashing. Un altro terzo, invece, analizza l'IA come *oggetto di studio*, misurando se le aziende che adottano queste tecnologie facciano meno greenwashing. Infine, le sette revisioni sistematiche (indicate nella tabella 3) costituiscono una categoria a sé. Essendo fonti secondarie, sono servite per inquadrare il contesto e confrontare i risultati.

Tabella 6 - Natura metodologica dei contributi inclusi

Categoria	N	%
IA come strumento di analisi testuale della rendicontazione	23	45,1%
— NLP tradizionale e tecniche lessicali e statistiche	7	13,7%
— Modelli linguistici pre-addestrati (famiglia BERT e transformer)	6	11,8%
— LLM generativi e deep learning	8	15,7%

Categoria	N	%
— Sistemi agentici e architetture multi-agente	2	3,9%
IA come variabile esplicativa in studi econometrici	15	29,4%
Revisioni sistematiche e analisi bibliometriche	7	13,7%
Machine learning applicato a dati strutturati	2	3,9%
Contributi concettuali	2	3,9%
Studio sperimentale	1	2,0%
Studio qualitativo	1	2,0%

Fonte: elaborazione personale

Capitolo 4 Evidenze dalla letteratura sul ruolo dell'intelligenza artificiale nella rendicontazione ESG

Nel capitolo 3 viene descritto come è stato costruito il campione.

Il presente capitolo, invece, ne discute i contenuti, organizzando le evidenze attorno alle tre RQ, con l'obiettivo di stabilire se l'intelligenza artificiale, applicata ai report di sostenibilità, agisca da *deterrente* o da *agevolatore* del greenwashing. Le tre RQ scompongono, infatti, l'obiettivo e i paragrafi che seguono affrontandole una per volta.

La letteratura attuale non esplora queste domande in modo uniforme, si concentra ampiamente sulla prima, lasciando in ombra la seconda e, soprattutto, la terza. L'estensione dei paragrafi che seguono riflette questo squilibrio. Finora, infatti, gli studi si sono focalizzati soprattutto sulla capacità dei modelli di riconoscere il greenwashing e sugli effetti della loro adozione, trascurando il modo in cui questa stessa tecnologia possa essere usata per produrlo.

4.1 Premessa

Prima di analizzare i risultati della SLR, è opportuno chiarire e richiamare le diverse tecniche di IA utilizzate negli studi che fanno parte del campione.

Il *Natural Language Processing* (NLP) è il ramo dell'intelligenza artificiale che elabora testi scritti in lingua corrente anziché in codice. Nella sua forma tradizionale l'analisi non interpreta il testo, ma lo converte in numeri e ne misura alcune proprietà (Lagasio, 2024). La *prima proprietà* è il peso dei singoli termini, di cui la tecnica più diffusa è la ponderazione *Term Frequency – Inverse Document Frequency* (TF-IDF): a ogni parola viene assegnato un punteggio tanto più alto, quanto più essa è frequente nel documento esaminato e rara nell'insieme degli altri documenti. I termini che compaiono spesso pesano poco, quelli distintivi pesano molto (Lagasio, 2024).

La *seconda proprietà* sono gli argomenti trattati. Le tecniche di *topic modeling* raggruppano le parole che tendono a comparire insieme e ricavano da questi raggruppamenti i temi ricorrenti del testo: temi che il documento non dichiara, ma che emergono dalla sua struttura statistica. L'algoritmo più usato è *Latent Dirichlet Allocation* (LDA) (Lagasio, 2024).

La *terza proprietà* è il tono. La *sentiment analysis* attribuisce al testo un punteggio che ne misura l'orientamento positivo o negativo.

Dalla combinazione di queste misure si ricava un “indicatore di scarto”, che andrà quindi a costituire l'eventuale segnale di ESG-washing.

Il limite dell'intera famiglia è che queste tecniche riconoscono le parole, non il contesto in cui compaiono (Xu et al., 2026). Per un dizionario, l'impresa che afferma di aver ridotto le proprie emissioni e quella che afferma di volerle ridurre producono il medesimo segnale. La generazione successiva di strumenti supera questo limite ricorrendo alle *reti neurali profonde*, sistemi che non applicano regole scritte da un programmatore, ma apprendono dagli esempi attraverso molti strati di elaborazione successivi (Zhao et al., 2026). Il loro addestramento avviene in due fasi: quella di pre-addestramento, in cui il modello impara la lingua in generale, su grandi quantità di testo di qualunque provenienza; successivamente, viene specializzato su un dominio ristretto, mediante un insieme di esempi classificati a mano, operazione detta *fine-tuning*.

I modelli più usati in questa letteratura appartengono alla *famiglia Bidirectional Encoder Representations from Transformers (BERT)*, dal nome del primo di essi. Il loro tratto caratteristico è la lettura bidirezionale: il significato di ciascuna parola viene determinato tenendo conto sia di quelle che la precedono, sia di quelle che la seguono, e l'unità di analisi diventa la frase anziché il termine isolato. È proprio la considerazione del contesto a consentire lo scioglimento delle ambiguità lessicali e l'estrazione di informazioni dettagliate, cosa che i metodi fondati sui dizionari non permettono (Wang et al., 2025a).

I *Large Language Model (LLM)*, rappresentano l'evoluzione più recente. Si tratta di reti neurali costruite sull'architettura *transformer*, che valuta la relazione di ogni parola con tutte le altre del testo, e pre-addestrate su quantità di testo enormemente maggiori (Mousavian Anaraki et al., 2025).

A differenza dei modelli precedenti non si limitano a classificare, bensì comprendono, generano linguaggio e colgono relazioni contestuali profonde all'interno del testo (Mousavian Anaraki et al., 2025). Sono gli strumenti su cui si fondano i sistemi commerciali di uso comune, come ChatGPT, Claude o Gemini.

È necessario citare due nozioni legate a questi modelli. La prima è l'allucinazione, cioè la produzione di un'affermazione plausibile ma non supportata dai dati (Colesanti Senni et al., 2025). In un'attività di verifica è un rischio serio, perché l'errore non si riconosce come tale (Dimes et al., 2026).

La seconda è la *Retrieval-Augmented Generation (RAG)*. Con questa tecnica il modello, prima di rispondere, recupera da fonti esterne i documenti pertinenti e basa la risposta

su di essi, anziché su ciò che ha appreso durante l'addestramento (Colesanti Senni et al., 2025).

L'ultimo sviluppo è costituito dai *sistemi agentici*. Un agente non si limita a rispondere a una singola richiesta: pianifica una sequenza di operazioni, la esegue e la adatta ai risultati che ottiene lungo il percorso (Gu et al., 2026). Nelle architetture multi-agente il compito viene suddiviso fra più modelli, ciascuno con un ruolo assegnato, i cui esiti parziali vengono ricomposti attraverso uno scambio interno al sistema (Song et al., 2026).

Applicate ai report di sostenibilità, queste tecniche intervengono su tre piani, che corrispondono alle tre RQ.

Il primo è l'impiego dell'intelligenza artificiale come strumento di analisi e di verifica, per individuare i segnali di greenwashing e per contrastarlo (RQ1).

Il secondo è l'impiego dell'intelligenza artificiale generativa da parte delle imprese, per redigere o rifinire i propri report, con l'effetto di ampliare lo spazio del greenwashing (RQ2).

Il terzo è il modo in cui la supervisione umana, aziendale e istituzionale si combina con la tecnologia per limitare i rischi che essa stessa introduce (RQ3).

4.2 L'IA come strumento di rilevazione e contrasto del greenwashing

La RQ1 contiene due verbi, *individuare* e *contrastare*, e a ciascuno corrisponde una letteratura con oggetto e metodo propri.

La prima impiega l'intelligenza artificiale come strumento di analisi. I report di sostenibilità vengono esaminati da un soggetto esterno all'impresa, di norma il ricercatore, per individuare le incoerenze testuali e stimare il rischio di comunicazione ingannevole (Lagasio, 2024; Gu et al., 2026).

La seconda cambia oggetto. Questo filone misura se l'adozione dell'intelligenza artificiale da parte delle imprese ne modifichi il comportamento dichiarativo e riduca le pratiche opportunistiche (Zhang, 2024; Tian e Shi, 2025; Cheng e Yang, 2026).

La distinzione non è solo formale, in quanto dimostrare che una tecnologia sa riconoscere il greenwashing è cosa diversa dal dimostrare che ne riduce l'incidenza. Nel primo caso conta l'efficacia degli strumenti di rilevazione, mentre nel secondo, contano gli incentivi dell'impresa, i controlli interni e il contesto istituzionale in cui opera.

Nei paragrafi successivi vengono presentate e analizzate le tecniche impiegate: dai metodi tradizionali di tipo lessicale e statistico ai modelli pre-addestrati, dai modelli generativi e di deep learning ai sistemi agentici.

4.2.1 L'NLP tradizionale

Le prime applicazioni dell'intelligenza artificiale alla rilevazione del greenwashing impiegano le tecniche tradizionali di *NLP*, che analizzano un testo contando o ponderando le parole che contiene. L'assunto comune è che l'opportunismo comunicativo lasci tracce misurabili nella superficie lessicale dei report. Queste tracce possono essere confrontate con altre misure ricavate dallo stesso testo oppure con riferimenti esterni, come la copertura dei media o le sanzioni ambientali (Davidescu et al., 2026; Gorovaia e Makrominas, 2025).

Gli studi raccolti in questo sottoparagrafo sono tutti lavori empirici, condotti su documenti realmente pubblicati. Un primo gruppo costruisce indicatori di scarto, che misurano la distanza fra quanto un'impresa dichiara e un termine di confronto.

Altri verificano l'affidabilità di questi strumenti o ne collegano i risultati agli esiti economici.

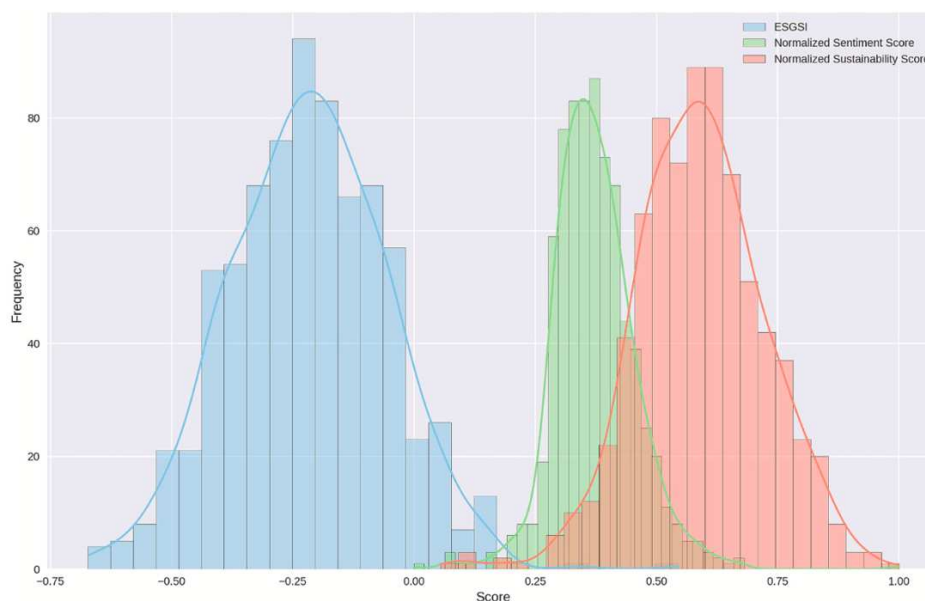
Un primo esempio è l'*ESG-washing Severity Index (ESGSI)* proposto da Lagasio (2024) e applicato ai report di sostenibilità pubblicati nel 2023 da 749 società quotate di diversi paesi, e confronta il modo in cui un report è scritto con quanto il report dice.

Il tono (*sentiment score*) viene misurato con la libreria *TextBlob*; la sostanza (*sustainability score*) con il TF-IDF, a partire dai termini che corrispondono agli indicatori degli standard GRI e SASB. Entrambi i punteggi sono riportati su una scala da 0 a 1 e l'indice è la loro differenza (vedi figura 24).

Quando il tono positivo supera la sostanza l'indice cresce, e valori alti segnalano quindi un rischio maggiore di ESG-washing. All'indice si affianca un'analisi LDA dei temi ricorrenti, che però non entra nel calcolo.

I valori medi dell'indice risultano negativi in tutti i settori: in media il tono dei report è più misurato della sostanza. L'autrice riconosce che i dati si riferiscono a un solo anno e non consentono inferenze causali, e indica la necessità di studi longitudinali.

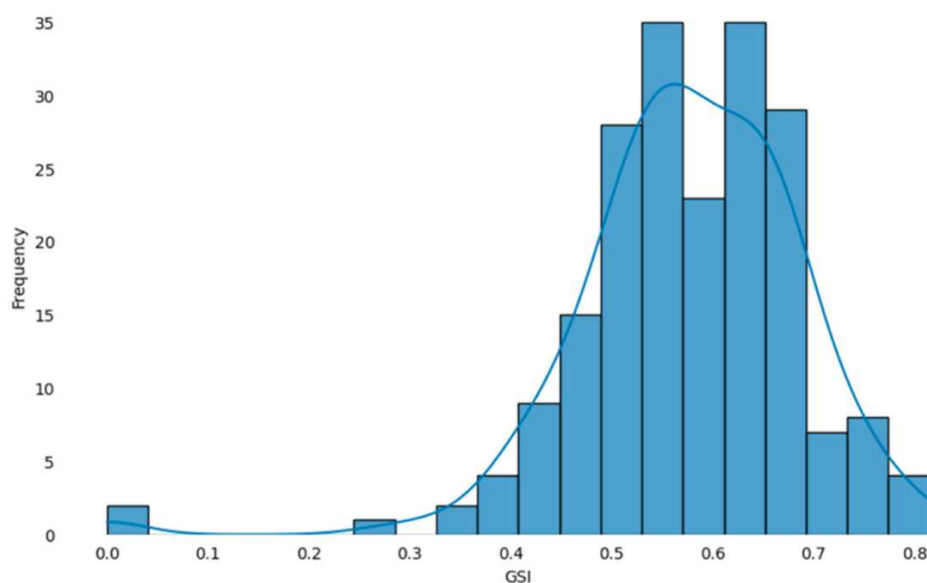
Figura 24 - Distribuzione dell'ESGSI: l'indice (in blu) come differenza tra punteggio di tono (in verde) e punteggio di sostanza (in rosso)



Fonte: Lagasio, 2024

Davidescu et al. (2026) affrontano lo stesso problema da un'angolatura diversa, confrontando il report con ciò che i giornali scrivono di quell'impresa. Il campione comprende 204 grandi imprese dell'Europa centro-orientale e riguarda i report pubblicati nel 2023. L'analisi procede su due livelli. Il primo è il *Greenwashing Severity Index (GSI)*, riportato in figura 25, calcolato sul solo testo dei report.

Figura 25 - Distribuzione dei valori del GSI



Fonte: Davidescu et al., 2026

Gli autori costruiscono tre dizionari, uno per ciascuna dimensione ESG. Con il TF-IDF misurano quanto spazio il documento dedica ai termini di ciascun dizionario e attribuiscono punteggi normalizzati tra 0 e 1. Un valore alto indica un racconto costruito quasi per intero su un linguaggio di sostenibilità, e segnala quindi una maggiore probabilità di greenwashing.

Il secondo livello presuppone il confronto con la stampa. Per ogni impresa il tono del report viene messo a confronto con quello degli articoli che la riguardano, e la correlazione di Pearson³⁵ tra i due misura quanto l'autorappresentazione aziendale si allinei con la percezione esterna, dove valori vicini a 0 indicano dissonanza. I due indicatori vanno letti insieme: un GSI alto, da solo, non basta a qualificare un settore come poco credibile. Nel complesso i risultati descrivono un greenwashing diffuso, ma di intensità moderata. Gli autori dichiarano tre limiti. L'analisi si basa su soli dati testuali, e misura quindi la coerenza della comunicazione, non la performance ESG effettiva. Le fonti usate sono in lingua inglese, con il rischio di trascurare il dibattito nelle lingue locali. Infine, lo studio osserva un solo anno e non permette di seguire l'evoluzione del fenomeno.

Gorovaia e Makrominas (2025) mettono il testo a confronto con un dato di condotta. Il materiale è costituito dai report di responsabilità sociale d'impresa di società quotate statunitensi pubblicati fra il 2008 e il 2022, per un campione appaiato di 1.120 documenti. L'approccio è a dizionario e si appoggia al lessico ambientale DiCoEnviro, con cui gli autori misurano l'attenzione ai temi ambientali, il tono e la leggibilità dei report. Queste misure vengono poi confrontate con un indicatore binario del comportamento, cioè l'aver ricevuto o meno una sanzione pecuniaria per violazioni ambientali. Le imprese sanzionate pubblicano report più lunghi e più positivi, con un contenuto ambientale più marcato ma meno leggibile. Nel loro insieme, questi tre studi non valgono tanto per i singoli risultati. Contano perché mostrano che la distanza fra dichiarazioni e comportamenti si può rilevare anche con strumenti semplici.

L'approccio di dos Santos et al. (2025), invece, è confrontare il report, oltre che con altri testi, con un dato oggettivo. La domanda di partenza è se le dichiarazioni di sostenibilità delle imprese brasiliane corrispondano a ciò che le imprese fanno davvero, e il quadro

³⁵ Indice statistico che misura l'intensità della relazione lineare fra due variabili. Assume valori compresi fra -1 e +1: valori vicini a +1 indicano che le variabili crescono insieme, valori vicini a -1 che si muovono in direzione opposta, valori vicini a 0 l'assenza di relazione lineare.

teorico di riferimento è quello del *decoupling*³⁶. Gli autori partono dalle 100 società con la maggiore capitalizzazione quotate alla borsa brasiliana B3, di cui 90 pubblicano un report di sostenibilità, 74 dichiarano le proprie emissioni di gas serra e 40 sono valutate dall'indice di reputazione Merco³⁷.

Su questa base, riferita agli esercizi 2022 e 2023, incrociano tre fonti: il *discorso interno dell'impresa*, misurato sui report; la *percezione esterna*, rappresentata dal punteggio Merco; e la *pratica effettiva*, cioè le emissioni (vedi tabella 7).

Tabella 7 - le tre variabili chiave

Variable name	Acronym	Description	Scale/range	Data source
ESG score	ESG	AI-generated composite score measuring environmental, social and governance performance based on corporate disclosures	0–1 (continuous)	Proprietary AI algorithm analysing sustainability reports
Greenhouse gas emissions	GHG	Total carbon dioxide equivalent emissions reported by the company	Metric tons (tCO ₂ e)	B3 stock exchange filings and Brazilian National Emissions Registry
Merco ESG responsibility index	Merco	Comprehensive evaluation of corporate ESG performance and reputation	0–10,000 (ordinal)	Merco Corporate Reputation Rankings (https://www.merco.info)

Fonte: dos Santos et al., 2025

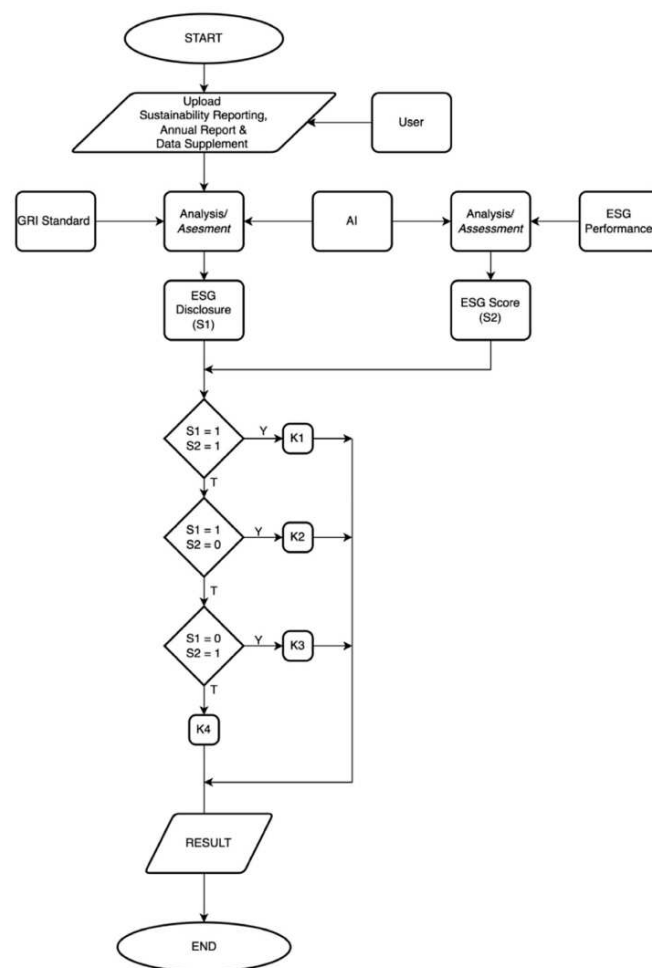
Gli autori costruiscono un dizionario ESG in portoghese, calcolando per ogni report quanto spesso ricorrono i termini di ciascuna delle tre categorie, rapportandoli al totale delle parole del documento. I punteggi così ottenuti vengono poi messi in relazione, prima con la reputazione esterna e poi con le emissioni. Il primo confronto non dà risultati significativi. Quanto un'impresa scrive di sostenibilità, quindi, non ha nulla a che vedere con la reputazione che le riconoscono dall'esterno. Il secondo confronto restituisce un risultato più netto e di segno opposto a quello atteso. A punteggi ESG più alti corrispondono emissioni maggiori. Sono quindi proprio le imprese che parlano di più di ambiente a inquinare di più. Per gli autori questo indica che la rendicontazione ambientale funziona come strumento di gestione dell'immagine più che come riflesso di riduzioni effettive delle emissioni. Il limite che riconoscono è il metodo: contare le parole non permette di cogliere il contesto in cui compaiono, e per questo presentano il proprio lavoro come una prova di concetto e i risultati come indizio di una probabilità di

³⁶ Vedi paragrafo 2.1.2.
³⁷ Il *Merco ESG* è un indice di reputazione aziendale che misura la percezione esterna delle performance ESG di un'impresa.

greenwashing, non come dimostrazione. Segnalano inoltre l'assenza di variabili di controllo, come dimensione, settore e performance finanziaria, e la necessità di campioni più ampi.

Sari et al. (2025) si pongono l'obiettivo di verificare se l'*AI-based greenwashing detection (AI-GW)*, uno strumento basato sull'intelligenza artificiale e sviluppato dagli stessi autori in un lavoro precedente, offra risultati attendibili quanto i tradizionali database commerciali, come Thomson Reuters ASSET4 e i dati ESG di Bloomberg. La verifica passa da due domande: se i punteggi di greenwashing prodotti da AI-GW siano diversi da quelli dei database, e se le due fonti portino alle stesse conclusioni sull'effetto del greenwashing sui risultati aziendali. Il modello AI-GW procede per confronto. Dai report annuali e di sostenibilità ricava un punteggio di *comunicazione ambientale (S1)*, che misura la conformità agli standard GRI, e uno di *performance effettiva (S2)*, costruito su indicatori di risultato. Se il primo supera il secondo, l'impresa viene classificata come greenwasher (=K2, vedi figura 25).

Figura 26 - Framework AI- GW



Fonte: Sari et al., 2025

Un approccio analogo, che distingue tra impegni dichiarati e risultati misurati, è alla base dei database commerciali scelti come termine di paragone.

Il campione di riferimento comprende 174 imprese indonesiane quotate che, tra il 2017 e il 2022, hanno pubblicato un'informativa ESG completa.

Per rispondere alla prima domanda, gli autori confrontano i valori medi prodotti dall'IA con quelli dei database. La differenza non è statisticamente significativa, le due misurazioni si muovono nella stessa direzione, pur mostrando una correlazione debole (0,235). Come sottolineano gli autori, un accordo così basso non è anomalo nel settore: studi precedenti evidenziano che persino tra sei note agenzie di rating la concordanza media si ferma a 0,54.

Passando alla seconda domanda, lo studio stima l'impatto del greenwashing sulle performance aziendali e confronta l'indice basato sull'AI-GW con i punteggi estratti dai database tradizionali. Il modello tiene conto anche dei risultati storici delle imprese. La differenza che emerge è netta: se il greenwashing viene misurato con l'IA, l'effetto sul *ROE* (*return on equity*) risulta sempre negativo e statisticamente robusto, anche inserendo variabili di controllo come dimensioni, indebitamento e liquidità.

Al contrario, utilizzando i dati tradizionali, questo effetto svanisce all'inserimento dei controlli. In sintesi, sebbene entrambe le fonti suggeriscano che il greenwashing penalizzi i risultati economici, solo la misurazione basata sull'IA si dimostra solida in tutti gli scenari analizzati.

Tra i limiti: la scarsa disponibilità di dati ESG completi in Indonesia, che riduce il campione, e i tempi di elaborazione, circa un'ora per ogni impresa, che rendono difficile usare lo strumento su larga scala. Le stesse tecniche sono state applicate anche a contesti settoriali specifici.

Bennaceur et al. (2026) hanno analizzato i report ESG delle banche degli Emirati Arabi Uniti (islamiche e convenzionali) tra il 2021 e il 2024. Utilizzando la *sentiment analysis* e il *topic modeling*, lo studio evidenzia un tono mediamente positivo, soprattutto nelle banche islamiche. Emerge una chiara tendenza a enfatizzare i risultati favorevoli trascurando le criticità, sebbene gli autori precisino che si tratta di un'indagine esplorativa su un campione ridotto.

Pant et al. (2026) spostano invece l'attenzione sull'impatto economico di queste comunicazioni, analizzando le informative sulla catena di fornitura delle principali aziende indiane (2021-2023). Combinando l'elaborazione del linguaggio naturale (NLP) e l'Intelligenza Artificiale Spiegabile (*Explainable AI*), lo studio dimostra che l'uso di un tono ingannevole penalizza la redditività aziendale (misurata tramite ROA e ROCE). Tuttavia, l'appartenenza a un gruppo industriale attenua questo effetto negativo, grazie ai vantaggi derivanti dalla reputazione condivisa e dai controlli interni. L'unico limite evidenziato è che l'analisi automatica del tono non riesce a catturare del tutto come gli stakeholder percepiscano tali documenti.

Tutti questi studi condividono un limite, che dos Santos et al. (2025) dichiarano espressamente: il conteggio delle parole dipende dalle liste di termini scelte a monte e non tiene conto del contesto in cui le parole compaiono. Per un dizionario basato sulle frequenze, le frasi «abbiamo ridotto le emissioni» e «ridurremo le emissioni» producono lo stesso segnale.

4.2.2 I modelli pre-addestrati

Il loro impiego si articola in due fasi. Nella prima il modello viene pre-addestrato su grandi quantità di testo non etichettato e apprende la struttura della lingua. Nella seconda, chiamata *fine-tuning*, viene addestrato una seconda volta su un numero ridotto di frasi che i ricercatori hanno letto e classificato a mano, così da adattarlo al compito specifico (Bingler et al., 2022). È il contesto a permettere di sciogliere le ambiguità del linguaggio e di distinguere un impegno futuro da un risultato già conseguito (Wang et al., 2025a).

Gli studi esaminati di seguito applicano questi modelli a contesti molto diversi, e permettono quindi di osservarne il funzionamento in condizioni differenti.

Bingler et al. (2022) si chiedono se l'adesione volontaria alle raccomandazioni della *Task Force on Climate-related Financial Disclosures (TCFD)*³⁸ migliori davvero l'informativa sui rischi climatici delle imprese. Le raccomandazioni, pubblicate nel 2017, chiedono di organizzare l'informativa climatica in quattro categorie. Due riguardano il modo in cui l'impresa si organizza, cioè la *governance* e la *gestione del rischio*, e raccolgono informazioni di carattere qualitativo e procedurale. Le altre due, la *strategia* e le *metriche e obiettivi*, riguardano l'esposizione effettiva dell'impresa al rischio climatico e i risultati

³⁸ La *Task Force on Climate-related Financial Disclosures* è stata istituita nel 2015 dal Financial Stability Board del G20, l'organismo che monitora la stabilità del sistema finanziario internazionale.

che dichiara di voler raggiungere. La distinzione fra i due gruppi è il perno su cui si regge il risultato dello studio.

Per misurare che cosa le imprese scrivano davvero, gli autori costruiscono ClimateBERT, un modello BERT specializzato sul linguaggio climatico dei documenti aziendali.

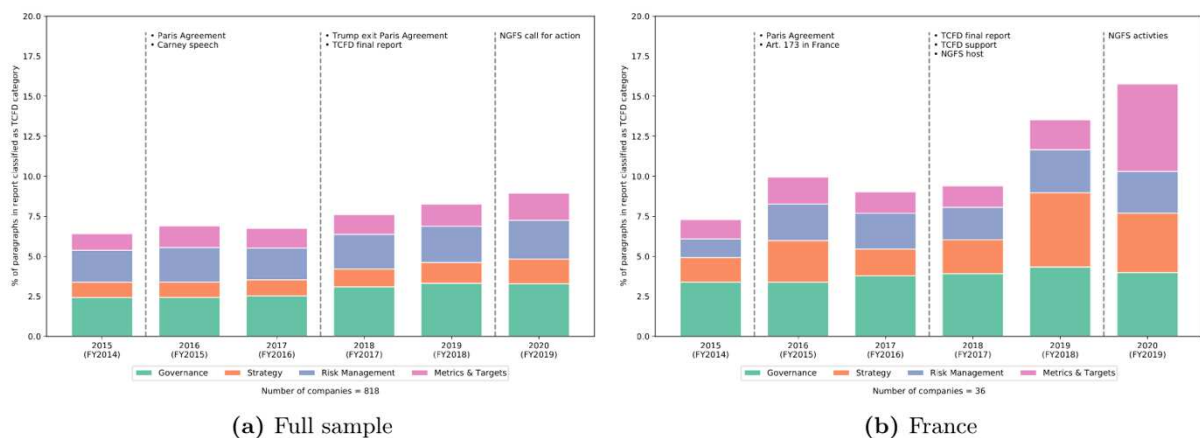
Il compito assegnato al modello è duplice: separare le frasi di contenuto climatico dal resto del documento e attribuire ciascuna di esse a una delle quattro categorie.

L'analisi riguarda 818 imprese che sostengono la TCFD, sugli esercizi dal 2014 al 2019, e comprende relazioni annuali, report di sostenibilità, report climatici o TCFD e pagine web aziendali.

Questo studio fissa il vocabolario con cui l'intera letteratura successiva descrive il fenomeno, introducendo le nozioni di *cheap talk* e *cherry-picking*.

Il *cheap talk* indica che l'adesione dichiarata alla TCFD non si traduce in un aumento effettivo dell'informativa. Il *cherry-picking* indica la scelta di comunicare soprattutto le categorie meno rilevanti sul piano materiale. I due fenomeni sono strettamente collegati: l'informativa cresce sulla *governance* e sulla *gestione del rischio*, mentre la *strategia* e le *metriche e obiettivi*, cioè le categorie che direbbero qualcosa sull'esposizione reale, restano indietro (vedi figura 27).

Figura 27 - Distribuzione dell'informativa climatica fra le quattro categorie TCFD

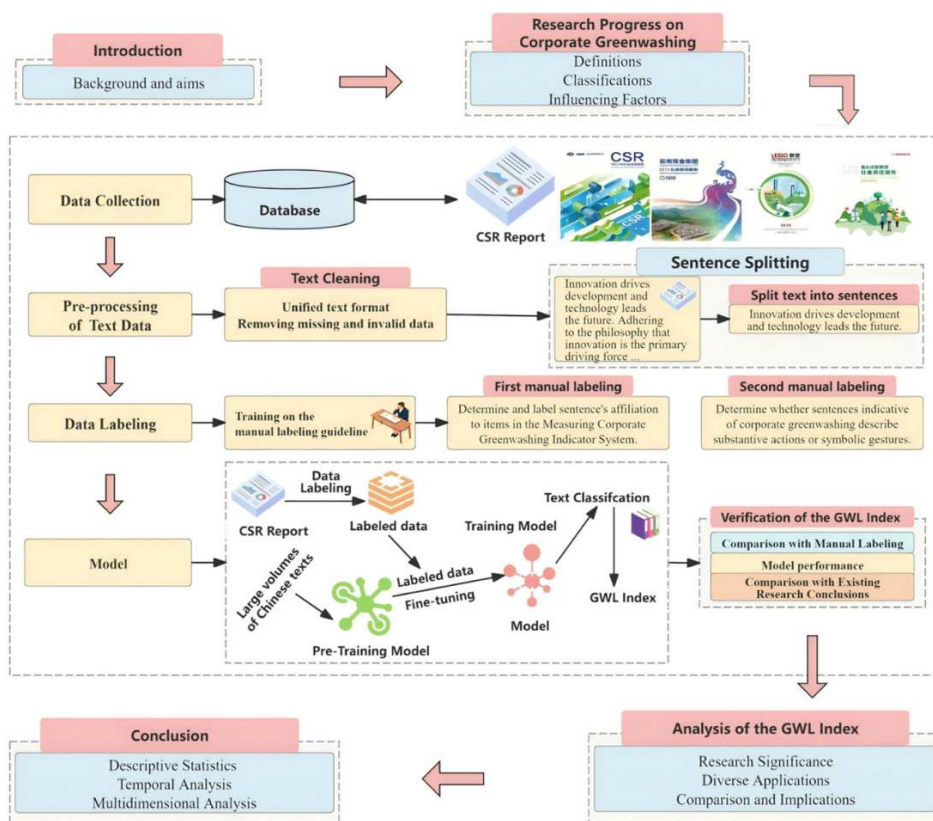


Fonte: Bingler et al., 2022

In un successivo lavoro di validazione, Bingler et al. (2024) mostrano che il *cheap talk* individuato dall'algoritmo non è un artefatto linguistico, perché corrisponde a comportamenti osservabili.

Gli autori costruiscono un indice di *cheap talk* con *ClimateBertCTI*³⁹ e lo applicano a 14.618 relazioni annuali delle imprese dell'indice *MSCI World*⁴⁰, dal 2010 al 2020, che rappresentano una quota rilevante delle emissioni mondiali di anidride carbonica. L'indice viene poi messo in relazione con le iniziative climatiche intraprese, con le emissioni di gas serra e con le notizie ambientali negative. Le imprese con livelli più alti di *cheap talk* registrano una maggiore crescita delle emissioni e una maggiore esposizione a notizie ambientali negative. Solo l'impegno climatico mirato si associa a un *cheap talk* più contenuto, mentre l'informativa volontaria si associa a un *cheap talk* più elevato. La letteratura successiva ha applicato questi modelli a contesti istituzionali diversi. Wang et al. (2025a) costruiscono un indice di greenwashing, il *GWL*, sui report di responsabilità sociale delle imprese cinesi quotate alle borse di Shanghai e Shenzhen (mercato A-share), dal 2008 al 2023, con il modello bidirezionale MacBERT (figura 28).

Figura 28 - Quadro metodologico per la costruzione dell'indice di greenwashing (GWL)



Fonte: Wang et al., 2025a

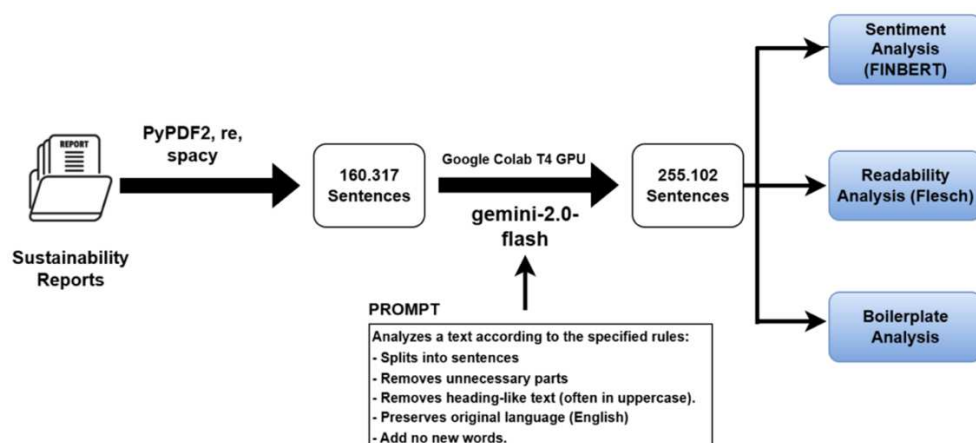
³⁹ Algoritmo di deep learning ottenuto specializzando il modello ClimateBERT su più compiti di classificazione. L'indice di cheap talk (CTI) è la quota di impegni formulati sul totale degli impegni dichiarati.

⁴⁰ L'MSCI World è un indice azionario che raccoglie le società a maggiore e media capitalizzazione quotate nei mercati sviluppati.

L'analisi procede frase per frase e misura due comportamenti distinti: la rivelazione selettiva (*selective disclosure*), cioè la scelta di riportare soltanto gli aspetti ambientali favorevoli, e la manipolazione espressiva (*expressive manipulation*), cioè l'abbellimento simbolico dell'immagine aziendale. Sul piano descrittivo il greenwashing risulta più contenuto nelle imprese di grandi dimensioni e nei settori ad alta intensità inquinante, dove l'obbligo di comunicare i dati sulle emissioni riduce lo spazio per la comunicazione strategica.

Kılınç et al. (2026) esaminano 105 report di sostenibilità delle imprese incluse nel *BIST Sustainability Index* turco⁴¹, pubblicati fra il 2021 e il 2023, per oltre 250.000 frasi complessive. Lo studio combina tre misure testuali (figura 29): leggibilità (*readability*), tono (*sentiment*) e ricorso a formule standardizzate (*boilerplate content*).

Figura 29 - Architettura del modello per l'analisi delle informative di sostenibilità



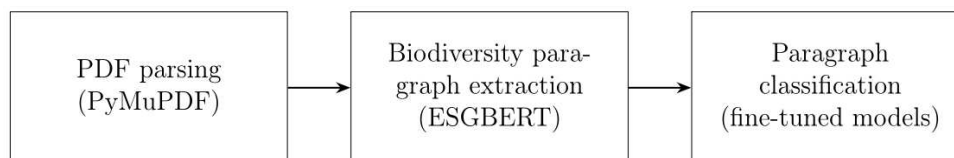
Fonte: Kılınç et al., 2026

La classificazione del tono frase per frase è affidata a *FinBERT*, un modello pre-addestrato su testi finanziari, ed è questa la ragione per cui lo studio appartiene a questa generazione di strumenti. I report risultano nel tempo più leggibili e più positivi, ma in alcuni casi il miglioramento si accompagna a un uso crescente di formule ricorrenti. Un piccolo gruppo di imprese presenta con continuità un profilo di comunicazione ad alto rischio.

⁴¹ Comprende le società quotate che superano soglie prestabilite nella valutazione ESG, condotta dal 2021 dal London Stock Exchange Group su dati pubblicamente disponibili.

Suárez-Giri et al. (2026) estendono l'analisi alla comunicazione sulla biodiversità nei report delle imprese Fortune Global 500⁴², fra il 2017 e il 2023, classificando le frasi con modelli pre-addestrati specializzati sul linguaggio ESG (figura 30).

Figura 30 - Schema del processo di analisi e classificazione del testo dei report



Fonte: Suárez-Giri et al., 2026

Gli autori distinguono due forme di greenwashing: il *cheap talk*, cioè il ricorso a impegni generici, e il *positive talk*, cioè l'enfasi sulle informazioni favorevoli. Il cheap talk si associa all'esposizione dell'impresa al rischio finanziario legato alla natura, il positive talk alla severità delle politiche di tutela della biodiversità nel paese in cui l'impresa ha sede. Entrambe le pratiche si intensificano attorno agli accordi di Kunming-Montreal, il quadro globale per la biodiversità negoziato fra il 2021 e il 2022, ma non nello stesso momento. Il cheap talk arriva per primo, quando gli investitori cominciano a interrogare le imprese sull'esposizione al rischio legato alla natura: un impegno generico non costa nulla, e si concede subito. Il positive talk si rafforza più tardi, quando le aspettative sulla tutela della biodiversità si consolidano e le imprese passano a mettere in evidenza i risultati favorevoli. Gli autori avvertono però di un limite. Le due variabili esplicative principali sono definite una a livello di settore e l'altra a livello di paese, e restano quindi costanti nel tempo per ciascuna impresa. Questo impedisce di attribuire loro un effetto: i risultati vanno letti come associazioni, non come rapporti di causa ed effetto.

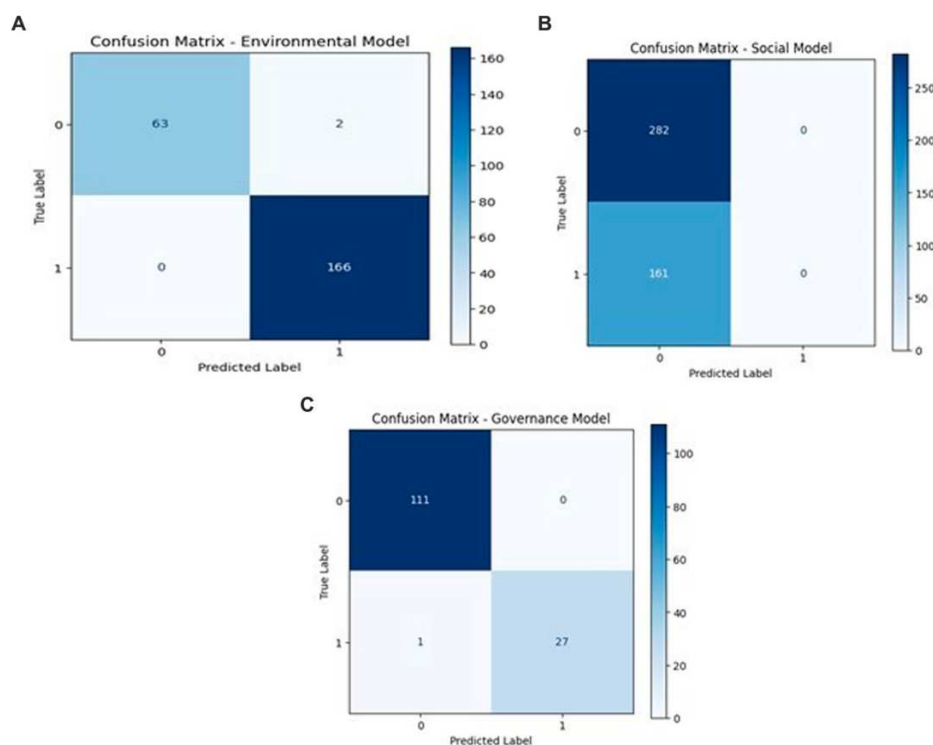
Telukdarie et al. (2026) specializzano un modello transformer, *DistilRoBERTa*, su un compito di classificazione: assegnare ciascun passaggio dei report all'indicatore ESG di cui tratta. L'analisi riguarda i report di 440 imprese quotate alla borsa di Johannesburg e alla borsa nazionale indiana.

L'accuratezza misura quanto spesso il modello assegna il passaggio all'indicatore corretto, rispetto a una classificazione di riferimento stabilita dai ricercatori.

⁴² Il Fortune Global 500 è la classifica annuale delle 500 maggiori imprese mondiali per ricavi, pubblicata dalla rivista *Fortune*.

Qui interessa soprattutto il divario fra i tre pilastri (vedi figura 31): 99,1% sugli indicatori ambientali (E) e 99,3% su quelli di governance (G), contro il 63,6% sul pilastro sociale (S).

Figura 31 - Classificazione per i pilastri ambientale (A), sociale (B) e di governance (C)⁴³



Fonte: Telukdarie et al., 2026

In altri termini, su tre passaggi di contenuto sociale il modello ne classifica correttamente circa due. Gli autori attribuiscono lo scarto alla natura della rendicontazione sociale, che è qualitativa, discorsiva e poco uniforme fra un'impresa e l'altra, mentre emissioni e assetti di governo si esprimono in grandezze e formule ricorrenti. La capacità di rilevazione automatica dipende quindi dal linguaggio a cui viene applicata: dove la rendicontazione è standardizzata il modello funziona, dove è narrativa la sua affidabilità scende.

I modelli esaminati in questo sottoparagrafo leggono il contesto e classificano gli enunciati, ma il confronto resta interno al testo: nessuno di essi verifica se quanto l'impresa afferma trovi riscontro in fonti esterne al documento. È il passo compiuto dagli strumenti esaminati nel sottoparagrafo successivo.

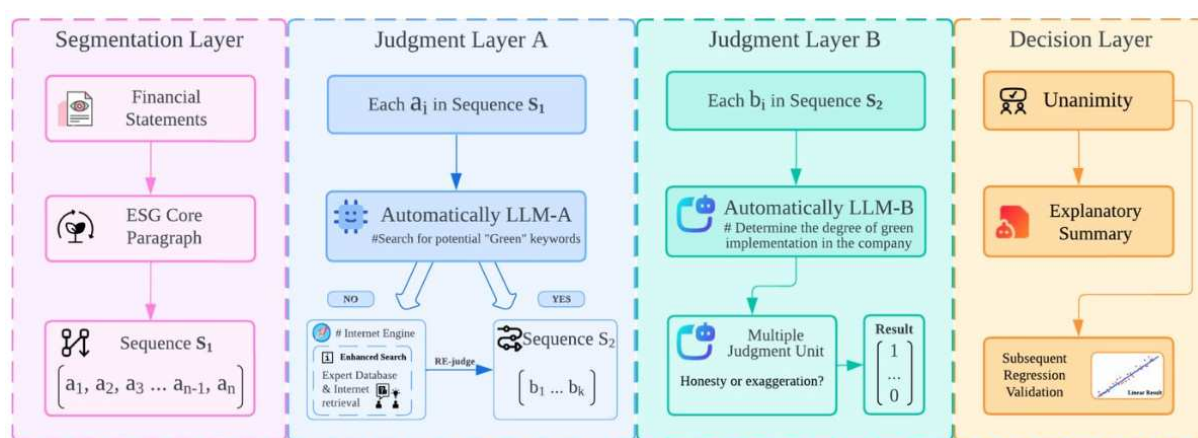
⁴³ La matrice mette a confronto la classificazione prodotta dal modello con quella di riferimento: sulla diagonale si trovano i casi classificati correttamente, fuori dalla diagonale gli errori.

4.2.3 I modelli generativi e il deep learning

Con i modelli generativi l'obiettivo dell'analisi cambia. Le generazioni precedenti classificano gli enunciati o ne pesano i termini, ma restano dentro il testo, confrontando parole con altre parole. I modelli generativi permettono invece di verificare, cioè di mettere a confronto quanto la relazione aziendale afferma con informazioni tratte da fonti esterne al documento.

Il caso più chiaro è *DeepGreen* (figura 32), il sistema di monitoraggio sviluppato da Xu et al. (2026).

Figura 32 - L'architettura a due stadi del sistema DeepGreen.



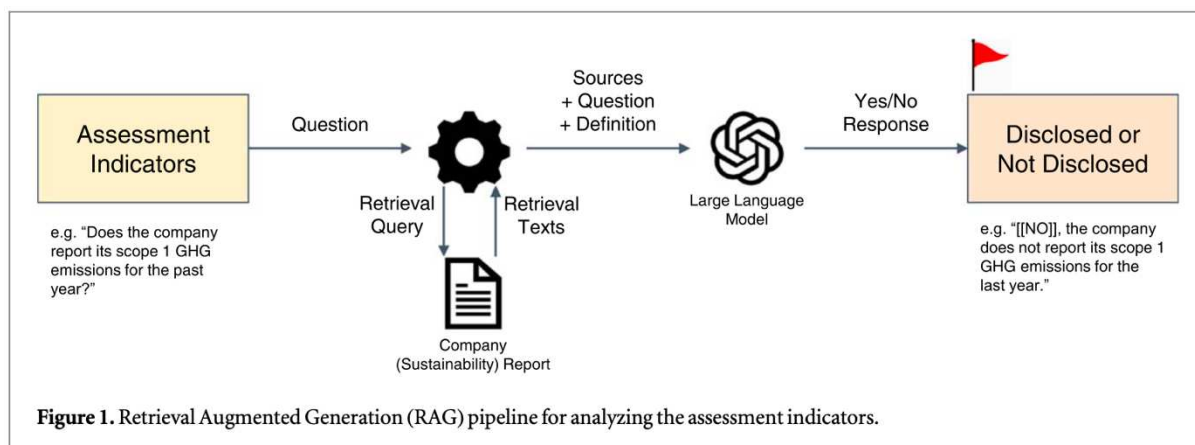
Fonte: Xu et al., 2026

L'architettura procede in due stadi: nel primo il modello individua nel testo i termini e le dichiarazioni di contenuto ambientale, nel secondo verifica se quelle dichiarazioni trovino riscontro in contenuti sostanziali all'interno del documento. Il sistema è applicato a 9.369 relazioni annuali di società cinesi quotate sul mercato A-share, pubblicate fra il 2021 e il 2023, e produce due indicazioni. La prima è tecnica. Ripetendo l'analisi e disattivando una funzione per volta, con quello che gli autori chiamano "esperimento di ablazione", si vede che a fare la differenza è la *Retrieval-Augmented Generation (RAG)*: fornendo al modello documenti di riferimento esterni, riduce le allucinazioni e porta l'accuratezza al 90%. Allungare il testo che il modello elabora in una sola volta, invece, non porta quasi alcun beneficio. La seconda indicazione riguarda l'efficacia predittiva del sistema: le imprese che mostrano alti segnali di greenwashing nei report sono le stesse che, negli anni successivi, ricevono più sanzioni ambientali. Per assicurarsi che questo legame non dipenda da fattori come le dimensioni o il settore dell'azienda, gli autori lo hanno

sottoposto a tre rigorosi test statistici (variabili strumentali, *matching* e test placebo), che ne hanno confermato la solidità. Questo studio rappresenta uno dei rari casi in letteratura capace di collegare ciò che un report dichiara a un esito ambientale concreto.

Colesanti Senni et al. (2025) applicano invece un modello generativo alla valutazione dei piani di transizione verso un'economia a basse emissioni. Gli autori esaminano 28 schemi di rendicontazione esistenti e ne ricavano una batteria di 64 indicatori, sottoposta poi al giudizio di esperti provenienti da 26 istituzioni, fra cui autorità di vigilanza finanziaria, investitori e organizzazioni non governative. Per ciascun indicatore il sistema formula una domanda, recupera dal report i passaggi pertinenti e solo su quelli fonda la risposta, che stabilisce se l'indicatore risulti comunicato o meno (figura 33).

Figura 33 - Il funzionamento della RAG nella valutazione degli indicatori dei piani di transizione



Fonte: Colesanti Senni et al., 2025

Lo strumento è applicato ai report di sostenibilità dell'esercizio 2022 di 143 imprese ad alta intensità di emissioni aderenti all'iniziativa Climate Action 100+⁴⁴. Il risultato è il più netto del campione sul divario fra dichiarazioni (*talk*) e azioni (*walk*). Le imprese comunicano molto sugli indicatori che riguardano la definizione degli obiettivi e le politiche generali, e assai meno su quelli che riguardano l'attuazione concreta delle strategie di decarbonizzazione, come la dismissione degli impianti alimentati da fonti fossili, il collegamento fra remunerazione dei dirigenti e risultati climatici o il ricorso ai crediti di carbonio. Gli autori non discutono limiti espliciti e indicano come sviluppo l'estensione dell'analisi a singoli settori.

⁴⁴ Iniziativa promossa da investitori istituzionali, rivolta alle imprese responsabili delle maggiori emissioni di gas serra a livello mondiale.

Accanto alla verifica dei testi, i modelli generativi vengono impiegati anche in modo strumentale, cioè per costruire variabili che senza di essi resterebbero fuori dalla portata della ricerca empirica.

Zhao et al. (2026) affrontano un problema diverso: nessuna singola fonte basta a riconoscere il greenwashing. Il loro modello, *DLM-ESG*, mette insieme quattro tipi di materiale: i testi dei report ESG, quelli dei bilanci, le linee guida di autoregolamentazione pubblicate dalle borse di Shanghai e Shenzhen e gli indicatori numerici tratti dalle banche dati finanziarie. Un modello generativo estrae da ciascun testo i contenuti pertinenti; un meccanismo di attenzione stabilisce poi quanto pesi ciascuna fonte, riconoscendo che un bilancio e un report ESG non hanno lo stesso valore informativo per questo scopo.

L'analisi riguarda 3.662 società quotate sui due mercati cinesi fra il 2018 e il 2022, di cui 470 classificate come greenwasher sulla base delle sanzioni ambientali ricevute, cioè imprese che esaltano le proprie iniziative ambientali nei report mentre sono oggetto di provvedimenti. Il modello riconosce questi casi meglio dei metodi di riferimento usati finora, e gli esperimenti mostrano che il contributo maggiore viene proprio dall'uso congiunto delle fonti. Gli autori indicano come sviluppo l'arricchimento del sistema di etichettatura con le comunicazioni delle autorità di vigilanza e con le verifiche di terze parti.

Ke et al. (2026) affrontano un problema di misurazione: valutare se i tribunali ambientali cinesi riducano il greenwashing richiede prima di tutto una misura del greenwashing. Gli autori la costruiscono interrogando tre modelli generativi di architettura diversa, DeepSeek-R1, Qwen-3 e OpenPangu-Pro⁴⁵, con la stessa domanda sui passaggi ambientali dei report, e classificando ciascun passaggio secondo la risposta prevalente fra i tre. Il ricorso a più modelli serve a contenere le allucinazioni: quando i tre concordano, il giudizio è più affidabile di quello di un modello singolo. Gli autori precisano che l'indicatore così ottenuto misura la tendenza dell'impresa all'esagerazione e alla retorica eccessivamente positiva, non il greenwashing effettivo.

Zhou C. et al. (2026) impiegano *GPT-4o*⁴⁶ per costruire lo *Specific Environmental Disclosure Index (SEDI)*. Il modello classifica una per una le frasi ambientali estratte dai report, distinguendo quelle *specifiche e verificabili*, che riportano dati, azioni o obiettivi

⁴⁵ Tre modelli linguistici di grandi dimensioni a codice aperto, sviluppati rispettivamente da DeepSeek, Alibaba e Huawei.

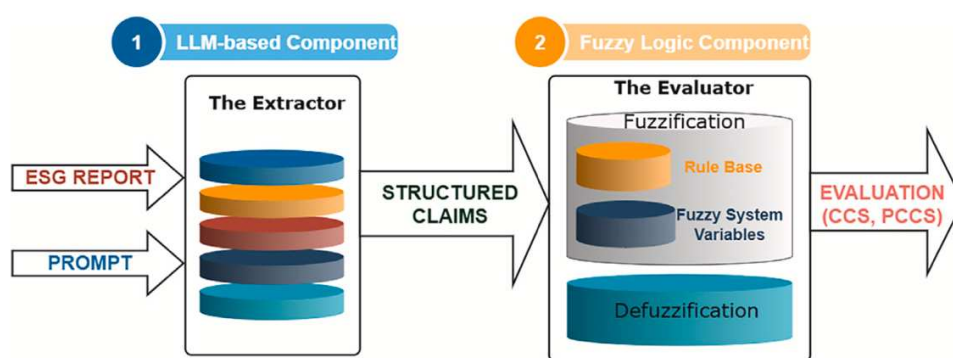
⁴⁶ GPT-4o è un modello linguistico di grandi dimensioni sviluppato da OpenAI e reso disponibile nel 2024.

controllabili, da quelle *vaghe*, che si limitano a dichiarazioni di intenti. L'indice è il rapporto fra le prime e il totale delle frasi ambientali. Lo strumento è applicato a 2.075 osservazioni impresa-anno di società cinesi quotate in settori ad alta intensità inquinante, fra il 2013 e il 2020, per verificare se obiettivi ambientali più severi fissati dalle amministrazioni cittadine inducano le imprese a comunicare in modo più verificabile. In questi tre contributi il modello generativo serve a rendere misurabile una variabile altrimenti inaccessibile.

I risultati sostantivi di Ke et al. (2026) e Zhou C. et al. (2026) riguardano gli effetti del contesto istituzionale, che vengono ripresi in seguito.

Lorkiewicz e Wasilewski (2026) segnano il limite attuale di questi strumenti. Il loro oggetto sono le dichiarazioni sulla gestione degli imballaggi contenute nei report di sostenibilità di 11 grandi imprese del commercio elettronico. Il sistema procede in due passaggi. Nel primo un modello generativo estrae dai report le affermazioni pertinenti, per esempio l'annuncio di una quantità di imballaggi risparmiata. Nel secondo un sistema di inferenza a *fuzzy logic* ne valuta la credibilità, guardando a quanto la dichiarazione sia specifica, a quali prove porti e a quale portata abbia. La fuzzy logic permette di assegnare gradi intermedi invece di un giudizio netto fra vero e falso, e si adatta così all'ambiguità tipica di queste affermazioni. I punteggi confluiscono infine nel *Packaging Circularity Commitment Score (PCCS)*, che rende confrontabili imprese e settori (figura 34).

Figura 34 - Lo schema del sistema ibrido a due passaggi



Fonte: Lorkiewicz e Wasilewski, 2026.

Ai fini di questa revisione, il risultato più rilevante è l'esito della prova a cui gli autori sottopongono il primo passaggio: le affermazioni estratte dal modello vengono confrontate con quelle individuate manualmente da esperti sugli stessi report. Il confronto mostra che quanto il modello individua è affidabile, ma che una parte rilevante

del materiale pertinente gli sfugge. Gli autori lo dichiarano apertamente: «*i modelli linguistici più avanzati estraggono le dichiarazioni con alta precisione, ma il loro richiamo complessivo è limitato*»⁴⁷. Gli stessi autori confrontano inoltre quattro modelli commerciali, Gemini, ChatGPT, Grok e Claude, e rilevano prestazioni diverse fra loro sugli stessi report.

È un'ammissione rara in questa letteratura, e indica dove si fermano gli strumenti esaminati fin qui. In tutti questi studi il lavoro è affidato a un solo modello, che riceve un testo e restituisce una classificazione o un punteggio in un unico passaggio. Quando quel passaggio fallisce, come nel caso del materiale che sfugge all'estrazione, non esiste un secondo controllo. Il sottoparagrafo successivo esamina i sistemi che dividono il compito fra più modelli specializzati, chiamati a cooperare e a verificarsi a vicenda.

4.2.4 I sistemi agentici

Il principio dei sistemi agentici è la divisione del lavoro. Invece di affidare la valutazione di un report a un'unica interrogazione del modello, il compito viene scomposto in operazioni distinte: estrarre le dichiarazioni, cercarne riscontro in fonti esterne, contestarle, ricomporre il giudizio, e ciascuna viene assegnata a un agente specializzato, che opera in sequenza e in dialogo con gli altri (Gu et al., 2026; Song et al., 2026).

Gu et al. (2026) costruiscono un sistema di questo tipo per la rilevazione del rischio di greenwashing. Il sistema valuta l'impresa lungo sette indicatori, che coprono obiettivi ambientali dichiarati, iniziative avviate, risultati conseguiti e attività delle diverse funzioni aziendali. L'elemento distintivo è la *ricerca in rete*: gli agenti confrontano quanto il report afferma con fonti esterne al documento, fra cui i dati della *Environmental Protection Agency (EPA)*⁴⁸ e i casi di greenwashing contestati dalla *Securities and Exchange Commission (SEC)*⁴⁹, oltre alle trascrizioni delle conferenze con gli analisti e alle notizie di stampa. I risultati confluiscono in un cruscotto informativo, pensato per un monitoraggio continuo e non per l'analisi di un singolo esercizio.

⁴⁷ Testo originale: «*While modern LLMs can extract claims with high precision, their comprehensive recall is limited*» (Lorkiewicz e Wasilewski, 2026, p. 14).

⁴⁸ La Environmental Protection Agency (EPA) è l'agenzia federale statunitense competente in materia di tutela ambientale. Fra le sue attività rientrano la raccolta dei dati sulle emissioni delle imprese e l'applicazione della normativa ambientale.

⁴⁹ La Securities and Exchange Commission (SEC) è l'autorità federale statunitense di vigilanza sui mercati finanziari.

Song et al. (2026) seguono una strada diversa. La loro architettura, *Agent-Greenwashing*, si regge su un confronto simulato fra quattro ruoli: un esaminatore ESG, un responsabile della comunicazione aziendale, un giornalista d'inchiesta e un arbitro che decide. Gli agenti argomentano in più tornate, e la contestazione reciproca serve a far emergere le omissioni che una lettura lineare del report non coglie. Il sistema è progettato per funzionare con modelli a codice aperto, e gli autori hanno reso pubblici sia il codice sia l'insieme di dati costruito per l'occasione, il primo del genere. L'accuratezza dichiarata è del 96,67%, ma va letta per quello che è, ossia una misura ottenuta sull'insieme di dati che gli autori stessi hanno costruito e annotato.

Questi risultati vanno confrontati con una verifica di segno opposto.

Jakubczak et al. (2026) sottopongono venti report ESG di imprese polacche a tre sistemi multimodali di uso corrente, ChatGPT, Claude e Gemini, applicando a tutti la stessa istruzione e chiedendo loro di valutare sia il testo sia le immagini. I giudizi vengono confrontati con i punteggi di una classifica nazionale costruita secondo gli standard europei di rendicontazione.

Tabella 8 - Punteggi assegnati da tre modelli multimodali agli stessi report

Company	Benchmark— Total Score	GPT	Claude	Gemini
Amica	40	2	6	7
Bogdanka	78	4	4	8
Castorama	87	4	6	7
CD Projekt	50	4	3	1
Coca-Cola HBC	99	7	4	6
Edge1 Solution	74	6	6	4
Ergo Hestia	98	4	3	7
Grupa Luxmed	71	6	3	7
GS1 Polska	41	3	5	4
Hegelmann Transporte	74	6	7	5
Holcim	82	6	5	4
Ikano Bank	47	4	6	4
mBank	82	4	3	3
Orange Polska	98	6	6	5
PKO BP	49	6	4	3
Polpharma	76	5	5	7
PSE	82	4	2	5
Santander Bank Polska	100	5	4	6
Unimot	75	4	6	7
Żywiec Zdrój	98	6	7	4

Fonte: Jakubczak et al., 2026

La concordanza fra i tre modelli è prossima allo zero⁵⁰: gli stessi report ricevono valutazioni diverse, in alcuni casi opposte, a seconda del sistema impiegato (tabella 8).

⁵⁰ Misurata con il coefficiente α di Krippendorff, che vale 1 quando i giudizi coincidono e 0 quando la concordanza non supera quella attribuibile al caso.

Gli autori attribuiscono il risultato a una confusione di fondo, quella fra cura della comunicazione e intento ingannevole: i modelli penalizzano le imprese che presentano i propri dati in modo curato, scambiando la qualità della veste grafica per un tentativo di abbellimento. Lo studio sceglie inoltre un contesto non anglofono, dove i modelli, addestrati in prevalenza su testi in inglese, sono messi alla prova più duramente.

Lo stesso errore si presenta in una forma diversa nel sistema *AuraESG*, descritto da Zhou Z. et al. (2026), un contributo che appartiene alla letteratura informatica più che a quella aziendale. Gli autori mostrano che i modelli generativi generalisti assegnano punteggi elevati alle dichiarazioni aziendali sulla base dei soli segnali favorevoli presenti nel testo: la formula di impegno, la cifra investita, la percentuale di riduzione annunciata. Quello che non fanno è chiedersi se il meccanismo dichiarato regga. L'esempio riportato è istruttivo: un'impresa attribuisce il calo della propria impronta di carbonio all'acquisto di certificati di energia rinnovabile e alla partecipazione a programmi volontari di compensazione, strumenti di mercato che non riducono le emissioni prodotte direttamente dall'impresa. Un modello generalista premia la dichiarazione, mentre *AuraESG* la segnala.

L'architettura proposta poggia su un grafo della conoscenza, che rende esplicite le relazioni fra le grandezze in gioco. A questo si aggiunge un addestramento per confronto fra scenari: al modello si chiede che cosa sia accaduto e, insieme, che cosa sarebbe accaduto in condizioni diverse. Il sistema ha 117 milioni di parametri e ottiene prestazioni superiori. Il guadagno, concludono gli autori, viene dalla capacità di ragionare sulle relazioni causali e non dalla dimensione del modello.

Nel loro insieme, i due studi indicano lo stesso limite: i sistemi automatici tendono a giudicare la forma della comunicazione anziché la sostanza di ciò che viene dichiarato, e la forma è precisamente ciò che un'impresa può curare a basso costo.

La divisione del compito fra più agenti migliora dunque la rilevazione, ma introduce un problema nuovo. Ogni passaggio automatico che si interpone fra il documento e il giudizio finale, allontana la decisione dal testo che dovrebbe giustificarla, e rende più difficile ricostruire perché il sistema sia arrivato a quella conclusione.

4.2.5 Che cosa misurano davvero le prestazioni dichiarate

Molti degli studi esaminati finora dichiarano un'accuratezza elevata nel classificare i contenuti dei report: il 96,67% di Song et al. (2026), il 99,1% di Telukdarie et al. (2026) sugli indicatori ambientali, il punteggio del 90% di Xu et al. (2026). Sono cifre che vanno

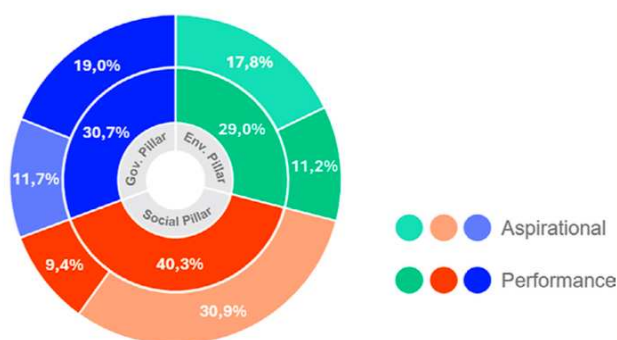
lette per quello che sono. Misurano quanto spesso il modello concorda con una classificazione di riferimento, costruita dai ricercatori stessi o ricavata da una fonte esterna. Dicono quindi se il sistema riproduce un giudizio che qualcuno ha già formulato, non se quel giudizio colga il greenwashing effettivo dell'impresa. Conviene allora chiedersi che cosa quelle percentuali rappresentino. I quattro elementi che seguono non provengono da una critica esterna alla letteratura: sono tutti dichiarati o documentati dagli studi del campione.

Il *primo elemento* riguarda il *termine di paragone*. Molti studi non misurano il greenwashing direttamente, perché non è osservabile: lo stimano come distanza fra ciò che l'impresa dichiara e un punteggio ESG assegnato da un'agenzia di mercato. Se l'impresa comunica molto e il punteggio è basso, la differenza viene letta come segnale di greenwashing.

Ferro et al. (2025) mostrano che questo termine di paragone è meno solido di quanto si supponga. Con tecniche di machine learning gli autori ricostruiscono la metodologia con cui LSEG, l'ex Refinitiv, assegna i punteggi ESG. La banca dati su cui lavorano passa da 2.352 imprese nel 2010 a 11.443 nel 2023. La domanda è quale peso abbiano, nel punteggio finale, gli impegni per il futuro rispetto ai risultati già conseguiti. Circa il 60% del punteggio dipende da elementi prospettici, cioè obiettivi dichiarati, politiche adottate e impegni presi, mentre solo il 40% riflette prestazioni passate. Gli autori parlano di una dipendenza eccessiva dalle promesse (figura 35). Ne segue una difficoltà per tutta la letteratura che usa quel punteggio come riferimento. Un'impresa che annuncia molti impegni ottiene un punteggio alto senza aver ancora fatto nulla, e proprio per questo la differenza fra le sue parole e il suo punteggio risulta contenuta.

Lo strumento che dovrebbe rilevare le promesse non mantenute le premia a sua volta.

Figura 35 - Composizione del punteggio ESG per pilastro



Fonte: Ferro et al., 2025.

Il *secondo elemento* riguarda ciò che il modello non vede. Le prestazioni dichiarate si calcolano quasi sempre sul materiale che il sistema ha estratto dal documento. Nulla dicono su quello che ha lasciato fuori.

È il punto che Lorkiewicz e Wasilewski (2026) rilevano confrontando l'estrazione automatica con il lavoro di annotatori esperti: il modello sbaglia poco su ciò che seleziona, ma seleziona meno di quanto dovrebbe. La conseguenza è seria per questo specifico oggetto di studio. Il greenwashing si manifesta anche in ciò che un report tace. Uno strumento che non individua una parte del materiale pertinente è per definizione cieco rispetto al *selective disclosure*, cioè a una delle forme in cui il fenomeno si manifesta.

Il *terzo elemento* riguarda la *qualità del giudizio*. Zhou Z. et al. (2026), nel presentare AuraESG, mostrano come i *general-purpose LLMs* valutino le dichiarazioni aziendali: il punteggio sale quando il testo è ricco di segnali favorevoli, cioè formule di impegno, cifre di investimento, percentuali di riduzione annunciate. Manca però il passaggio successivo, quello in cui si verifica se la dichiarazione regga. Un calo delle emissioni può derivare da un cambiamento nei processi produttivi oppure dall'acquisto di titoli sul mercato, e per il modello le due cose si equivalgono, perché entrambe si presentano nel testo come un risultato positivo. Il modello valuta dunque la forma della comunicazione anziché la sostanza di ciò che viene affermato. È un problema particolare, perché la forma è precisamente la dimensione su cui un'impresa può intervenire a costo quasi nullo, e quindi quella che uno strumento di rilevazione dovrebbe trattare con più diffidenza.

Il *quarto elemento* riguarda la stabilità del giudizio. Jakubczak et al. (2026) sottopongono gli stessi venti report a tre sistemi di uso corrente. La concordanza che trovano è prossima allo zero: le valutazioni dei tre modelli non si assomigliano più di quanto farebbero tre giudizi assegnati a caso. Se sistemi ugualmente accreditati valutano diversamente lo stesso documento, almeno due di essi si stanno sbagliando, e non è possibile sapere quale sia quello giusto. L'accuratezza che ciascuno dichiara descrive allora il rapporto fra quel modello e la propria base di riferimento, non la credibilità del report analizzato.

Nessuno di questi elementi, preso da solo, invalida la letteratura esaminata: gli studi restano rigorosi e i risultati riproducibili. Considerati insieme, però, indicano che cosa questi strumenti possano fare oggi. Sanno individuare, in un volume di documenti che nessun revisore leggerebbe per intero, quelli che meritano attenzione. Non sanno stabilire se un'impresa stia ingannando i propri interlocutori. Sono, in altri termini, strumenti di segnalazione e non verdetti.

4.2.6 Dall'individuazione al contrasto

Gli studi raccolti in questo sottoparagrafo si distinguono da tutti i precedenti per l'uso che fanno dell'IA. Nei sottoparagrafi precedenti è lo strumento con cui il ricercatore analizza i report; qui è l'impresa ad adottarla, nei propri processi produttivi, gestionali e informativi, e la domanda è che cosa accada alle sue pratiche di rendicontazione.

L'oggetto della misurazione, quindi, passa dalla capacità di riconoscere il greenwashing alla sua incidenza. Cambia anche il tipo di ricerca: invece di costruire strumenti di analisi, questi studi verificano se le imprese che adottano l'IA riducano la distanza fra ciò che dichiarano (*talk*) e ciò che fanno (*walk*). Su tredici studi di questo tipo, dodici concludono che l'adozione dell'IA riduce il greenwashing; uno solo trova l'effetto opposto.

Zhang (2024) analizza le imprese cinesi quotate fra il 2014 e il 2021 e rileva che l'adozione dell'IA migliora la qualità dell'informativa di sostenibilità, riducendo la distanza fra quanto viene comunicato e il punteggio ESG attribuito.

Cheng e Yang (2026) studiano le quotate del mercato A-share fra il 2010 e il 2023 con il *double machine learning*, una tecnica che impiega algoritmi di apprendimento automatico per stimare quanto dell'andamento del greenwashing, sia attribuibile all'utilizzo dell'IA e quanto invece alle altre caratteristiche dell'impresa. Da un lato l'IA, integrata nei sistemi di vigilanza, rende più efficiente il controllo esterno e quindi più rischioso dichiarare il falso. Dall'altro, nelle mani dell'impresa, la stessa tecnologia può facilitare la manipolazione dell'informazione. Sul piano delle risorse gli autori ragionano allo stesso modo: la tecnologia può favorire gli investimenti ambientali reali, ma anche limitarsi a rendere più efficiente l'allocazione del capitale. Nel complesso prevale l'*effetto inibitorio*, che gli autori attribuiscono alla maggiore efficienza della vigilanza e alla riduzione degli investimenti inefficienti.

Bai et al. (2025) ricavano il grado di adozione dell'IA dai dati sull'installazione di robot industriali raccolti dalla *International Federation of Robotics*, riferiti alle quotate cinesi fra il 2011 e il 2019. La scelta si spiega con l'oggetto: nei processi produttivi l'IA si traduce prima di tutto in automazione, e il numero di robot installati è un dato verificabile, che l'impresa non dichiara di sé. Gli autori individuano tre meccanismi, che agiscono in due modi diversi. I primi due riguardano le risorse: l'automazione riduce i costi e accresce i profitti, e migliora la produttività. Un'impresa che dispone di margini più ampi può permettersi interventi ambientali effettivi, e ha meno bisogno di sostituirli con annunci.

Il terzo riguarda invece l'informazione: processi automatizzati generano dati sistematici sul funzionamento dell'impresa, riducono la distanza fra ciò che l'impresa sa di sé e ciò che il mercato può osservare, e restringono così lo spazio in cui la comunicazione può discostarsi dai fatti. La differenza fra i due gruppi si nota già prima delle stime: le imprese con un livello di adozione dell'IA più alto presentano valori di greenwashing inferiori lungo tutto il periodo osservato (Figura 36).

Figura 36 - Andamento greenwashing nelle imprese cinesi per livello di adozione dell'IA, 2011-2019



Fonte: Bai et al., 2025

Khan et al. (2024), su dati 2009–2020, arrivano allo stesso risultato e lo sottopongono a tre controlli: *variabili strumentali*, *propensity score matching (PSM)* e *minimi quadrati a due stadi*, tutte tecniche volte a escludere che siano le imprese già virtuose ad adottare per prime l'IA, e non l'IA a renderle tali.

He e Wang (2025) costruiscono un indice di adozione specializzando un modello linguistico, *ERNIE 3.0*, sul riconoscimento delle dichiarazioni relative all'IA nelle relazioni annuali delle quotate cinesi fra il 2011 e il 2022. L'indice non tratta l'IA come un blocco unico: la suddivide in categorie tecnologiche, dalla visione artificiale al riconoscimento del linguaggio all'apprendimento automatico. Diventa così possibile chiedersi quale componente produca l'effetto. Sono i sistemi di pianificazione e decisione, cioè gli strumenti con cui l'impresa programma le attività e sostiene le scelte gestionali, mentre le altre componenti incidono meno. A ridurre il greenwashing, quindi, sembra essere soprattutto l'impiego dell'IA nei processi che governano le decisioni aziendali.

Ren et al. (2025) affrontano un problema che riguarda tutti questi studi: le imprese che adottano per prime l'IA potrebbero essere anche quelle già più attente all'ambiente, e in

quel caso l'effetto osservato non dipenderebbe dalla tecnologia. Per aggirare la difficoltà gli autori sfruttano i programmi pilota per le città a basse emissioni, con cui il governo cinese ha selezionato alcune città per sperimentare politiche di riduzione delle emissioni. Le imprese che hanno sede in quelle città si trovano così esposte a condizioni diverse per una decisione presa dall'esterno, e non per una propria scelta. Confrontando il loro comportamento con quello delle altre si ottiene una stima più attendibile.

Su dati 2012–2022, gli autori trovano che la diffusione dell'IA riduce il greenwashing, e che la riduzione è maggiore nelle imprese impegnate in attività di innovazione ambientale.

Tian e Shi (2025) misurano l'adozione contando quanto spesso i termini relativi all'IA compaiono nelle relazioni annuali, sul presupposto che un'impresa che ne parla molto la stia effettivamente impiegando. Il loro contributo riguarda il percorso attraverso cui l'effetto si produce. L'innovazione verde (*green innovation*) ne spiega una parte: l'IA favorisce lo sviluppo di soluzioni ambientali effettive come nuovi processi, tecnologie, brevetti in campo ambientale, e sono queste a ridurre il divario fra dichiarazioni e comportamenti. Una parte dell'effetto, però, resta anche tenendo conto di questo canale, e va quindi attribuita ad altro. L'effetto risulta inoltre più forte in due condizioni: quando l'impresa è circondata da concorrenti che adottano la stessa tecnologia, e quando si trova sotto pressione sul piano finanziario.

Agli studi discussi si aggiungono, con risultati coerenti, Wang et al. (2025b), che sfruttano le zone pilota per l'innovazione nell'IA come esperimento di politica pubblica, Yu et al. (2026), Bergougui et al. (2026), Lei et al. (2026) e Xi et al. (2026).

La convergenza è quasi completa. Proprio per questo merita di essere guardata da vicino. Gli studi non si limitano a stimare l'effetto, propongono anche una spiegazione del perché l'adozione dell'IA dovrebbe modificare il comportamento dichiarativo.

Ricorrono *cinque meccanismi*.

Il primo è il *miglioramento del controllo interno*: sistemi informativi più strutturati rendono più difficile che un dato ambientale venga riportato senza riscontro (Yu et al., 2026).

Il secondo è l'*allentamento dei vincoli finanziari*: un'impresa che ottiene finanziamenti esterni, dalle banche o dal mercato, a condizioni migliori, può permettersi investimenti ambientali effettivi, e ha meno bisogno di sostituirli con annunci di facciata (Zhang, 2024; Lei et al., 2026).

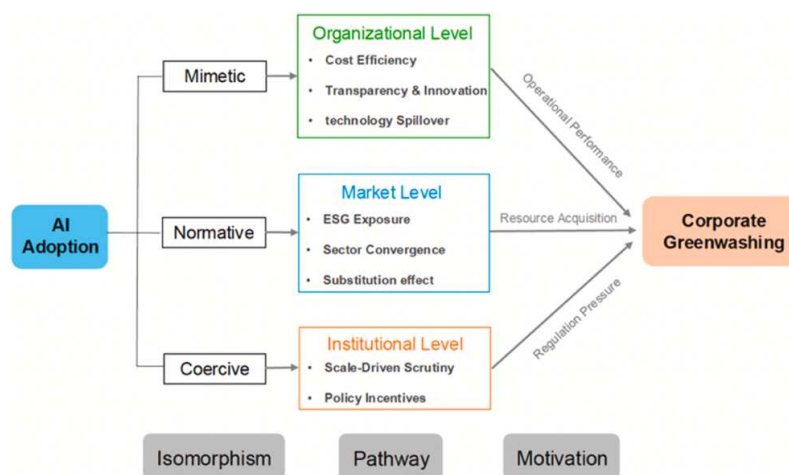
Il terzo è la *riduzione dell'asimmetria informativa fra impresa e interlocutori esterni*, che restringe lo spazio in cui la comunicazione può discostarsi dai fatti (Bai et al., 2025; Yu et al., 2026).

Il quarto è l'*aumento dell'attenzione degli analisti finanziari*, che accresce la probabilità che una discrepanza venga notata (Lei et al., 2026).

Il quinto è l'*efficienza della vigilanza esterna*, che con strumenti automatici controlla un numero maggiore di imprese (Cheng e Yang, 2026).

Xi et al. (2026) ricompongono questi elementi in un quadro unitario e leggono l'effetto come una forma di isomorfismo istituzionale: le organizzazioni che operano nello stesso ambiente tendono a rendere simili le proprie strutture e i propri comportamenti. L'adozione della tecnologia agisce su tre fronti contemporaneamente (figura 37). Sul piano interno rafforza l'efficienza organizzativa e la capacità di trasformazione ambientale dell'impresa. Sul piano esterno aumenta il controllo e le aspettative di legittimazione. A questo si aggiunge un migliore allineamento agli incentivi e agli obblighi formali posti dalle istituzioni. Il comportamento dichiarativo si adegua perché le tre tipologie di pressioni (*coercitiva, mimetica e normativa*) spingono nella stessa direzione.

Figura 37 - I tre percorsi attraverso cui l'adozione dell'IA agisce sul comportamento dichiarativo



Fonte: Xi et al., 2026

Una voce discordante arriva da Łudzińska (2025), rassegna richiamata perché raccoglie in modo indipendente i risultati della stessa letteratura. Gli studi che l'autrice sintetizza portano a conclusioni opposte su due punti. Il primo riguarda il controllo interno. L'automazione dei flussi informativi riduce i passaggi in cui una persona verifica il dato

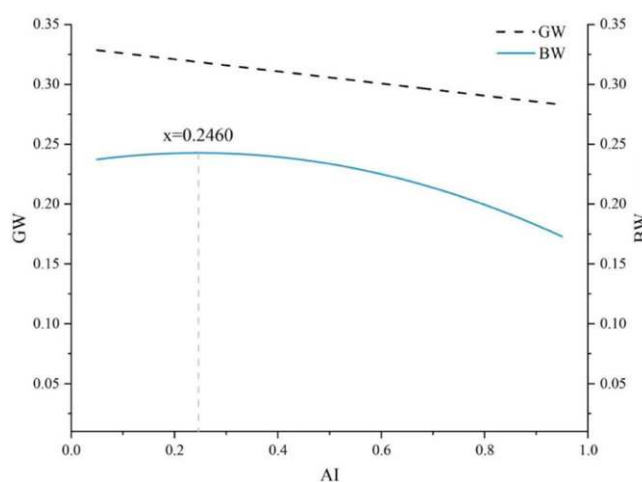
prima che venga comunicato: se il sistema produce il numero e lo trasmette senza che nessuno lo esamini, il controllo non si rafforza, si assottiglia. È il contrario del primo meccanismo rivendicato dagli studi sull'adozione, che considerano il miglioramento del controllo interno la ragione principale della riduzione del greenwashing. Il secondo punto riguarda l'impatto ambientale della tecnologia stessa. Addestrare e far funzionare i modelli richiede grandi quantità di energia, e i centri di calcolo che li ospitano hanno un'impronta di carbonio propria. Un'impresa può quindi ridurre il divario fra ciò che dichiara e ciò che fa e, allo stesso tempo, aumentare le proprie emissioni per effetto degli strumenti che ha adottato. Il valore di questo contrappunto sta soprattutto nella sua provenienza: è una sintesi della stessa letteratura condotta da chi non ha costruito nessuno dei modelli discussi.

Tre studi complicano il quadro.

Yu et al. (2026) osservano che un'impresa può distorcere la propria informativa ambientale in due direzioni opposte. La prima è il *greenwashing*, cioè dichiarare più di quanto si faccia. La seconda è il *brownwashing*, cioè dichiarare meno: un'impresa può tacere risultati ambientali che ha effettivamente conseguito, per non attirare l'attenzione di chi vigila, per non innalzare le aspettative future o per non esporsi a richieste di ulteriori impegni. In entrambi i casi la comunicazione non corrisponde ai fatti.

Su dati relativi alle quotate cinesi fra il 2012 e il 2023, gli autori trovano che l'IA riduce il greenwashing in modo lineare. Con il brownwashing, invece, la relazione cambia direzione oltre un certo livello di adozione (fissato a 0,246 dell'indice impiegato): è ciò che si intende per relazione a U rovesciata (figura 38).

Figura 38 - Relazioni stimate fra adozione dell'IA, greenwashing e brownwashing



Fonte: Yu et al., 2026

La spiegazione che ne danno è che i due effetti dell'IA non si manifestano insieme. Nelle fasi iniziali prevale la possibilità di occultare la propria performance ambientale, favorita anche dall'opacità degli algoritmi. Quando l'adozione si estende, prevalgono invece il monitoraggio continuo, la tracciabilità dei dati e la verifica esterna, che rendono il silenzio difficile da sostenere.

Vale la pena notare che gli effetti negativi richiamati, come opacità degli algoritmi, distorsione dei dati, risorse sottratte alla ricerca ambientale, sono gli stessi che Łudzińska (2025) raccoglie dalla letteratura: qui però vengono riconosciuti da autori che concludono comunque per un effetto di contenimento.

I meccanismi individuati sono tre: il miglioramento della qualità del controllo interno, l'allentamento dei vincoli finanziari e la riduzione dell'asimmetria informativa. L'effetto di contenimento risulta rafforzato dalla presenza di una regolazione pubblica più stringente.

Bergougui et al. (2026) si chiedono se l'effetto dipenda dal tipo di IA adottata e dalle condizioni in cui l'impresa opera. Sulle quotate cinesi fra il 2012 e il 2022, costruiscono misure di adozione distinte per sottocampo tecnologico e trovano che l'apprendimento automatico e i sistemi di pianificazione e decisione esercitano il vincolo più forte sul greenwashing, mentre le altre componenti incidono poco. È lo stesso risultato a cui arrivano He e Wang (2025).

Il secondo contributo riguarda le condizioni di contesto. Gli autori applicano una *regressione a soglia*, cioè una tecnica che verifica se la relazione fra due variabili cambi struttura oltre un determinato valore di una terza variabile. Nel loro caso le variabili di soglia sono l'intensità della regolazione ambientale e il grado di sviluppo delle tecnologie ambientali dell'impresa. Il risultato è che il legame fra adozione dell'IA e greenwashing si modifica al variare di queste due condizioni. La stessa tecnologia, quindi, non produce lo stesso effetto ovunque.

Lei et al. (2026) separano due momenti che gli altri studi trattano insieme: l'annuncio dell'adozione dell'IA e la sua effettiva implementazione. Sulle quotate cinesi fra il 2013 e il 2023, gli autori confrontano l'andamento del greenwashing prima e dopo l'annuncio.

Il metodo impiegato è lo *staggered difference-in-differences (DID)*. Si osserva la variazione nelle imprese che annunciano l'adozione, la si confronta con quella delle imprese che non hanno ancora annunciato nulla, e la differenza fra le due variazioni viene attribuita all'annuncio.

La riduzione del greenwashing si manifesta già a partire dall'annuncio, e non dal momento in cui la tecnologia entra effettivamente nei processi. Un annuncio di adozione attira osservatori, e con essi cresce la probabilità che una discrepanza fra dichiarazioni e risultati venga notata: è il canale dell'attenzione degli analisti finanziari. Il secondo canale è l'allentamento dei vincoli finanziari, che rende praticabili investimenti ambientali reali. L'effetto è più forte nelle imprese grandi, in quelle dei settori inquinanti e in quelle statali. A disciplinare l'impresa, quindi, sembra essere l'attenzione che l'annuncio dell'adozione richiama sul suo operato.

L'unico studio che arriva a una conclusione opposta è quello di Yadav e Asongu (2026). Su 1.260 osservazioni impresa-anno relative a sei economie asiatiche emergenti, l'adozione dell'IA si associa a una maggiore intensità di greenwashing. È anche l'unico del gruppo che non impiega dati cinesi. Il risultato viene discusso nel 4.3.4.

Perché la convergenza non equivale a conferma: i dodici studi che trovano l'effetto *deterrente* condividono quattro condizioni:

- il *paese*: tutti analizzano imprese quotate nella Cina continentale. Il contesto è quindi uno solo, e per di più caratterizzato da una regolazione ambientale in rapida evoluzione, da un forte intervento pubblico e da una quota rilevante di imprese a controllo statale;
- la *fonte dei dati*: il database CSMAR, che raccoglie i dati finanziari e societari delle quotate cinesi, compare in nove studi su dodici. Le variabili di controllo, le classificazioni settoriali e i criteri con cui i dati sono costruiti sono perciò in larga parte gli stessi;
- la *misura del greenwashing*: è calcolata quasi sempre come differenza fra un punteggio di informativa ambientale e un punteggio di performance ESG. È esattamente la misura che il 4.2.5 ha mostrato dipendere da rating costruiti in prevalenza sugli impegni dichiarati anziché sui risultati conseguiti. Dodici studi arrivano quindi alla stessa conclusione impiegando uno strumento che condivide la distorsione che dovrebbe rilevare;
- le *misure di adozione*: sono più varie delle precedenti, ma sono tutte ricavate dagli stessi documenti societari e dalle stesse banche dati, e alcune dipendono da ciò che l'impresa dichiara di sé.

Una convergenza ottenuta in queste condizioni misura anche la somiglianza dei disegni di ricerca, non solo la stabilità del fenomeno.

Questi studi restano validi, ma la base empirica su cui poggia l'effetto deterrente è più stretta di quanto il numero dei lavori suggerisca.

4.2.7 Sintesi

Il quadro che emerge dal 4.2 è duplice, e i due versanti si trovano in condizioni molto diverse. Sul versante dell'individuazione (*detection*) la letteratura è matura e procede per accumulazione: ogni generazione tecnologica nasce dal limite di quella precedente. Le tecniche lessicali fondate su dizionari, *TF-IDF* e *sentiment analysis* riconoscono le parole, ma non il contesto in cui compaiono. I modelli pre-addestrati della famiglia *BERT* superano questo limite leggendo la frase, e distinguono l'impegno dichiarato dal risultato conseguito. Restano però dentro il testo. I modelli generativi fanno un passo ulteriore e confrontano le dichiarazioni con fonti esterne, dove la RAG riduce il rischio di risposte prive di riscontro. Anche questi lavorano in un unico passaggio; i sistemi agentici dividono invece il compito fra più modelli che si controllano a vicenda, fino a simulare il confronto fra ruoli contrapposti.

La tabella 9 riassume questa progressione per generazione tecnologica e misura prodotta.

Tabella 9 - Quadro degli studi di rilevazione

Generazione tecnologica	Contributi	Misura prodotta
NLP tradizionale (4.2.1)	Lagasio (2024); Davidescu et al. (2026); Gorovaia e Makrominas (2025); dos Santos et al. (2025); Sari et al. (2025); Bennaceur et al. (2026); Pant et al. (2026)	Indici di scarto fra tono e contenuto del testo, o fra testo e un riferimento esterno (stampa, sanzioni, emissioni, punteggi di database)
Modelli pre-addestrati (4.2.2)	Bingler et al. (2022); Bingler et al. (2024); Wang et al. (2025a); Kılınç et al. (2026); Suárez-Giri et al. (2026); Telukdarie et al. (2026)	Classificazione delle singole frasi e indici che ne derivano: <i>cheap talk</i> , <i>cherry-picking</i> , <i>positive talk</i> , indice di <i>greenwashing</i> , leggibilità e <i>boilerplate</i>

Modelli generativi e deep learning (4.2.3)	Xu et al. (2026); Colesanti Senni et al. (2025); Zhao et al. (2026); Ke et al. (2026); Zhou C. et al. (2026); Lorkiewicz e Wasilewski (2026)	Verifica della fondatezza delle dichiarazioni: quota di indicatori comunicati, specificità dell'informativa, credibilità delle dichiarazioni
Sistemi agentici (4.2.4)	Gu et al. (2026); Song et al. (2026); Jakubczak et al. (2026); Zhou Z. et al. (2026)	Profili di rischio e punteggi di credibilità costruiti attraverso più passaggi di verifica

Fonte: elaborazione personale

I risultati più solidi, però, vengono dagli studi che verificano il segnale testuale rispetto a un riferimento esterno al documento, indipendentemente dalla sofisticazione dei sistemi impiegati. Xu et al. (2026) mostrano che i segnali estratti dai testi anticipano le sanzioni ambientali ricevute negli anni successivi. Colesanti Senni et al. (2025) sottopongono la propria griglia al giudizio di esperti provenienti da ventisei istituzioni, mentre dos Santos et al. (2025) confrontano il discorso aziendale con le emissioni di gas serra e con un indice reputazionale di mercato. Uno strumento risulta dunque affidabile quando dispone di un termine di paragone che l'impresa non controlla.

Sul versante del contrasto (*deterrence*) il quadro si rovescia. L'evidenza appare compatta, perché dodici studi su tredici concludono che l'adozione dell'IA riduce la comunicazione ingannevole, attraverso il rafforzamento dei controlli interni, l'allentamento dei vincoli finanziari e la maggiore attenzione degli analisti. La tabella 10 mostra però su che cosa poggia quella compattezza: dodici studi analizzano lo stesso mercato, ricorrono in larga parte alla stessa banca dati e adottano una misura di greenwashing che il 4.2.5 ha mostrato dipendere da rating costruiti sulle promesse più che sui risultati.

L'unico studio condotto fuori dalla Cina, su sei economie asiatiche emergenti, inverte il segno della relazione: dove i presidi di verifica e la regolazione sono più deboli, l'adozione dell'IA si associa a una maggiore intensità di greenwashing (Yadav e Asongu, 2026).

Tabella 10 - Studi sull'effetto dell'adozione dell'IA

Studio	Contesto	Direzione dell'effetto
Zhang (2024)	Quotate cinesi, 2014–2021	Riduzione

Khan et al. (2024)	Quotate cinesi A-share, 2009–2020	Riduzione
Bai et al. (2025)	Quotate cinesi, 2011–2019	Riduzione
He e Wang (2025)	Quotate cinesi A-share, 2011–2022	Riduzione
Ren et al. (2025)	Quotate cinesi A-share, 2012–2022	Riduzione
Tian e Shi (2025)	Quotate cinesi A-share, 2012–2022	Riduzione
Wang et al. (2025b)	Quotate cinesi A-share, 2008–2023	Riduzione
Cheng e Yang (2026)	Quotate cinesi A-share, 2010–2023	Riduzione
Yu et al. (2026)	Quotate cinesi A-share, 2012–2023	Riduzione del greenwashing; relazione a U rovesciata con il brownwashing
Bergougui et al. (2026)	Quotate cinesi, 2012–2022	Riduzione, con effetti di soglia
Lei et al. (2026)	Quotate cinesi A-share, 2013–2023	Riduzione
Xi et al. (2026)	Quotate cinesi	Riduzione
Yadav e Asongu (2026)	Sei economie asiatiche emergenti, 1.260 osservazioni impresa-anno	Aumento

Fonte: elaborazione personale

La prima domanda di ricerca riceve quindi due risposte di forza diversa. Che l'IA sappia individuare i segnali del greenwashing è dimostrato da una letteratura ampia e coerente. Che ne riduca l'incidenza è affermato da molti studi, ma in condizioni troppo simili fra loro perché la convergenza possa valere come conferma.

4.3 L'IA generativa come agevolatore

Il paragrafo inverte la prospettiva e affronta la RQ2: se e in che modo l'IA generativa possa agire da *agevolatore* del greenwashing, permettendo alle imprese di rendere più efficace e più estesa la propria comunicazione ingannevole.

La questione è speculare a quella del paragrafo precedente, ma la letteratura le ha dedicato un'attenzione molto minore. Dei sedici studi primari del campione che toccano questa domanda, dodici la affrontano in senso negativo: misurano l'effetto dell'adozione

dell'IA da parte delle imprese e concludono che il greenwashing si riduce, e sono stati discussi nel 4.2.6. Restano quattro contributi che documentano l'ipotesi opposta.

4.3.1 Il rischio teorizzato

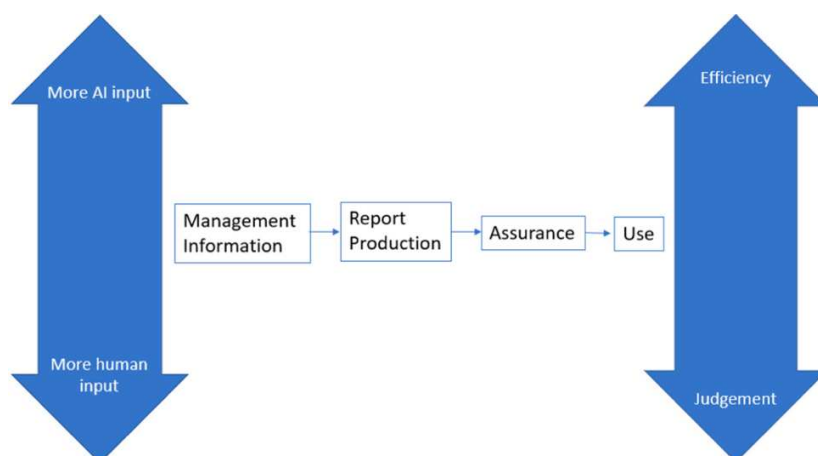
Il contributo di de Villiers et al. (2024b) è il riferimento concettuale di questo paragrafo, dal momento che imposta la questione che gli studi successivi tentano di misurare.

Gli autori adottano l'*impianto critico di Alvesson e Deetz*⁵¹, articolato in tre momenti.

Il primo, l'*insight*, ricostruisce lo stato della letteratura e delle prassi. Il secondo, la *critique*, esamina le influenze che l'IA generativa esercita su chi produce l'informativa non finanziaria, su chi la certifica e su chi la utilizza. Il terzo, la *transformative redefinition*, ne ricava le implicazioni e propone un'agenda di ricerca. Per sostenere l'analisi gli autori confrontano testi prodotti da un modello generativo con testi scritti da esseri umani.

Il quadro concettuale che gli autori propongono colloca l'intera filiera della rendicontazione, dall'informazione gestionale alla redazione del report, alla revisione, alla certificazione, fino all'analisi da parte di chi lo legge, fra due estremi: quello in cui ogni fase è automatizzata e quello in cui tutto è scritto da esseri umani. La pratica corrente si colloca in mezzo, e il compromesso in gioco è fra l'efficienza che l'automazione consente e il giudizio che solo l'intervento umano garantisce (figura 39).

Figura 39 - Quadro concettuale per la valutazione delle implicazioni dell'IA sulla rendicontazione di sostenibilità



Fonte: de Villiers et al., 2024b

⁵¹ Quadro di analisi proposto da Alvesson e Deetz nel 2000 per la ricerca critica in ambito manageriale, articolato in tre compiti: comprendere il fenomeno (*insight*), interrogarne i presupposti e gli interessi che lo orientano (*critique*), proporne una ridefinizione e indicare nuove direzioni di ricerca (*transformative redefinition*).

Il meccanismo che individuano dipende dal modo in cui questi modelli vengono costruiti, a prescindere dalle intenzioni dell'impresa. Un modello generativo apprende da grandi quantità di documenti esistenti, e fra questi ci sono anche i report di sostenibilità già pubblicati. Il greenwashing è un fenomeno diffuso: nei dati di addestramento finiscono quindi sia resoconti accurati sia documenti che esagerano o selezionano ciò che comunicano. Quando al modello viene chiesto di redigere una sezione di report, il testo che restituisce riproduce lo stile e i contenuti appresi, greenwashing compreso.

La comunicazione ingannevole può così propagarsi senza che nessuno la scelga deliberatamente. L'impresa che affida al modello la stesura di un passaggio non sta decidendo di esagerare: riceve un testo simile a quelli che il modello ha visto, e quei testi sono in parte già esagerati. È il greenwashing per commissione (*by commission*).

In questo senso l'IA rende la retorica ambientale più uniforme e più difficile da distinguere: il linguaggio del greenwashing diventa lo standard a cui il modello si conforma. Gli autori indicano però anche il rovescio della medaglia. La stessa tecnologia può essere impiegata per confrontare l'informativa narrativa con altre fonti, e quindi per verificarla: è la funzione documentata nel 4.2, e i due usi coesistono nello stesso strumento.

Una parte dell'analisi riguarda, infine, chi certifica i report. Gli autori osservano che le procedure di *assurance* tradizionali si adattano male alla natura qualitativa dell'informativa di sostenibilità, invitano chi le pratica a sperimentare l'uso dell'intelligenza artificiale e segnalano, fra i temi aperti, la questione della responsabilità quando nel processo intervengono sistemi automatici. Il tema viene ripreso nel 4.4.

4.3.2 L'evidenza sperimentale

Dimes et al. (2026) sono gli unici, all'interno del campione, a verificare sperimentalmente se l'IA faciliti effettivamente il greenwashing, anziché limitarsi ad argomentarlo sul piano teorico. L'obiettivo dichiarato è osservare come le persone si comportino davanti a un testo prodotto dall'IA, sia quando lo redigono sia quando lo leggono per formulare un giudizio. Gli autori si chiedono se l'intervento umano che precede e segue la generazione sia sufficiente a correggere le distorsioni del testo prodotto dall'IA.

I partecipanti sono quattordici studenti di un corso magistrale di *sustainability accounting*, impiegati come approssimazione dei due attori della filiera informativa, cioè chi redige il report di sostenibilità e chi lo utilizza. L'esperimento si svolge in una

successione di prove. Nelle prime due gli studenti agiscono come redattori, nelle successive come utilizzatori; in una prova conclusiva riflettono per iscritto sui risultati dell'intera sequenza.

Nella prima prova ciascuno studente deve produrre con ChatGPT il messaggio del CEO agli azionisti, la lettera di apertura che accompagna il report di sostenibilità, con un vincolo di cinquecento parole, per un'impresa forestale neozelandese inesistente.

Gli autori motivano la scelta di questa sezione con due ragioni: riassume l'orientamento complessivo del documento ed è composta quasi soltanto di testo discorsivo, il che consente di isolare la capacità di generazione linguistica del modello senza l'interferenza di tabelle e dati numerici. Il settore forestale è selezionato per la sua ambivalenza: presenta impatti ambientali e sociali sia positivi sia negativi, ed è al tempo stesso causa del cambiamento climatico e possibile rimedio. Anche l'impresa inesistente risponde a una ragione precisa. Il modello non può attingere a un documento reale, e produce quindi contenuti generici, che riflettono le sue tendenze di fondo anziché le caratteristiche di un'azienda specifica. Ogni studente consegna quattro elementi: la conversazione integrale con il modello, il testo generato senza modifiche, le correzioni apportate a mano e la versione finale. Il confronto fra il secondo e il quarto elemento misura quanto l'intervento umano cambi il risultato.

Per quantificare il greenwashing contenuto in ciascun testo gli autori adottano una versione adattata del *Cheap talk index (CTI)*, lo strumento elaborato da Lagasio (2024).

Il punteggio cresce al crescere del divario fra ciò che il testo afferma e ciò che è possibile controllare.

Il risultato della prima fase è netto. Le quattordici versioni finali, già riviste dagli studenti, ottengono un punteggio medio di greenwashing pari a 16,2, contro 7,3 dei tre report reali usati come riferimento, selezionati fra imprese forestali di dimensione e complessità simili. Il confronto decisivo è però quello interno. I testi generati dal modello senza alcun intervento ottengono 16,8, le versioni corrette a mano 16,2, e fra i due valori non c'è differenza statisticamente significativa. Gli studenti, insomma, non hanno riformulato i prompt né modificato in misura apprezzabile il testo ricevuto, e in alcuni casi le correzioni hanno perfino accentuato la distorsione. La revisione umana, nel punto in cui si colloca abitualmente, non ha funzionato come filtro.

Nella seconda fase gli stessi partecipanti cambiano ruolo. I trentadue testi complessivi, cioè ventotto versioni prodotte dagli studenti, tre report reali e un testo di controllo

generato dal modello, vengono anonimizzati con nomi di imprese forestali inventati. Sono poi uniformati nel carattere e nel formato, privati delle immagini e di ogni riferimento a fatti o persone reali. A ciascuno studente ne vengono assegnati dieci: i tre report reali, il testo di controllo e sei testi prodotti dai compagni, mai il proprio.

Gli studenti non sanno che nel gruppo sono stati inseriti documenti autentici e devono valutare ogni testo con un punteggio da uno a sette rispetto a sette affermazioni. Le prime tre riguardano il *contenuto*:

- il report contiene un equilibrio adeguato fra informazione qualitativa e quantitativa;
- il report contiene un equilibrio adeguato fra informazione positiva e negativa;
- il report menziona un'ampia gamma di portatori di interesse.

Le tre successive riguardano il *tono*:

- il report suona ottimista;
- il report suona sicuro;
- il report suona affidabile.

L'ultima è una domanda diretta: investirei in questa impresa.

Le prime tre affermazioni riguardano dunque ciò che il report dichiara, le tre successive il modo in cui lo dichiara, e l'ultima serve a capire come il lettore reagisca nel complesso. I punteggi mostrano una preferenza per i testi generati dall'IA, ma non su tutte le dimensioni. Sul *contenuto* la differenza è significativa solo per l'equilibrio fra informazione qualitativa e quantitativa; sulle altre due affermazioni non lo è. Sul *tono* è significativa in tutti e tre i casi: i testi generati suonano più ottimisti, più sicuri e più affidabili di quelli autentici. Sulla disponibilità a investire, invece, i due gruppi non si distinguono. Il risultato è dunque duplice: i lettori giudicano più credibili testi che il *CTI* colloca al doppio del livello di greenwashing dei documenti reali, e li giudicano tali per come suonano, non per ciò che dichiarano. Alla fine, gli studenti vengono informati dell'inganno e analizzano essi stessi i dati raccolti.

I limiti che gli autori dichiarano sono proporzionati alla natura dell'esperimento. Il campione è piccolo, quattordici partecipanti, e questo riduce la possibilità di

generalizzare. Incide anche l'età: tutti hanno meno di trent'anni, e la letteratura indica atteggiamenti più fiduciosi verso l'IA nelle generazioni più giovani, il che comprime la variabilità. Gli studenti, poi, sono un proxy accettato dei manager, ma non subiscono conseguenze reali per le proprie scelte, e davanti a un documento destinato al mercato la revisione sarebbe forse più cauta. Pesano infine due scelte di disegno: la concentrazione sul solo messaggio del CEO, sezione che la letteratura ha già collegato alla gestione dell'immagine, e la selezione degli indicatori di valutazione.

Il contributo che gli autori rivendicano è duplice. Da un lato lasciano un protocollo sperimentale replicabile, che può essere esteso a gruppi più ampi e diversificati e ad altre sezioni del report. Dall'altro propongono un quadro concettuale che distribuisce le responsabilità fra chi redige, chi utilizza e chi certifica il documento, e per ciascun ruolo indica il punto in cui il controllo umano diventa necessario, perché l'IA non prolunghi le strategie di gestione dell'immagine.

4.3.3 Il greenwashing algoritmico osservato

Vilanova e Bazanini (2026) osservano una filiera reale anziché misurare testi: è l'unico studio di caso del campione. Il loro oggetto è la governance di sostenibilità nella filiera bovina del sud-ovest del Mato Grosso, in Brasile, e la domanda riguarda la distribuzione delle tecnologie digitali, cioè chi le possiede lungo la catena e chi ne resti escluso.

Gli autori confrontano tre fonti di evidenza. La prima sono i documenti pubblici delle imprese. La seconda è costituita da nove interviste condotte fra giugno e ottobre 2025 con allevatori, rappresentanti sindacali, enti di formazione e funzionari del controllo sanitario. La terza è un questionario ai consumatori, con centotrentuno risposte valide su duecentocinquanta inviti. Dal confronto emergono le divergenze fra ciò che i documenti dichiarano e ciò che accade alla base della filiera.

La divergenza principale è un'asimmetria tecnologica. A valle della filiera, presso le imprese di trasformazione, la distribuzione e gli esportatori, si trovano il monitoraggio satellitare, gli algoritmi di tracciabilità e gli strumenti di redazione automatizzata dei report ESG, che servono anche a legittimare l'immagine dell'impresa. Presso gli allevatori, invece, questi strumenti non ci sono. Il satellite rileva la deforestazione su larga scala, ma il singolo pascolo non lo misura nessuno. Ciò che le imprese a valle dichiarano di poter verificare, a monte non è verificabile.

Un sistema di controllo, però, esiste. La governance sanitaria della filiera è solida e riconosciuta sui mercati esteri, ma è costruita per la sanità animale e non comunica con la verifica socio-ambientale. Neppure la domanda esercita pressione: il 53,1% dei consumatori indica il prezzo come criterio prioritario di acquisto e il 40% continuerebbe ad acquistare carne bovina associata alla deforestazione se costasse meno.

Da qui gli autori ricavano il concetto di *greenwashing algoritmico per omissione* (*algorithmic greenwashing by omission*). La distorsione non nasce da un dato falso. Le imprese a valle dichiarano sistemi di monitoraggio che esistono davvero, ma che coprono soltanto i segmenti in cui sono installati. Poiché il documento non precisa dove quella copertura si interrompa, il lettore ne ricava un'immagine di tracciabilità che riguarda l'intera filiera. Le tecnologie si concentrano dove misurare è facile, e ciò che non viene misurato resta invisibile senza che nessuno debba negarlo. Gli autori chiamano questo scarto miraggio digitale della sostenibilità (*digital mirage of sustainability*).

I limiti dichiarati riguardano l'impianto: ricerca qualitativa ed esplorativa, situata in un solo territorio, con poche interviste e un campione di consumatori non probabilistico, senza pretesa di generalizzazione. Gli autori precisano inoltre che la loro evidenza documenta il *greenwashing by omission*.

È questa distinzione a collegare lo studio ai due sottoparagrafi precedenti. Il versante per commissione è il meccanismo spiegato nel 4.3.1 e osservato sperimentalmente nel 4.3.2: lì l'IA produce il testo che abbellisce, qui è la sua semplice presenza dichiarata a legittimare l'impresa. I due meccanismi sono distinti e possono convivere nello stesso report.

4.3.4 Complementarità istituzionale

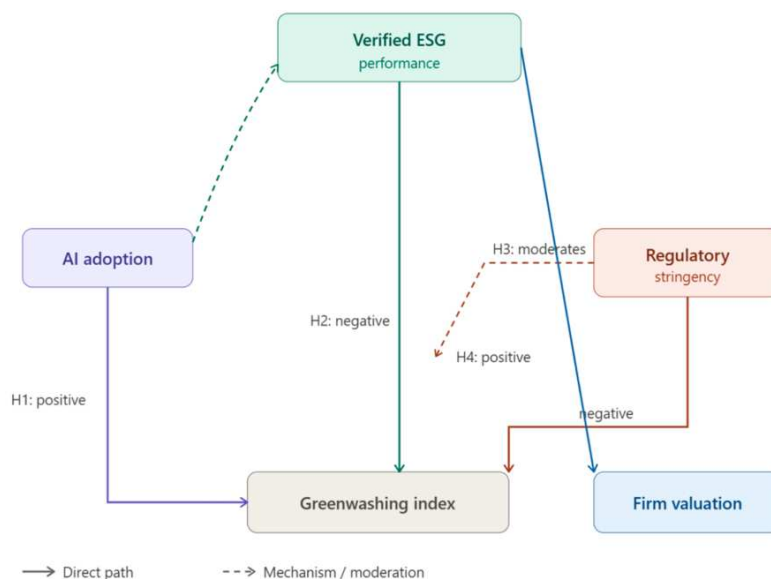
Lo studio di Yadav e Asongu (2026), già richiamato fra i lavori sull'effetto dell'adozione, è l'unico contributo quantitativo che documenta un'associazione positiva fra l'impiego dell'IA e il greenwashing. Oltre alla direzione della relazione, gli autori indagano le condizioni che la determinano: in quali contesti istituzionali le dichiarazioni di sostenibilità restano ancorate ai risultati, e in quali se ne allontanano. Il campione è composto da 180 imprese quotate in sei economie asiatiche emergenti, Cina, India, Malesia, Filippine, Thailandia e Vietnam, osservate ogni anno dal 2017 al 2023, per un totale di 1.260 osservazioni impresa-anno. I dati provengono da una banca dati commerciale. La scelta di sei paesi anziché di uno solo è la caratteristica che distingue

questo lavoro da quelli esaminati nel 4.2.6: consente, infatti, di confrontare imprese che operano sotto regimi di verifica e di controllo diversi, e quindi di misurare l'effetto del contesto anziché tenerlo fermo. Le variabili su cui si regge l'analisi sono quattro. L'*adozione dell'IA* è rilevata con un indicatore binario, costruito dagli autori perché nelle banche dati commerciali non esiste una misura standardizzata di questo fenomeno. Il *greenwashing* è misurato con un indice a livello di impresa, ripreso dalla letteratura esistente, che contrappone l'intensità delle dichiarazioni ai segnali di risultato effettivamente osservabili. La performance ESG verificata considera solo i dati ambientali e sociali che un soggetto esterno ha controllato e confermato. Restano fuori gli obiettivi che l'impresa si dà da sé e i risultati che dichiara senza che nessuno li abbia verificati. La severità regolamentare riguarda invece il paese, non l'impresa: misura quanto siano stringenti le regole sulla rendicontazione ambientale e quanto le autorità siano in grado di farle rispettare. Cambia da paese a paese e, all'interno dello stesso paese, da un anno all'altro.

Su queste variabili gli autori costruiscono quattro ipotesi (figura 40):

- all'adozione dell'IA corrisponde un greenwashing maggiore;
- a una performance ESG verificata corrisponde un greenwashing minore;
- questo secondo effetto si rafforza dove la regolazione è più severa;
- il mercato premia la performance verificata, non l'adozione dell'IA in quanto tale.

Figura 40 - Quadro concettuale dell'adozione dell'IA, del greenwashing e delle forze disciplinanti



Fonte: Yadav e Asongu, 2026

I risultati confermano tutte e quattro le ipotesi. Le imprese che adottano l'IA mostrano un livello di greenwashing più alto: il coefficiente è positivo, pari a 0,434. Le imprese con una performance ESG verificata mostrano invece un livello più basso, con un coefficiente negativo pari a -0,753. Entrambi i valori sono statisticamente significativi, cioè difficilmente attribuibili al caso⁵².

Il terzo risultato riguarda il modo in cui le due forze interagiscono, ed è il più importante. Il coefficiente che le mette in relazione è negativo, pari a -0,090: significa che l'effetto di contenimento della verifica indipendente diventa più forte man mano che la regolazione del paese si fa più severa. La stessa verifica, in altre parole, produce risultati diversi a seconda dell'ambiente normativo in cui viene svolta.

Il quarto risultato riguarda il mercato. Il valore attribuito all'impresa cresce quando la performance ESG è verificata, ma non cresce quando l'impresa adotta l'IA. Dichiarare di usare questi sistemi, di per sé, non produce alcun premio in termini di valutazione.

Gli autori ripetono l'analisi con nove modelli statistici diversi, che trattano i dati con criteri alternativi. Segno e grandezza dei coefficienti restano gli stessi in tutti i casi, il che rende improbabile che i risultati dipendano dal metodo scelto. Verificano inoltre che l'associazione fra adozione dell'IA e greenwashing non sia prodotta da un piccolo gruppo di imprese anomale; infatti, vale allo stesso modo per le imprese con livelli bassi, medi e alti di greenwashing. Un solo risultato non trova conferma. La proprietà statale non introduce differenze significative, mentre diversi studi sul mercato cinese trovano comportamenti distinti fra imprese pubbliche e private. Da qui la proposta teorica del lavoro, che gli autori chiamano *complementarità istituzionale*. Verifica indipendente e regolazione sono strumenti che funzionano insieme: la verifica produce le prove dettagliate, impresa per impresa, su cui l'autorità può poi intervenire; la regolazione obbliga le imprese a farsi verificare e stabilisce a quali condizioni. Dove agisce uno solo dei due, il contenimento del greenwashing si indebolisce.

Lo studio, però, non spiega perché l'adozione dell'IA si accompagni a un maggiore greenwashing. La misura utilizzata registra soltanto se l'impresa impieghi o meno questi sistemi, e il testo dei documenti non viene mai esaminato. Gli autori ne sono consapevoli, e per il futuro suggeriscono di analizzare direttamente i testi, così da riconoscere le parti

⁵² I coefficienti sono calcolati su variabili standardizzate, quindi non vanno letti come percentuali: indicano di quanto cambia l'indice di greenwashing, misurato in deviazioni standard, al variare della condizione considerata.

generate dall'IA. I limiti dichiarati sono tre. Il primo riguarda il tipo di risultato ottenuto: i dati mostrano che i due fenomeni si presentano insieme, non che il primo causi il secondo. Per stabilire un nesso causale servirebbe un evento esterno capace di modificare le condizioni, per esempio una riforma delle regole, oppure un'adozione dell'IA avvenuta in momenti diversi per imprese diverse. C'è poi la misura dell'adozione, che non distingue né quanto l'IA venga usata né con quali controlli: un'impresa che la impiega per un singolo processo riceve lo stesso valore di una che ne ha fatto l'infrastruttura della propria rendicontazione. L'ultimo limite riguarda la composizione del campione, che include solo le imprese con dati completi per tutti e sette gli anni, e favorisce quindi le imprese più stabili, quelle che pubblicano informazioni ESG con regolarità. I risultati, avvertono gli autori, vanno estesi con cautela alle imprese che escono dal mercato o che pubblicano in modo discontinuo. Resta da spiegare il contrasto con il 4.2.6, dove dodici studi su tredici concludono che l'adozione dell'IA riduca la comunicazione ingannevole. Più che a un errore di uno dei due gruppi di ricerca, la differenza sembra dovuta al contesto istituzionale, che cambia il segno della relazione. I dodici studi osservano tutti il mercato cinese, dove i controlli sono consolidati e le regole sono percepite come vincolanti: lì l'adozione dell'IA si accompagna a una migliore qualità dell'informazione.

Yadav e Asongu osservano sei mercati, in buona parte con controlli più deboli e regole applicate in modo meno uniforme: lì la stessa tecnologia si accompagna a un divario maggiore fra ciò che l'impresa dichiara e ciò che fa. L'IA, in altri termini, non ha una direzione propria: assume quella che l'assetto istituzionale le imprime.

Su questa dipendenza dal contesto si tornerà nella discussione conclusiva.

4.3.5 Sintesi

I quattro contributi che documentano l'ipotesi opposta a quella discussa a quella della sezione 4.2 non formano un corpo omogeneo: ciascuno porta un tipo di evidenza diverso (tabella 11).

Tabella 11 - I quattro contributi sulla RQ2

Studio	Tipo di evidenza	Meccanismo documentato
de Villiers et al. (2024b)	Concettuale	Riproduzione del greenwashing appreso dai dati di addestramento

Dimes et al. (2026)	Sperimentale	Commissione: il modello genera il testo
Vilanova e Bazanini (2026)	Caso studio qualitativo	Omissione: l'IA dichiarata legittima l'impresa
Yadav e Asongu (2026)	Quantitativa	Associazione fra adozione e greenwashing, condizionata dal contesto istituzionale

Fonte: elaborazione personale

Nessuno dei quattro misura ciò che misurano gli altri, e quindi non possono confermarsi a vicenda. Quello che stabiliscono insieme è che l'IA può facilitare il greenwashing in due modi distinti. Il primo è per *commissione*, quando il modello genera il testo che attenua, uniforma e abbellisce: è il meccanismo che Dimes et al. osservano in condizioni controllate e che de Villiers et al. spiegano risalendo ai dati di addestramento, perché un modello che ha appreso da report già pubblicati riproduce anche il greenwashing che quei report contenevano. Il secondo è per *omissione*, quando la sola dichiarazione di impiegare sistemi di monitoraggio legittima l'impresa, a prescindere da ciò che quei sistemi coprono davvero. Perché si produca il primo, il modello deve scrivere; per il secondo basta che l'impresa ne annunci l'esistenza.

Yadav e Asongu aggiungono che nessuno dei due meccanismi opera allo stesso modo ovunque. La loro associazione positiva fra adozione dell'IA e greenwashing si attenua dove la verifica indipendente e la regolazione agiscono insieme: l'esito, quindi, sta nell'assetto istituzionale più che nella tecnologia. Manca invece un'osservazione diretta del fenomeno. Nessuno studio del campione verifica, su ampi campioni di report reali, se un documento sia stato redatto con l'IA e con quali conseguenze sulla sua attendibilità. Gli strumenti per collegare i segnali testuali agli esiti esistono già, come mostra il 4.2.3; quello che manca è il riconoscimento della paternità del testo. È la direzione che gli stessi autori dell'esperimento indicano per la ricerca futura.

La risposta alla seconda domanda di ricerca è quindi asimmetrica rispetto alla prima. La letteratura ha documentato i meccanismi con cui l'IA può facilitare il greenwashing, mentre la loro diffusione resta ancora da misurare.

4.4 Il controllo umano e istituzionale

Il paragrafo affronta la RQ3, che si chiede in che modo l'IA, affiancata dal controllo *umano* e *istituzionale*, possa limitare i rischi algoritmici e preservare la credibilità dell'informativa di sostenibilità. La domanda nasce dai due paragrafi precedenti. Nel 4.2 si è visto che l'IA riconosce i segnali della comunicazione ingannevole meglio di qualunque strumento precedente, ma che i risultati più solidi arrivano solo quando il testo viene confrontato con un riferimento che l'impresa non controlla. Nel 4.3 si è visto che la stessa tecnologia può produrre il testo distorto e legittimare chi lo pubblica, e quel paragrafo si è chiuso sulla constatazione che l'esito dipende dall'assetto istituzionale più che dalla tecnologia. Le due conclusioni portano allo stesso punto: se l'IA riduca o amplifichi il divario fra dichiarazioni e risultati dipende soprattutto dal sistema di controlli entro cui viene impiegata. I contributi del corpus selezionato che rispondono direttamente a questa domanda sono cinque. La materia, però, non si esaurisce in essi: diversi studi discussi nei paragrafi precedenti si chiudono con indicazioni sui controlli necessari, richiamate qui quando integrano l'evidenza dei cinque contributi principali. L'esposizione segue la distanza dal documento. Si parte dal punto in cui il testo viene materialmente prodotto e ci si allontana progressivamente, fino alle autorità che stabiliscono le regole entro cui l'impresa comunica.

I livelli sono cinque:

- il flusso di lavoro con cui il report viene redatto, e il punto in cui vi si inserisce l'intervento umano (*human-in-the-loop*), combinando responsabilità algoritmica e supervisione umana;
- il consiglio di amministrazione e i controlli interni all'impresa (supervisione interna);
- il controllo esterno del mercato che ne valuta l'attendibilità, in particolare le banche finanziatrici e gli investitori;
- i soggetti terzi che certificano l'informativa, attraverso le procedure di assurance e la revisione di sostenibilità (il presidio dell'assurance);
- le istituzioni pubbliche e le autorità di regolamentazione (il contesto istituzionale).

L'ordine non implica una gerarchia di efficacia. Indica soltanto la posizione rispetto al documento, e serve a distinguere i controlli che agiscono prima che il report venga pubblicato da quelli che intervengono dopo, quando il testo è già nelle mani di chi lo legge. I cinque livelli non sono coperti dalla letteratura nella stessa misura.

4.4.1 Responsabilità algoritmica e supervisione umana

Il contributo di Harke (2026) è l'unico capitolo di volume del campione. Non porta dati propri: confronta in modo sistematico le prestazioni dei modelli linguistici e quelle degli analisti umani nella rilevazione del greenwashing, basandosi sulla letteratura esistente e su due applicazioni documentate.

La tesi dell'autore è che l'automazione integrale non sia praticabile, e che l'unico assetto difendibile sia quello in cui il modello sorveglia i grandi volumi mentre l'esperto umano decide sui casi controversi. Il rischio da cui l'autore muove riguarda la responsabilità algoritmica (*algorithmic accountability*). Sostituire il giudizio soggettivo dell'analista con l'esito di un sistema opaco non garantisce di per sé una maggiore integrità del mercato: introduce un rischio nuovo, che Harke chiama *AI audit washing*. Si tratta del fenomeno per cui il ricorso ad algoritmi complessi fornisce «una falsa impressione di garanzia matematica su dati ESG fundamentalmente viziati»⁵³. Perché il fenomeno si configuri è sufficiente che un'istituzione finanziaria, un'agenzia di rating o un intermediario dichiari la validazione tramite IA senza rendere noti i dati su cui il modello è stato addestrato, i tassi di allucinazione riscontrati e i parametri di taratura. L'opacità che lo strumento doveva ridurre si sposta così dal documento al sistema che lo esamina.

Un secondo rischio riguarda le distorsioni incorporate nel modello. Harke osserva che un algoritmo assorbe le disuguaglianze presenti nei dati su cui è stato addestrato. Se l'addestramento avviene in prevalenza su informative di grandi imprese tecnologiche occidentali, il modello apprende le convenzioni linguistiche di quei documenti: il lessico, la sintassi, le strutture argomentative. Applicato poi su scala globale, rischia di penalizzare le imprese manifatturiere dei mercati emergenti, i cui report non presentano la stessa forma levigata, indipendentemente dalla loro condotta ambientale effettiva.

Il modello scambierebbe cioè una *differenza di stile* per una *differenza di sostanza* .

⁵³ Testo originale: «a false sense of mathematical assurance over fundamentally flawed ESG data» (Harke, 2026).

Un riscontro parziale viene dallo studio di Telukdarie et al. (2026), già esaminato nel 4.2.2. Il loro modello ha il compito di riconoscere e classificare gli indicatori contenuti nei report di sostenibilità, e la sua prestazione cambia molto a seconda del tipo di indicatore. Sugli *indicatori ambientali* e di *governance* l'accuratezza supera il 99%: sono grandezze espresse in forma standardizzata (emissioni, consumi o composizione degli organi sociali) e il modello le individua quasi sempre correttamente. Sugli *indicatori sociali* scende al 63,6%, perché in quell'ambito le imprese scrivono in forma discorsiva, con formule che variano da un documento all'altro, e non esiste un modo convenzionale di esprimere il dato. La prestazione del modello dipende dunque da quanto il contenuto sia standardizzato.

Il sostegno, però, si ferma qui. Telukdarie et al. collegano il divario al tipo di indicatore e non alla provenienza dell'impresa, e il loro campione è composto da imprese quotate in India e in Sudafrica, cioè da quei mercati emergenti che, secondo l'ipotesi di Harke, il modello dovrebbe penalizzare. Su due dimensioni su tre l'accuratezza è invece quasi perfetta. L'argomento a favore di un assetto ibrido Harke lo costruisce confrontando due esperienze opposte, che rappresentano i due estremi possibili.

La prima mette alla prova modelli generativi generalisti, cioè sistemi di uso comune impiegati senza alcun adattamento al compito. Tre di questi modelli devono assegnare un punteggio di sostenibilità a società calcistiche europee leggendone soltanto le informative pubbliche. Come termine di paragone si usa il punteggio che esperti umani avevano attribuito alle stesse società, e l'accuratezza misura quanto il giudizio del modello vi si avvicini. Il migliore dei tre, GPT-4o, raggiunge un'accuratezza compresa fra il 56% e il 58%, mentre gli altri due restano sotto il 50%. Più degli errori di valutazione, però, preoccupa la tendenza dei modelli a inventare dati. Tutti e tre i modelli producono dati ambientali inesistenti ma verosimili, per esempio quantità di materiale riciclato che nel documento non compaiono, con tassi compresi fra il 10% e il 16%, e continuano a farlo anche quando vengono predisposti passaggi di verifica. In un contesto di certificazione formale, osserva Harke, un tasso di fabbricazione del 10% comporta responsabilità legali rilevanti, perché il modello può penalizzare un'impresa conforme o assolvere un inquinatore.

La seconda esperienza è l'opposto: un sistema costruito su misura per il compito. Il sistema *DeepGreen* (Xu et al., 2026), già esaminato nel 4.2.3, analizza 9.369 relazioni annuali del mercato cinese pubblicate in tre anni. Oltre a leggere il testo, il sistema

confronta ogni dichiarazione con fonti esterne verificate, e questo riduce le invenzioni che affliggono i modelli generalisti. I segnali che produce, inoltre, anticipano le sanzioni ambientali che le imprese riceveranno negli anni successivi, il che mostra che la rilevazione coglie qualcosa di reale e non solo una particolarità del linguaggio. Il volume elaborato, d'altra parte, è fuori dalla portata di qualunque revisione manuale.

Il confronto porta a una conclusione simmetrica. La revisione affidata ai soli esseri umani non regge i volumi perchè nessun gruppo di analisti può leggere novemila relazioni in tre anni. L'automazione integrale non regge l'inaffidabilità del modello, che inventa dati in una parte non trascurabile dei casi e non coglie il contesto in cui una dichiarazione va valutata. Nessuno dei due estremi, quindi, è praticabile da solo.

L'autore riassume il percorso in un confronto fra quattro approcci alla revisione, dal più tradizionale a quello che propone, indicando per ciascuno il meccanismo di funzionamento, i punti di forza e i limiti (tabella 12).

Tabella 12 - Confronto fra gli approcci alla revisione del greenwashing

Auditing Framework	Primary Mechanism	Core Strengths	Critical Limitations
Traditional Manual Auditing	Human expert review and subjective scoring of corporate sustainability reports.	Deep contextual and ethical understanding; high accuracy on edge cases.	Severe scalability limits; susceptible to cognitive bias and institutional capture.
Early NLP (Keyword/Dictionary)	Basic keyword frequency tracking and mathematical readability scoring.	High processing scale; objective and consistent rule application.	Semantic blindness; easily bypassed by strategic deflection and complex syntax.
Modern LLM (Semantic Analysis)	Deep contextual comprehension and verification against external datasets.	Contextual comprehension at the macroeconomic scale; resistant to basic polishing.	Susceptible to probabilistic hallucinations; operates as an opaque black box.
Proposed HITL Governance	LLM-driven market surveillance coupled with human expert adjudication for anomalies.	Balances algorithmic scale with human ethical judgment; neutralizes hallucinations.	High initial implementation costs; requires complex regulatory integration.

Fonte: Harke, 2026

Da qui la proposta dell'architettura che l'autore chiama *human-in-the-loop*, letteralmente “con l'uomo nel circuito”. Il sistema automatico elabora i documenti, estrae le dichiarazioni ambientali, le confronta con banche dati esterne e segnala le discrepanze. Quando incontra una dichiarazione ambigua, però, non formula una stima probabilistica;

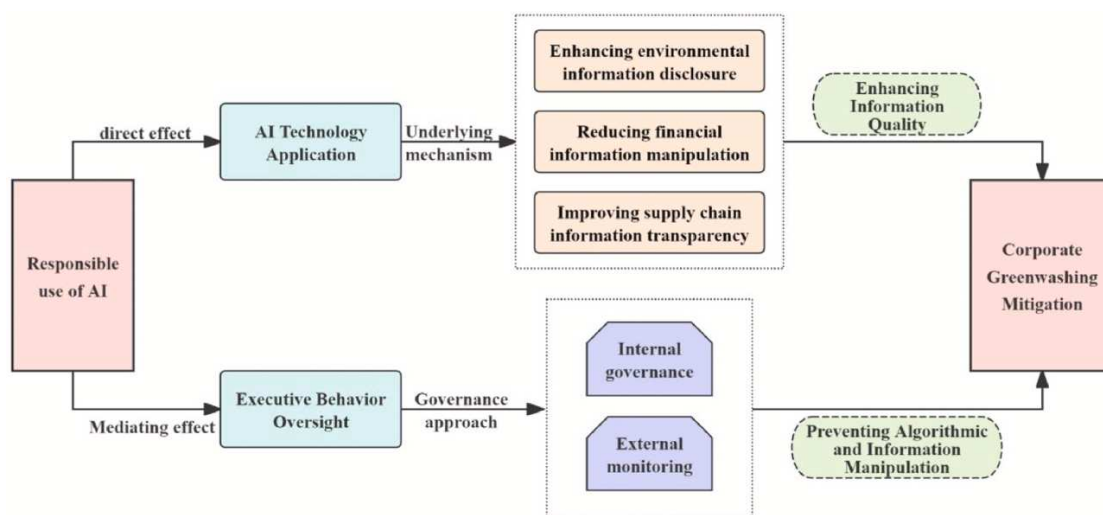
sospende l'elaborazione e trasmette il passaggio specifico a un analista, che applica un giudizio contestuale. La decisione finale resta quindi ancorata a un accertamento e non a un calcolo di probabilità. L'architettura prevede inoltre una retroazione: quando l'analista stabilisce che il modello ha allucinato o ha frainteso una strategia legittima, la correzione rientra nel sistema e ne affina l'accuratezza. Harke non presenta questo assetto come una soluzione priva di costi, e indica tre difficoltà di attuazione. La prima è *economica*. Far operare modelli di grandi dimensioni, con verifica continua contro fonti esterne e su milioni di documenti, richiede una capacità di calcolo costante, e quel costo è una barriera per gli studi di revisione più piccoli. La seconda riguarda la *tenuta del passaggio umano*. Un sistema tarato in modo troppo sensibile produce un numero elevato di falsi positivi: gli analisti vengono sommersi di segnalazioni, e il processo torna esattamente al problema di volume che l'automazione doveva risolvere. La terza è *organizzativa*. La procedura ibrida deve innestarsi in strutture di conformità rigide, e l'alternanza fra estrazione automatica e giudizio umano deve reggere gli standard di documentazione richiesti dalle autorità. Perché il controllo sia verificabile anche dall'esterno, infine, l'autore ritiene necessario che il sistema, insieme al punteggio di rischio, produca un registro consultabile delle proprie decisioni con i pesi applicati, i passaggi testuali estratti dal documento e i dati esterni usati nel confronto. Le lacune che Harke segnala riguardano le condizioni necessarie affinché tutto questo funzioni. La prima è l'assenza di una definizione operativa condivisa di greenwashing: i modelli vengono addestrati su insiemi di dati frammentati e su valutazioni proprietarie che incorporano le distorsioni di chi le produce, e servirebbero invece raccolte aperte, verificate da esperti, che colleghino il testo dei documenti a indicatori fisici di prestazione. La seconda è la mancanza di studi longitudinali: le valutazioni disponibili fotografano un momento, mentre le imprese adattano il proprio linguaggio non appena un meccanismo di rilevazione diventa noto. È un'osservazione che riguarda l'intero capitolo, perché implica che l'efficacia di uno strumento di rilevazione tenda a deteriorarsi man mano che il suo funzionamento si diffonde.

4.4.2 La supervisione interna

Il secondo livello si colloca sul flusso operativo, quindi riguarda chi decide come l'IA venga impiegata. Lo studio di Jiang et al. (2026) è l'unico del campione che misura questo passaggio. Il suo risultato principale coincide con quello dei lavori esaminati nel 4.2.6,

perché anch'esso stima l'effetto dell'adozione dell'IA su imprese quotate cinesi; viene trattato in questo paragrafo perché il suo contributo distintivo riguarda la condizione che rende possibile l'effetto. Il campione comprende le imprese quotate nei mercati di Shanghai e Shenzhen fra il 2012 e il 2022, per 8.336 osservazioni impresa-anno. Il greenwashing è misurato come scarto fra il punteggio che l'impresa ottiene per quantità di *informativa ESG pubblicata* e il punteggio che riceve per la *performance ESG effettiva*: quanto più il primo supera il secondo, tanto più l'impresa dichiara più di quanto faccia. L'adozione dell'IA è rilevata contando nelle relazioni annuali la frequenza di settantatré termini, ricavati a partire da cinquantadue parole chiave iniziali con una tecnica che individua i vocaboli semanticamente vicini. Rispetto a un indicatore che registri solo la presenza o l'assenza della tecnologia, questa misura coglie anche l'intensità dell'impiego. I risultati reggono a una serie di verifiche, fra cui il *double machine learning*. Il risultato principale è che l'IA riduce il greenwashing in misura statisticamente significativa. Gli autori individuano tre canali attraverso cui ciò accade: l'informativa ambientale migliora, la manipolazione dell'informazione finanziaria si riduce e la trasparenza della catena di fornitura aumenta. Tutti e tre agiscono nella stessa direzione, cioè innalzano la qualità dell'informazione che l'impresa produce, e su quella qualità si fonda la riduzione del divario fra dichiarazioni e fatti (figura 41).

Figura 41 – Framework sull'uso responsabile dell'IA e della supervisione degli amministratori



Fonte: Jiang et al., 2026

Questo effetto, però, non è costante. Gli autori partono dall'ipotesi che l'IA possa essere piegata anche a fini opportunistici, attraverso la manipolazione algoritmica e la

distorsione dell'informazione, e verificano che cosa accada quando i dirigenti hanno incentivi in quella direzione. Il dato che ottengono è più sfumato dell'ipotesi. L'effetto di contenimento non si rovescia: si indebolisce, fino a perdere significatività statistica. L'IA, insomma, smette di funzionare come correttivo, senza per questo trasformarsi in uno strumento di manipolazione.

I contesti in cui l'effetto si attenua sono tre: le imprese fortemente inquinanti, quelle in difficoltà finanziaria e quelle i cui dirigenti ragionano sul breve periodo. L'orientamento temporale della dirigenza è misurato contando, nella relazione degli amministratori, quarantatré parole associate a un orizzonte ravvicinato: dove queste sono più frequenti, l'effetto dell'IA sul greenwashing non è significativo. Sono tutte situazioni in cui il vantaggio immediato di apparire sostenibili supera il costo atteso di essere smentiti.

Gli autori verificano allora che cosa possa vincolare il comportamento dirigenziale. La supervisione degli amministratori viene scomposta in cinque strumenti, due interni all'impresa e tre esterni.

Quattro funzionano: la frequenza del ricambio dei dirigenti, la quota di capitale in mano a investitori istituzionali di tipo professionale, il posizionamento della società di revisione incaricata e il numero di questioni chiave segnalate in sede di revisione.

Il quinto, la proporzione di amministratori indipendenti in consiglio, non produce alcun effetto significativo. È il risultato meno atteso: l'amministratore indipendente è lo strumento di governo societario che la letteratura invoca più spesso, e qui la sua presenza non cambia nulla. Due dei quattro strumenti efficaci riguardano invece la revisione esterna, di cui si tratterà nel 4.4.3.

La formula con cui gli autori riassumono la propria conclusione è che l'IA vada accompagnata dal controllo sul comportamento di chi la usa. La tecnologia, da sola, non orienta nulla: mette a disposizione una capacità che può servire a rendere l'informativa più accurata o a renderla più difficile da contestare, e a decidere quale dei due esiti si realizzi è l'assetto di governo dell'impresa.

I limiti che gli autori dichiarano sono due. Il primo riguarda la misura dell'adozione, ricavata dal testo delle relazioni annuali: la frequenza con cui un'impresa nomina l'IA può non corrispondere all'uso che ne fa davvero, e per accertarlo servirebbero fonti di dati diverse. Il secondo riguarda l'oggetto stesso dello studio. Gli autori precisano di aver esaminato l'IA come strumento di governo, senza affrontare i rischi che essa introduce, cioè la distorsione algoritmica, l'opacità e la responsabilità per le decisioni automatizzate.

Un'indicazione convergente arriva da Cheng e Yang (2026), che nel 4.2.6 abbiamo incontrato fra gli studi sull'effetto dell'adozione. Nelle implicazioni pratiche, non fra i risultati, gli autori raccomandano di istituire meccanismi decisionali fondati sulla collaborazione fra uomo e macchina, in cui esperti di settore riesaminino gli esiti prodotti dai modelli prima che vengano utilizzati; l'indicazione è rivolta soprattutto alle imprese di piccole e medie dimensioni, mentre alle maggiori suggeriscono di integrare l'IA nell'intera struttura di governo ESG. A tutte raccomandano di sottoporre i dati ESG elaborati automaticamente a una verifica indipendente di terza parte.

4.4.3 Il controllo esterno del mercato

Lo studio di Zheng et al. (2026) è l'unico del campione in cui l'IA si trova nelle mani di chi controlla anziché di chi è controllato. Qui la usa la banca che deve decidere a quali condizioni prestare denaro.

Il contesto è quello del credito verde, cioè dei finanziamenti che le banche concedono a condizioni agevolate alle imprese con buone credenziali ambientali. Lo strumento nasce per premiare chi inquina meno, ed è proprio per questo esposto alla manipolazione: all'impresa conviene apparire ambientalmente meritevole, perché quell'apparenza si converte in denaro a costo minore. Gli autori si chiedono se l'IA, in mano al creditore, riduca questo margine.

Il campione comprende 1.209 contratti di finanziamento stipulati in Cina fra il 2019 e il 2023. Per misurare il greenwashing dell'impresa si ricorre a una tecnica di apprendimento automatico che scompone l'informativa ambientale, separa le dichiarazioni sostanziali da quelle puramente simboliche e calcola lo scarto fra le due.

L'adozione dell'IA da parte delle banche è ricavata a mano dai bilanci degli istituti, distinguendo fra chi si è limitato a dichiarare una strategia e chi la applica effettivamente. Ciò che lo studio vuole spiegare è il prezzo del singolo finanziamento, cioè il ricarico che la banca chiede rispetto al tasso di riferimento per prestare denaro a quella specifica impresa.

Il risultato è che una stessa impresa con segnali di greenwashing paga di più se la banca che la finanzia impiega l'IA, e non paga di più se la banca non la impiega. Non è dunque l'adozione della tecnologia a rendere il credito più caro, né il greenwashing a essere penalizzato ovunque: l'effetto nasce dalla combinazione delle due condizioni.

Guardate nel complesso, le banche dotate di IA applicano anzi condizioni migliori, con ricarichi più bassi e durate più lunghe, e sono in prevalenza istituti grandi e con rating elevato. Il maggior costo colpisce le imprese sospettate di greenwashing, e solo quando chi presta ha gli strumenti per riconoscerle.

Gli autori verificano infine che il risultato non dipenda da altro. Le banche che adottano l'IA sono diverse dalle altre per dimensione e solidità, e anche le imprese con alti livelli di greenwashing hanno caratteristiche proprie: il rischio è che a spiegare il maggior costo siano quelle differenze, e non la tecnologia. Per escluderlo, il confronto viene ristretto a banche e imprese simili sotto ogni altro profilo, e si ricorre a una tecnica statistica che isola la parte di adozione dell'IA indipendente dalle scelte dell'istituto. Il risultato resta. Due analisi di approfondimento mostrano dove l'effetto si concentra.

La prima riguarda le banche. La penalizzazione è più forte negli istituti che hanno già esperienza nel credito verde. Per gli autori l'IA funziona come un acceleratore e non come un sostituto della competenza: la banca esperta sa già dove guardare, e il sistema le conferma o le smentisce il sospetto. Quella inesperta non riconosce gli indizi, e gli strumenti che costruisce servono a poco.

La seconda riguarda le imprese, e dà il risultato meno prevedibile. L'effetto si osserva solo per le imprese non inquinanti. Per quelle che appartengono ai sedici settori classificati come inquinanti dal ministero cinese, invece, non c'è differenza fra le banche che usano l'IA e quelle che non la usano. Gli autori spiegano che le imprese inquinanti sono già sotto osservazione: il settore in cui operano segnala da solo un movente al greenwashing, e le banche ne controllano le dichiarazioni comunque. L'IA serve dunque dove nessun segnale esteriore invita a controllare. Non sostituisce il controllo dove già esiste: lo estende dove mancava.

Gli autori individuano due ragioni per cui le banche si comportano così. La prima è che l'IA permette di riconoscere meglio il rischio. Riportando a una forma comparabile informazioni ambientali di natura diversa, la banca capisce più facilmente quali dichiarazioni siano gonfiate e quanto sia probabile che quel prestito non venga restituito. La seconda è che la banca protegge sé stessa: se finanzia un'impresa che viene poi smentita sui propri impegni ambientali, rischia sanzioni e danno di reputazione.

L'implicazione non riguarda solo il credito. Finché il greenwashing viene individuato in uno studio o in una procedura di controllo, resta un fatto documentale. Quando invece chi presta denaro lo riconosce e lo traduce in un tasso più alto, diventa un costo. L'impresa

che appare più sostenibile di quanto sia non rischia più soltanto una sanzione futura e incerta: paga subito.

I limiti dichiarati sono due. Il primo conta oltre questo studio: gli autori non possono dire che cosa accada quando anche l'impresa usa l'IA per scrivere le proprie dichiarazioni, mentre la banca la usa per verificarle. Riconoscono che è la domanda successiva, ma i dati non permettono di rispondere. È lo stesso problema del §4.4.1, dove si osserva che le imprese cambiano linguaggio appena scoprono come vengono controllate. Il secondo limite è che l'analisi usa soltanto dati statistici, e uno studio qualitativo chiarirebbe meglio come le singole decisioni vengano prese.

A questo livello appartiene anche un risultato di Yadav e Asongu (2026), esaminati nel 4.3.4: il valore che il mercato attribuisce all'impresa cresce quando la performance ESG è verificata, ma non quando l'impresa dichiara di usare l'IA. I due studi dicono quindi la stessa cosa su mercati diversi. Chi presta denaro e chi compra azioni distinguono entrambi fra ciò che l'impresa dimostra e ciò che annuncia, e il premio va alla prova.

4.4.4 Il presidio dell'assurance

Il quarto livello è quello di chi certifica il documento prima che circoli. La certificazione dell'informativa non finanziaria, l'*assurance*, è un servizio che l'impresa acquista da un soggetto esterno perché attesti l'attendibilità di ciò che pubblica, e si è diffusa proprio come risposta alla scarsa credibilità dei report di sostenibilità. Che serva davvero a questo, però, non è ovvio.

Xing et al. (2026) impostano la questione come un'alternativa fra due ipotesi.

Nella prima, l'impresa che si sottopone alla certificazione ne ricava indicazioni su come migliorare, e riduce di conseguenza il divario fra ciò che dichiara e ciò che fa.

Nella seconda, la certificazione serve solo a ottenere un giudizio favorevole. L'impresa cerca il certificatore disposto a concederle e continua a comunicare come prima, con in più la firma di un terzo.

Lo studio mette alla prova le due ipotesi su un campione di 1.406 imprese quotate cinesi osservate dal 2007 al 2022, per 8.535 osservazioni impresa-anno, che scendono a 6.795 nelle stime perché la certificazione di un esercizio viene messa in relazione con il greenwashing dell'esercizio successivo. I dati sulla certificazione sono raccolti a mano dai singoli report, uno per uno. Il greenwashing, invece, è misurato come scarto fra le parole e i fatti: un algoritmo di apprendimento automatico classifica tutte le frasi dei report di

sostenibilità e separa le dichiarazioni che esprimono intenti e concetti, il *green talk*, da quelle che riferiscono azioni compiute, il *green walk*. Quanto più la prima componente supera la seconda, tanto più il documento è simbolico. È necessario specificare che in Cina la rendicontazione di sostenibilità non è obbligatoria, e la certificazione è a maggior ragione volontaria. Le imprese che la richiedono potrebbero quindi essere già le più attente, e l'effetto misurato rifletterebbe la loro natura anziché il servizio ricevuto.

Gli autori affrontano la questione confrontando fra loro imprese simili sotto ogni altro profilo e ricorrendo a tecniche che isolano la parte della scelta indipendente dalle caratteristiche dell'impresa; il risultato regge a entrambe le verifiche.

Il risultato è che la certificazione riduce il greenwashing negli esercizi successivi, in misura significativa e persistente. Gli autori definiscono il proprio esito sorprendente, perché contrasta con una parte della letteratura precedente secondo cui la certificazione potrebbe essere sfruttata per mascherare il greenwashing anziché per correggerlo.

Il canale che produce l'effetto è quello che gli autori chiamano *suggestion effect*. Le due vie attraverso cui passa sono misurabili, e gli autori le verificano entrambe: l'impresa certificata adotta con maggiore frequenza pratiche di gestione ambientale, e allinea la propria rendicontazione agli standard internazionali di riferimento. Sono questi due passaggi intermedi a produrre la riduzione del divario.

L'analisi per componenti conferma che il riequilibrio avviene dai due lati. Dopo la certificazione le espressioni simboliche diminuiscono, e le azioni sostanziali aumentano, misurate attraverso la spesa ambientale sostenuta e l'innovazione verde realizzata. L'impresa, in altre parole, non si limita a scrivere meno, fa di più.

L'effetto, però, non è uniforme, e le quattro analisi di eterogeneità sono la parte più utile dello studio. Le prime due riguardano chi certifica. La riduzione è più marcata quando l'incarico è affidato a una società di revisione contabile anziché a una società di consulenza specializzata in certificazioni, e quando il certificatore rientra fra le prime cento posizioni di una graduatoria di settore anziché fra gli operatori minori. Fin qui il risultato è quello che ci si attenderebbe: conta la competenza di chi guarda.

Le altre due riguardano invece il momento in cui la certificazione avviene, e sono meno prevedibili. L'effetto è più forte per la prima certificazione che l'impresa richiede rispetto a quelle successive, e anche quando la certificazione è discontinua, cioè non rinnovata ogni anno. Letti insieme al canale individuato dagli autori, i due risultati si spiegano. Se ciò che riduce il greenwashing è il suggerimento, il primo certificatore ha molto da

suggerire e il quinto assai meno, perché i miglioramenti più accessibili sono già stati realizzati. La certificazione che si ripete identica ogni anno diventa una procedura, e come tale perde efficacia. La conclusione che se ne ricava è quindi più circostanziata di un giudizio favorevole alla certificazione in quanto tale. Il controllo di terza parte riduce il greenwashing quando chi lo esercita ha competenza e reputazione da difendere, e quando ha ancora qualcosa da dire all'impresa che lo ha incaricato. I limiti dichiarati sono due. Il primo riguarda la natura dell'analisi: il campione è ampio, ma non consente di esaminare i casi estremi, per i quali gli autori indicano come strada la ricerca qualitativa o lo studio di casi singoli. Il secondo riguarda l'estensione dei risultati, che coprono le sole società quotate cinesi, mentre la certificazione si sta diffondendo in altri mercati e presso imprese non quotate, che i dati disponibili non permettono di studiare. Una precisazione è necessaria sul rapporto fra questo contributo e l'oggetto della tesi. Nello studio di Xing et al. l'apprendimento automatico serve ai ricercatori per misurare il greenwashing, mentre l'oggetto della valutazione è la certificazione umana. Il contributo risponde quindi alla terza domanda di ricerca per l'oggetto del proprio giudizio, non per la tecnica impiegata. Il collegamento con il resto del capitolo, però, esiste ed è stretto. Xing et al. mostrano che il controllo di terza parte funziona, e a quali condizioni. Harke (2026) avverte che, quando quel controllo viene automatizzato, rischia di conservarne la forma perdendone la sostanza, perché un giudizio prodotto da un sistema opaco non è verificabile. A questo si aggiunge l'osservazione di de Villiers et al. (2024b): le procedure di certificazione tradizionali si adattano male alla natura qualitativa dell'informativa di sostenibilità, e gli autori invitano chi le pratica a sperimentare l'IA, segnalando la questione ancora aperta della responsabilità quando nel processo intervengono sistemi automatici. Il livello della certificazione si trova dunque in una fase di passaggio: quello che oggi funziona è un controllo umano, e l'automazione che lo attende non ha ancora le garanzie necessarie a renderlo affidabile.

4.4.5 Il contesto istituzionale

L'ultimo livello è quello delle istituzioni pubbliche. Due studi del campione lo documentano, attraverso due forme diverse di intervento: la via *giudiziaria*, con i tribunali specializzati in materia ambientale, e quella *amministrativa*, con gli obiettivi ambientali assegnati ai governi locali.

Ke et al. (2026) sfruttano una circostanza: in Cina i tribunali ambientali sono stati istituiti in momenti e luoghi diversi. La progressione consente di trattare l'istituzione del tribunale come un esperimento quasi naturale. Si osserva il comportamento delle imprese prima e dopo, e lo si confronta con quello delle imprese dei territori in cui il tribunale non è ancora stato istituito. Il campione comprende le imprese quotate nei mercati A-share dal 2001 al 2024, e il greenwashing è misurato con un insieme di modelli linguistici diversi, che esaminano separatamente le informative ambientali e i cui giudizi vengono poi combinati. Il risultato è che, dopo l'istituzione del tribunale, le imprese riducono la retorica ambientale eccessiva. Gli autori sono però prudenti su due punti. Il primo è la causa: i due fatti si presentano insieme, ma non si può escludere del tutto che a produrre il cambiamento sia stato qualcos'altro avvenuto nello stesso periodo. Il secondo è ciò che l'indicatore misura, cioè la tendenza a esagerare e non una frode accertata; un punteggio alto segnala un sospetto, non una prova.

Gli autori individuano tre canali attraverso cui il tribunale produce quell'effetto, che non si riducono al semplice timore della condanna.

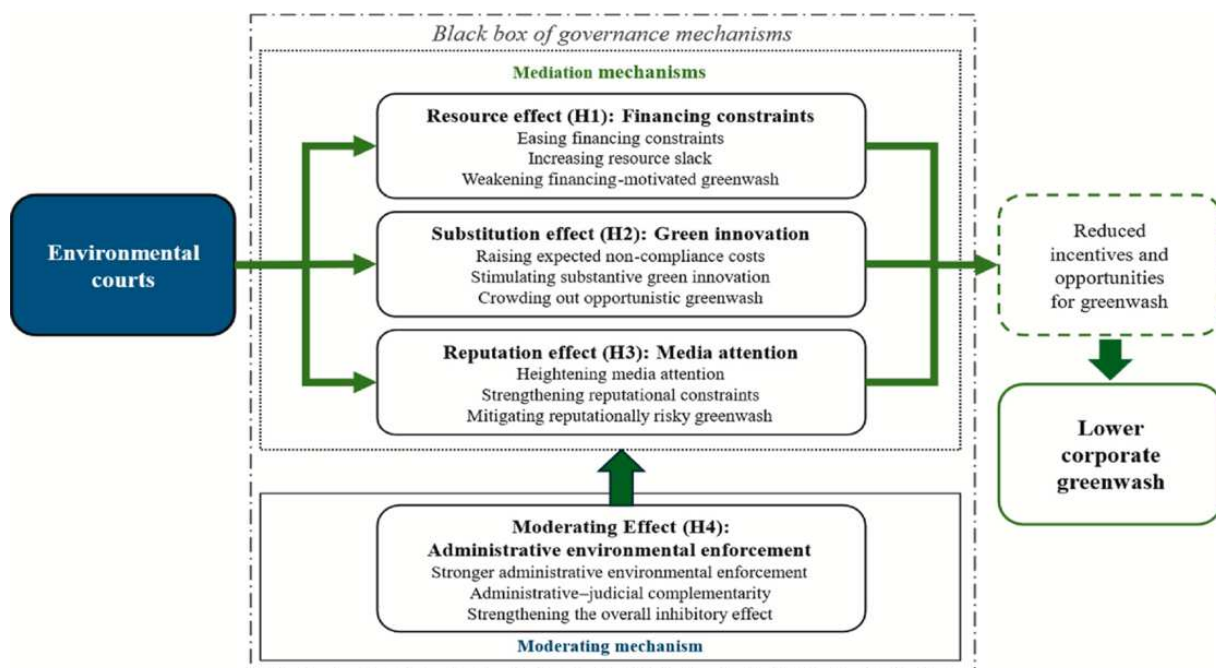
Il *primo* è l'allentamento dei vincoli finanziari: dove esiste un tribunale specializzato, il rischio ambientale diventa più leggibile per chi finanzia, e le imprese che migliorano davvero ottengono credito più facilmente.

Il *secondo* è l'aumento dell'innovazione verde: investire in tecnologie pulite diventa più conveniente che investire in comunicazione, perché il secondo tipo di investimento non protegge più dal contenzioso.

Il *terzo* è il rafforzamento della sorveglianza dei media, che dispongono di un interlocutore a cui le segnalazioni possono essere portate. I tre canali agiscono insieme e modificano il rapporto fra costi e benefici della comunicazione ambientale.

Gli autori rappresentano l'insieme in uno schema che parte dall'istituzione del tribunale, distingue i tre canali e li fa convergere nella riduzione degli incentivi e delle occasioni di greenwashing; nella parte inferiore compare l'intensità dell'applicazione amministrativa, il cui ruolo è discusso di seguito (figura 42).

Figura 42 - Quadro dei meccanismi attraverso cui i tribunali ambientali incidono sul greenwashing



Fonte: Ke et al., 2026

Un ulteriore risultato riguarda il rapporto fra i due tipi di intervento pubblico.

L'effetto dei tribunali è più forte dove l'applicazione amministrativa delle norme ambientali è più intensa: la via giudiziaria e quella amministrativa si rafforzano a vicenda anziché sostituirsi.

Zhou C. et al. (2026) esaminano invece uno strumento amministrativo, il sistema di responsabilità sugli obiettivi ambientali (*Environmental Target Responsibility system - ETR*). I governi locali cinesi ricevono obiettivi ambientali da raggiungere, e dal 2013 il loro conseguimento entra nella valutazione dei funzionari. Il sistema combina quindi due leve: una *disciplinare*, perché il mancato raggiungimento comporta conseguenze, e una di *carriera*, perché il raggiungimento favorisce l'avanzamento.

Gli autori studiano che cosa accada alle imprese quando la città in cui operano riceve obiettivi più ambiziosi, su un campione di imprese quotate appartenenti ai settori a forte impatto ambientale, fra il 2013 e il 2020.

Per misurare la qualità dell'informativa costruiscono un *indice di specificità*. Le frasi dei report di sostenibilità vengono classificate in due categorie, *informazione generica* e *informazione specifica*, e la classificazione è messa alla prova su diecimila frasi codificate

a mano. L'indice misura, dunque, quanto di ciò che l'impresa scrive sia circostanziato e controllabile, e non quanto scriva.

Il risultato è che obiettivi comunali più severi si accompagnano a un'informativa più specifica: le imprese sostituiscono le dichiarazioni generiche con informazioni verificabili. L'effetto è più marcato in tre gruppi, cioè i grandi inquinatori sottoposti a monitoraggio prioritario, le imprese politicamente più deboli e quelle che faticano ad accedere al credito. Per loro l'informativa specifica serve a conservare legittimità agli occhi dell'amministrazione e ad accedere alle risorse che l'amministrazione controlla. Chi ha un indice più alto ottiene infatti maggiori sussidi pubblici l'anno successivo, e il legame si rafforza dove gli obiettivi ambientali sono più severi. Chi è politicamente forte o già finanziato ha meno bisogno di dimostrare, e reagisce meno.

Il limite che gli autori dichiarano riguarda proprio questo: il sistema di responsabilità sugli obiettivi è un assetto di governo particolare, e i risultati non si estendono automaticamente ad altre economie emergenti.

In entrambi gli studi, come già in Xing et al., i modelli linguistici sono lo strumento con cui i ricercatori misurano il fenomeno e non la tecnologia sottoposta a giudizio. Ciò che i due contributi dimostrano è che il comportamento comunicativo delle imprese cambia quando esiste un'autorità capace di accertare e di sanzionare, e cambia di più dove quell'autorità è più vicina.

Il legame con l'oggetto della tesi emerge accostando questi risultati a due studi già esaminati. Yadav e Asongu (2026) mostrano che l'effetto dell'adozione dell'IA sul greenwashing si attenua dove la verifica indipendente e la regolazione operano insieme, e si manifesta in pieno dove sono deboli.

Zheng et al. (2026) mostrano che l'IA in mano a una banca produce un costo per l'impresa perché quella banca opera dentro un sistema di credito verde che le impone di guardare. Gli strumenti di rilevazione, in altre parole, producono conseguenze soltanto quando esiste qualcuno che ha il potere di trarne le conseguenze.

I tribunali ambientali e gli obiettivi comunali mostrano che cosa significhi, in concreto, disporre di quel potere.

4.4.6 Sintesi

Il paragrafo ha esaminato il controllo sull'IA e sul greenwashing su cinque livelli.

Nel flusso di lavoro, Harke (2026) propone di affidare al modello i grandi volumi e all'analista umano i casi ambigui, ma si tratta di una proposta, di cui l'autore stesso indica le difficoltà di attuazione.

Nel governo interno, Jiang et al. (2026) mostrano che l'IA riduce il greenwashing solo dove il comportamento dei dirigenti è vincolato, e non grazie agli amministratori indipendenti, che nel loro campione non producono effetti.

Sul mercato, le banche che adottano l'IA applicano tassi più alti alle imprese sospettate di greenwashing (Zheng et al., 2026), e gli investitori premiano la performance verificata, non l'annuncio dell'adozione (Yadav e Asongu, 2026). La certificazione umana riduce il greenwashing, ma perde efficacia quando si ripete identica (Xing et al., 2026).

Tribunali ambientali e obiettivi assegnati alle amministrazioni locali, infine, spingono le imprese verso un'informativa più verificabile (Ke et al., 2026; Zhou C. et al. 2026).

Letti insieme, questi studi fanno emergere uno schema che si ripete tre volte, in contesti e con metodi diversi. In Jiang et al. l'IA contiene il greenwashing dove esiste un controllo sui dirigenti; in Ke et al. i tribunali funzionano di più dove l'applicazione amministrativa è intensa; in Yadav e Asongu la verifica indipendente rende di più dove la regolazione è severa. In tutti e tre i casi un controllo produce effetti nella misura in cui ne esiste un altro che lo sostiene: l'efficacia di ciascun livello dipende dagli altri.

Il secondo risultato conferma lo squilibrio annunciato in apertura. L'evidenza sull'IA come strumento di controllo si concentra sul governo interno e sul mercato. Mancano proprio i due livelli in cui l'IA interverrebbe più direttamente sul documento: il flusso di lavoro, per il quale esiste solo una proposta, e la certificazione, per la quale esistono uno studio sulla certificazione umana e due riflessioni teoriche sui rischi della sua automazione, di Harke e di de Villiers et al. (2024b). Nessuno studio osserva come chi certifica usi l'IA.

Gu et al. (2026) progettano il proprio sistema come supporto per i revisori e Jakubczak et al. (2026) mettono alla prova strumenti accessibili ai professionisti, ma in entrambi i casi a usarli sono i ricercatori. Questa lacuna si accosta a quella del 4.3.5, dove si è osservato che nessuno studio riconosce se un report reale sia stato scritto con l'IA. Le due mancanze hanno un'origine comune: la letteratura osserva l'IA sempre nelle mani di un solo attore alla volta (l'impresa, il ricercatore, la banca) e mai nella situazione in cui una parte la usa per scrivere e l'altra per verificare. La risposta alla terza domanda di ricerca è quindi

duplice. Il controllo umano e istituzionale limita i rischi dell'IA, in modo documentato quando più livelli agiscono insieme. Ma i controlli che dovrebbero governare l'IA stessa, cioè l'intervento umano nella redazione e la certificazione assistita da sistemi automatici, restano proposte: la letteratura sa indicare dove dovrebbero collocarsi, non ancora quanto funzionino.

4.5 Considerazioni finali

L'analisi della letteratura ha permesso di rispondere alle tre RQ con un grado di approfondimento diverso. La RQ1 è quella su cui la letteratura è più ampia, con trentasei studi primari.

Sul versante della rilevazione, ventitré studi descrivono una progressione in cui ogni generazione di strumenti corregge un limite di quella precedente: i dizionari contano le parole senza coglierne il contesto, i modelli pre-addestrati leggono la frase ma restano dentro il documento, i modelli generativi e i sistemi agentici lo confrontano con fonti esterne. I risultati più affidabili, però, vengono dagli strumenti che confrontano ciò che il testo dichiara con un riferimento che l'impresa non controlla, anche quando sono tecnicamente meno avanzati: le sanzioni ambientali ricevute negli anni successivi (Xu et al., 2026), il giudizio di esperti esterni (Colesanti Senni et al., 2025), le emissioni effettivamente prodotte (dos Santos et al., 2025). Sul versante del contrasto, dodici dei tredici studi sull'effetto dell'adozione concludono che l'IA riduca il greenwashing. Tutti e dodici, però, osservano imprese quotate cinesi, ricorrono in larga parte alla stessa banca dati e misurano il greenwashing a partire da rating ESG che premiano più gli impegni dichiarati che i risultati conseguiti. L'unico di questi studi condotto su più paesi, quello di Yadav e Asongu (2026), trova l'effetto opposto.

La RQ2 raccoglie sedici studi, ma dodici sono gli stessi studi econometrici della RQ1, che concludono che l'IA limiti il greenwashing.

I quattro che sostengono che l'IA agevoli il greenwashing portano ciascuno un tipo di evidenza diverso: un quadro concettuale (de Villiers et al., 2024b), un esperimento su quattordici studenti (Dimes et al., 2026), un caso di studio su una singola filiera (Vilanova e Bazanini, 2026), un'analisi econometrica su sei economie asiatiche emergenti (Yadav e Asongu, 2026). Quello che documentano insieme sono due modi distinti di facilitare il greenwashing: per commissione, quando il modello genera il testo che abbellisce, e per

omissione, quando la sola dichiarazione del suo impiego legittima l'impresa. Nessuno di essi, però, osserva su report reali se un documento sia stato scritto con l'IA.

La RQ3 è la meno coperta della letteratura. Il suo risultato più solido riguarda il rapporto fra i diversi controlli: in Jiang et al. (2026), in Ke et al. (2026) e in Yadav e Asongu (2026), con contesti e metodi diversi, ciascun controllo produce effetti nella misura in cui ne esiste un altro che lo sostiene. Restano invece senza verifica i controlli che dovrebbero governare l'IA stessa. L'intervento umano nel processo di redazione è una proposta (Harke, 2026), e nessuno studio osserva come chi certifica usi l'IA.

4.5.1 I fattori che distinguono il ruolo deterrente da quello agevolatore

Nessuno degli studi esaminati spiega da solo perché l'IA riduca il greenwashing in alcuni casi e lo faciliti in altri. Letti insieme, però, fanno emergere quattro condizioni:

- *chi impiega la tecnologia.* L'effetto dell'IA cambia con il soggetto che la usa. Quando è impiegata dall'esterno per analizzare i report, come negli studi del 4.2, serve a individuare il greenwashing. Quando è impiegata da chi deve decidere se finanziare un'impresa, si traduce in un costo per chi comunica in modo ingannevole. Quando è l'impresa stessa a usarla, l'esito resta aperto: la stessa tecnologia può rendere l'informativa più accurata o più persuasiva, e gli studi del 4.3 mostrano che cosa accada quando prevale il secondo uso;
- *la credibilità dei controlli.* Il secondo fattore è il contesto in cui l'impresa opera. Dove i controlli sono deboli, l'adozione dell'IA si accompagna a più greenwashing; dove sono credibili, a meno. Ma un controllo più intenso non è necessariamente più efficace: quando gli obblighi diventano molto onerosi, l'IA può servire a soddisfarli soltanto nella forma. Poiché questo cambiamento di segno si osserva anche all'interno di uno stesso mercato, il contrasto fra gli studi condotti in Cina e quello condotto su più paesi si spiega con il contesto istituzionale più che con la geografia;
- *la presenza di un controllo umano.* Il terzo fattore è l'esistenza, nel processo, di qualcuno che eserciti un giudizio sugli esiti del modello. La sola presenza non basta, perché chi può correggere un testo generato non è detto che lo faccia. Il controllo umano funziona quando chi lo esercita ha ragioni e strumenti per dubitare, e quando anche il suo comportamento è sottoposto a vincoli;

- *il grado di standardizzazione del linguaggio*. Il quarto fattore riguarda la materia su cui l'IA lavora. Dove l'informazione è già codificata in metriche e schemi comuni la rilevazione automatica è molto accurata; dove il racconto è discorsivo e ogni impresa scrive a modo proprio, perde efficacia. È in quello spazio che il greenwashing trova il suo margine maggiore.

I quattro fattori non agiscono separatamente. L'IA limita il greenwashing quando è usata da chi controlla, in un contesto che rende il controllo credibile, con un giudizio umano capace di dubitare e su informazioni abbastanza codificate da poter essere verificate. Quando una di queste condizioni manca, la stessa tecnologia può servire a scrivere meglio anziché a dire il vero.

4.5.2 Un confronto con altre revisioni della letteratura

Le sette SLR incluse nel campione di articoli selezionati non sono state usate come evidenza nei paragrafi precedenti. Qui vengono discusse con finalità di comparazione dei risultati ottenuti.

Moodaley e Telukdarie (2023) hanno evidenziato come l'impiego dell'IA per contrastare il greenwashing nella rendicontazione di sostenibilità costituisca un ambito di ricerca ancora limitatamente esplorato, sottolineandone il potenziale per gli studi futuri.

Tre anni dopo, Vilchez Olivares e Astorga De La Cruz (2026) descrivono la rilevazione del greenwashing come uno dei fronti di ricerca più attivi nell'applicazione dell'IA all'informativa ESG, e ne attribuiscono la crescita anche agli obblighi di verifica introdotti dalla CSRD. Entrambe le revisioni confermano quanto emerso nella sezione 4.2, cioè che la letteratura si è sviluppata soprattutto intorno alla capacità dell'IA di individuare il greenwashing. Nessuna delle due, però, si chiede se l'IA possa anche facilitarlo.

Mousavian Anaraki et al. (2025) riconoscono ai modelli linguistici un'elevata accuratezza nell'individuare le dichiarazioni fuorvianti, in linea con i risultati della sezione 4.2.

Lublóy et al. (2025) mettono però in luce il punto debole: la maggior parte delle misure di greenwashing usate negli studi empirici segnala un greenwashing possibile, non accertato, e la più diffusa, ossia lo scarto fra quantità di informativa ESG e performance ESG, presenta carenze rilevanti. È la stessa fragilità individuata in questa revisione negli studi sull'effetto dell'adozione dell'IA per la rendicontazione di sostenibilità.

Janik e Ryszko (2026) aggiungono che supporre che le imprese seguano strategie di rendicontazione uniformi è una semplificazione eccessiva: un'osservazione coerente con l'idea che l'esito dell'IA dipenda dalle condizioni in cui l'impresa opera. Le teorie a cui riconducono le strategie di greenwashing, infine, sono le teorie organizzative, le teorie economico – relazionali e quelle comportamentali, nonché le stesse discusse nella sezione 2.5. Le due SLR che si avvicinano ai risultati di questa revisione, sono quelle di Liu et al. (2026) e di Łudzińska (2025). I primi concludono che manca un'evidenza causale solida sul fatto che l'adozione dell'IA da parte delle imprese migliori la credibilità dell'informativa, e segnalano che l'IA può introdurre rischi nuovi: narrazioni levigate ma superficiali e opacità degli algoritmi nei processi di revisione. Sono, in forma sintetica, i meccanismi documentati nella sezione 4.3 e il rischio discusso nella sezione 4.4.1.

Łudzińska trova una relazione prevalentemente positiva fra l'adozione dell'IA e performance ESG, e include la riduzione del greenwashing fra i benefici, ma conclude che la sua efficacia dipenda fortemente da fattori contestuali: è la conclusione del 4.5.1.

Le revisioni confermano quindi i principali risultati di questo capitolo: la rilevazione è la parte più sviluppata della letteratura, l'effetto dell'adozione poggia su misure fragili, e l'esito dell'IA dipende dal contesto. Una sola di esse, quella di Liu, Yuan e Zhu, accenna ai rischi che l'IA introduce. Nessuna si chiede esplicitamente se l'IA scoraggi il greenwashing o lo faciliti.

4.5.3 I limiti della letteratura

La letteratura evidenzia anche una serie di limiti riguardo lo studio del greenwashing e dell'IA:

- *la concentrazione geografica.* Quasi tutti gli studi che stimano l'effetto dell'adozione dell'IA sul greenwashing osservano imprese quotate cinesi. L'unico condotto su più paesi arriva al risultato opposto. Il 4.5.1 ha mostrato che l'esito dipende dal contesto istituzionale. Per questo il risultato prevalente non si può estendere ad altri mercati senza nuove verifiche;
- *la circolarità della misura.* Molti studi misurano il greenwashing con i rating ESG. Questi rating però danno un peso rilevante agli impegni che l'impresa dichiara. Una parte della misura dipende quindi dalle stesse dichiarazioni che dovrebbe valutare. Risulta, quindi, difficile stabilire se l'IA riduca il greenwashing o cambi soltanto il modo in cui l'impresa comunica;

- *lo squilibrio fra rilevazione ed effetto.* La letteratura ha prodotto molti strumenti per riconoscere il greenwashing, ma molti meno per verificare che cosa accade quando le imprese usano l'IA nella redazione dei report. L'ipotesi che l'IA possa facilitare il greenwashing si basa su pochi studi, ciascuno con un tipo di evidenza diverso. Il 4.3.5 ne ha indicato la ragione: i processi interni di redazione non sono osservabili dall'esterno;
- *un ridotto focus sui processi organizzativi.* Quasi tutti gli studi analizzano i documenti già pubblicati. Pochi osservano come il testo viene scritto, chi interviene e come chi controlla usa l'IA;
- *l'interazione inesplorata tra redazione e controllo.* Le imprese possono usare l'IA per scrivere i report, mentre revisori e autorità la usano per controllarli. Nessuno studio osserva cosa accade quando i due usi si incontrano. Non è ancora noto se gli strumenti di controllo riconoscano il greenwashing nei testi scritti con l'IA, né se le imprese imparino a scrivere testi che sfuggono ai controlli.

Conclusioni

La presente tesi si è proposta di indagare, attraverso una revisione sistematica della letteratura condotta su un corpus di 51 contributi, se l'IA applicata ai report di sostenibilità contribuisca a contrastare il greenwashing oppure, al contrario, ne favorisca la diffusione. Nel complesso, i risultati indicano che l'IA non esercita sul greenwashing un effetto univoco: l'esito dipende dal soggetto che la impiega e dal sistema di controlli entro cui viene utilizzata, e al variare di tali condizioni la medesima tecnologia può ridurre il fenomeno oppure amplificarlo. Il lavoro ha cercato di rispondere a tre domande di ricerca. La *prima* riguardava l'impiego dell'IA per individuare e contrastare il greenwashing, e rappresenta l'ambito in cui la letteratura risulta più sviluppata. Sul piano della rilevazione, le ricerche mostrano che gli strumenti si sono evoluti per generazioni successive, ciascuna nata per superare i limiti della precedente. Le tecniche lessicali si limitano a contare le occorrenze dei termini senza coglierne il contesto. I modelli pre-addestrati analizzano la frase nel suo insieme e distinguono gli impegni dai risultati conseguiti. I modelli generativi e i sistemi agentici, infine, confrontano il contenuto del report con fonti esterne. Alcuni studi, inoltre, confrontano il segnale testuale con un riferimento esterno all'impresa, come le sanzioni ambientali ricevute negli anni successivi (Xu et al., 2026), il giudizio di esperti esterni (Colesanti Senni et al., 2025) o le emissioni prodotte (dos Santos et al., 2025). Sul piano del contrasto, dodici dei tredici studi sull'effetto dell'adozione dell'IA concludono che essa riduce il greenwashing, grazie a sistemi di controllo interno più efficaci, una maggiore efficienza e attenzione da parte degli analisti. La portata di questo risultato è tuttavia circoscritta, poichè gli studi esaminano esclusivamente imprese quotate cinesi, ricorrono in larga parte alla medesima banca dati e misurano per lo più il greenwashing come differenza tra l'informativa ambientale e un rating ESG. Si tratta tuttavia di una misura soggetta a possibili bias, poichè i rating ESG attribuiscono un peso maggiore agli impegni dichiarati che ai risultati conseguiti (Ferro et al., 2025). L'unico studio condotto su più paesi (Yadav e Asongu, 2026) rileva, al contrario, un'associazione positiva tra adozione dell'IA e greenwashing. La capacità dell'IA di riconoscere il greenwashing risulta quindi documentata, mentre la sua capacità di ridurlo poggia su una base empirica ancora ristretta.

La *seconda domanda di ricerca* si chiede se, e attraverso quali meccanismi, l'IA possa agire da agevolatore del greenwashing, rendendo la comunicazione ingannevole più efficace e più difficile da individuare. La letteratura vi ha dedicato un'attenzione molto minore. Dei

sedici studi primari che toccano la questione, dodici sono gli studi appena richiamati a proposito della prima domanda di ricerca. Soltanto quattro documentano l'ipotesi opposta, ciascuno con un tipo di evidenza differente: un'elaborazione concettuale (de Villiers et al., 2024b), un esperimento (Dimes et al., 2026), uno studio di caso (Vilanova e Bazanini, 2026) e un'analisi econometrica su sei economie asiatiche emergenti (Yadav e Asongu, 2026). Poiché nessuno degli studi misura ciò che misurano gli altri, essi non possono confermarsi a vicenda; letti insieme, consentono però di individuare due distinti meccanismi di agevolazione. Il primo meccanismo opera per commissione (*by commission*) e riguarda l'IA generativa impiegata nella redazione dei report. Un modello generativo apprende da ampi insiemi di documenti, tra cui i report di sostenibilità già pubblicati. Poiché il greenwashing è diffuso, i dati di addestramento comprendono sia resoconti accurati sia documenti che enfatizzano o selezionano le informazioni in modo opportunistico. Quando redige una sezione di report, il modello linguistico IA restituisce un testo conforme a ciò che ha appreso e riproduce così anche il greenwashing di quei documenti, senza che l'impresa lo scelga deliberatamente (de Villiers et al., 2024b). L'IA attenua dunque le criticità e rende la comunicazione più persuasiva agendo sul tono più che sul contenuto. Il secondo meccanismo opera per omissione (*by omission*) e non richiede che l'IA produca alcun testo. Nel caso studiato da Vilanova e Bazanini (2026), riguardante la filiera bovina brasiliana, le imprese a valle della filiera dichiarano di impiegare sistemi di monitoraggio satellitare e algoritmi di tracciabilità che esistono realmente, ma che coprono soltanto i segmenti in cui sono installati e mancano presso gli allevatori. Poiché i documenti non precisano dove la copertura si interrompa, il lettore ne ricava un'immagine di tracciabilità estesa all'intera filiera. La distorsione non nasce da un dato falso: è la sola dichiarazione della tecnologia a legittimare l'impresa, mentre ciò che non viene misurato resta invisibile. I meccanismi di agevolazione risultano quindi descritti e in parte osservati, ma la loro diffusione rimane poco esplorata, perché nessuno studio del campione verifica su report reali se un documento sia stato redatto con l'IA e con quali conseguenze sulla sua attendibilità. Gli strumenti per collegare i segnali testuali agli esiti esistono già; manca la possibilità di attribuire con certezza la paternità del testo. La *terza domanda di ricerca* riguarda, invece, la capacità del controllo umano e istituzionale di limitare i rischi algoritmici e di preservare la credibilità dei report, ed è quella su cui la letteratura offre meno evidenze. Il risultato più solido emerge da tre studi con contesti e metodi differenti (Jiang et al., 2026; Ke et al., 2026; Yadav e Asongu, 2026):

ciascun controllo produce effetti quando è sostenuto da altri controlli, come accade quando la supervisione sui dirigenti accompagna l'adozione dell'IA o quando la verifica indipendente si affianca a una regolazione stringente. Restano invece prive di verifica empirica le forme di controllo destinate a governare l'IA stessa.

La definizione di procedure strutturate di verifica umana dei contenuti generati dall'IA a supporto della redazione dei report di sostenibilità costituisce, allo stato attuale, un tema ancora poco studiato (Harke, 2026). Inoltre, mancano ancora ricerche su come i soggetti incaricati dell'assurance impieghino l'IA nella propria attività.

Dall'insieme dei risultati emergono quattro condizioni da cui dipende il ruolo dell'IA: il soggetto che la impiega, la credibilità dei controlli nel contesto istituzionale in cui l'impresa opera, la presenza di un giudizio umano capace di mettere in discussione gli esiti del modello e il grado di standardizzazione delle informazioni analizzate. L'IA tende a limitare il greenwashing quando queste condizioni ricorrono insieme; se anche una sola viene meno, la stessa tecnologia può servire a rendere i report più persuasivi senza renderli più veritieri.

Il confronto con le altre revisioni sistematiche del corpus conferma questi risultati: la rilevazione è l'area più sviluppata della letteratura (Mousavian Anaraki et al., 2025; Vilchez Olivares e Astorga De La Cruz, 2026), le misure usate per stimare l'effetto dell'adozione dell'IA presentano limiti metodologici (Lublóy et al., 2025) e l'esito dipende in misura rilevante dal contesto (Łudzińska, 2025).

Liu et al. (2026) accennano ai rischi che l'IA può introdurre nella rendicontazione, ma nessuna di queste revisioni si chiede esplicitamente se l'IA scoraggi o faciliti il greenwashing. Rispetto a esse, la presente tesi esamina anche il ruolo agevolatore della tecnologia e individua le condizioni che spiegano la divergenza dei risultati.

La revisione, infine, mette in luce alcuni gap che delineano possibili direzioni per la ricerca futura. In primo luogo, emerge la necessità di studi focalizzati su più paesi e contesti istituzionali, per verificare se i risultati ottenuti sulle imprese quotate cinesi valgano anche altrove, e misure di greenwashing indipendenti dai rating ESG.

In secondo luogo, sarebbe opportuno affiancare all'analisi dei documenti pubblicati lo studio dei processi con cui i report vengono redatti e certificati. Ciò permetterebbe di comprendere che cosa accada quando l'IA è impiegata sia da chi redige i report sia da chi li controlla, e se gli strumenti di controllo sappiano riconoscere il greenwashing nei testi generati dall'IA.

Bibliografia

Adadi, A. & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, Vol. 6, pp. 52138-52160.

Alhaddi, H. (2015). Triple Bottom Line and Sustainability: A Literature Review. *Business and Management Studies*, Vol. 1, No. 2, pp. 6-10.

Bai, C., Yao, D. & Xue, Q. (2025). Does Artificial Intelligence Suppress Firms' Greenwashing Behavior? Evidence from Robot Adoption in China. *Energy Economics*, Vol. 142, art. 108168.

Bandura, A. (1999). Moral Disengagement in the Perpetration of Inhumanities. *Personality and Social Psychology Review*, Vol. 3, No. 3, pp. 193-209.

Barney, J. B. (2001). Resource-Based Theories of Competitive Advantage: A Ten-Year Retrospective on the Resource-Based View. *Journal of Management*, Vol. 27, No. 6, pp. 643-650.

Barney, J. B., Ketchen, D. J. & Wright, M. (2021). Resource-Based Theory and the Value Creation Framework. *Journal of Management*, Vol. 47, No. 7, pp. 1936-1955.

Bennaceur, F., Bendob, A. & Othman, A. H. A. (2026). Leveraging AI for ESG Disclosures: A Deep Dive into UAE's Islamic and Conventional Banking Sectors. *Journal of Financial Reporting and Accounting*.

Bergougui, B., Sulimany, H. G. H. & Alharbi, A. A. (2026). Combating ESG Greenwashing Through AI Models: Evidence from Disaggregated AI Technologies, Mechanisms, and Thresholds. *Corporate Social Responsibility and Environmental Management*.

Bingler, J. A., Kraus, M., Leippold, M. & Webersinke, N. (2022). Cheap Talk and Cherry-Picking: What ClimateBert Has to Say on Corporate Climate Risk Disclosures. *Finance Research Letters*, Vol. 47, art. 102776.

Bingler, J. A., Kraus, M., Leippold, M. & Webersinke, N. (2024). How Cheap Talk in Climate Disclosures Relates to Climate Initiatives, Corporate Emissions, and Reputation Risk. *Journal of Banking & Finance*, Vol. 164, art. 107191.

Bromley, P. & Powell, W. W. (2012). From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World. *Academy of Management Annals*, Vol. 6, No. 1, pp. 483-530.

Carreño, I. (2023). To Address “Greenwashing” and Misleading Environmental Claims, the European Commission Publishes a Proposal on “Green Claims” and Their Substantiation. *European Journal of Risk Regulation*, Vol. 14, No. 3, pp. 607-611.

Chamola, V., Hassija, V., Sulthana, A. R., Ghosh, D., Dhingra, D. & Sikdar, B. (2023). A Review of Trustworthy and Explainable Artificial Intelligence (XAI). *IEEE Access*, Vol. 11, pp. 78994-79015.

Cheng, Z. & Yang, G. (2026). The Impact of Artificial Intelligence on Corporate ESG Greenwashing. *Socio-Economic Planning Sciences*, Vol. 103, art. 102351.

Cho, C. H. & Patten, D. M. (2007). The Role of Environmental Disclosures as Tools of Legitimacy: A Research Note. *Accounting, Organizations and Society*, Vol. 32, No. 7-8, pp. 639-647.

Colesanti Senni, C., Schimanski, T., Bingler, J., Ni, J. & Leippold, M. (2025). Using AI to Assess Corporate Climate Transition Disclosures. *Environmental Research Communications*, Vol. 7, No. 2, art. 021010.

Connelly, B. L., Certo, S. T., Ireland, R. D. & Reutzel, C. R. (2011). Signaling Theory: A Review and Assessment. *Journal of Management*, Vol. 37, No. 1, pp. 39-67.

Connelly, B. L., Certo, S. T., Reutzel, C. R., DesJardine, M. R. & Zhou, Y. S. (2025). Signaling Theory: State of the Theory and Its Future. *Journal of Management*, Vol. 51, No. 1, pp. 24-61.

Cors, M. S. & Thiébaut, R. (2025). Artificial Intelligence and the Impact of the EU AI Act in Business Organizations. *AI Magazine*, Vol. 46, No. 4, art. e70039.

Das, A. & Rad, P. (2020). Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey. *arXiv preprint* arXiv:2006.11371.

Davidescu, A. A., Manta, E. M., Bîrlan, I., Miler, A.-M. & Niță, S.-C. (2026). Detecting Greenwashing in ESG Disclosure: An NLP-Based Analysis of Central and Eastern European Firms. *Sustainability*, Vol. 18, No. 3, art. 1486.

de Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B. & Soares, G. R. d. L. (2020). Concepts and Forms of Greenwashing: A Systematic Review. *Environmental Sciences Europe*, Vol. 32, art. 19.

de Villiers, C., Dimes, R., La Torre, M. & Molinari, M. (2024a). The International Sustainability Standards Board's (ISSB) Past, Present, and Future: Critical Reflections and a Research Agenda. *Pacific Accounting Review*, Vol. 36, No. 2, pp. 255-273.

de Villiers, C., Dimes, R. & Molinari, M. (2024b). How Will AI Text Generation and Processing Impact Sustainability Reporting? Critical Analysis, a Conceptual Framework and Avenues for Future Research. *Sustainability Accounting, Management and Policy Journal*, Vol. 15, No. 1, pp. 96-118.

Delmas, M. A. & Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, Vol. 54, No. 1, pp. 64-87.

Dimes, R., de Villiers, C., Denny, P. & Leinonen, J. (2026). The Impact of Generative AI on the Perpetuation and Detection of Greenwashing in Sustainability Reports: Insights from an Experiment. *Business Strategy and the Environment*, pp. 1-16.

Donaldson, T. & Preston, L. E. (1995). The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *The Academy of Management Review*, Vol. 20, No. 1, pp. 65-91.

dos Santos, I. A. A., de Araújo Moxotó, A. C. & de Souza Mota, E. (2025). Greenwashing in Brazilian Corporations: A Machine Learning Approach to Unmask the Discourse-Practice Gap. *Corporate Social Responsibility and Environmental Management*, Vol. 32, No. 6, pp. 8214-8224.

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., et al. (2023). Explainable AI (XAI): Core Ideas, Techniques, and Solutions. *ACM Computing Surveys*, Vol. 55, No. 9, pp. 1-33.

Dyllick, T. & Hockerts, K. (2002). Beyond the Business Case for Corporate Sustainability. *Business Strategy and the Environment*, Vol. 11, No. 2, pp. 130-141.

EFRAG (2025a). *Basis for Conclusions: Draft Amended ESRS*. EFRAG, dicembre 2025.

EFRAG (2025b). *Draft ESRS 1: General Requirements*. EFRAG, novembre 2025.

EFRAG (2025c). *Draft ESRS 2: General Disclosures*. EFRAG, novembre 2025.

Eze, M. (2025). Digital Greenwashing in the Age of Sustainability Marketing: A Meta-Analysis of Consumer Perception, Detection, and Backlash. *International Journal of Marketing Studies*, Vol. 17, No. 2, pp. 32-50.

Ferro, A., Marazzina, D. & Stocco, D. (2025). Uncovering ESG Ratings: The (Im)Balance of Aspirational and Performance Features. *Corporate Social Responsibility and Environmental Management*, Vol. 32, No. 5, pp. 5895-5917.

Forliano, C., Battisti, E., de Bernardi, P. & Kliestik, T. (2025). Mapping the Greenwashing Research Landscape: A Theoretical and Field Analysis. *Review of Managerial Science*, Vol. 19, No. 11, pp. 3407-3456.

Fornasari, T. & Traversi, M. (2024). The Impact of the CSRD and the ESRS on Non-Financial Disclosure. *Symphonya. Emerging Issues in Management*, No. 1, pp. 117-133.

Freeman, R. E., Phillips, R. & Sisodia, R. (2020). Tensions in Stakeholder Theory. *Business & Society*, Vol. 59, No. 2, pp. 213-231.

Freeman, R. E., Wicks, A. C. & Parmar, B. (2004). Stakeholder Theory and “The Corporate Objective Revisited”. *Organization Science*, Vol. 15, No. 3, pp. 364-369.

Friedman, M. (1970). The Social Responsibility of Business Is to Increase Its Profits. *The New York Times Magazine*, 13 settembre 1970, pp. 32-33.

Global Reporting Initiative (2021). *GRI 1: Foundation 2021*. Global Sustainability Standards Board.

Gorovaia, N. & Makrominas, M. (2025). Identifying Greenwashing in Corporate-Social Responsibility Reports Using Natural-Language Processing. *European Financial Management*, Vol. 31, pp. 427-462.

Gu, Y., Jiang, L., Dai, J. & Vasarhelyi, M. (2026). An LLM-Based Agentic System for Greenwashing Detection. *International Journal of Accounting Information Systems*, Vol. 57, art. 100780.

Harke, G. (2026). Algorithmic Accountability in Greenwashing Detection: Hybrid Governance in ESG Auditing. In Phan, T. T. N. (A cura di), *Impact of Market Sentiment on Green Valuations*, IGI Global, pp. 167-202.

He, L.-Y. & Wang, L. (2025). Can Artificial Intelligence Curb Greenwashing? Firm-Level Evidence Based on Large Language Model. *Energy Economics*, Vol. 152, art. 108954.

Jakubczak, J. K., Chmielewska-Muciek, D. & Iwanicka, K. (2026). Evaluating Multimodal AI for Greenwashing Detection: A Comparative Analysis of ChatGPT, Claude, and Gemini in ESG Reports. *Sustainability*, Vol. 18, No. 1, art. 236.

Janik, A. & Ryszko, A. (2026). Greenwashing in Sustainability Reporting: A Systematic Literature Review of Strategic Typologies and Content-Analysis-Based Measurement Approaches. *Sustainability*, Vol. 18, No. 1, art. 17.

Jiang, N., Wu, G., Zou, W. & Chen, X. (2026). Responsible Use of Artificial Intelligence and the Mitigation of Corporate Greenwashing: The Role of Executive Oversight. *Socio-Economic Planning Sciences*, Vol. 105, art. 102436.

Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, Vol. 47, No. 2, pp. 263-292.

Karasek, R. & Bryant, P. (2012). Signaling Theory: Past, Present, and Future. *Academy of Strategic Management Journal*, Vol. 11, No. 1, pp. 91-99.

Ke, Z., Gao, D., Zhong, X. & Wang, X. (2026). Can Environmental Courts Inhibit Corporate Greenwash? Evidence from Heterogeneous LLMs. *International Review of Economics & Finance*, Vol. 110, art. 105502.

Khan, M. K., Huo, C., Zahid, R. M. A. & Maqsood, U. S. (2024). The Automated Sustainability Auditor: Does Artificial Intelligence Curtail Greenwashing Behavior in Chinese Firms? *Business Strategy and the Environment*, Vol. 33, No. 8, pp. 9015-9039.

Kılınç, Y., İnce, M. R. & Badem, A. C. (2026). A Multi-Dimensional Textual Framework for Detecting Greenwashing in Sustainability Reporting. *Discover Sustainability*, Vol. 7, art. 482.

Kim, B. J., Jeong, S., Cho, B. K. & Chung, J. B. (2025). AI Governance in the Context of the EU AI Act. *IEEE Access*, Vol. 13, pp. 144126-144142.

Kücükgül, E., Cerin, P. & Liu, Y. (2022). Enhancing the Value of Corporate Sustainability: An Approach for Aligning Multiple SDGs Guides on Reporting. *Journal of Cleaner Production*, Vol. 333, art. 130005.

Kunisch, S., Denyer, D., Bartunek, J. M., Menz, M. & Cardinal, L. B. (2023). Review Research as Scientific Inquiry. *Organizational Research Methods*, Vol. 26, No. 1, pp. 3-45.

Lagasio, V. (2024). ESG-Washing Detection in Corporate Sustainability Reports. *International Review of Financial Analysis*, Vol. 96, art. 103742.

Lei, X., Kocoglu, M. & Nikbakht, E. (2026). The Announcement Effect: How AI Adoption Reduces Corporate ESG Greenwashing. *Finance Research Letters*, Vol. 99, art. 109890.

Li, T.-T., Wang, K., Sueyoshi, T. & Wang, D. D. (2021). ESG: Research Progress and Future Prospects. *Sustainability*, Vol. 13, No. 21, art. 11663.

Liu, J., Yuan, Y. & Zhu, Z. (2026). The Role of Artificial Intelligence in Enhancing ESG Disclosure Quality in Accounting. *Journal of Risk and Financial Management*, Vol. 19, No. 1, art. 58.

Lockett, A., O'Shea, R. P. & Wright, M. (2008). The Development of the Resource-Based View: Reflections from Birger Wernerfelt. *Organization Studies*, Vol. 29, No. 8-9, pp. 1125-1141.

Lockett, A., Thompson, S. & Morgenstern, U. (2009). The Development of the Resource-Based View of the Firm: A Critical Appraisal. *International Journal of Management Reviews*, Vol. 11, No. 1, pp. 9-28.

Lorkiewicz, W. & Wasilewski, A. (2026). A Hybrid Intelligence Approach for Assessing Packaging Sustainability Claims in E-Commerce. *Applied Soft Computing*, Vol. 201, art. 115504.

Lublóy, Á., Keresztúri, J. L. & Berlinger, E. (2025). Quantifying Firm-Level Greenwashing: A Systematic Literature Review. *Journal of Environmental Management*, Vol. 373, art. 123399.

Łudzińska, K. (2025). The Firm-Level AI-ESG Performance Nexus: A Rapid Literature Review and Future Research Agenda. *Economics and Environment*, Vol. 95, No. 4, art. 1170.

Luisi, F. & Cucaro, O. M. S. (2026). Talking the Walk: ESG Discourse as Strategic Dynamic Capability. *Business Strategy and the Environment*, pp. 1-17.

Lyon, T. P. & Maxwell, J. W. (2006). *Greenwash: Corporate Environmental Disclosure under Threat of Audit*. Ross School of Business Working Paper Series, No. 1055. University of Michigan.

Maraolo, A. E. (2021). Una bussola per le revisioni sistematiche: la versione italiana della nuova edizione del PRISMA statement. *Giornale Italiano di Health Technology Assessment and Delivery (GIHTAD)*, Vol. 14, No. 1.

Moneva, J. M., Archel, P. & Correa, C. (2006). GRI and the Camouflaging of Corporate Unsustainability. *Accounting Forum*, Vol. 30, No. 2, pp. 121-137.

Moodaley, W. & Telukdarie, A. (2023). Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review. *Sustainability*, Vol. 15, No. 2, art. 1481.

Moore, C. (2008). Moral Disengagement in Processes of Organizational Corruption. *Journal of Business Ethics*, Vol. 80, No. 1, pp. 129-139.

Mousavian Anaraki, S. A., Croce, D. & Basili, R. (2025). Large Language Models for Sustainability Reporting: A Systematic Review and Research Agenda. *Sustainable Futures*, Vol. 10, art. 101494.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *Journal of Clinical Epidemiology*, Vol. 134, pp. 178-189.

Pant, P., Muhammed Niyas, K. P., Yadav, R., Dadsena, K. & Verma, R. (2026). Deceptive Tone in Sustainable Supply Chain Disclosure and Its Impact on Firm Performance: The Role of Organizational Structure. *Business Strategy and the Environment*.

Parmar, B. L., Freeman, R. E., Harrison, J. S., Wicks, A. C., Purnell, L. & De Colle, S. (2010). Stakeholder Theory: The State of the Art. *Academy of Management Annals*, Vol. 4, No. 1, pp. 403-445.

Patten, D. M. (1992). Intra-Industry Environmental Disclosures in Response to the Alaskan Oil Spill: A Note on Legitimacy Theory. *Accounting, Organizations and Society*, Vol. 17, No. 5, pp. 471-475.

Randers, J., Rockström, J., Stoknes, P. E., Golüke, U., Collste, D. & Cornell, S. (2018). *Transformation is Feasible: How to Achieve the Sustainable Development Goals within Planetary Boundaries*. Stockholm Resilience Centre & BI Norwegian Business School.

Ren, X., Hu, S., Sun, X. & Zhou, D. (2025). The Impact of Artificial Intelligence on Corporate Greenwashing: Evidence from the Chinese Listed Firms. *Journal of Accounting Literature*.

Rogers, K. & Hudson, B. (2011). The Triple Bottom Line. *OD Practitioner*, Vol. 43, No. 4, pp. 3-9.

Sari, R. C., Sholihin, M., Cahaya, F. R., Ishola, A., Pranesti, A., Sari, A. R., et al. (2025). Artificial Intelligence-Based ESG Greenwashing Detection: Road to Net Zero Carbon and Its Impact on Corporate Performance. *Business Strategy & Development*, Vol. 8, No. 4, art. e70228.

Sari, Y. M., Saraswati, E., Atmini, S. & Prastiwi, A. (2026). A Systematic Literature Review of Assurance and Greenwashing in Sustainability Reporting. *Cogent Business & Management*, Vol. 13, No. 1, art. 2672039.

Scott, W. R. (1987). The Adolescence of Institutional Theory. *Administrative Science Quarterly*, Vol. 32, No. 4, pp. 493-511.

Scott, W. R. (2014). Institutions and Organizations. Ideas, Interests and Identities. *M@n@gement*, Vol. 17, No. 2, pp. 136-140.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, Vol. 69, No. 1, pp. 99-118.

Song, X., Dong, Y., Dong, Y. & Liu, Z. (2026). Enhancing Global Value Chain Transparency: A Multi-Agent AI Framework for Detecting ESG Greenwashing with Open-Source Innovations. *Journal of Business Research*, Vol. 215, art. 116242.

Suárez-Giri, F., Mayoral Blaya, S. & Rodríguez López, R. (2026). Greenwashing in Biodiversity Reporting: The Role of Financial Risk Exposure and Conservation Policy Stringency. *Global Finance Journal*, Vol. 70, art. 101243.

Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *Academy of Management Review*, Vol. 20, No. 3, pp. 571-610.

Telukdarie, A., Nyathi, M. H. L. & Fabchi, R. J. (2026). Environmental, Social, and Governance Sustainability: An AI-Centric Approach Driving Data Standardization and Automation. *Frontiers in Sustainability*, Vol. 7, art. 1793155.

Tian, X. & Shi, D. (2025). Facilitating or Inhibiting: A Study on the Impact of Artificial Intelligence on Corporate Greenwashing. *Sustainability*, Vol. 17, No. 5, art. 2154.

Todaro, N. M., Testa, F. & Frey, M. (2024). *Integrare la sostenibilità in azienda: un percorso strategico, organizzativo e culturale*. FrancoAngeli.

Tranfield, D., Denyer, D. & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, Vol. 14, No. 3, pp. 207-222.

Tylec, A. (2025). From NFRD to CSRD: Non-Financial Reporting "Yesterday and Today". *Scientific Papers of Silesian University of Technology. Organization and Management Series*, No. 234, pp. 573-589.

Vilanova, M. E. M. & Bazanini, R. (2026). Digital Mirage of Sustainability: Algorithmic Greenwashing and Informational Asymmetries in the Brazilian Beef Supply Chain. *Cleaner Waste Systems*, Vol. 15, art. 100538.

Vilchez Olivares, P. A. & Astorga De La Cruz, B. J. (2026). AI and Data Analytics in Sustainable Financial Reporting and ESG Disclosure: A Systematic Literature Review. *Sustainability*, Vol. 18, No. 11, art. 5393.

Wang, X., Gao, X. & Sun, M. (2025a). Construction and Analysis of Corporate Greenwashing Index: A Deep Learning Approach. *EPJ Data Science*, Vol. 14, art. 44.

Wang, X., Gao, X., Sun, M. & Zheng, C. (2025b). The Impact of Artificial Intelligence on Corporate Greenwashing: Evidence from the New Generation of Artificial Intelligence Innovation and Development Pilot Zones Policy. *Environmental Sciences Europe*, Vol. 37, art. 216.

Wilson Sonsini (2024). 10 Things You Should Know About the EU Artificial Intelligence Act. Wilson Sonsini Goodrich & Rosati, aprile 2024.

World Commission on Environment and Development (1987). *Report of the World Commission on Environment and Development: Our Common Future*. United Nations.

Xi, M., Xia, R. & Pan, W. (2026). AI Adoption and Corporate Green Behavior: Institutional Isomorphism in the Algorithmic Age. *Journal of Digital Economy*, Vol. 5, pp. 243-256.

Xing, C., Shan, Y. G., Yang, F. & Zhang, Y. (2026). The Effect of CSR Assurance on Subsequent Corporate Greenwashing: Suggestion Acquisition or Opinion Shopping? *The British Accounting Review*, Vol. 58, No. 4, art. 101744.

Xu, C., Liu, J., Li, Z. & Lin, C. (2026). DeepGreen: Effective LLM-Driven Greenwashing Monitoring System Designed for Empirical Testing – Evidence from China. *Computational Economics*, pp. 1-35.

Yadav, A. & Asongu, S. A. (2026). The AI Sustainability Paradox: How Verification and Regulation Synergize to Curb Greenwashing in Emerging Markets. *Corporate Social Responsibility and Environmental Management*.

Yang, Z., Nguyen, T. T. H., Nguyen, H. N., Nguyen, T. T. N. & Cao, T. T. (2020). Greenwashing Behaviours: Causes, Taxonomy and Consequences Based on a Systematic Literature Review. *Journal of Business Economics and Management*, Vol. 21, No. 5, pp. 1486-1507.

Yu, X., Murad, M., Chen, C., Wang, T. & Jia, X. (2026). “Good Citizen” or “Good Actor”: The Asymmetric Impact of Artificial Intelligence on Firms' Tactical Environmental Information Disclosure. *Business Strategy and the Environment*, Vol. 35, No. 4, pp. 5030-5049.

Zhang, D. (2024). The Pathway to Curb Greenwashing in Sustainable Growth: The Role of Artificial Intelligence. *Energy Economics*, Vol. 133, art. 107562.

Zhang, Y. (2011). The Analysis of Shareholder Theory and Stakeholder Theory. *2011 Fourth International Conference on Business Intelligence and Financial Engineering*, Wuhan, Cina, pp. 90-92.

Zhao, Q., Zhang, J., Sun, X., Ma, J. & Wang, G. (2026). Green or Greenwashing? A Novel Deep Learning Method Integrating Multi-Source Data for ESG Greenwashing Detection. *Journal of Cleaner Production*, Vol. 555, art. 148147.

Zheng, M., Zhang, L., Tripe, D. & Zhang, Y. (2026). Can Artificial Intelligence Mitigate Greenwashed Green Credit? Evidence from Loan Contracts of Chinese Listed Firms. *International Review of Financial Analysis*, Vol. 110, art. 104948.

Zhou, C., Liu, L. & Wang, H. (2026). How Government Environmental Targets Shape Firms' Environmental Disclosure Strategy: Evidence from LLM-Based Textual Analysis. *Corporate Social Responsibility and Environmental Management*.

Zhou, Z., Gao, S. & Wang, Z. (2026). AuraESG: A Cognitive Architecture for Auditable ESG Analysis through Knowledge-Infused Causal Reasoning and Hybrid Deliberation. *Information Sciences*, Vol. 754, art. 123647.

Zioło, M., Bąk, I. & Spoz, A. (2024). Literature Review of Greenwashing Research: State of the Art. *Corporate Social Responsibility and Environmental Management*, Vol. 31, No. 6, pp. 5343-5356.

Fonti normative

Commissione Europea (2023). *Proposta di direttiva del Parlamento europeo e del Consiglio sull'attestazione e sulla comunicazione delle asserzioni ambientali esplicite (direttiva sulle asserzioni ambientali)*. COM(2023) 166 final, 22 marzo 2023.

European Commission (2026a). *Annexes 1 to 2 to the Commission Delegated Regulation (EU) .../... of 3.7.2026 amending Delegated Regulation (EU) 2023/2772 as regards the simplification of certain sustainability reporting standards*. C(2026) 5010 final – Annexes 1 to 2.

European Commission (2026b). *Annexes 1 to 2 to the Commission Delegated Regulation (EU) .../... supplementing Directive 2013/34/EU of the European Parliament and of the Council by establishing sustainability reporting standards for voluntary use by undertakings protected by the value chain cap*. C(2026) 5011 final – Annexes 1 to 2.

European Commission (2026c). *Commission Delegated Regulation (EU) .../... of 3.7.2026 amending Delegated Regulation (EU) 2023/2772 as regards the simplification of certain sustainability reporting standards*. C(2026) 5010 final.

European Commission (2026d). *Commission Staff Working Document accompanying the documents Commission Delegated Regulation (EU) .../... amending Delegated Regulation (EU) 2023/2772 ... and Commission Delegated Regulation (EU) .../... supplementing Directive 2013/34/EU* SWD(2026) 500 final.

Unione Europea (2014). *Direttiva 2014/95/UE del Parlamento Europeo e del Consiglio del 22 ottobre 2014 recante modifica della direttiva 2013/34/UE per quanto riguarda la comunicazione di informazioni di carattere non finanziario e di informazioni sulla diversità da parte di talune imprese e di taluni gruppi di grandi dimensioni*. Gazzetta ufficiale dell'Unione europea, L 330/1, 15.11.2014.

Unione Europea (2022). *Direttiva (UE) 2022/2464 del Parlamento Europeo e del Consiglio del 14 dicembre 2022 che modifica il regolamento (UE) n. 537/2014, la direttiva 2004/109/CE, la direttiva 2006/43/CE e la direttiva 2013/34/UE per quanto riguarda la*

rendicontazione societaria di sostenibilità. Gazzetta ufficiale dell'Unione europea, L 322/15, 16.12.2022.

Unione Europea (2024a). *Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale (regolamento sull'intelligenza artificiale).* Gazzetta ufficiale dell'Unione europea, Serie L, 2024/1689, 12.7.2024.

Unione Europea (2024b). *Direttiva (UE) 2024/825 del Parlamento Europeo e del Consiglio del 28 febbraio 2024 che modifica le direttive 2005/29/CE e 2011/83/UE per quanto riguarda la responsabilizzazione dei consumatori per la transizione verde mediante il miglioramento della tutela dalle pratiche sleali e dell'informazione.* Gazzetta ufficiale dell'Unione europea, Serie L 2024/825, 6.3.2024.

Sitografia

European Commission (2026e). Commission Adopts Revised Sustainability Reporting Standards. Directorate-General for Financial Stability, Financial Services and Capital Markets Union. https://finance.ec.europa.eu/news/commission-adopts-revised-sustainability-reporting-standards-2026-07-03_en [data di ultima consultazione: 23/09/2026].

European Commission (2026f). Corporate Sustainability Reporting. Directorate-General for Financial Stability, Financial Services and Capital Markets Union. https://finance.ec.europa.eu/financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en [data di ultima consultazione: 23/09/2026].

European Commission (2026g). Implementing and Delegated Acts – Corporate Sustainability Reporting Directive (CSRD). Directorate-General for Financial Stability, Financial Services and Capital Markets Union. https://finance.ec.europa.eu/regulation-and-supervision/financial-services-legislation/implementing-and-delegated-acts/corporate-sustainability-reporting-directive_en [data di ultima consultazione: 23/09/2026].

