



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze e Tecnologie Informatiche

Corso di Laurea in Informatica per il Management

Analisi di influenza politica tramite tecniche di analisi testuale

Analisi dell'influenza politica sull'Artificial Intelligence Act, sulla
Sustainable Corporate Governance, sul Digital Services Act e sul
Digital Markets Act

Relatore:
Chiar.mo Prof.
Edoardo Mollona

Presentata da:
Lorenzo Magrini

II Sessione

Anno Accademico 2024/2025

Indice

Introduzione	6
1 Processo legislativo europeo	7
1.1 Unione Europea	7
1.1.1 Parlamento europeo	8
1.1.2 Consiglio europeo	8
1.1.3 Consiglio dell'Unione europea	9
1.1.4 Commissione europea	9
1.1.5 Corte di giustizia dell'Unione europea	10
1.1.6 Banca centrale europea	10
1.1.7 Corte dei conti europea	11
1.2 Procedura legislativa europea	11
1.2.1 Consultazione pubblica	11
1.2.2 Proposta legislativa	12
1.2.3 Prima lettura	13
1.2.4 Seconda lettura	14
1.2.5 Conciliazione	15
1.2.6 Terza lettura	15
2 Temi delle direttive analizzate	16
2.1 Artificial Intelligence Act	16
2.2 Sustainable Corporate Governance	17
2.3 Digital Services Act	17
2.4 Digital Markets Act	18
3 Metodologia	20
3.1 Origini e sviluppo del concetto di similarità	20

3.2	Tecniche adottate	22
3.2.1	TF-IDF	23
3.2.2	Indice di Jaccard	24
3.2.3	Sentence-BERT all-mpnet-base-v2	24
3.2.4	Sentence-BERT all-MiniLM-L6-v2	25
3.3	Raccolta e struttura dei dati	26
3.4	Procedura operativa	26
3.5	Fuzzy matching per l'uniformazione dei nomi	28
3.6	Dalla tabella finale all'analisi dei risultati	29
4	Risultati	30
4.1	Artificial Intelligence Act	31
4.2	Sustainable Corporate Governance	37
4.3	Digital Services Act	42
4.4	Digital Markets Act	47
4.5	Analisi trasversale	53
	Conclusioni	55
	Artificial Intelligence Act	55
	Sustainable Corporate Governance	57
	Digital Services Act	58
	Digital Markets Act	60
	Analisi trasversale e considerazioni conclusive	61
	Bibliografia	64
	Ringraziamenti	68

Elenco delle figure

4.1	Top 10 Organizzazioni per Similarità	31
4.2	Top 10 Organizzazioni per Variazione Totale	32
4.3	Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.	33
4.4	Similarità media per tipologia di organizzazione nell'AI Act.	34
4.5	Influenza media per tipologia di organizzazione nell'AI Act.	34
4.6	Distribuzione delle categorie in base a Similarità e Variazione Totale.	35
4.7	Variazione media per finestra di consultazione – Artificial Intelligence Act	36
4.8	Top 10 Organizzazioni per Similarità	37
4.9	Top 10 Organizzazioni per Variazione Totale	37
4.10	Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.	38
4.11	Similarità media per tipologia di organizzazione nella direttiva Sustainable Corporate Governance.	39
4.12	Influenza media per tipologia di organizzazione nella direttiva Sustainable Corporate Governance.	39
4.13	Distribuzione delle categorie in base a Similarità e Variazione Totale.	40
4.14	Variazione media per finestra di consultazione – Sustainable Corporate Governance	41
4.15	Top 10 Organizzazioni per Similarità	42
4.16	Top 10 Organizzazioni per Variazione Totale	42
4.17	Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.	43
4.18	Similarità media per tipologia di organizzazione nel Digital Services Act.	44
4.19	Influenza media per tipologia di organizzazione nel Digital Services Act.	44
4.20	Distribuzione delle categorie in base a Similarità e Variazione Totale.	45
4.21	Variazione media per finestra di consultazione – Digital Services Act	46
4.22	Top 10 Organizzazioni per Similarità	47
4.23	Top 10 Organizzazioni per Variazione Totale	48
4.24	Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.	49

4.25	Similarità media per tipologia di organizzazione nel Digital Markets Act.	50
4.26	Influenza media per tipologia di organizzazione nel Digital Markets Act.	50
4.27	Distribuzione delle categorie in base a Similarità e Variazione Totale.	51
4.28	Variazione media per finestra di consultazione – Digital Markets Act	52
4.29	Correlazioni tra direttive	53
4.30	Correlazioni tra categorie di organizzazioni	54

Introduzione

La presente tesi si concentra sul ruolo dei *positioning paper*, ossia i feedback attraverso i quali attori economici e organizzazioni esprimono le proprie posizioni in merito a nuove iniziative normative dell'Unione Europea.

L'obiettivo principale è quello di analizzare come tali contributi possano incidere sull'elaborazione delle proposte legislative, mettendoli a confronto con i documenti ufficiali della Commissione europea e valutando le dinamiche di influenza che emergono da questo processo. A tale scopo sono state adottate diverse metodologie di analisi testuale per misurare il grado di somiglianza tra i vari feedback inviati e le versioni dei testi legislativi nelle varie fasi del loro sviluppo.

L'analisi ha permesso di individuare non solo le organizzazioni più influenti, ma anche le principali categorie di attori capaci di incidere in modo significativo sui contenuti delle quattro direttive esaminate: l'*Artificial Intelligence Act*, la *Sustainable Corporate Governance*, il *Digital Services Act* e il *Digital Markets Act*.

È stato inoltre valutato se esista un momento più favorevole per esercitare influenza, verificando se i feedback risultino più incisivi nella fase iniziale di consultazione o in quella successiva all'adozione della proposta.

Il primo capitolo ("Processo legislativo europeo") offre una panoramica generale del funzionamento delle istituzioni dell'Unione Europea e descrive in dettaglio le fasi del procedimento legislativo ordinario, dalla consultazione pubblica fino alla terza lettura.

Il secondo capitolo ("Temi delle direttive analizzate") si concentra sul contenuto delle quattro iniziative legislative considerate, descrivendone il contesto di origine, gli obiettivi perseguiti e le soluzioni normative proposte, con particolare attenzione alle modalità e ai tempi delle consultazioni pubbliche.

Il terzo capitolo ("Metodologia") illustra gli strumenti tecnici utilizzati per l'analisi, soffermandosi sui metodi di calcolo della similarità testuale e sulle procedure di confronto tra i contributi delle organizzazioni e i documenti ufficiali della Commissione Europea. Viene inoltre descritto il processo di raccolta, pulizia e strutturazione dei dati, insieme alle modalità di uniformazione

dei nomi delle organizzazioni.

Il quarto capitolo (“Risultati”) presenta le evidenze empiriche emerse dall’applicazione delle tecniche descritte, illustrando attraverso numerosi grafici i risultati ottenuti per ciascuna delle quattro direttive — *Artificial Intelligence Act*, *Sustainable Corporate Governance*, *Digital Services Act* e *Digital Markets Act*. Per ogni direttiva vengono mostrati i grafici relativi alle organizzazioni più influenti, alle principali categorie di attori e all’evoluzione temporale dell’influenza, mettendo in evidenza i valori di similarità, variazione e le differenze tra le varie fasi del processo legislativo.

Il capitolo si conclude con un’analisi trasversale che confronta i risultati delle quattro direttive, evidenziando correlazioni e tendenze comuni nei diversi ambiti normativi.

Infine, le Conclusioni raccolgono e interpretano i risultati emersi, proponendone inoltre una lettura comparata. In questa parte finale vengono formulate riflessioni sulle dinamiche osservate, individuando eventuali pattern ricorrenti e le differenze più significative tra le direttive analizzate.

Capitolo 1

Processo legislativo europeo

In questo capitolo viene analizzato il processo legislativo dell'Unione Europea, con particolare attenzione alla procedura di consultazione pubblica e al ruolo dei portatori di interesse. Verrà illustrato il contesto generale in cui operano e collaborano le diverse istituzioni europee, e descritto il percorso attraverso cui nascono e vengono approvate le direttive, evidenziando i passaggi fondamentali che conducono dalla proposta iniziale fino all'adozione finale.

1.1 Unione Europea

Nata nel dopoguerra in risposta alla necessità di garantire stabilità e cooperazione tra gli Stati europei, l'Unione Europea ha progressivamente ridotto il rischio di nuovi conflitti trasformandosi da semplice comunità economica, volta alla gestione condivisa di risorse strategiche, a un'unione politica ed economica che oggi riunisce 27 Paesi. Essa rappresenta attualmente il più grande mercato integrato al mondo e fonda la propria azione sui valori di dignità, libertà e solidarietà sanciti dalla Carta dei diritti fondamentali dell'Unione.¹

È costituita da sette istituzioni principali:

- **Parlamento europeo**²
- **Consiglio europeo**³
- **Consiglio dell'Unione europea**⁴

¹Carta dei diritti fondamentali dell'Unione Europea. Disponibile su: <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:12012P/TXT>

²Sito ufficiale: <https://www.europarl.europa.eu/portal/it>

³Sito ufficiale: <https://www.consilium.europa.eu/it/european-council/>

⁴Sito ufficiale: <https://www.consilium.europa.eu/it/council-eu/>

- Commissione europea⁵
- Corte di giustizia dell'Unione europea⁶
- Banca centrale europea⁷
- Corte dei conti europea⁸

1.1.1 Parlamento europeo

Il Parlamento europeo è l'istituzione che rappresenta i cittadini dell'Unione Europea. È composto da un massimo di 750 deputati, distribuiti tra gli Stati membri in proporzione alla popolazione, con un minimo di 6 e un massimo di 96 seggi per Paese.

I parlamentari sono eletti a suffragio universale diretto e si organizzano in gruppi politici trans-nazionali. Il lavoro parlamentare si articola in commissioni specializzate e in sessioni plenarie, che si tengono principalmente a Strasburgo e Bruxelles. Il Parlamento condivide il potere legislativo e di bilancio con il Consiglio dell'Unione europea e svolge un ruolo di controllo democratico sulle altre istituzioni.

Le principali competenze di questa istituzione sono le seguenti:

1. adottare, insieme al Consiglio, la legislazione europea;
2. approvare il bilancio dell'Unione;
3. eleggere il Presidente della Commissione e approvare i Commissari;
4. esercitare funzioni di controllo politico e democratico.

1.1.2 Consiglio europeo

Il Consiglio europeo riunisce i capi di Stato o di governo degli Stati membri, il Presidente del Consiglio europeo e il Presidente della Commissione. Il suo Presidente è eletto dai membri per un mandato di due anni e mezzo, rinnovabile una volta. Non ha funzioni legislative dirette, ma definisce le priorità e gli orientamenti politici generali dell'UE.

Le responsabilità principali di quest'organo si articolano in:

1. fissare l'agenda e le priorità politiche generali dell'UE;

⁵Sito ufficiale: https://commission.europa.eu/index_it

⁶Sito ufficiale: https://curia.europa.eu/jcms/jcms/j_6/

⁷Sito ufficiale: <https://www.ecb.europa.eu/home/html/index.it.html>

⁸Sito ufficiale: <https://www.eca.europa.eu/it>

2. risolvere questioni istituzionali e costituzionali;
3. nominare alte cariche come il Presidente della Commissione o l'Alto rappresentante per gli affari esteri.

1.1.3 Consiglio dell'Unione europea

Il Consiglio dell'Unione europea, comunemente chiamato anche Consiglio dei ministri, è l'organo che rappresenta i governi degli Stati membri. È composto dai ministri nazionali competenti per la materia trattata (ad esempio economia, agricoltura, affari esteri) e si riunisce in diverse configurazioni a seconda dell'argomento all'ordine del giorno.

La presidenza del Consiglio è esercitata a turno da ciascuno Stato membro per un periodo di sei mesi, secondo un sistema di rotazione prestabilito. Insieme al Parlamento europeo, il Consiglio condivide il potere legislativo e quello di bilancio, assumendo un ruolo centrale nell'elaborazione e nell'approvazione delle politiche dell'Unione.

Tra le funzioni più rilevanti di questa istituzione si possono indicare:

1. adottare le leggi europee insieme al Parlamento;
2. approvare il bilancio dell'Unione;
3. coordinare le politiche economiche e sociali degli Stati membri;
4. concludere accordi internazionali a nome dell'UE.

1.1.4 Commissione europea

La Commissione europea è l'organo esecutivo dell'Unione e rappresenta l'interesse generale dell'UE al di sopra di quelli dei singoli Stati membri. Ha il compito di proporre nuove iniziative legislative, vigilare sull'attuazione delle politiche e garantire il rispetto dei trattati. È composta da 27 commissari, uno per ciascuno Stato membro, incluso il Presidente e otto vicepresidenti.

Il Presidente della Commissione è designato dal Consiglio europeo e deve ottenere l'approvazione del Parlamento insieme all'intero collegio. Oltre a svolgere la funzione di iniziativa legislativa, la Commissione gestisce il bilancio dell'Unione, coordina i programmi e le politiche nei diversi settori e rappresenta l'UE nei negoziati internazionali in materie di sua competenza. Essa costituisce quindi il motore politico e amministrativo del sistema europeo, con il ruolo di promuovere l'interesse comune e assicurare l'applicazione uniforme delle regole nei 27 Stati

membri.

I compiti principali attribuiti a quest'organo comprendono:

1. proporre nuove leggi europee;
2. gestire le politiche comuni e i programmi di finanziamento;
3. garantire l'applicazione del diritto dell'Unione;
4. rappresentare l'UE sulla scena internazionale nelle sue competenze.

1.1.5 Corte di giustizia dell'Unione europea

La Corte di giustizia assicura il rispetto del diritto dell'Unione nell'interpretazione e nell'applicazione dei trattati. È composta da un giudice per ogni Stato membro, affiancato da avvocati generali che forniscono pareri indipendenti. I giudici sono nominati di comune accordo dai governi per un mandato di sei anni, rinnovabile.

Le funzioni essenziali di questa istituzione possono essere sintetizzate come segue:

1. vigilare sull'applicazione uniforme del diritto dell'UE;
2. risolvere controversie tra istituzioni e Stati membri;
3. esaminare la legittimità degli atti europei;
4. rispondere ai rinvii pregiudiziali dei giudici nazionali.

1.1.6 Banca centrale europea

La Banca centrale europea (BCE) è responsabile della politica monetaria dell'area euro. Ha sede a Francoforte e la sua struttura decisionale include il Consiglio direttivo (Presidente, Vice-presidente e governatori delle banche centrali nazionali dell'eurozona) e il Comitato esecutivo. Le prerogative fondamentali di quest'organo includono:

1. mantenere la stabilità dei prezzi nell'eurozona;
2. definire e attuare la politica monetaria comune;
3. gestire le riserve ufficiali e condurre operazioni di cambio;
4. supervisionare le banche e i mercati finanziari.

1.1.7 Corte dei conti europea

La Corte dei conti europea è l'istituzione di revisione esterna delle finanze dell'UE. È composta da un membro per ogni Stato membro, nominato dal Consiglio previa consultazione del Parlamento. Non ha potere giurisdizionale, ma svolge un ruolo cruciale di trasparenza e controllo. Tra i compiti più significativi di questa istituzione rientrano:

1. verificare entrate e spese dell'Unione;
2. controllare la legittimità e regolarità delle operazioni finanziarie;
3. redigere relazioni annuali e speciali sul bilancio;
4. contribuire al miglioramento della gestione finanziaria europea.

1.2 Procedura legislativa europea

La procedura legislativa ordinaria si articola in fasi successive: dalla consultazione pubblica e preparazione della proposta, fino alla prima e seconda lettura in Parlamento e Consiglio, all'eventuale conciliazione e infine alla terza lettura con l'approvazione.

1.2.1 Consultazione pubblica

La consultazione pubblica è uno strumento centrale del quadro di *better regulation* e consente a cittadini, imprese e organizzazioni di contribuire alla definizione delle politiche europee. Essa mette a disposizione più finestre, ognuna con proprie caratteristiche e tempistiche standard.

Prima finestra: Call for Evidence La *Call for Evidence* costituisce la fase iniziale di consultazione. Si tratta di un documento sintetico che descrive:

- obiettivi dell'iniziativa;
- opzioni politiche in valutazione;
- campo di applicazione della futura proposta.

In questa fase:

1. è possibile inviare contributi scritti (standard: 4 settimane);

2. parallelamente può essere compilato un questionario di consultazione pubblica (standard: 12 settimane, disponibile in tutte le lingue ufficiali dell'UE).

I contributi ricevuti sono pubblicati sul portale *Have Your Say*.⁹ Successivamente la Commissione:

- redige un *factual summary report* (entro 8 settimane dalla chiusura della consultazione);
- pubblica un *synopsis report* che illustra come le osservazioni siano state prese in considerazione.

Seconda finestra: Feedback dopo l'adozione della proposta Una volta adottata formalmente la proposta legislativa:

- il testo viene pubblicato su *Have Your Say*;
- si apre una fase di feedback aperta per 8 settimane;
- i contributi ricevuti sono resi pubblici e trasmessi al Parlamento europeo e al Consiglio, affinché possano essere considerati nelle successive fasi del processo legislativo.

È tuttavia importante sottolineare che le tempistiche sopra indicate rappresentano uno standard generale. Esse possono variare a seconda della sensibilità politica e della complessità dell'iniziativa: alcuni casi, come sarà mostrato nei capitoli successivi, hanno previsto consultazioni più lunghe o più brevi rispetto ai periodi ordinari.

1.2.2 Proposta legislativa

Il processo legislativo ordinario inizia con la presentazione di una proposta da parte della Commissione europea, unica istituzione competente ad avviare nuove iniziative legislative. La Commissione elabora il testo sulla base di analisi tecniche, consultazioni pubbliche e valutazioni d'impatto, con l'obiettivo di individuare la soluzione normativa più adeguata al problema da affrontare.

La proposta può assumere la forma di un regolamento, di una direttiva o di una decisione, a seconda della materia e degli obiettivi ricercati. Una volta approvata dal collegio dei commis-

⁹Piattaforma ufficiale della Commissione per la partecipazione dei cittadini: https://ec.europa.eu/info/law/better-regulation/have-your-say_it

sari¹⁰, essa viene trasmessa al Parlamento europeo e al Consiglio dell'Unione europea, che sono le due istituzioni titolari del potere legislativo. Contemporaneamente, il documento viene reso pubblico e messo a disposizione di cittadini e stakeholder attraverso i canali ufficiali dell'Unione, tra cui la piattaforma *Have Your Say*. Questo garantisce trasparenza e permette di monitorare le fasi di avanzamento della proposta.

La presentazione formale da parte della Commissione segna quindi l'inizio dell'iter legislativo e consente l'avvio delle successive fasi di valutazione e decisione da parte del Parlamento e del Consiglio, a partire dalla prima lettura.

1.2.3 Prima lettura

A seguito della presentazione formale da parte della Commissione, il processo legislativo entra nella fase della prima lettura, in cui il Parlamento europeo è chiamato a esaminare per la prima volta il testo.

Il Parlamento discute la proposta all'interno delle proprie commissioni competenti per materia, che hanno il compito di analizzare il contenuto, suggerire eventuali modifiche e redigere un rapporto. Una volta concluso il lavoro in commissione, il testo passa alla sessione plenaria, dove gli eurodeputati possono approvarlo senza modifiche, proporre emendamenti oppure respingerlo. Se il Parlamento approva la proposta senza cambiamenti, il Consiglio dell'Unione europea può a sua volta adottarla direttamente, portando così alla conclusione del processo già in prima lettura.

Più frequentemente, invece, il Parlamento introduce modifiche al testo della Commissione. In questo caso, la proposta emendata viene trasmessa al Consiglio, che ha la possibilità di:

1. accogliere integralmente le modifiche approvate dagli eurodeputati, con conseguente adozione;
2. adottare una propria posizione, respingendo in parte o del tutto le modifiche, con il procedimento che proseguirà così alla fase successiva (seconda lettura).

La prima lettura è quindi un passaggio cruciale perché stabilisce il primo orientamento politico delle istituzioni nei confronti della proposta. Gran parte del dibattito politico prende forma durante la prima lettura, spesso accompagnata da negoziati informali tra Parlamento e

¹⁰Il collegio è composto da 27 commissari, uno per ciascun paese dell'UE. Insieme, i 27 membri del collegio rappresentano la leadership politica della Commissione e restano in carica per 5 anni. La presidente della Commissione assegna a ciascun commissario competenze per settori d'intervento specifici. Sito ufficiale: https://commission.europa.eu/about/organisation/college-commissioners_it

Consiglio. Pur non essendo l'ultimo passo, la prima lettura spesso orienta in modo decisivo il contenuto finale dell'atto, dato che un ampio consenso già in questa fase può accelerare il processo legislativo.

1.2.4 Seconda lettura

La seconda lettura rappresenta il momento in cui Parlamento europeo e Consiglio tornano a confrontarsi sulla proposta legislativa dopo che entrambi hanno espresso una prima posizione. Essa si apre soltanto se durante la prima lettura Parlamento e Consiglio non sono riusciti a raggiungere un accordo sullo stesso testo.

Il procedimento riprende dalla posizione espressa dal Consiglio a seguito del voto del Parlamento nella prima lettura. Questa viene inoltrata al Parlamento, che entro un periodo fissato deve valutarla e decidere come proseguire.

A questo punto, gli eurodeputati hanno diverse opzioni:

1. approvare senza modifiche la posizione del Consiglio, con la conseguenza che l'atto legislativo è adottato;
2. respingere l'intera proposta, chiudendo così il procedimento;
3. proporre emendamenti mirati alla posizione del Consiglio.

Se il Parlamento introduce nuove modifiche, il testo viene rinviato al Consiglio, che a sua volta deve decidere se:

1. accogliere gli emendamenti, portando all'adozione dell'atto;
2. respingerli, con conseguente passaggio del procedimento alla fase di conciliazione.

La seconda lettura è spesso considerata una fase delicata, perché riduce gli spazi di manovra politica rispetto alla prima. Le istituzioni hanno infatti margini di tempo ridotti e si trovano a dover scegliere in modo rapido se virare verso un compromesso o continuare a sostenere le proprie divergenze. In questa fase, inoltre, entrano in gioco con maggiore intensità i negoziati informali tra le delegazioni del Parlamento e del Consiglio, con l'obiettivo di evitare di arrivare allo stallo della conciliazione.

In sintesi, la seconda lettura rappresenta un momento di verifica dell'allineamento tra Parlamento e Consiglio. Qualora si arrivi a un testo condiviso, l'adozione dell'atto è immediata; in caso contrario, il procedimento proseguirà alla fase di conciliazione.

1.2.5 Conciliazione

Se dopo la seconda lettura Parlamento europeo e Consiglio non riescono a raggiungere un accordo, il procedimento legislativo entra nella fase di conciliazione. In questo momento viene istituito un comitato paritetico, composto da un numero uguale di rappresentanti delle due istituzioni:

- da un lato, i membri del Consiglio (o loro rappresentanti permanenti);
- dall'altro, una delegazione del Parlamento europeo scelta proporzionalmente tra i gruppi politici.

Il compito del comitato è quello di trovare un testo comune accettabile da entrambe le parti. Le opzioni possibili sono due:

1. se viene raggiunto un compromesso, il comitato approva un testo congiunto che viene trasmesso sia al Parlamento che al Consiglio;
2. se non si raggiunge alcun accordo, l'atto legislativo decade e il procedimento si chiude senza approvazione.

La fase di conciliazione è quindi cruciale, poiché rappresenta l'ultima possibilità per evitare il fallimento del processo legislativo. Essa comporta intense negoziazioni politiche e richiede compromessi significativi da entrambe le istituzioni.

1.2.6 Terza lettura

La terza lettura si svolge soltanto se la fase di conciliazione ha avuto esito positivo, ossia se il comitato paritetico ha prodotto un testo comune. In questa fase il documento viene sottoposto nuovamente sia al Parlamento europeo sia al Consiglio dell'Unione europea.

Ciascuna istituzione ha davanti a sé due possibilità:

1. approvare il testo concordato in conciliazione: in questo caso l'atto legislativo è definitivamente adottato;
2. respingere il testo: in tal caso l'atto decade e il procedimento legislativo si conclude senza successo.

La terza lettura non consente ulteriori modifiche: il Parlamento e il Consiglio possono solo approvare o respingere il testo elaborato in conciliazione. Quest'ultimo passaggio assicura la conclusione del processo, ma solo se la conciliazione ha prodotto un accordo fra le istituzioni; in caso contrario, l'intero procedimento rischia di concludersi senza successo.

Capitolo 2

Temi delle direttive analizzate

In questo capitolo vengono presentati i temi trattati nelle normative su cui è stata svolta l'analisi. Si tratta di iniziative legislative europee di grande rilevanza, che hanno sollevato un ampio dibattito politico e sociale e che sono state accompagnate da consistenti processi di consultazione pubblica. Le proposte considerate sono: l'Artificial Intelligence Act, l'iniziativa sulla Sustainable Corporate Governance, il Digital Services Act e il Digital Markets Act. Per ciascuna di esse verranno descritti il contesto di nascita e i problemi a cui intendevano dare risposta, gli obiettivi perseguiti e le soluzioni normative individuate. Verranno evidenziati i tempi e le modalità delle consultazioni pubbliche, che costituiscono la base del confronto svolto in questa ricerca tra i testi ufficiali e i contributi inviati da organizzazioni e portatori di interesse, con l'obiettivo di valutare la loro effettiva influenza sul processo legislativo.

2.1 Artificial Intelligence Act

Il contesto di nascita dell'Artificial Intelligence Act è legato alla crescente diffusione delle tecnologie basate sull'intelligenza artificiale e alla necessità di garantire che il loro sviluppo avvenga nel rispetto dei diritti fondamentali e della sicurezza dei cittadini europei. Il problema principale individuato dalla Commissione europea riguarda i rischi legati a un utilizzo non regolamentato dell'IA, che potrebbe generare discriminazioni, violazioni della privacy o danni alla sicurezza. Gli obiettivi perseguiti dalla normativa sono quelli di creare un quadro giuridico armonizzato, promuovere l'innovazione in un contesto di fiducia e stabilire obblighi proporzionati al livello di rischio dei sistemi di intelligenza artificiale. Le soluzioni individuate consistono in una classificazione dei sistemi di IA in base al rischio (inaccettabile, alto, limitato, minimo) e nell'introduzione di requisiti stringenti per quelli ad alto rischio. Per quanto riguarda le consultazioni

pubbliche¹, la Commissione ha avviato una consultazione aperta dal 20 febbraio al 14 giugno 2020 (circa 16 settimane), alla quale ha fatto seguito una Call for Evidence pubblicata dal 23 luglio al 10 settembre 2020 (4 settimane per i feedback). Infine, dopo l'adozione della proposta, è stata aperta una finestra di 8 settimane di consultazione dal 26 aprile al 6 agosto 2021.

2.2 Sustainable Corporate Governance

L'iniziativa sulla Sustainable Corporate Governance nasce in un contesto in cui il dibattito sul ruolo delle imprese nella transizione verso un'economia più sostenibile e inclusiva diventava sempre più acceso. Il problema individuato era la tendenza delle aziende a privilegiare i risultati e la massimizzazione dei profitti di breve periodo a scapito della sostenibilità sociale e ambientale, trascurando spesso gli impatti di lungo termine. Gli obiettivi principali sono quelli di orientare la governance societaria verso una prospettiva di lungo termine, estendere il ruolo e le responsabilità degli amministratori e assicurare che le scelte aziendali tengano conto degli impatti ambientali, sociali e sui diritti umani. Le misure previste comprendono l'introduzione di obblighi per le grandi imprese, tra cui l'adozione di piani di sostenibilità e il controllo della loro applicazione lungo l'intera catena del valore. Le attività di consultazione² anche in questo caso si sono articolate in più fasi. In primo luogo, la Commissione ha lanciato una Call for Evidence, rimasta aperta dal 30 luglio all'8 ottobre 2020 (circa 10 settimane). Successivamente è stata organizzata una consultazione pubblica, svoltasi tra il 26 ottobre 2020 e l'8 febbraio 2021 (circa 15 settimane). Infine, una volta adottata la proposta, è stato previsto un ulteriore periodo di feedback della durata di 8 settimane, compreso tra il 28 marzo e il 23 maggio 2022.

2.3 Digital Services Act

Il Digital Services Act nasce come parte del Digital Services Act package, un pacchetto legislativo presentato dalla Commissione europea nel dicembre 2020 per aggiornare le regole dello spazio digitale europeo. L'iniziativa aveva lo scopo di affrontare le nuove sfide poste dalle piattaforme online e dai mercati digitali, articolandosi fin dall'inizio in due proposte complementari: da un lato il Digital Services Act, dedicato alla sicurezza e ai contenuti online, dall'altro il Digital Markets Act, volto a garantire condizioni di concorrenza eque nei mercati dominati da grandi

¹Consultazione pubblica sull'*Artificial Intelligence Act*, disponibile su [Have Your Say - AI Act](#)

²Consultazione pubblica sulla *Sustainable Corporate Governance*, disponibile su [Have Your Say - Sustainable Corporate Governance](#).

operatori digitali.

Il Digital Services Act nasce dall'esigenza di regolare la crescente diffusione delle piattaforme online e i problemi connessi alla pubblicazione di contenuti illegali, alla disinformazione e alla scarsa trasparenza nei meccanismi digitali che gestiscono le informazioni, colmando le lacune lasciate dalla Direttiva e-commerce del 2000.

Gli obiettivi della normativa sono quelli di rendere lo spazio digitale più sicuro, rafforzare la protezione dei diritti fondamentali e garantire un mercato più equilibrato. Le soluzioni introdotte comprendono obblighi di trasparenza per le piattaforme, regole più chiare per la rimozione dei contenuti illegali e un sistema di vigilanza più stringente per le piattaforme di dimensioni molto grandi.

Per il Digital Services Act, il percorso di consultazione³ è iniziato con una Call for Evidence aperta dal 2 al 30 giugno 2020 (4 settimane). In parallelo, la Commissione ha avviato una consultazione pubblica che si è svolta dal 2 giugno all'8 settembre 2020 (circa 14 settimane). Successivamente, dopo l'adozione della proposta, è stato attivato un ulteriore periodo di feedback compreso tra il 16 dicembre 2020 e il 31 marzo 2021 (circa 15 settimane).

2.4 Digital Markets Act

In parallelo al Digital Services Act, la Commissione ha sviluppato il Digital Markets Act, che rappresenta l'altro pilastro del pacchetto legislativo sui servizi digitali. Se il Digital Services Act si concentra sulla gestione dei contenuti e sulla sicurezza online, il Digital Markets Act affronta gli squilibri concorrenziali creati dai grandi operatori del settore, i cosiddetti gatekeeper, in grado di condizionare la libertà di scelta dei consumatori.

Gli obiettivi della normativa sono quelli di rendere i mercati digitali più equi e contendibili, favorire la crescita delle imprese innovative e introdurre tutele più efficaci per i consumatori. Le soluzioni individuate prevedono obblighi e divieti specifici per i gatekeeper, come l'obbligo di consentire la disinstallazione delle applicazioni preinstallate, quello di permettere agli utenti di scegliere liberamente i servizi complementari (ad esempio motori di ricerca, sistemi di pagamento o browser) e il divieto di favorire i propri servizi rispetto a quelli dei concorrenti, accompagnati da un sistema sanzionatorio rigoroso in caso di violazioni.

Le consultazioni pubbliche⁴ hanno seguito un percorso parallelo a quello del Digital Services Act, con una consultazione tra il 2 giugno e l'8 settembre 2020, una Call for Evidence pubblica-

³Consultazione pubblica sul *Digital Services Act*, disponibile su [Have Your Say - Digital Services Act](#).

⁴Consultazione pubblica sul *Digital Markets Act*, disponibile su [Have Your Say - Digital Markets Act](#).

ta sempre dal 2 al 30 giugno 2020 e una fase di feedback successiva all'adozione della proposta tra il 16 dicembre 2020 e il 5 maggio 2021.

Capitolo 3

Metodologia

Come già anticipato nell’Introduzione, lo scopo di questa tesi è analizzare in che misura i feedback inviati nell’ambito delle consultazioni pubbliche abbiano inciso sull’elaborazione dei testi legislativi dell’Unione Europea. Per rispondere a questa domanda è stato adottato un approccio basato sulla similarità testuale, che permette di confrontare documenti diversi misurandone il grado di vicinanza linguistica o semantica.

In termini generali, la similarità testuale misura quanto due testi possano essere considerati simili dal punto di vista del contenuto. Tale somiglianza può essere calcolata in molti modi: considerando la sovrapposizione di termini, la frequenza delle parole, oppure, nei modelli più recenti, il significato delle frasi attraverso rappresentazioni vettoriali.

3.1 Origini e sviluppo del concetto di similarità

L’idea di confrontare testi diversi per valutarne la somiglianza ha una storia che affonda le proprie radici negli studi di *information retrieval*¹ degli anni Sessanta. In quel periodo la sfida principale era quella di permettere ai sistemi informatici di recuperare documenti pertinenti a partire da una query inserita dall’utente. Si trattava, quindi, di capire quali testi, all’interno di grandi archivi, fossero “più simili” alla richiesta espressa. In questa fase iniziale, la similarità veniva calcolata sulla base della presenza o assenza delle stesse parole, trattando i documenti come insiemi di termini.

Con l’evoluzione della ricerca, negli anni Ottanta e Novanta si affermarono tecniche più strutturate, che trasformavano i testi in vettori numerici, ossia liste ordinate di valori che rappresentano

¹Il termine *information retrieval* indica l’insieme delle tecniche per il recupero automatico di informazioni rilevanti da grandi archivi testuali o multimediali. Risorse: <https://www.ibm.com/it-it/think/topics/information-retrieval>.

la presenza o la frequenza delle parole all'interno di ciascun documento. Uno degli approcci più diffusi era il modello *bag-of-words*², in cui ogni documento veniva rappresentato attraverso le frequenze dei termini che lo componevano. In questo modo, la somiglianza poteva essere stimata con strumenti matematici relativamente semplici: dal conteggio delle parole comuni, alla distanza tra i vettori, fino al calcolo dell'angolo che li separa. Ad esempio, se un documento A e un documento B vengono rappresentati come vettori (2,1,0) e (3,1,0), dove ciascun numero indica la frequenza di una specifica parola, il fatto che i due vettori presentino valori quasi identici comporta una distanza ridotta e un angolo molto piccolo, segnalando una forte affinità tra i due testi. Anche se basati solo sul lessico, questi metodi si rivelarono pratici e immediati, e per molto tempo hanno rappresentato la soluzione più utilizzata nelle applicazioni di analisi testuale.

Dagli anni Duemila, lo sviluppo del *Natural Language Processing* (NLP)³ e dei modelli statistici ha introdotto una svolta importante, con l'obiettivo di superare i limiti degli approcci puramente lessicali, incapaci di cogliere il significato e il contesto delle parole. Si svilupparono così metodi più avanzati, basati sull'idea che il significato di una parola dipenda dal contesto in cui viene usata. Secondo questo approccio, riassunto dalla nota espressione di John Rupert Firth "you shall know a word by the company it keeps", il significato di una parola può essere colto osservando i termini con cui essa tende a comparire. Da questa intuizione nacquero i primi modelli di *word embedding*⁴, capaci di proiettare le parole in spazi vettoriali in cui la vicinanza geometrica corrisponde a una vicinanza semantica. Ad esempio, in un modello di questo tipo i termini *algoritmo* e *intelligenza artificiale* tendono a essere rappresentati da vettori molto vicini, poiché compaiono spesso in contesti simili, mentre una parola come *musica* risulterebbe collocata molto più lontano.

Negli anni più recenti, lo sviluppo di architetture neurali sempre più sofisticate ha portato a una nuova generazione di modelli, come quello di *BERT*⁵ e le sue numerose varianti, in grado

²Il modello *bag-of-words* rappresenta un testo come un insieme di parole, ignorando la struttura sintattica e l'ordine, concentrandosi solo sulla frequenza dei termini. Risorse: <https://www.ibm.com/it-it/think/topics/bag-of-words>.

³Il *Natural Language Processing* (NLP) è un'area dell'intelligenza artificiale che consente ai computer di comprendere, interpretare e generare linguaggio umano. Risorse: <https://www.ibm.com/it-it/think/topics/natural-language-processing>.

⁴I *word embedding* rappresentano le parole come vettori in uno spazio numerico, in cui la distanza geometrica riflette la loro somiglianza semantica. Risorse: <https://www.ibm.com/it-it/think/topics/word-embeddings>.

⁵*BERT* (*Bidirectional Encoder Representations from Transformers*) è un modello di linguaggio sviluppato da Google AI Research per l'elaborazione del linguaggio naturale. Risorse: *BERT - NAACL 2019*.

di generare rappresentazioni non solo di singole parole ma anche di intere frasi e documenti. Questi modelli hanno reso possibile analizzare la similarità testuale in modo più profondo, cogliendo sfumature di significato e relazioni contestuali che sfuggono completamente agli approcci tradizionali.

In sintesi, il concetto di similarità testuale si è evoluto da semplici conteggi di parole a sofisticati sistemi di rappresentazione semantica. Oggi possiamo distinguere due grandi famiglie di tecniche:

- **tecniche semplici**, che trattano i testi come insiemi o vettori di termini e calcolano la somiglianza in base alla sovrapposizione o alla distanza lessicale (ad esempio *bag-of-words*, TF-IDF, indice di Jaccard);
- **tecniche complesse**, che utilizzano modelli neurali e *embedding* semantici per rappresentare i testi come punti in uno spazio multidimensionale, in cui la distanza riflette la somiglianza di significato (ad esempio BERT e i suoi derivati).

In altre parole, questa evoluzione mostra come la ricerca abbia nel tempo spostato l'attenzione dal livello puramente lessicale a quello semantico, rendendo possibili analisi più raffinate e più vicine al modo in cui gli esseri umani riconoscono la somiglianza tra testi.

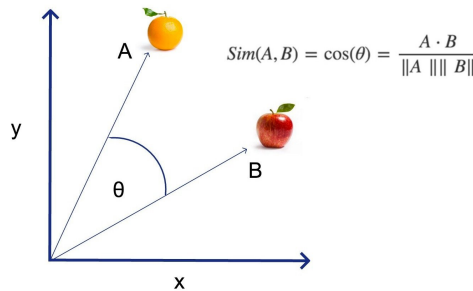
3.2 Tecniche adottate

Le tecniche adottate in questo studio sono state selezionate con l'obiettivo di confrontare approcci diversi, sia per livello di complessità che per costo computazionale. Da un lato sono state incluse metodologie più semplici, basate su misurazioni puramente lessicali (TF-IDF e indice di Jaccard); dall'altro sono stati utilizzati modelli più sofisticati, in grado di cogliere le relazioni semantiche (Sentence-BERT all-mpnet-base-v2 e Sentence-BERT all-MiniLM-L6-v2), ma più onerosi dal punto di vista computazionale.

Questa scelta ha permesso non solo di valutare le differenze nei risultati prodotti dalle varie tecniche, ma anche di osservare il compromesso tra accuratezza e costi di elaborazione, aspetto cruciale quando gli elementi diventano numerosi. In fase di analisi finale, i valori di similarità ottenuti con i diversi metodi sono stati aggregati tramite il calcolo della media, così da ricavare un indicatore unico della vicinanza tra feedback e testi legislativi.

Per tutte le tecniche che restituiscono una rappresentazione numerica dei testi, come i vettori di pesi della TF-IDF e gli embedding dei modelli BERT, la similarità è stata calcolata utilizzando

la *cosine similarity*⁶. Questa misura, come suggerisce il nome, calcola il coseno dell'angolo che separa i due vettori: un valore vicino a 1 indica che i vettori sono quasi paralleli, e quindi che i testi sono molto simili; un valore vicino a 0 (o negativo) indica invece una forte divergenza.



La scelta di questa metrica si basa su due motivi principali. Da un lato, la *cosine similarity* non dipende dalla lunghezza dei documenti ma confronta soltanto la direzione dei vettori, risultando quindi adatta a testi di dimensioni diverse come quelli considerati in questa ricerca. Dall'altro, ha fornito prestazioni migliori rispetto ad altre misure comunemente utilizzate, come la distanza euclidea o il coefficiente di correlazione.

3.2.1 TF-IDF

La prima tecnica utilizzata è la TF-IDF⁷ (Term Frequency – Inverse Document Frequency), una delle più note nell'ambito dell'information retrieval. L'idea di base è che un termine sia tanto più rappresentativo di un documento quanto più ricorre al suo interno, ma allo stesso tempo tanto meno informativo quanto più è diffuso in tutti i documenti della raccolta.

Matematicamente, il peso di un termine t in un documento d si calcola come:

$$w_{t,d} = tf_{t,d} \times \log \frac{N}{df_t}$$

dove:

- $tf_{t,d}$ è la frequenza del termine t nel documento d ;
- df_t è il numero di documenti che contengono il termine t ;

⁶Risorse: IBM - Cosine Similarity.

⁷TF-IDF. Risorse: https://medium.com/@Sumeet_Agrawal/tf-idf-47f8f2ebd987

- N è il numero totale di documenti della raccolta.

Una volta trasformati i testi in vettori TF-IDF, cioè in sequenze numeriche che associano a ciascun termine presente nella raccolta un peso calcolato come prodotto tra la frequenza della parola nel documento e il logaritmo del rapporto tra il numero totale di documenti e il numero di documenti che contengono quella parola, la similarità tra due documenti viene calcolata con la *cosine similarity*, che misura il coseno dell'angolo tra i vettori. Un valore vicino a 1 indica che i testi sono molto simili, mentre un valore vicino a 0 indica che sono molto diversi.

Il limite principale di TF-IDF è che si basa su un livello puramente lessicale, senza tenere conto del significato delle parole. Due documenti che usano termini diversi per esprimere lo stesso concetto risulteranno lontani, mentre testi che condividono molte parole ma in contesti del tutto differenti potrebbero apparire artificialmente vicini.

3.2.2 Indice di Jaccard

La seconda tecnica è l'indice di Jaccard⁸, un metodo semplice e intuitivo che considera i testi come insiemi di parole. La similarità viene calcolata rapportando il numero di termini in comune all'insieme complessivo delle parole utilizzate nei due testi:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

dove A e B sono gli insiemi dei termini presenti rispettivamente nei due testi.

Il risultato varia tra 0 (nessuna parola in comune) e 1 (insiemi identici). Pur non tenendo conto della frequenza delle parole né del contesto semantico, l'indice di Jaccard è utile come misura di base per valutare la sovrapposizione lessicale e fornire un primo livello di confronto.

La debolezza di questa misura emerge soprattutto con documenti lunghi: poiché la probabilità di avere molte parole identiche diminuisce con l'aumentare del testo e i punteggi risultano quasi sempre bassi. Ne deriva una sottostima della similarità, anche quando i contenuti affrontano temi molto simili.

3.2.3 Sentence-BERT all-mpnet-base-v2

La terza tecnica adottata appartiene alla famiglia dei modelli neurali di tipo Sentence-BERT. In particolare, è stato utilizzato il modello *all-mpnet-base-v2*⁹, tra i più performanti per il calcolo della similarità semantica.

⁸ Jaccard. Risorse: <https://mayurdhvajsinhjadeja.medium.com/jaccard-similarity-34e2c15fb524>

⁹ *all-mpnet-base-v2*. Risorse: <https://huggingface.co/sentence-transformers/all-mpnet-base-v2>

A differenza dei metodi lessicali, BERT e le sue varianti trasformano i testi in *embedding* vettoriali, cioè rappresentazioni numeriche che catturano il significato del contenuto. L'idea alla base di BERT (*Bidirectional Encoder Representations from Transformers*) è di analizzare ogni elemento del testo considerando contemporaneamente il contesto che lo precede e quello che lo segue, ovvero tenendo conto sia delle parti che vengono prima sia di quelle che vengono dopo, così da poter interpretare al meglio il significato generale. In pratica, ogni frase o parola viene convertita in una sequenza di numeri e collocata in uno spazio vettoriale molto ampio: invece di avere solo due o tre dimensioni, come in un grafico tradizionale, ciascun testo viene rappresentato come un vettore di 768 dimensioni, valore che corrisponde alla dimensione dello spazio utilizzata da questo specifico modello per generare le sue rappresentazioni. In questo spazio, la vicinanza tra due frasi non è fisica ma concettuale: frasi con significati simili finiscono vicine, mentre frasi con significati diversi risultano lontane.

La similarità tra due frasi si calcola quindi come *cosine similarity* tra i rispettivi embedding. Questo approccio permette di cogliere relazioni semantiche molto più profonde: due testi possono risultare simili anche se non condividono molte parole, purché esprimano concetti affini. Il limite principale riguarda soprattutto l'impatto computazionale: l'elaborazione richiede tempo e risorse elevate. Inoltre, come tutti i modelli addestrati su grandi dataset, i risultati dipendono dalla qualità e dai bias presenti nei dati di training, che possono influenzare le analisi.

3.2.4 Sentence-BERT all-MiniLM-L6-v2

La quarta tecnica è anch'essa un modello Sentence-BERT, nella variante *all-MiniLM-L6-v2*¹⁰. Rispetto al precedente, questo modello è stato progettato per essere più leggero ed efficiente dal punto di vista computazionale, riducendo il numero di parametri e la profondità della rete neurale.

Pur essendo meno complesso, mantiene buone prestazioni nel calcolo della similarità semantica ed è quindi particolarmente utile in scenari in cui è necessario elaborare grandi quantità di testi in tempi contenuti. Anche in questo caso la similarità viene calcolata come *cosine similarity* sugli embedding generati dal modello.

La criticità principale riguarda la precisione: la maggiore velocità e leggerezza del modello si traduce in una minore capacità di cogliere tutte le sfumature di significato, specialmente nei

¹⁰ *all-MiniLM-L6-v2*. Risorse: <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>

testi più complessi o articolati. Si tratta quindi di un buon compromesso, ma non sempre raggiunge il livello di accuratezza dei modelli più grandi.

3.3 Raccolta e struttura dei dati

L'analisi per ogni direttiva si basa su un file CSV che raccoglie in formato strutturato i testi dei feedback inviati dalle organizzazioni nell'ambito delle consultazioni pubbliche. La costruzione di questo dataset ha previsto in una fase preliminare lo scaricamento dei documenti originali attraverso tecniche di *web scraping* dai portali ufficiali della Commissione europea, seguito da operazioni di pulizia e normalizzazione del testo. Queste fasi, non svolte direttamente nell'ambito di questa tesi, hanno permesso di ottenere un insieme di documenti pronti per essere utilizzati all'interno delle analisi svolte.

I file CSV utilizzati per l'analisi includono diverse colonne descrittive, ma per questo studio sono state considerate esclusivamente le variabili ritenute rilevanti, ovvero:

- **Organisation:** il nome dell'organizzazione che ha inviato il feedback;
- **User type:** la tipologia di attore (ad esempio impresa, associazione, cittadino, NGO);
- **Clean text:** il contenuto del feedback, dopo la fase di pulizia;
- **Act/Phase:** l'indicazione della finestra temporale in cui il contributo è stato inviato, distinguendo tra la fase iniziale di consultazione (Roadmap/Call for Evidence) e quella successiva all'adozione della proposta (Commission Adoption).

Questa struttura consente di associare in maniera univoca ogni feedback a un'organizzazione, a una tipologia di attore e a una precisa fase del processo legislativo, offrendo quindi il punto di partenza per il calcolo dei valori di similarità e per le successive comparazioni tra i diversi attori.

3.4 Procedura operativa

Dopo aver definito le tecniche da utilizzare, è stato necessario applicarle in modo sistematico a ciascuna direttiva analizzata. L'obiettivo era misurare, per ogni organizzazione e per ogni suo feedback, quanto i testi inviati risultassero simili ai documenti ufficiali pubblicati dalla Commissione nelle diverse fasi del processo legislativo.

Per ogni direttiva sono stati innanzitutto individuati i testi di riferimento, che corrispondono

alle principali tappe dell'iter: il documento di avvio (*Inception impact assessment*), la proposta legislativa della Commissione (*Proposal for a regulation*) e il testo finale approvato (*Final Version*). Questi testi costituiscono i punti di confronto rispetto ai quali valutare la somiglianza dei feedback ricevuti.

I contributi delle organizzazioni sono stati poi suddivisi in base alla fase in cui erano stati inviati: da un lato quelli relativi alla *Call for Evidence* o *Roadmap*, ossia la consultazione iniziale, e dall'altro quelli raccolti nella fase di *Commission Adoption*, avviata dopo la presentazione della proposta. Indipendentemente dalla fase di provenienza, tutti i feedback sono stati confrontati con l'insieme dei documenti ufficiali della Commissione — *Inception Impact Assessment*, *Proposal for a Regulation* e *Final Version* — al fine di ottenere una misura completa e comparabile della similarità.

Una volta definito l'insieme dei documenti di confronto, per ogni feedback sono stati calcolati i valori di similarità attraverso le quattro tecniche scelte: TF-IDF, indice di Jaccard, Sentence-BERT *all-MiniLM-L6-v2* e Sentence-BERT *all-mpnet-base-v2*. Ogni metodo restituisce un punteggio compreso tra 0 e 1, dove valori più alti indicano una maggiore vicinanza tra i testi. L'impiego combinato di tecniche semplici e complesse ha permesso di mettere a confronto i risultati ottenuti, così da poter valutare sia la semplicità degli approcci tradizionali sia la capacità dei modelli neurali di cogliere relazioni semantiche più profonde.

Per rendere più semplice l'analisi comparativa, i valori di similarità calcolati con le diverse tecniche sono stati infine aggregati in un indicatore unico, ottenuto tramite media. Questa scelta ha permesso di ridurre la complessità dei dati e di disporre, per ciascun confronto, di un solo valore su cui lavorare. In fase di aggregazione, tuttavia, non è stato considerato l'indice di Jaccard: i punteggi restituiti da questa misura risultavano infatti sistematicamente molto più bassi rispetto a quelli delle altre tecniche, e la loro inclusione avrebbe abbassato in modo drastico il valore medio.

Il risultato di questa procedura è una tabella finale che riporta, per ogni organizzazione, i valori medi di similarità dei feedback rispetto ai testi ufficiali. Tale tabella costituisce la base per le analisi successive, che, attraverso diverse rappresentazioni grafiche, mirano a individuare le organizzazioni e le categorie più influenti e più vicine alle posizioni della Commissione, oltre a evidenziare eventuali pattern ricorrenti.

3.5 Fuzzy matching per l'uniformazione dei nomi

Un passaggio preliminare fondamentale per garantire la coerenza dei risultati ha riguardato l'uniformazione dei nomi delle organizzazioni. Non di rado, infatti, lo stesso soggetto appare con denominazioni leggermente diverse nelle varie fasi del processo: abbreviazioni, sigle, traduzioni o semplici refusi possono rendere difficile riconoscere che si tratta della stessa organizzazione. Ad esempio, un'azienda potrebbe comparire come *Company Ltd.* in un feedback e come *Company Limited* in un altro, oppure un'associazione internazionale potrebbe presentarsi una volta con il nome in inglese e un'altra volta con quello nella lingua nazionale.

Per ovviare a questo problema è stato applicato un metodo di *fuzzy matching*¹¹, che consente di confrontare stringhe di testo non solo sulla base dell'identità esatta, ma anche della loro somiglianza. A differenza di un controllo tradizionale, che restituirebbe “diversi” due nomi anche se differiscono solo per una lettera o per una sigla, il fuzzy matching assegna un punteggio numerico di similarità compreso tra 0 e 100, che riflette quanto due stringhe possano essere considerate simili.

Dal punto di vista scientifico, molte tecniche di fuzzy matching si basano sulla cosiddetta *Levenshtein distance*¹² (o *edit distance*), che misura il numero minimo di operazioni elementari (inserzioni, cancellazioni, sostituzioni di caratteri) necessarie per trasformare una stringa nell'altra. La distanza così calcolata può essere trasformata in un indice di similarità normalizzato:

$$\text{Sim}(s_1, s_2) = \left(1 - \frac{D(s_1, s_2)}{\max(|s_1|, |s_2|)}\right) \times 100$$

dove $D(s_1, s_2)$ rappresenta la distanza di Levenshtein tra le due stringhe e $|s|$ indica la lunghezza della stringa s .

Con questa formula, due stringhe identiche ottengono un punteggio di similarità pari a 100, mentre due stringhe completamente diverse si avvicinano a 0. In questo caso la misurazione è stata implementata tramite la libreria *RapidFuzz*, che consente di calcolare in modo rapido i punteggi e di individuare, per ciascun nome di organizzazione presente nella fase di Adoption, il corrispondente più vicino tra quelli già comparsi nella Roadmap. In questo modo è stato possibile accoppiare organizzazioni che, pur essendo le stesse, venivano indicate con denominazioni differenti.

Per ridurre il rischio di associazioni errate è stata inoltre stabilita una soglia minima di simila-

¹¹*Fuzzy matching*. Risorse: Medium - Fuzzy Matching Algorithms.

¹²*Levenshtein distance*. Risorse: Medium - Levenshtein Distance.

rità, definita in maniera specifica per ciascuna analisi: al di sotto della quale i nomi sono stati mantenuti distinti.

3.6 Dalla tabella finale all'analisi dei risultati

La costruzione della tabella finale rappresenta il punto di arrivo della fase metodologica e, al tempo stesso, il punto di partenza per le analisi successive. Grazie all'aggregazione dei punteggi di similarità, infatti, per ogni organizzazione si dispone di un insieme di metriche che ne descrive il grado di vicinanza ai testi ufficiali nelle diverse fasi del processo legislativo.

Questa struttura consente di effettuare numerosi confronti e approfondimenti. È ad esempio possibile calcolare i *delta* di similarità tra una fase e l'altra, osservando quanto un'organizzazione si avvicini o si allontani dal testo ufficiale passando dall'*Inception Impact Assessment* alla *Proposal for a Regulation*, oppure dalla *Proposal for a Regulation* alla *Final Version*. Tali differenze forniscono un'indicazione preziosa sulla capacità dei feedback di incidere sui contenuti normativi o, al contrario, sulla stabilità delle posizioni istituzionali nel tempo.

La tabella consente inoltre di sviluppare analisi comparative tra diverse tipologie di attori, mettendo a confronto, ad esempio, le imprese con le organizzazioni non governative (NGO), le associazioni di categoria con le associazioni a tutela dei consumatori. In questo modo è possibile individuare quali gruppi risultino mediamente più vicini alle posizioni della Commissione e in quali fasi del processo legislativo si concentri maggiormente la loro influenza.

Tutti questi elementi verranno approfonditi nel capitolo successivo, dedicato ai risultati dell'analisi. Lì verranno presentati i delta calcolati e le comparazioni tra categorie e tipologie di organizzazioni, con l'obiettivo di valutare il ruolo dei feedback nel processo legislativo europeo.

Capitolo 4

Risultati

Dopo aver descritto nel capitolo precedente le tecniche utilizzate per misurare la similarità tra i testi, questo capitolo presenta i risultati dell'analisi condotta sulle quattro iniziative legislative considerate: *Artificial Intelligence Act*, *Sustainable Corporate Governance*, *Digital Services Act* e *Digital Markets Act*.

L'obiettivo è capire in che misura i contributi delle organizzazioni abbiano inciso sull'elaborazione dei testi normativi, mettendo in evidenza chi ha avuto più capacità di influenza e in quale fase del processo questa influenza si è fatta sentire maggiormente.

L'analisi si concentra su tre aspetti chiave:

- **Chi conta davvero?**

Sono stati individuati i soggetti che hanno mostrato i valori più elevati sia in termini di similarità con i testi ufficiali, sia in termini di variazioni complessive tra le diverse fasi. Questo doppio approccio permette di distinguere tra chi si è mantenuto costantemente vicino alle posizioni della Commissione e chi, invece, ha inciso maggiormente sui cambiamenti del testo.

- **Quali categorie spiccano?**

Le organizzazioni sono state raggruppate per tipologia – imprese, associazioni di categoria, NGO, mondo accademico, gruppi di consumatori, ecc. – e per ciascun gruppo è stata calcolata l'influenza media. Questo confronto ha permesso di evidenziare le categorie più incisive nel percorso legislativo.

- **Quando si riesce a influenzare di più?**

Infine sono state confrontate le due fasi principali della consultazione: la prima, che va dalla pubblicazione della *Call for Evidence* alla proposta della Commissione, e la seconda,

che va dalla proposta alla versione finale. In questo modo si può capire se i contributi risultano più efficaci nelle fasi iniziali o in quelle finali.

4.1 Artificial Intelligence Act

L'*Artificial Intelligence Act* è la prima iniziativa europea pensata per stabilire regole comuni sull'uso dell'intelligenza artificiale, con l'obiettivo di garantire sicurezza, trasparenza e tutela dei diritti fondamentali.

Per individuare gli attori più rilevanti, sono state analizzate due dimensioni: la similarità, che misura quanto le posizioni delle organizzazioni restano vicine a quelle della Commissione durante l'intero processo, e la variazione totale, che evidenzia quanto i loro contributi abbiano inciso sui testi ufficiali.

Le Figure seguenti mostrano la classifica delle dieci organizzazioni più influenti secondo entrambi i criteri.

Top 10 organizzazioni per Similarità - Artificial Intelligence Act

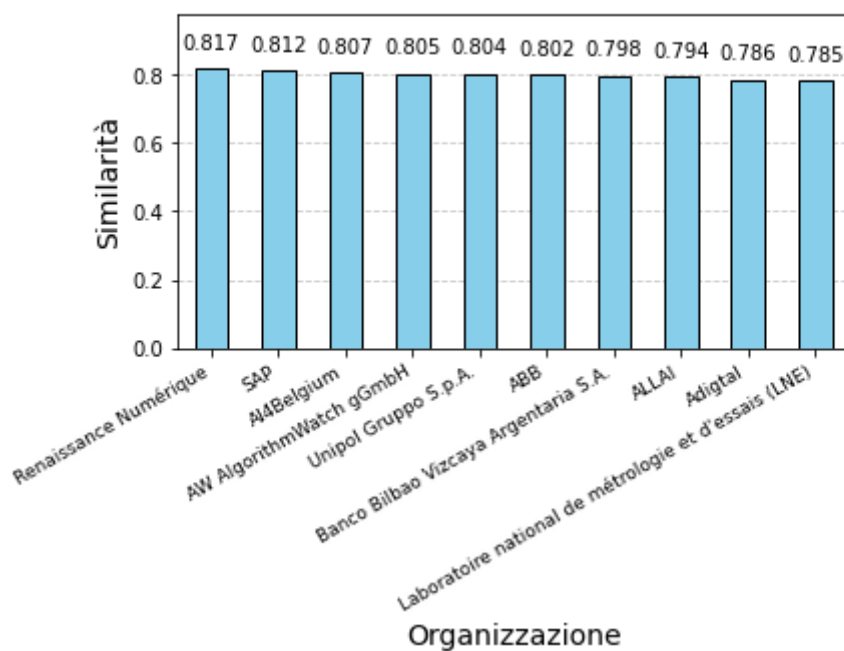


Figura 4.1: Top 10 Organizzazioni per Similarità

Top 10 organizzazioni per Variazione Totale - Artificial Intelligence Act

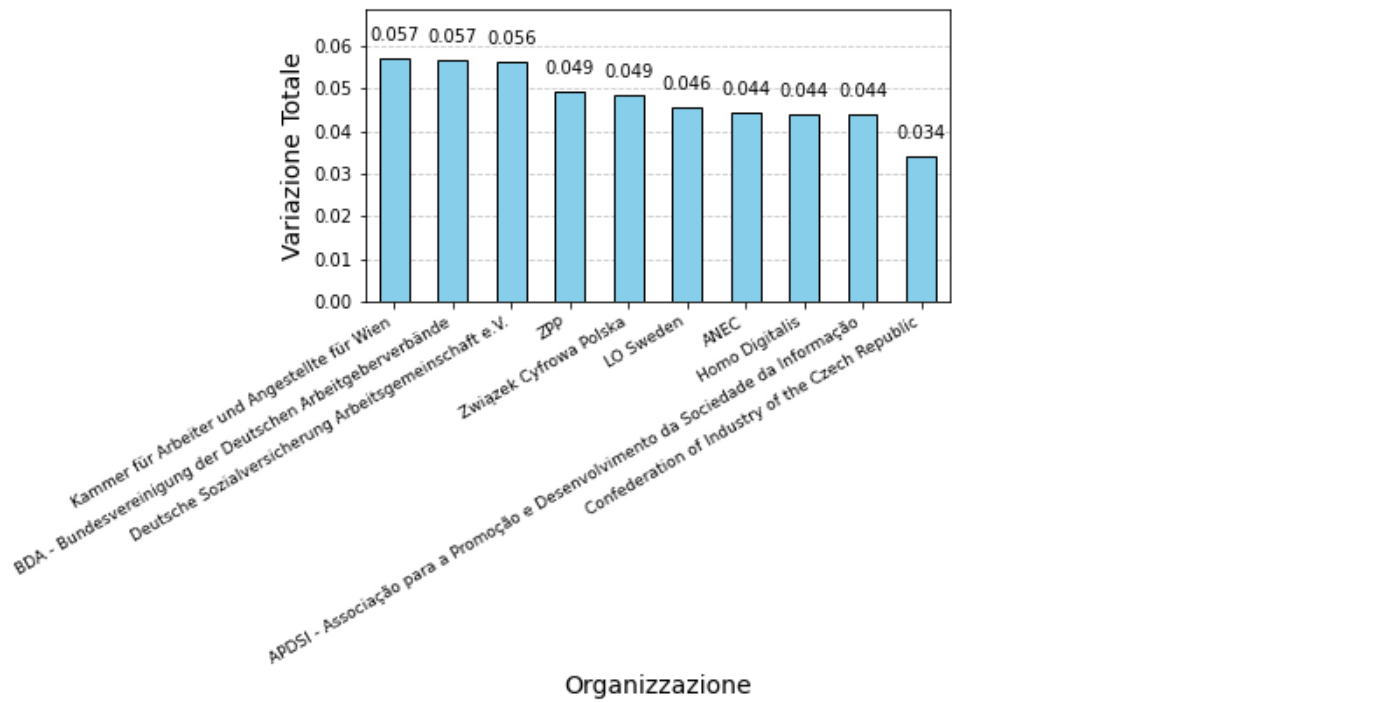


Figura 4.2: Top 10 Organizzazioni per Variazione Totale

Un'ulteriore rappresentazione è osservabile tramite lo scatter plot, che mostra la posizione relativa delle organizzazioni rispetto alle due dimensioni (similarità e variazione).



Figura 4.3: Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.

Lo scatter plot permette di osservare contemporaneamente le due dimensioni considerate: la similarità (asse verticale) e la variazione totale (asse orizzontale). Le linee rosse tratteggiate rappresentano i valori medi delle due variabili e suddividono lo spazio in quattro quadranti. Il quadrante in alto a destra evidenzia le organizzazioni che risultano sia molto vicine alle posizioni della Commissione (alta similarità), sia particolarmente capaci di incidere sul testo (alta variazione).

Considerando le diverse tipologie di organizzazioni, le figure seguenti mostrano la similarità media e l'influenza media calcolate per ciascuna categoria in relazione all'Artificial Intelligence Act, evidenziando i valori medi ottenuti per ogni gruppo di interesse.

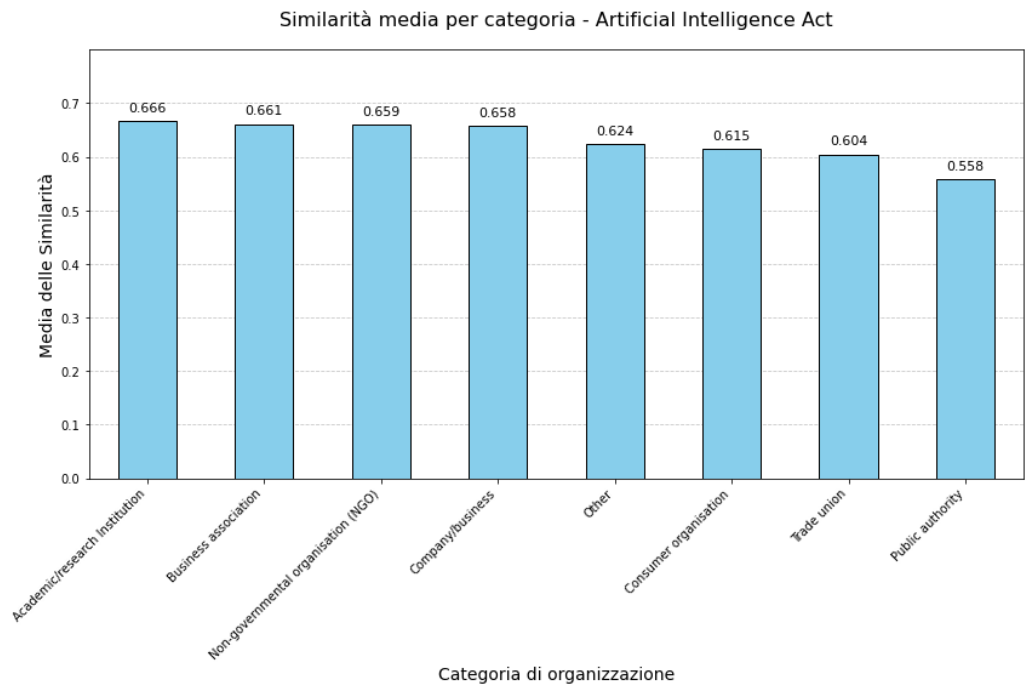


Figura 4.4: Similarità media per tipologia di organizzazione nell'AI Act.

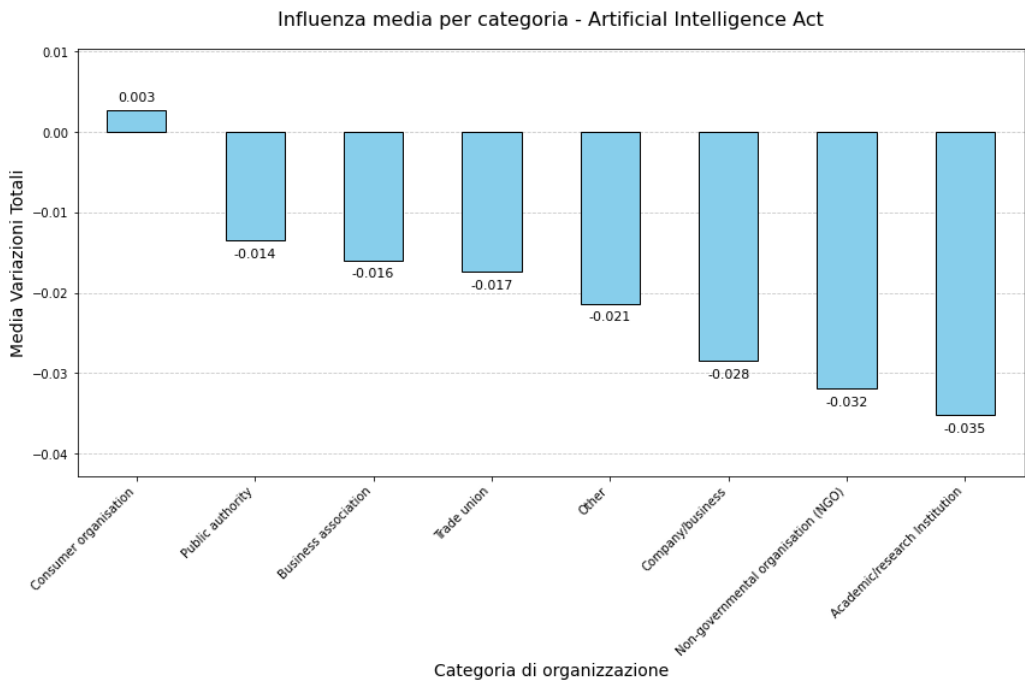


Figura 4.5: Influenza media per tipologia di organizzazione nell'AI Act.

Anche per le categorie, una rappresentazione aggiuntiva è fornita dallo scatter plot, che mostra la posizione relativa delle diverse categorie rispetto alle due dimensioni considerate: similarità e variazione.

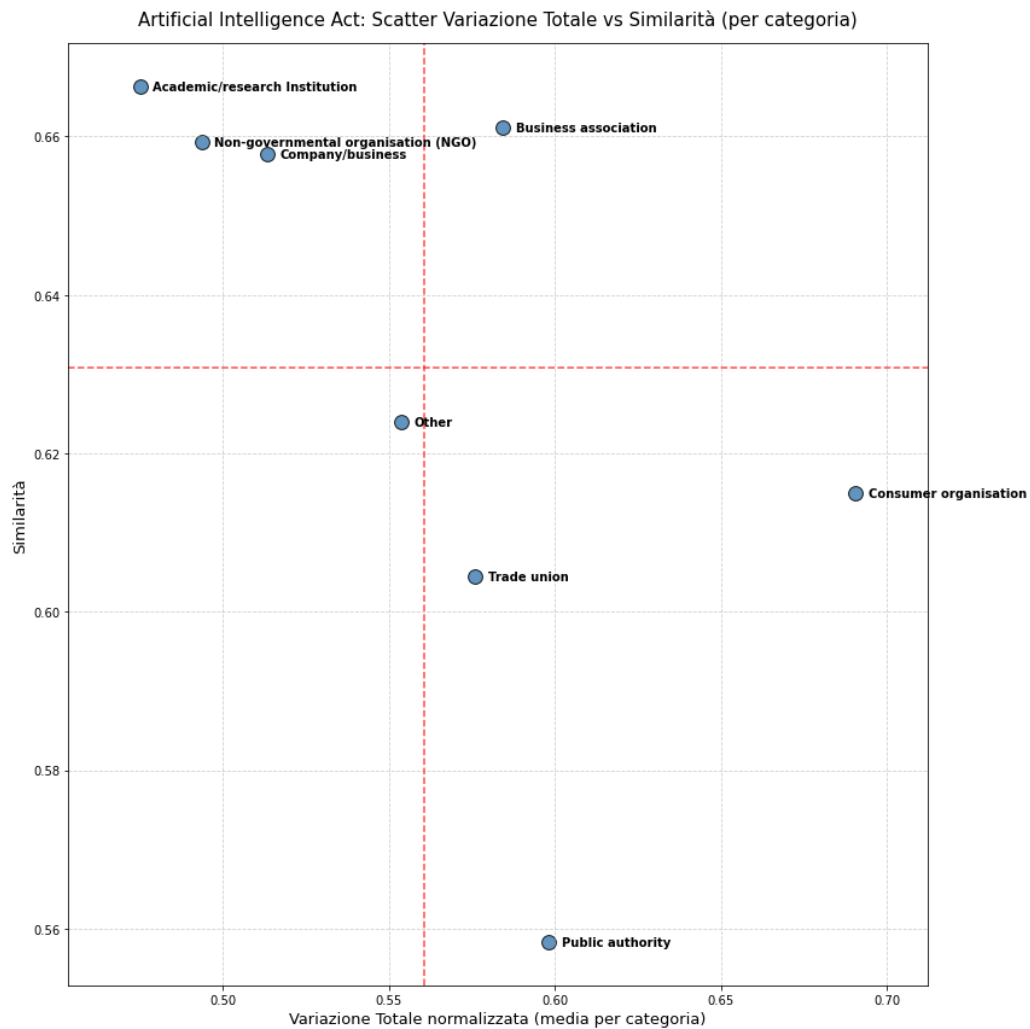


Figura 4.6: Distribuzione delle categorie in base a Similarità e Variazione Totale.

Anche la distribuzione temporale dell'influenza, come visibile nel grafico sottostante, è chiara: la fase iniziale della consultazione, dalla *Call for Evidence* alla proposta della Commissione, è quella in cui i contributi hanno inciso di più, con un incremento medio di circa **0,003**. Nella fase successiva, dalla proposta della Commissione alla versione finale, i valori si riducono leggermente (circa **-0,027**).

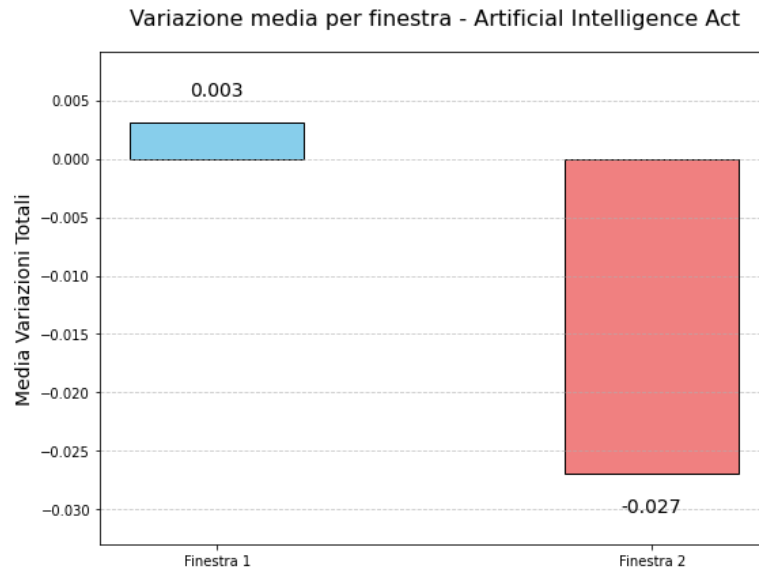


Figura 4.7: Variazione media per finestra di consultazione – Artificial Intelligence Act

4.2 Sustainable Corporate Governance

La direttiva sulla *Sustainable Corporate Governance* nasce con l'obiettivo di rafforzare la responsabilità delle imprese lungo le catene di fornitura, introducendo obblighi di T volti a garantire il rispetto dei diritti umani e la tutela dell'ambiente in tutte le fasi della produzione. L'analisi delle posizioni espresse dalle organizzazioni permette di identificare quali attori abbiano esercitato un'influenza significativa sul percorso legislativo. Le due classifiche che seguono riportano le dieci organizzazioni più vicine alle posizioni della Commissione e quelle che hanno inciso maggiormente sulle modifiche del testo.

Top 10 organizzazioni per Similarità - Sustainable Corporate Governance

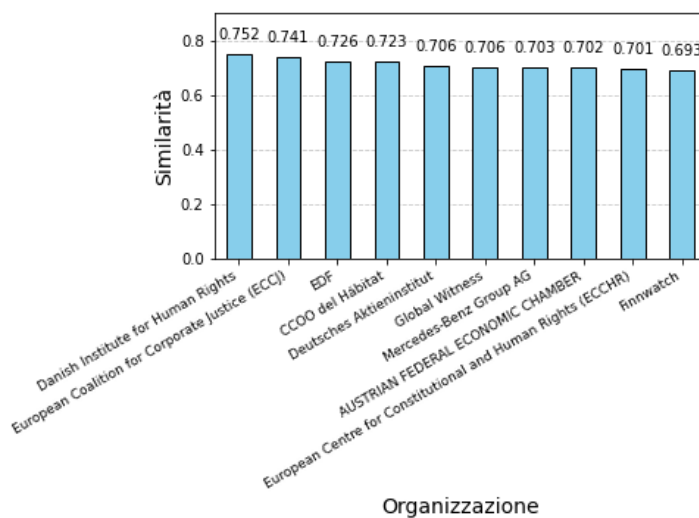


Figura 4.8: Top 10 Organizzazioni per Similarità

Top 10 organizzazioni per Variazione Totale - Sustainable Corporate Governance

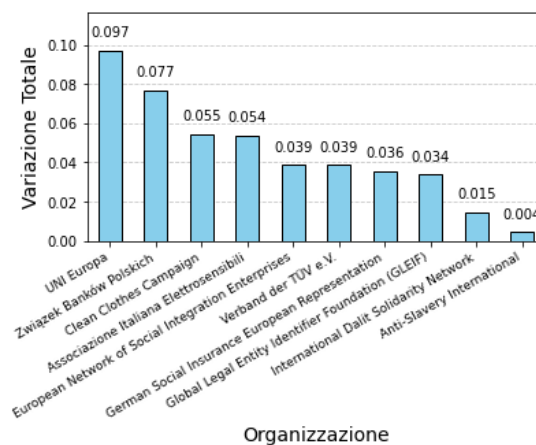


Figura 4.9: Top 10 Organizzazioni per Variazione Totale

Anche per questa direttiva è stato realizzato lo scatter plot che mette a confronto le due dimensioni considerate (similarità e variazione totale). Come già discusso nel caso precedente, le linee rosse tratteggiate suddividono lo spazio in quadranti e permettono di individuare gli attori più rilevanti nella parte in alto a destra del grafico.

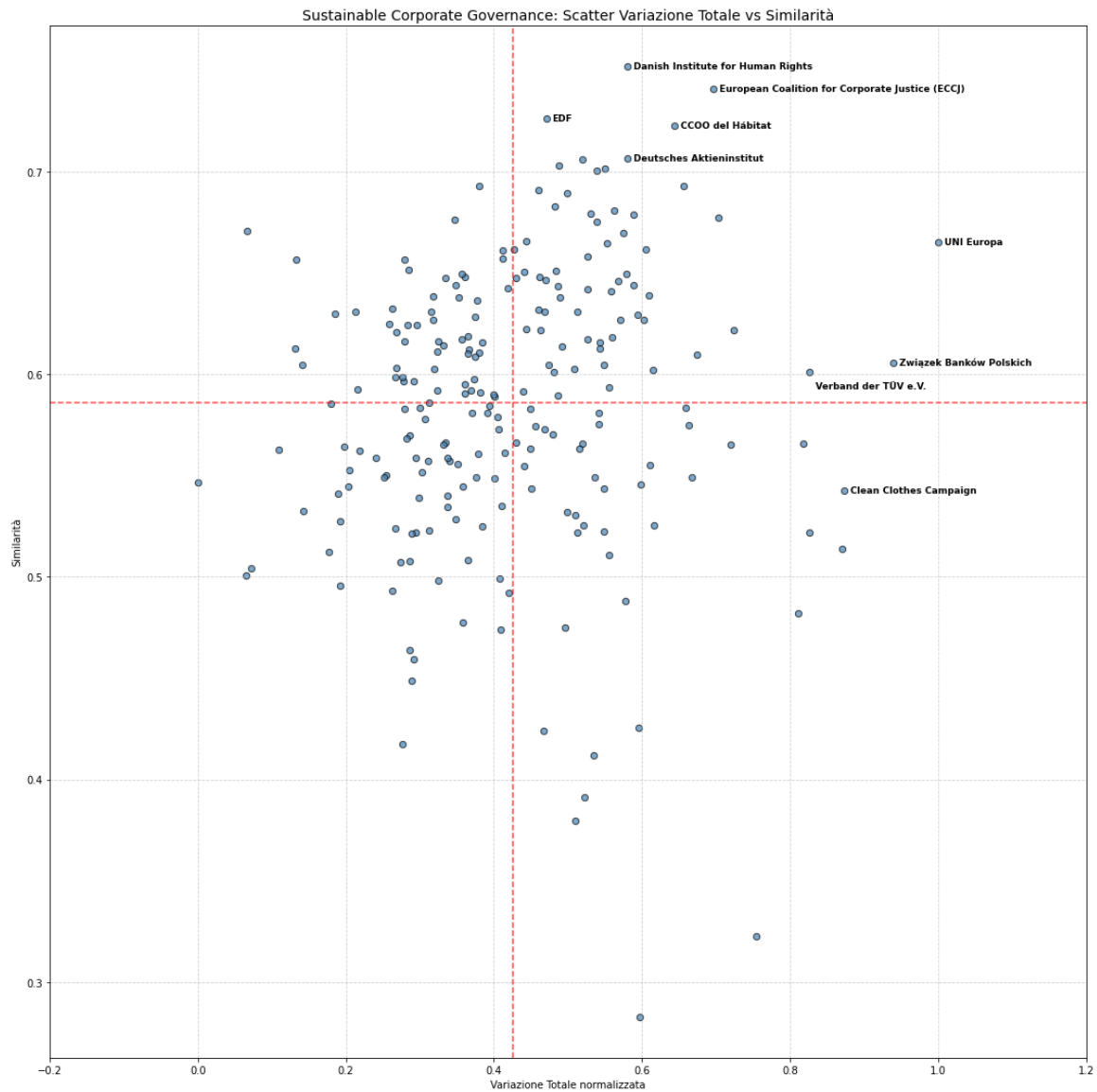


Figura 4.10: Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.

Per quanto riguarda le diverse tipologie di organizzazioni presenti nella Sustainable Corporate Governance, i grafici seguenti riportano i valori medi di similarità e influenza calcolati per ciascuna categoria, offrendo una panoramica delle tendenze dei diversi gruppi di interesse.

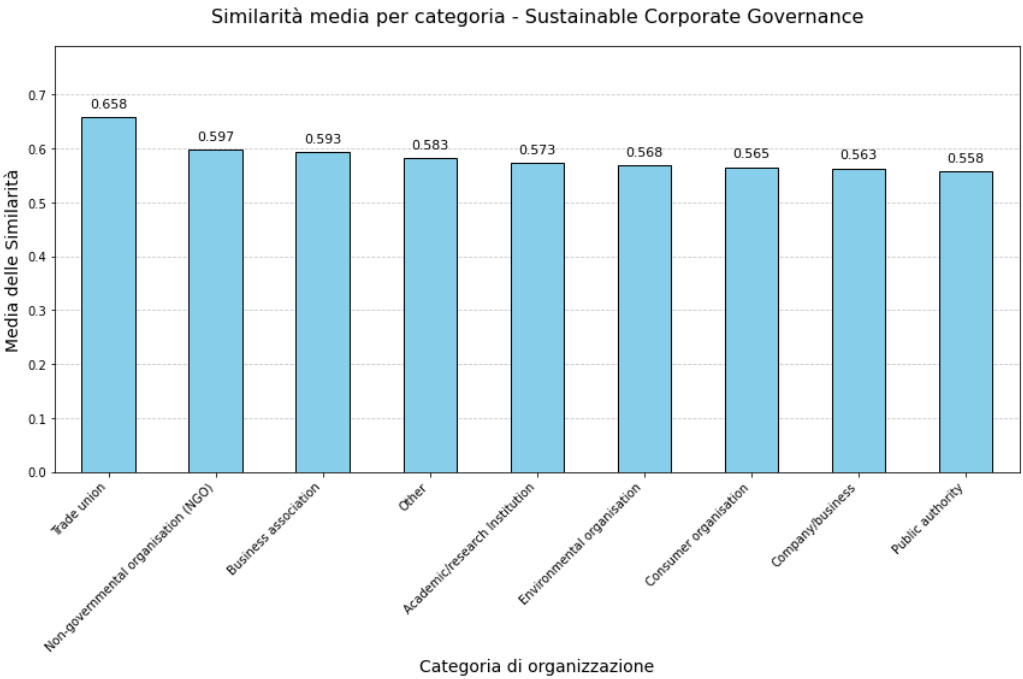


Figura 4.11: Similarità media per tipologia di organizzazione nella direttiva Sustainable Corporate Governance.

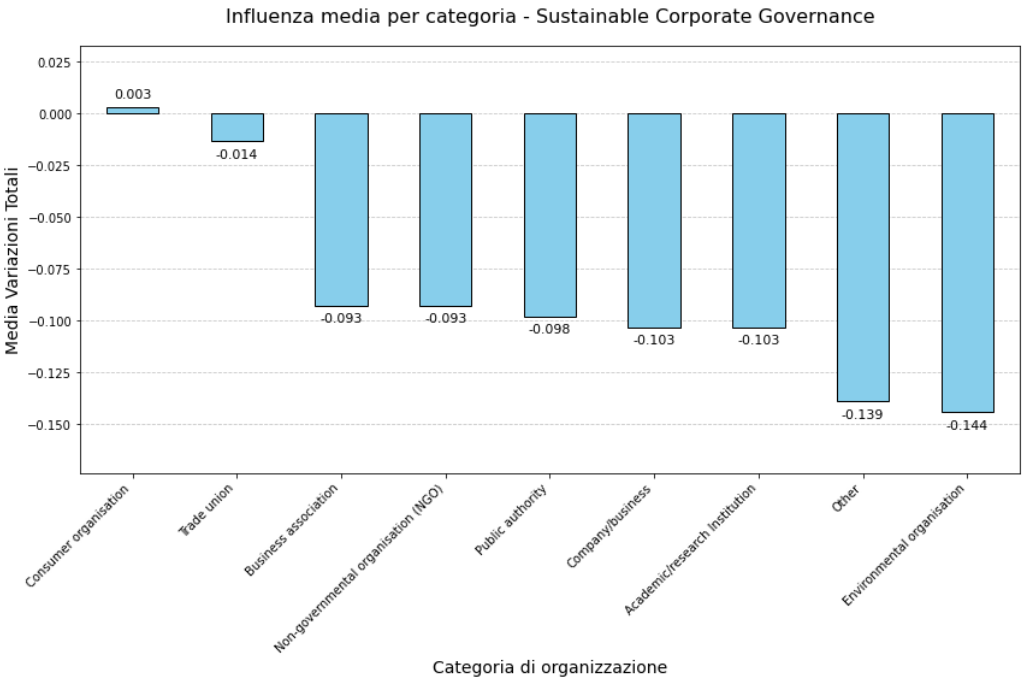


Figura 4.12: Influenza media per tipologia di organizzazione nella direttiva Sustainable Corporate Governance.

Anche in questo caso lo scatter plot offre un'ulteriore prospettiva, mostrando la posizione relativa delle diverse categorie rispetto alle due dimensioni considerate.

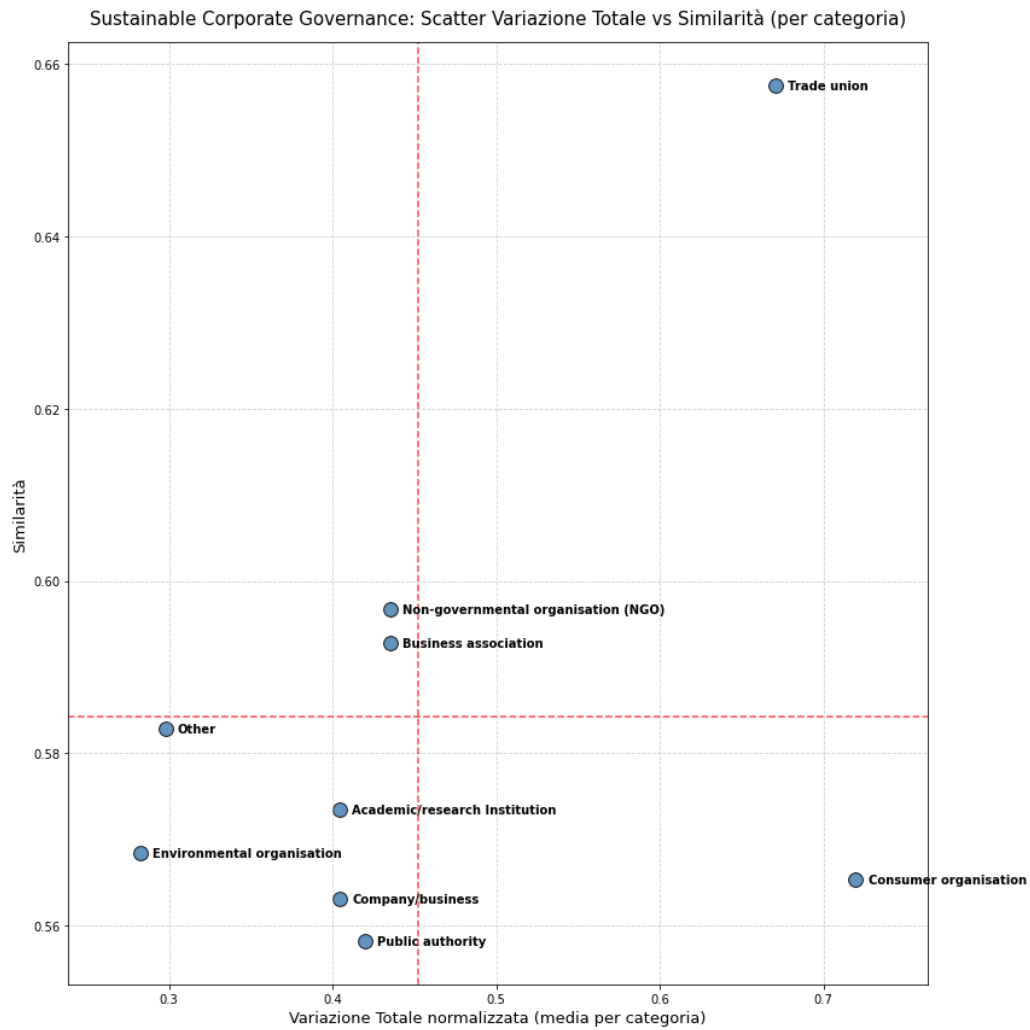


Figura 4.13: Distribuzione delle categorie in base a Similarità e Variazione Totale.

Dal punto di vista temporale, anche per questa direttiva la distribuzione dell'influenza, come riportato nel grafico che segue, mostra un andamento netto: la prima fase della consultazione (dalla *Call for Evidence* alla proposta della Commissione) è quella in cui i contributi hanno inciso di più, con un incremento medio positivo (circa **0,004**). Al contrario, nella seconda fase (dalla proposta al testo finale) la media diventa negativa (circa **-0,1**).

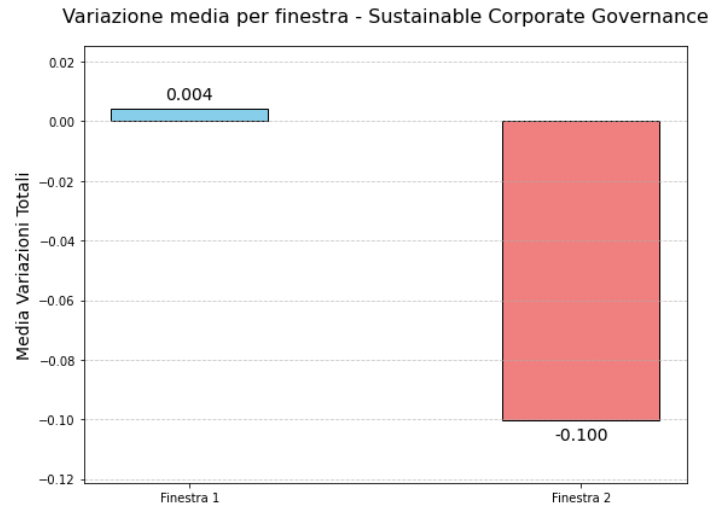


Figura 4.14: Variazione media per finestra di consultazione – Sustainable Corporate Governance

4.3 Digital Services Act

Il *Digital Services Act* affianca il *Digital Markets Act* e introduce regole comuni per la moderazione dei contenuti online, la gestione dei dati degli utenti e la trasparenza delle grandi piattaforme digitali.

Per il Digital Services Act i dati evidenziano la presenza di attori costantemente vicini alle posizioni della Commissione e di altri capaci di influenzare in modo significativo le variazioni del testo. Le Figure riportano le dieci organizzazioni più rilevanti per ciascun indicatore.

Top 10 organizzazioni per Similarità - Digital Services Act

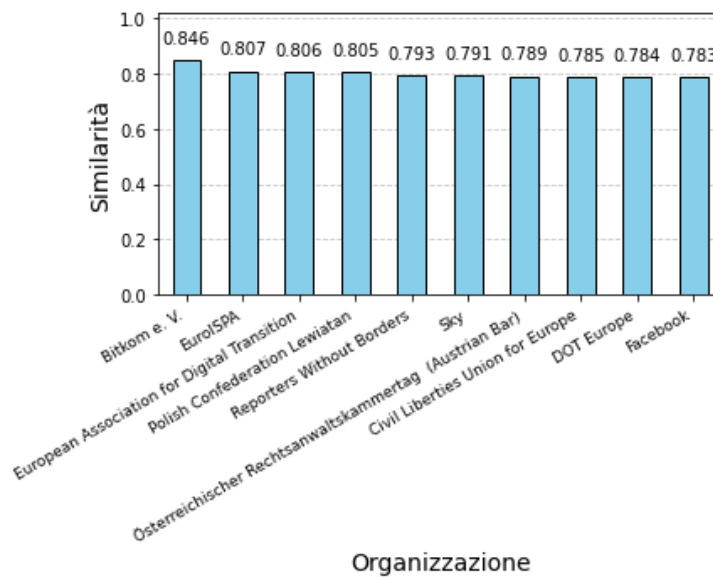


Figura 4.15: Top 10 Organizzazioni per Similarità

Top 10 organizzazioni per Variazione Totale - Digital Services Act

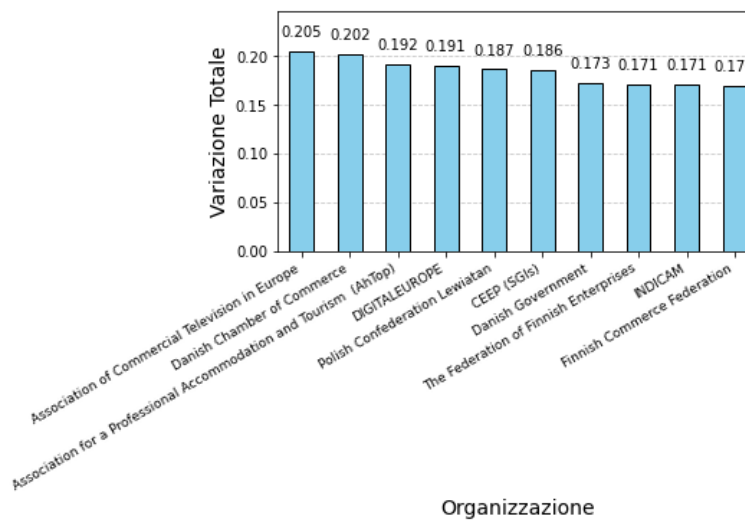


Figura 4.16: Top 10 Organizzazioni per Variazione Totale

Anche per questa direttiva è stato realizzato lo scatter plot, che mostra le organizzazioni in base alle due dimensioni considerate. Come già visto per i casi precedenti, il quadrante in alto a destra evidenzia gli attori che risultano sia molto vicini alle posizioni della Commissione, sia particolarmente capaci di incidere sulle modifiche del testo.

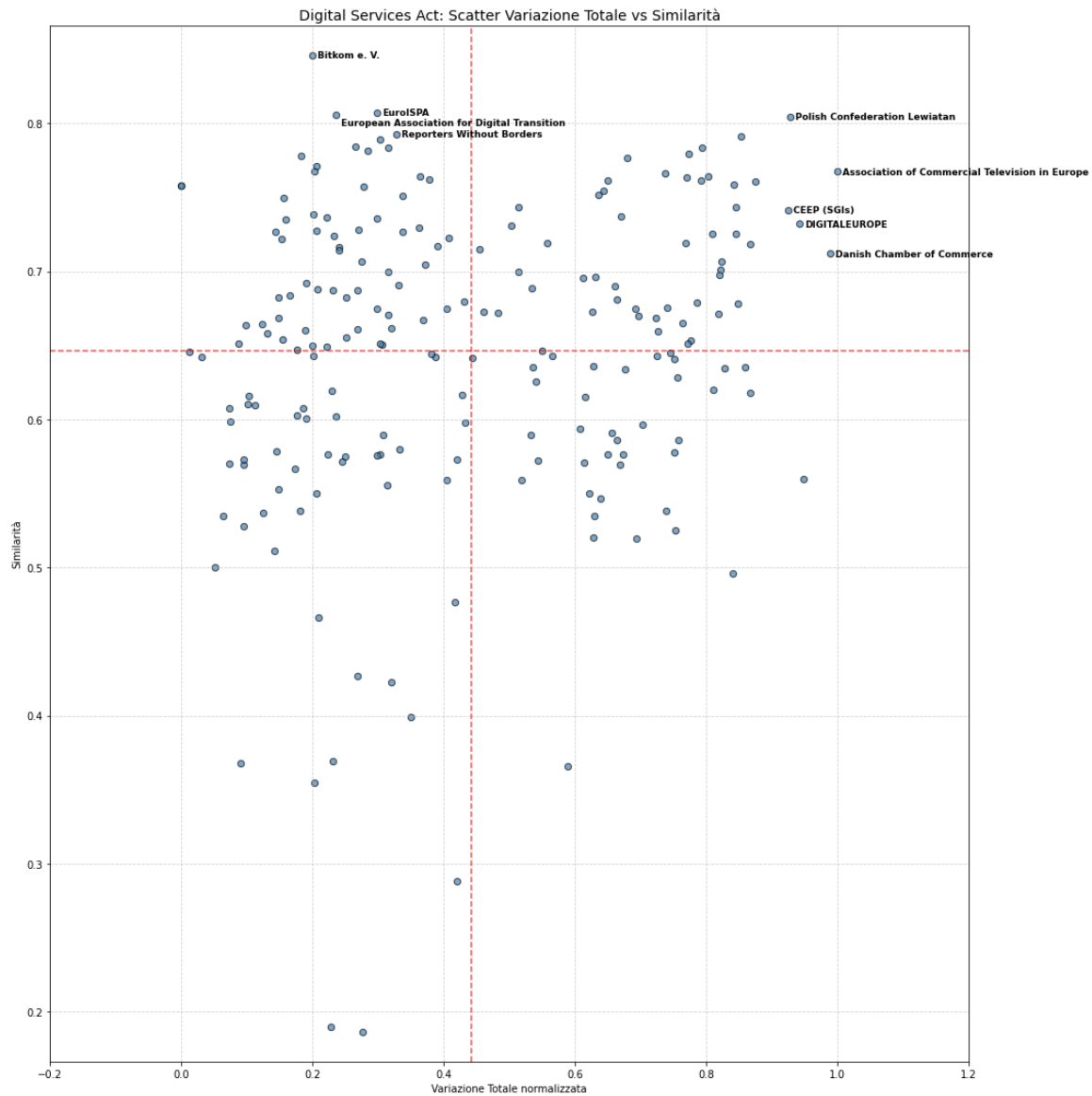


Figura 4.17: Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.

Analizzando le differenti categorie di organizzazioni, le figure sottostanti mostrano come variano la similarità media e l'influenza media rispetto al Digital Services Act, offrendo una panoramica comparativa tra i diversi gruppi di interesse.

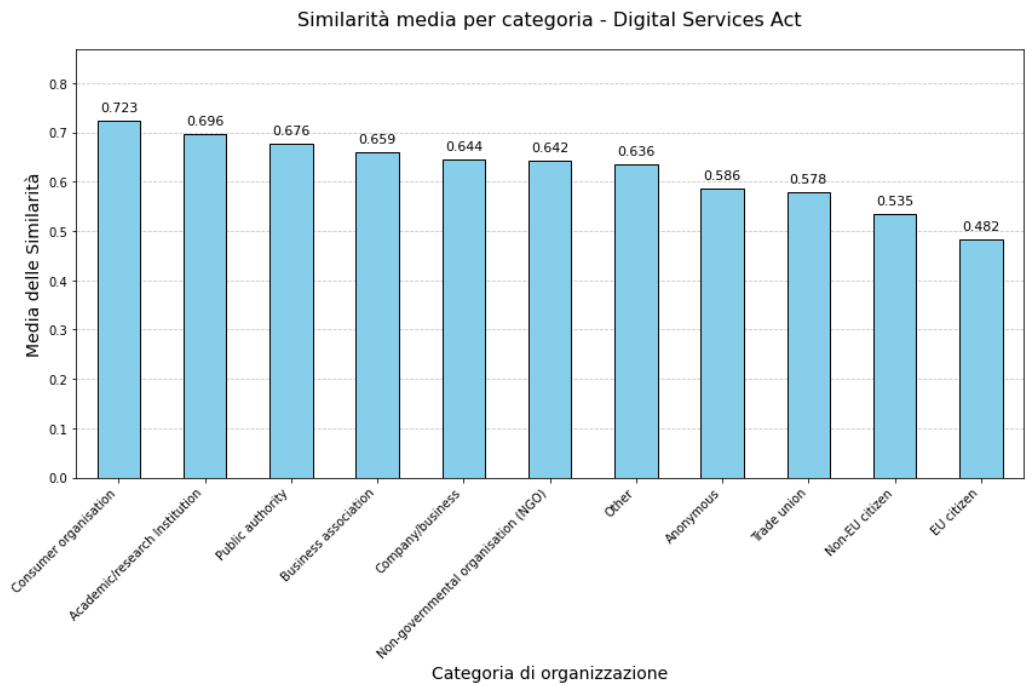


Figura 4.18: Similarità media per tipologia di organizzazione nel Digital Services Act.

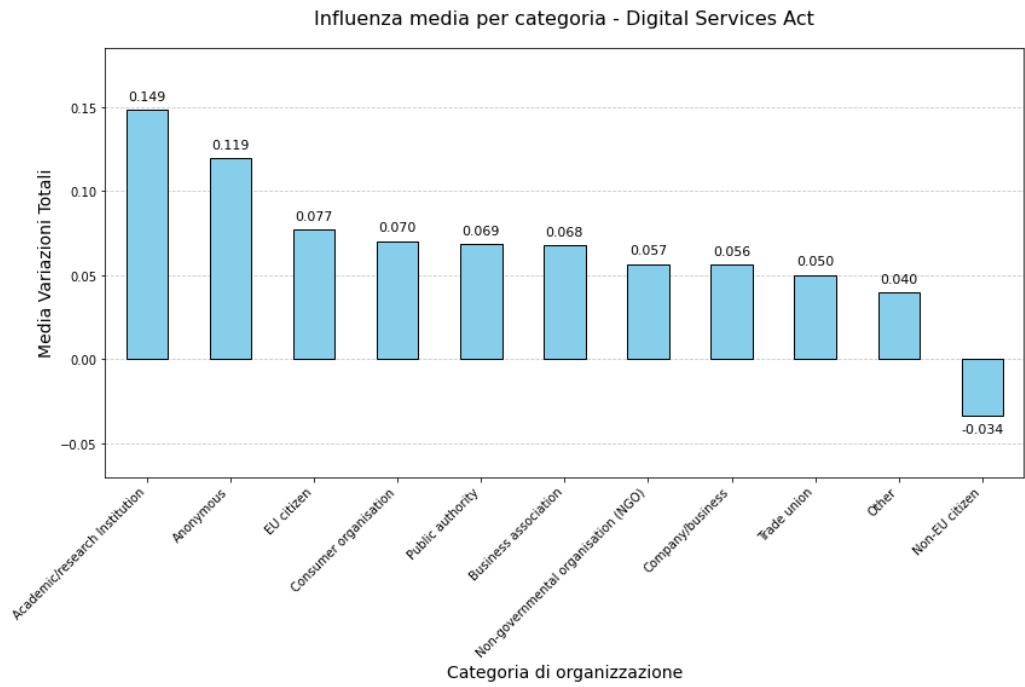


Figura 4.19: Influenza media per tipologia di organizzazione nel Digital Services Act.

Anche qui sempre grazie allo scatter plot possiamo osservare le posizioni relative alle diverse categorie presenti, rispetto alle dimensioni considerate.

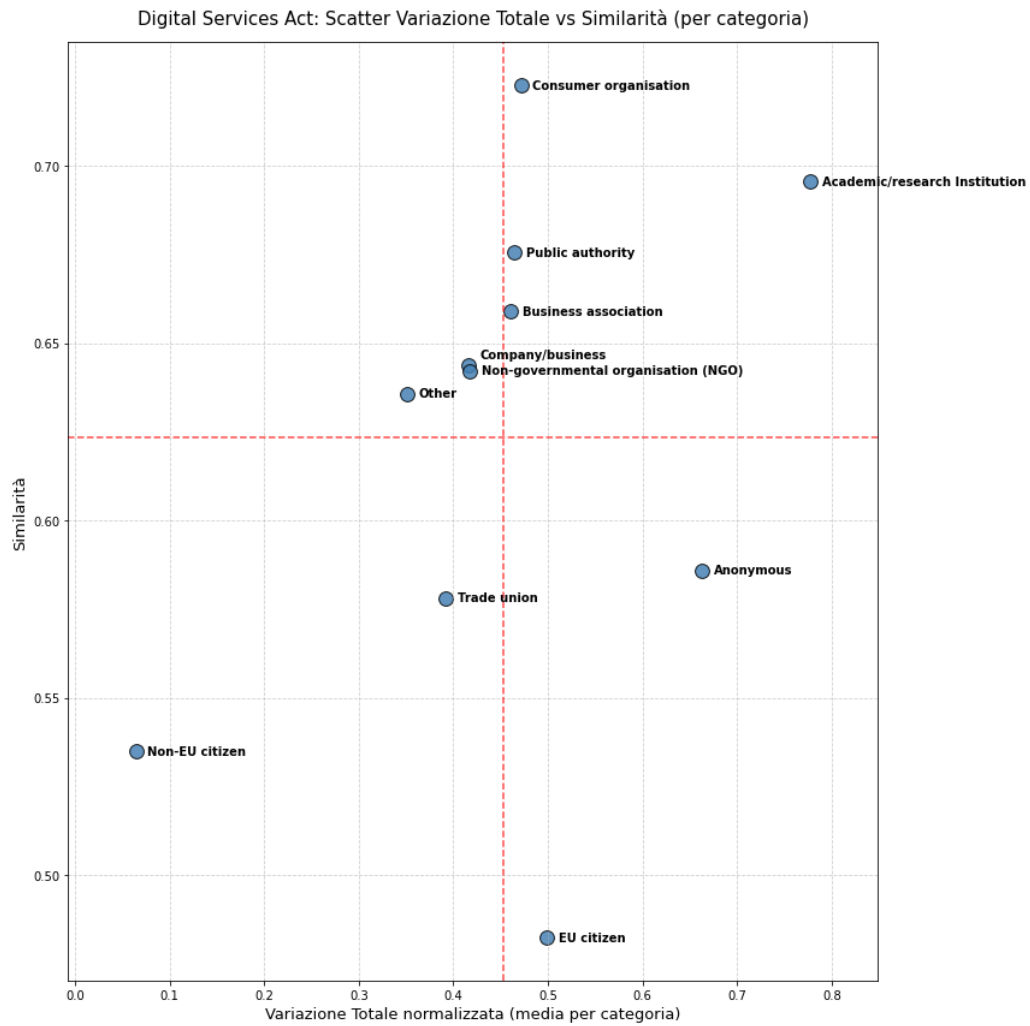


Figura 4.20: Distribuzione delle categorie in base a Similarità e Variazione Totale.

Dal punto di vista temporale, anche nel *Digital Services Act* la fase iniziale della consultazione (dalla *Call for Evidence* alla proposta della Commissione) si conferma cruciale, con un incremento medio di circa **0,056**, mentre nella seconda fase (dalla proposta al testo finale) il valore scende invece a circa **0,007**.

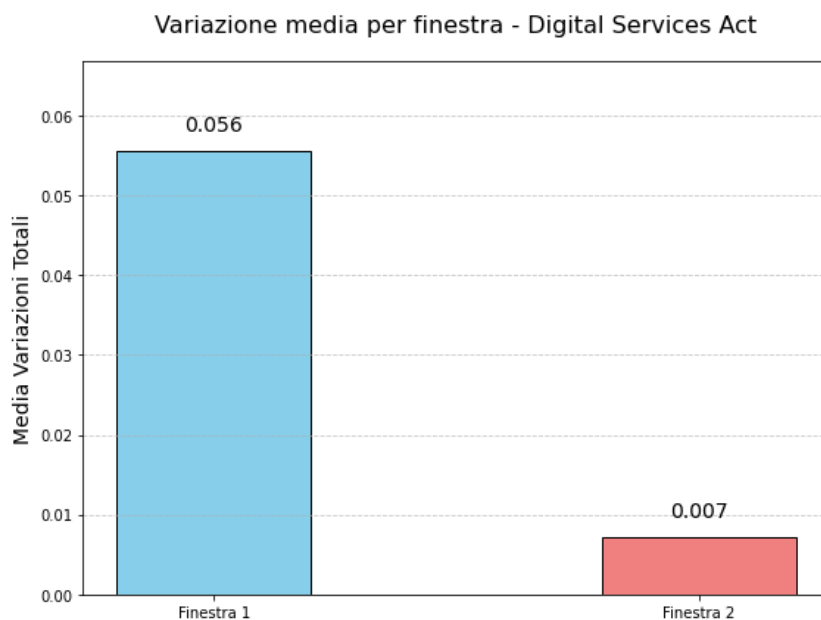


Figura 4.21: Variazione media per finestra di consultazione – Digital Services Act

4.4 Digital Markets Act

Il *Digital Markets Act* rappresenta la principale iniziativa europea volta a regolamentare i grandi operatori digitali, con l'obiettivo di garantire mercati più equi e contrastare le pratiche anticoncorrenziali delle cosiddette *gatekeeper platforms*.

Nel caso del Digital Markets Act, il confronto dei dati consente di individuare le organizzazioni che hanno svolto un ruolo di primo piano durante il processo legislativo. Le Figure riportano le dieci più rilevanti, distinte per similarità e per variazione totale.



Figura 4.22: Top 10 Organizzazioni per Similarità

Top 10 organizzazioni per Variazione Totale - Digital Markets Act

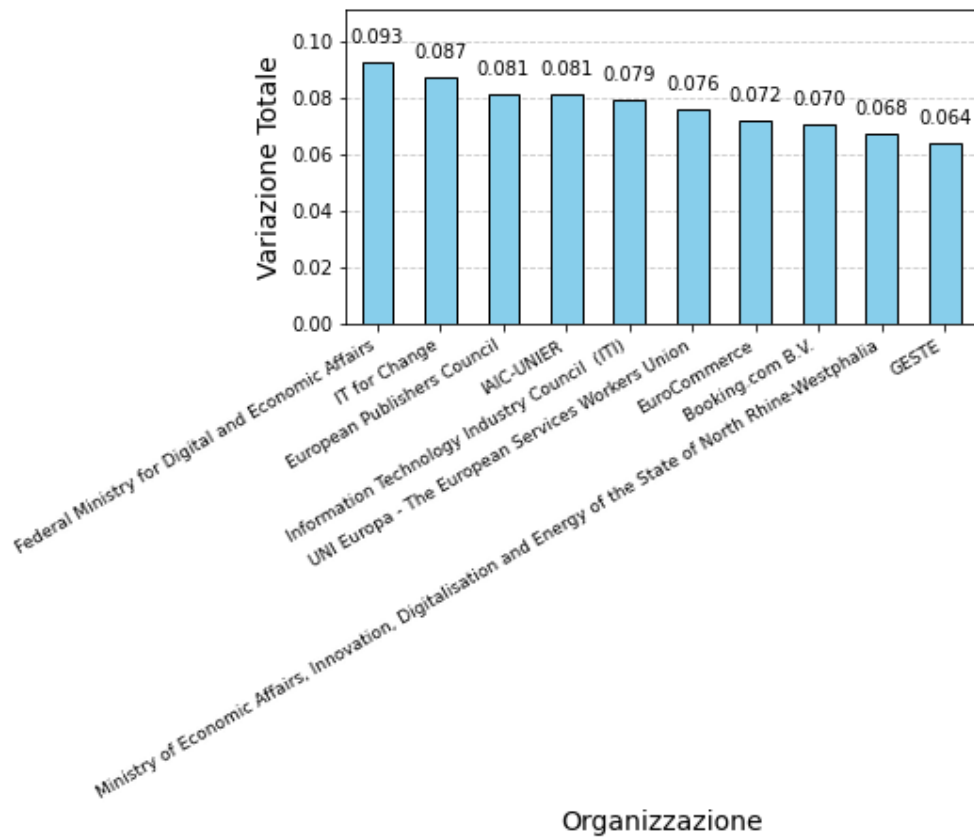


Figura 4.23: Top 10 Organizzazioni per Variazione Totale

La rappresentazione congiunta delle due dimensioni, offerta dallo scatter plot, mette in evidenza le organizzazioni che si collocano nel quadrante in alto a destra, ossia quelle caratterizzate da un'elevata similarità con le posizioni della Commissione e da un impatto significativo sui cambiamenti del testo.

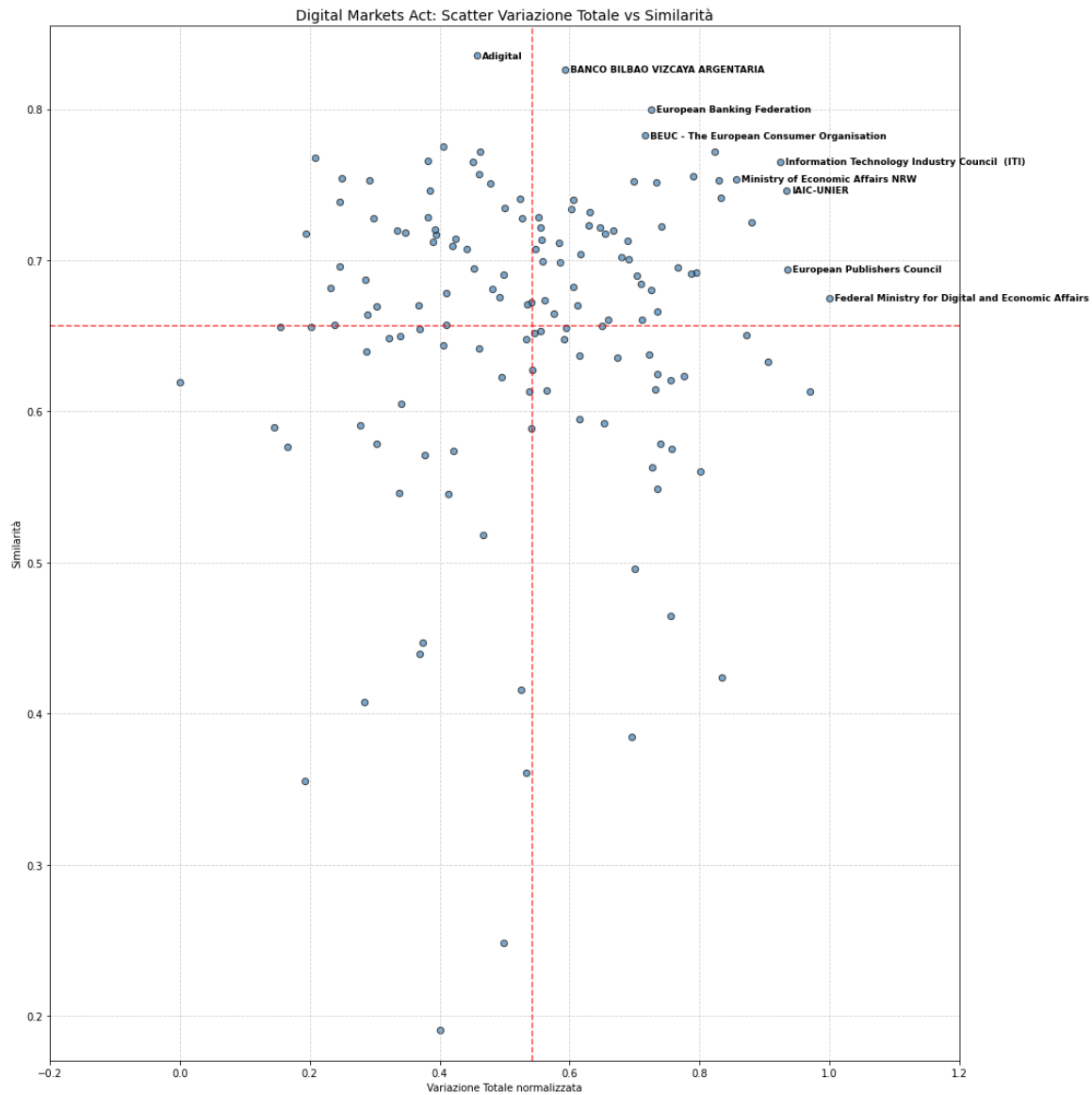


Figura 4.24: Distribuzione delle organizzazioni in base a Similarità e Variazione Totale.

Le figure al di sotto riportate relative al Digital Markets Act presentano la distribuzione dei valori medi di similarità e influenza per ciascuna tipologia di organizzazione e consentendo di osservare le differenze tra i vari gruppi di interesse coinvolti.

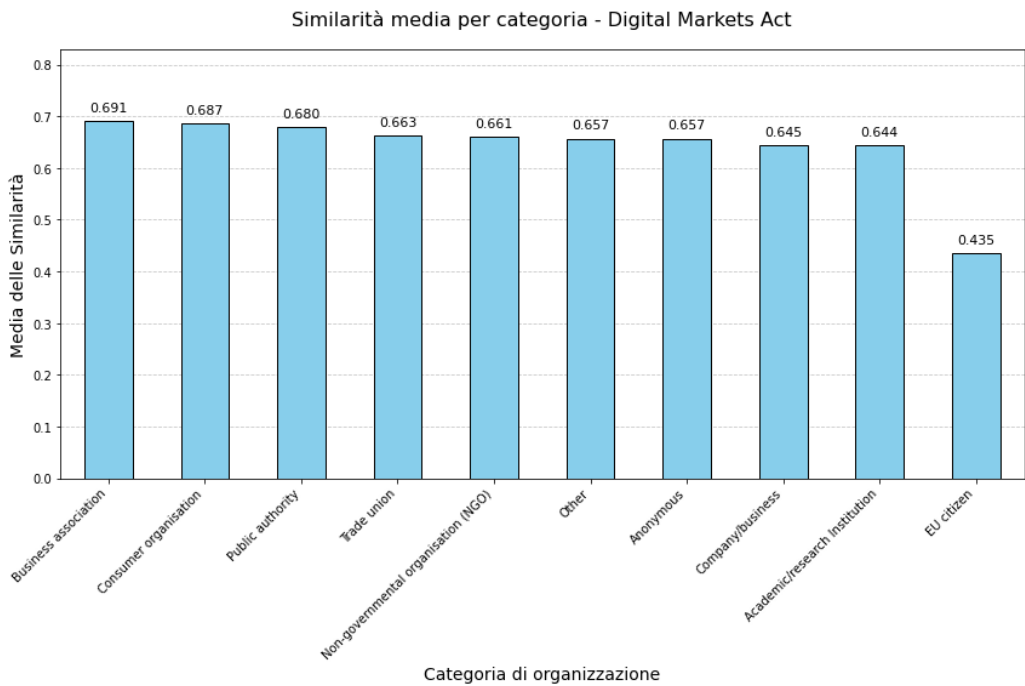


Figura 4.25: Similarità media per tipologia di organizzazione nel Digital Markets Act.

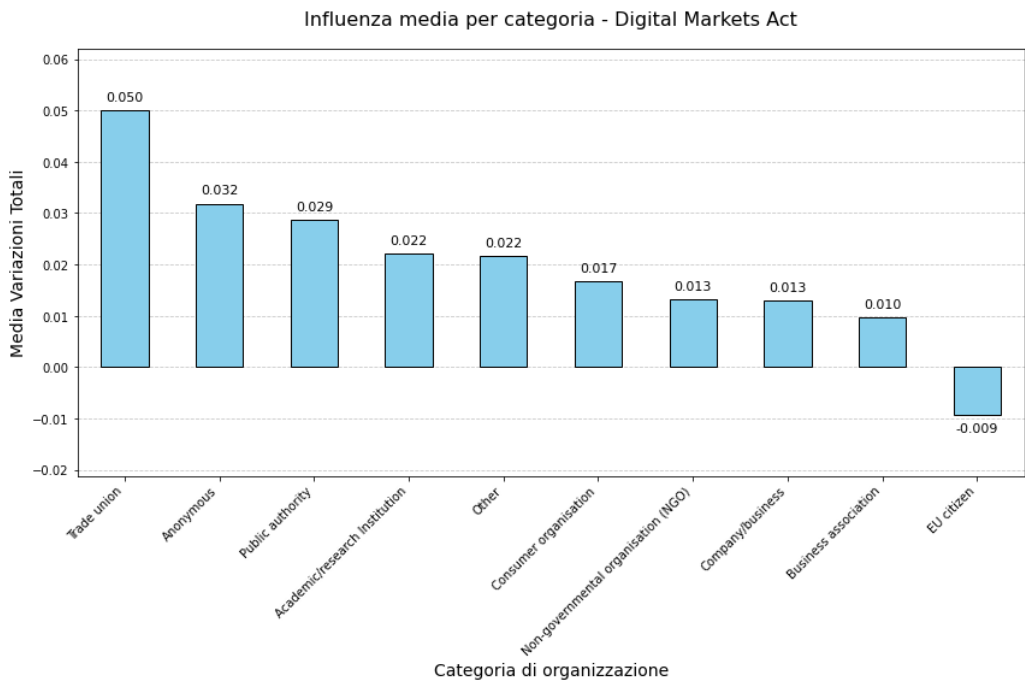


Figura 4.26: Influenza media per tipologia di organizzazione nel Digital Markets Act.

Come per i casi precedenti lo scatter plot ci aiuta ad osservare le posizioni relative alle diverse categorie presenti, rispetto alle due dimensioni considerate.

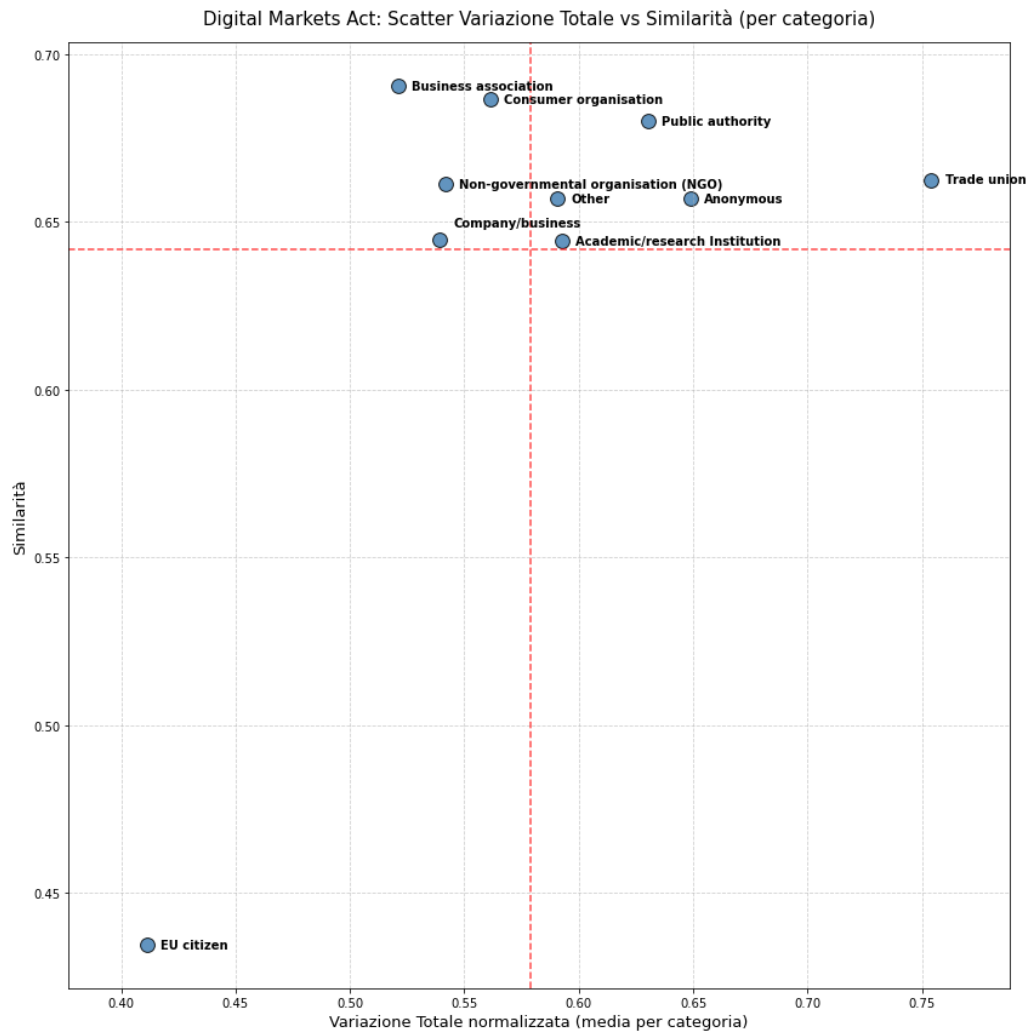


Figura 4.27: Distribuzione delle categorie in base a Similarità e Variazione Totale.

Infine, analizzando la dimensione temporale, anche nel caso del *Digital Markets Act* la fase iniziale della consultazione (dalla *Call for Evidence* alla proposta della Commissione) risulta essere quella più influenzabile, con un incremento medio di circa **0,029**. Al contrario, anche in questo caso, nella seconda fase (dalla proposta al testo finale) i valori registrati diventano negativi, scendendo a circa **-0,016**.

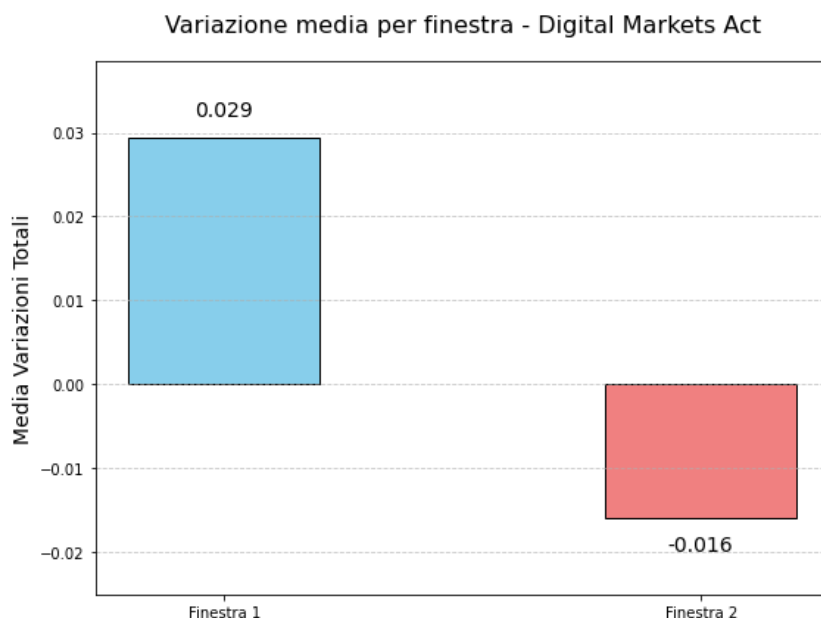


Figura 4.28: Variazione media per finestra di consultazione – Digital Markets Act

4.5 Analisi trasversale

Per approfondire le relazioni trasversali, è utile osservare le correlazioni tra le diverse direttive e le diverse categorie di organizzazioni, in modo da evidenziare eventuali associazioni o somiglianze nei comportamenti e nelle posizioni espresse. Le rappresentazioni grafiche al di sotto riportate, permettono di cogliere legami comuni o differenze significative tra i vari gruppi di interesse e i vari ambiti normativi analizzati.

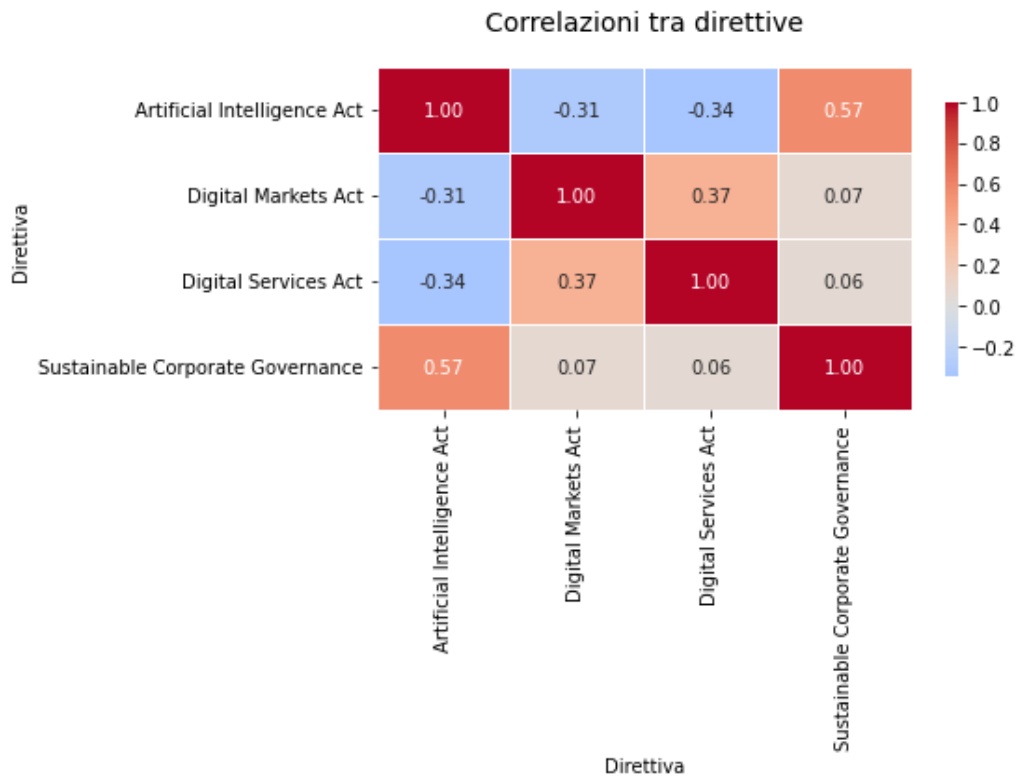


Figura 4.29: Correlazioni tra direttive

Questa prima rappresentazione mette in luce, ad esempio, la forte correlazione positiva tra l'*AI Act* e la *Sustainable Corporate Governance* ($r = 0.57$), segno che alcune tipologie di gruppi di interesse si sono mosse in modo simile su entrambe le normative. Si osserva inoltre una correlazione positiva, seppur più contenuta, tra il *Digital Services Act* e il *Digital Markets Act* ($r = 0.37$), coerente con la loro natura di iniziative parallele e complementari nel pacchetto sul digitale.

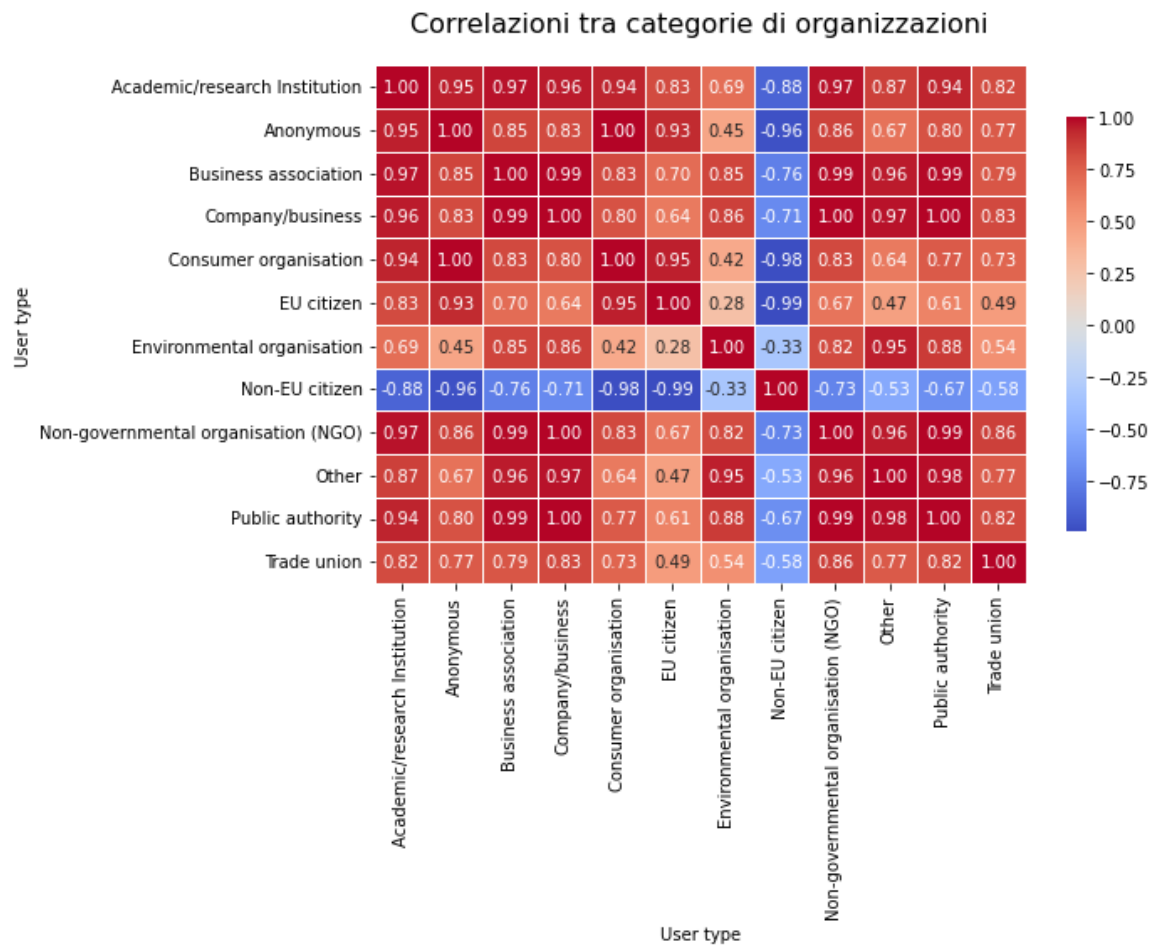


Figura 4.30: Correlazioni tra categorie di organizzazioni

Questa seconda matrice mostra invece le correlazioni tra i diversi gruppi di interesse, alcuni di essi tendono a muoversi in modo coordinato e trasversale, mentre in particolare la categoria Non-EU citizen, mostra un andamento divergente, riflettendo interessi e strategie differenti rispetto al resto delle categorie.

Conclusioni

Questo capitolo raccoglie e interpreta i principali risultati emersi dall'analisi delle quattro direttive esaminate — *Artificial Intelligence Act*, *Sustainable Corporate Governance*, *Digital Services Act* e *Digital Markets Act* — con l'obiettivo di evidenziare le dinamiche comuni e le specificità di ciascun processo di consultazione.

Dopo aver presentato le conclusioni relative a ogni singolo caso, la parte finale offrirà una riflessione trasversale sulle tendenze complessive, cercando di individuare la presenza di pattern ricorrenti e di eventuali correlazioni tra le diverse direttive e le varie categorie di organizzazioni coinvolte nei processi di consultazione.

Artificial Intelligence Act

Dinamiche di influenza e principali attori

Le classifiche delle migliori dieci organizzazioni per similarità e per variazione totale mettono in luce due forme di influenza complementari.

Da un lato, le organizzazioni più allineate alla Commissione — tra cui **Renaissance Numérique**, **SAP**, **AI4Belgium**, **Unipol Gruppo S.p.A.**, **ABB** e **BBVA** — hanno mostrato un approccio tecnico e istituzionale coerente con la linea della Commissione, favorendo una forte similarità con il testo legislativo.

Dall'altro, le organizzazioni con la maggiore variazione totale — come **Kammer für Arbeiter und Angestellte für Wien**, **Bundesvereinigung der Deutschen Arbeitgeberverbände (BDA)**, **Deutsche Sozialversicherung Arbeitsgemeinschaft**, **ZPP Polska**, **LO Sweden**, **ANEC** e **Homo Digitalis** — appartengono prevalentemente al mondo sindacale, industriale e dei consumatori. Questi attori hanno esercitato un'influenza più diretta sul contenuto, proponendo modifiche mirate alla tutela dei lavoratori, alla rappresentanza sociale e ai diritti dei cittadini.

Categorie di organizzazioni e modalità di influenza

L'analisi per categorie conferma una netta distinzione tra soggetti di tipo tecnico e attori sociali, evidenziando due modalità di partecipazione e influenza.

Le **istituzioni accademiche e di ricerca** registrano la similarità media più elevata (**0,666**), seguite da **associazioni di categoria** (**0,661**), **ONG** (**0,659**) e **imprese** (**0,658**). Queste tipologie di organizzazioni risultano quindi maggiormente allineate all'approccio tecnico-regolatorio della Commissione, basato su principi di gestione del rischio e trasparenza.

Le **organizzazioni dei consumatori**, pur mostrando una similarità media più bassa (**0,615**), sono l'unica categoria con una variazione positiva (**+0,003**). Ciò suggerisce che, pur partendo da posizioni meno vicine a quelle istituzionali, siano riuscite a ottenere modifiche mirate su temi cruciali come la protezione dei dati, la sicurezza e i diritti degli utenti.

Questa tendenza è chiaramente visibile nello scatter plot, che mostra la presenza di due grandi gruppi di attori:

- **Da un lato**, le **organizzazioni tecniche e accademiche**, con elevata similarità e variazione contenuta, più vicine alla visione della Commissione e caratterizzate da contributi volti a consolidare la struttura della proposta;
- **Dall'altro**, le **organizzazioni sociali e dei consumatori**, meno allineate alle posizioni istituzionali ma capaci di incidere in modo concreto e mirato sull'evoluzione del testo legislativo.

Evoluzione temporale dell'influenza

L'analisi della variazione media per finestra temporale evidenzia una chiara concentrazione dell'influenza nella fase iniziale della consultazione.

Nel passaggio dalla *Call for Evidence* alla proposta della Commissione si osserva una variazione positiva (**+0,003**), segno che le organizzazioni hanno avuto un ruolo attivo nel contribuire alla definizione del quadro normativo.

Al contrario, nella seconda fase — tra la proposta e il testo finale — la variazione diventa negativa (**-0,027**), indicando che, una volta consolidato il testo, le possibilità di incidere si riducono sensibilmente.

Questa dinamica conferma che l'efficacia dei contributi tende a concentrarsi nei momenti iniziali del processo decisionale, quando la Commissione è maggiormente aperta all'integrazione di proposte esterne.

Sustainable Corporate Governance

Dinamiche di influenza e principali attori

L'analisi dei livelli di similarità e variazione totale mette in evidenza un gruppo di organizzazioni fortemente allineate alla visione della Commissione, tra cui **Danish Institute for Human Rights, European Coalition for Corporate Justice (ECCJ), EDF, CCOO del Hábitat, Deutsches Aktieninstitut, Global Witness e Mercedes-Benz Group AG**. Questi soggetti appartengono in prevalenza al mondo della ricerca, dei diritti umani e delle grandi imprese, e si caratterizzano per un orientamento coerente con la visione della Commissione, contribuendo al consolidamento del testo normativo.

Le organizzazioni con la maggiore variazione totale — tra cui **UNI Europa, Związek Banków Polskich, Clean Clothes Campaign, Associazione Italiana Elettrotecnica, Verband der TÜV e.V. e la Global Legal Entity Identifier Foundation (GLEIF)** — appartengono invece al mondo sindacale, industriale e della società civile. Questi attori hanno avuto un ruolo più diretto e propositivo, introducendo modifiche concrete al testo, su tematiche a loro care come la responsabilità aziendale, la trasparenza delle catene di fornitura e la tutela dei diritti dei lavoratori.

Categorie di organizzazioni e modalità di influenza

L'analisi per categorie conferma la distinzione tra soggetti tecnici e sociali, mostrando approcci di influenza differenti. Le **organizzazioni sindacali e dei lavoratori** presentano la similarità media più elevata (**0,658**), seguite da **ONG (0,597)** e **associazioni di categoria (0,593)**. Le **imprese**, le **organizzazioni dei consumatori** e le **autorità pubbliche** mostrano invece valori più contenuti, compresi tra **0,563** e **0,558**.

L'elevata similarità delle organizzazioni sindacali riflette la natura della direttiva, fortemente orientata alla sostenibilità sociale, alla trasparenza e alla tutela dei diritti lungo le catene di fornitura globali.

L'analisi dell'influenza media per categoria mostra inoltre che le **organizzazioni dei consumatori** sono le uniche a registrare una variazione positiva (**+0,003**), mentre tutte le altre categorie presentano valori negativi. I valori più bassi si osservano invece per le **organizzazioni ambientali (-0,144)**. Questo valore fortemente negativo associato alle organizzazioni ambientali suggerisce che, pur avendo partecipato attivamente alla consultazione, la loro influenza sul testo finale sia stata limitata, indicando come, nel processo di definizione della

direttiva, gli aspetti ambientali abbiano ricevuto una minore attenzione rispetto a quelli sociali e di governance.

Lo scatter plot conferma la presenza di una chiara polarizzazione tra le categorie di attori:

- **Da un lato**, le **organizzazioni sindacali (Trade Union)** si distinguono per l'elevata similarità con la Commissione, evidenziando una posizione stabile e fortemente allineata alle linee guida istituzionali;
- **Dall'altro lato**, le **organizzazioni dei consumatori** mostrano la variazione più alta e positiva (+0,003), pur con una similarità più bassa, indicando un'influenza esercitata in modo mirato su aspetti specifici, soprattutto in materia di responsabilità delle imprese e tutela dei cittadini.

Le altre categorie — come **ONG, associazioni di categoria, istituzioni accademiche, imprese e autorità pubbliche** — si distribuiscono attorno ai valori medi.

Evoluzione temporale dell'influenza

La distribuzione temporale dell'influenza evidenzia una chiara concentrazione nella fase iniziale del processo consultivo. Nel passaggio dalla *Call for Evidence* alla proposta della Commissione si osserva una variazione positiva (+0,004), mentre nella fase successiva — tra la proposta e il testo finale — si registra un valore negativo (−0,100). Questo andamento indica che la capacità delle organizzazioni di incidere sul contenuto della direttiva si riduce sensibilmente una volta che la proposta viene formalizzata e la Commissione tende a consolidare la propria posizione iniziale.

Digital Services Act

Dinamiche di influenza e principali attori

Dall'analisi emerge un insieme di organizzazioni che condividono in larga parte le posizioni della Commissione, tra cui **Bitkom e.V., EuroISPA, European Association for Digital Transition, Reporters Without Borders, Sky, Austrian Bar, Civil Liberties Union for Europe, DOT Europe e Facebook**. Si tratta di soggetti con forti competenze tecniche e istituzionali nel settore digitale, che condividono la visione della Commissione volta a bilanciare innovazione, concorrenza e tutela dei diritti fondamentali.

Al contrario, tra le organizzazioni con la maggiore variazione totale emergono **Association of Commercial Television in Europe**, **Danish Chamber of Commerce**, **DIGITAL EUROPE**, **Polish Confederation Lewiatan**, **CEEP (SGIs)** e il **Danish Government**. Questi attori rappresentano soprattutto interessi economici e industriali e si sono concentrati su temi a loro cari legati alla responsabilità delle piattaforme, alla pubblicità online e alla moderazione dei contenuti, aree in cui le decisioni della Commissione hanno avuto un impatto diretto sul mercato digitale europeo.

Categorie di organizzazioni e modalità di influenza

L'analisi per categorie conferma un quadro di partecipazione ampio e diversificato. Le **organizzazioni dei consumatori (0,723)** e le **istituzioni accademiche (0,696)** risultano le più vicine alle posizioni della Commissione, seguite da **autorità pubbliche (0,676)** e **associazioni di categoria (0,659)**. I valori più bassi si registrano invece per i **cittadini europei (0,482)** e **non europei (0,535)**, segno di una maggiore distanza rispetto al linguaggio tecnico e politico del dibattito istituzionale.

L'analisi dell'influenza media per categoria evidenzia un quadro fortemente asimmetrico: le **istituzioni accademiche (0,149)** e gli **attori anonimi (0,119)** mostrano la variazione media più alta, seguite dai **cittadini dell'Unione (+0,077)** e dalle **organizzazioni dei consumatori (+0,070)**. Le categorie più tradizionali — come **sindacati**, **imprese** e **autorità pubbliche** — presentano invece valori più contenuti, mentre i **cittadini non europei** mostrano una media negativa (**-0,034**).

Lo scatter plot evidenzia una chiara differenziazione tra le categorie di attori:

- **Da un lato**, le **istituzioni accademiche e di ricerca** si collocano nella fascia più alta di similarità, combinando competenze tecniche e un forte allineamento con la visione della Commissione;
- **Dall'altro**, le **organizzazioni dei consumatori** mostrano la variazione più elevata, indicando una capacità di incidere in modo concreto, pur partendo da posizioni meno vicine a quelle istituzionali.

Le **autorità pubbliche**, le **associazioni di categoria** e le **imprese** si distribuiscono su valori intermedi, contribuendo in modo più tenue e bilanciato al processo regolatorio.

Evoluzione temporale dell'influenza

La distribuzione temporale dell'influenza mostra una netta concentrazione nella prima fase del processo consultivo. Nel passaggio dalla *Call for Evidence* alla proposta della Commissione si registra una variazione positiva e significativa (+0,056), mentre nella fase successiva — dalla proposta al testo finale — la variazione si riduce (+0,007). Ciò conferma che la capacità delle organizzazioni di incidere sul contenuto della direttiva è risultata più forte nelle fasi iniziali, quando il quadro normativo era ancora flessibile e aperto ai contributi esterni.

Digital Markets Act

Dinamiche di influenza e principali attori

L'analisi evidenzia un gruppo di organizzazioni particolarmente allineate alla visione della Commissione, tra cui **Adigital**, **BBVA**, **European Banking Federation**, **BEUC – The European Consumer Organisation**, **European Publishers Council**, **Austrian Bar**, **Google**, **Booking.com** e l'**Information Technology Industry Council (ITI)**. Si tratta di soggetti dotati di una solida competenza tecnica e di un ruolo consolidato nel mercato digitale europeo, che hanno contribuito a rafforzare la stabilità del testo, in particolare su temi legati alla concorrenza, alla trasparenza e alla regolamentazione dei *gatekeeper* centrali all'interno della direttiva.

Le organizzazioni con la maggiore variazione totale comprendono invece il **Federal Ministry for Digital and Economic Affairs**, **IT for Change**, **European Publishers Council**, **AIC-UNIER**, **Information Technology Industry Council (ITI)**, **UNI Europa** e **Booking.com**. Questi attori, appartenenti prevalentemente al mondo istituzionale e industriale, hanno avuto un ruolo attivo nel proporre modifiche mirate su questioni a loro affini come la tutela dei dati e l'accesso equo ai mercati digitali.

Categorie di organizzazioni e modalità di influenza

L'analisi della similarità media per categoria evidenzia un'ampia convergenza tra i principali attori economici e istituzionali. Le **associazioni di categoria** (0,691), le **organizzazioni dei consumatori** (0,687) e le **autorità pubbliche** (0,680) risultano le più vicine alle posizioni della Commissione, seguite da **sindacati** e **ONG** (tra 0,663 e 0,661). Le **istituzioni accademiche** (0,644) e le **imprese** (0,645) mostrano valori intermedi, mentre i **cittadini**

europei (0,435) si confermano la categoria più distante, riflettendo una minore familiarità con il linguaggio tecnico e politico del processo decisionale europeo.

L'influenza media per categoria vede invece i **sindacati (0,050)**, gli **attori anonimi (0,032)** e le **autorità pubbliche (0,029)** registrare i valori più elevati, seguiti da **accademici e organizzazioni dei consumatori** (entrambi **0,022**). Le **imprese** e le **associazioni di categoria** mostrano valori inferiori, mentre i **cittadini europei** presentano una media negativa (**-0,009**), segno di una capacità di influenza più debole rispetto ai soggetti istituzionalizzati.

Lo scatter plot evidenzia un quadro complessivamente omogeneo, con la maggior parte delle categorie posizionate nella fascia alta di similarità e con variazioni contenute:

- **Da un lato**, le **organizzazioni tecniche e istituzionali** — come associazioni di categoria, autorità pubbliche e organizzazioni dei consumatori — si collocano tra i valori più elevati di similarità, mostrando un allineamento stabile con la visione della Commissione;
- **Dall'altro**, i **cittadini europei** rappresentano un caso isolato, caratterizzato da valori significativamente più bassi di similarità e variazione, a indicare una partecipazione marginale e una minore capacità di incidere sul processo regolatorio.

Evoluzione temporale dell'influenza

La distribuzione temporale dell'influenza mostra una chiara concentrazione nella fase iniziale del processo consultivo. Nel passaggio dalla *Call for Evidence* alla proposta della Commissione si registra una variazione positiva (**+0,029**), mentre nella fase successiva — tra la proposta e il testo finale — la variazione diventa negativa (**-0,016**). Questo andamento conferma anche in questo caso che la capacità delle organizzazioni di incidere sul contenuto della normativa è risultata maggiore nei momenti iniziali, quando la Commissione era più aperta ai contributi esterni dei diversi attori.

Analisi trasversale e considerazioni conclusive

L'analisi congiunta delle quattro direttive — **Artificial Intelligence Act**, **Sustainable Corporate Governance**, **Digital Services Act** e **Digital Markets Act** — permette di individuare alcune tendenze comuni e relazioni significative tra categorie di attori e processi consultivi.

Dinamiche generali

In tutte le direttive, la prima fase del processo di consultazione (*Call for Evidence* → *Proposta*) risulta decisiva per incidere sul contenuto legislativo. Si osserva una variazione positiva in questa fase, seguita da una flessione o stabilizzazione successiva, a conferma che la Commissione europea tende a essere più aperta ai contributi esterni nelle fasi iniziali, mentre una volta formalizzata la proposta le possibilità di modifica si riducono sensibilmente.

Non emerge la presenza di un “super-attore” dominante in grado di influenzare trasversalmente tutti i processi decisionali. L’influenza varia infatti in funzione del tema trattato e del tipo di direttiva: le normative di natura digitale coinvolgono principalmente attori tecnici e industriali, mentre quelle legate a diritti e sostenibilità mostrano un maggiore protagonismo delle organizzazioni sociali e accademiche.

Correlazioni tra direttive

La prima matrice delle correlazioni presente nel capitolo dei risultati mostra come le direttive **Artificial Intelligence Act** e **Sustainable Corporate Governance** siano fortemente connesse ($r = 0,57$), suggerendo una convergenza tra attori orientati alla tutela dei diritti, all’etica e alla sostenibilità.

Inoltre, mostra come anche le direttive **Digital Services Act** e **Digital Markets Act** siano fortemente connesse ($r = 0,37$), segno della presenza di partecipanti comuni e strategie coordinate nel contesto del pacchetto europeo sul digitale. Un ulteriore aspetto particolarmente interessante riguarda la sovrapposizione dei partecipanti: circa la metà delle organizzazioni coinvolte nel DMA ha preso parte anche alla consultazione sul DSA, a conferma dell’esistenza di un ampio terreno comune tra i due processi.

L’analisi delle variazioni totali rafforza questa evidenza: per le organizzazioni presenti in entrambe le direttive, la correlazione tra i valori di influenza risulta pari a circa **0,325 (coefficiente di Pearson¹)**. In altri termini, gli attori che hanno inciso maggiormente sul **Digital Markets Act** tendono, in media, a risultare influenti anche nel **Digital Services Act**.

Questi risultati mostrano come, nel settore digitale, molte strategie di lobbying siano state non solo coordinate, ma anche trasversali tra le due iniziative. Non a caso, il DMA e il DSA sono stati concepiti fin dall’inizio come parte di uno stesso pacchetto legislativo e presentati in

¹Il coefficiente di *Pearson* misura la forza e la direzione della relazione lineare tra due variabili, assumendo valori compresi tra -1 e +1. Risorse: <https://www.scribbr.com/statistics/pearson-correlation-coefficient/>

parallelo dalla Commissione europea, spingendo i principali stakeholder a esprimere posizioni coerenti su entrambe le normative.

Correlazioni tra categorie di organizzazioni

La seconda matrice mette invece in evidenza le relazioni tra le diverse categorie di organizzazioni, a seguito di un'analisi complessiva delle quattro direttive. Le correlazioni più alte si osservano tra **istituzioni accademiche**, **associazioni di categoria**, **imprese** e **organizzazioni dei consumatori**, che mostrano un comportamento coordinato e trasversale tra le varie direttive.

Queste categorie tendono a muoversi in modo sinergico, condividendo linguaggi e strategie di intervento talvolta simili, in linea con la visione della Commissione. Al contrario, i **Non-EU citizens** si distinguono per valori di correlazione più bassi e in alcuni casi negativi, riflettendo un comportamento più indipendente e meno allineato al resto delle categorie. Ciò suggerisce che gli attori extraeuropei partecipano con logiche e priorità differenti, meno integrate nel contesto politico dell'Unione.

Bibliografia

Regolamenti e Direttive

1. *Artificial Intelligence Act* — Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 <https://eur-lex.europa.eu/eli/reg/2024/1689/oj?locale=it>
2. *Sustainable Corporate Governance* — Direttiva (UE) 2024/1760 del Parlamento europeo e del Consiglio del 13 giugno 2024 sulla due diligence aziendale in materia di sostenibilità (Corporate Sustainability Due Diligence Directive) <https://eur-lex.europa.eu/eli/dir/2024/1760/oj?locale=it>
3. *Digital Services Act* — Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 sul mercato unico dei servizi digitali <https://eur-lex.europa.eu/eli/reg/2022/2065/oj?locale=it>
4. *Digital Markets Act* — Regolamento (UE) 2022/1925 del Parlamento europeo e del Consiglio del 14 settembre 2022 su mercati contestabili e equi nel settore digitale <https://eur-lex.europa.eu/eli/reg/2022/1925/oj?locale=it>

Siti Web

1. *Carta dei diritti fondamentali dell'Unione Europea*. <https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:12012P/TXT>
2. *Parlamento europeo*, sito ufficiale. <https://www.europarl.europa.eu/portal/it>
3. *Consiglio europeo*, sito ufficiale. <https://www.consilium.europa.eu/it/european-council/>
4. *Consiglio dell'Unione europea*, sito ufficiale. <https://www.consilium.europa.eu/it/council-eu/>
5. *Commissione europea*, sito ufficiale. https://commission.europa.eu/index_it
6. *Corte di giustizia dell'Unione europea*, sito ufficiale. https://curia.europa.eu/jcms/jcms/j_6/
7. *Banca centrale europea*, sito ufficiale. <https://www.ecb.europa.eu/home/html/index.it.html>
8. *Corte dei conti europea*, sito ufficiale. <https://www.eca.europa.eu/it>
9. *Piattaforma "Have Your Say"*, Commissione europea – Portale ufficiale per la consultazione pubblica. https://ec.europa.eu/info/law/better-regulation/have-your-say_it
10. *Collegio dei commissari della Commissione europea*, descrizione e composizione. https://commission.europa.eu/about/organisation/college-commissioners_it
11. *Consultazione pubblica sull'Artificial Intelligence Act*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Artificial-intelligence-ethical-and-legal-requirements_it
12. *Consultazione pubblica sulla Sustainable Corporate Governance*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12548-Sustainable-corporate-governance_it

13. *Consultazione pubblica sul Digital Services Act*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12417-Legge-sui-servizi-digitali-approfondire-il-mercato-interno-e-chiarire-le-responsabilita-per-i-servizi-digitali_it
14. *Consultazione pubblica sul Digital Markets Act*. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12418-Digital-Services-Act-package-ex-ante-regulatory-instrument-of-very-large-online-platforms-acting-as-gatekeepers_en
15. IBM, *Information Retrieval – introduzione e concetti di base*. <https://www.ibm.com/it-it/think/topics/information-retrieval>
16. IBM, *Bag-of-Words – rappresentazione testuale*. <https://www.ibm.com/it-it/think/topics/bag-of-words>
17. IBM, *Natural Language Processing (NLP)*. <https://www.ibm.com/it-it/think/topics/natural-language-processing>
18. IBM, *Word Embeddings – rappresentazioni semantiche*. <https://www.ibm.com/it-it/think/topics/word-embeddings>
19. Devlin, J. et al. (2019), *BERT – Bidirectional Encoder Representations from Transformers*. NAACL 2019. <https://aclanthology.org/N19-1423.pdf>
20. IBM, *Cosine Similarity – metodi e applicazioni*. <https://www.ibm.com/it-it/think/topics/cosine-similarity>
21. *TF-IDF*. Medium. https://medium.com/@Sumeet_Agrawal/tf-idf-47f8f2ebd987
22. *Jaccard Similarity*. Medium. <https://mayurdhvajsinhjadeja.medium.com/jaccard-similarity-34e2c15fb524>
23. Hugging Face, *all-mpnet-base-v2 – Sentence Transformers*. <https://huggingface.co/sentence-transformers/all-mpnet-base-v2>
24. Hugging Face, *all-MiniLM-L6-v2 – Sentence Transformers*. <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2>
25. *Fuzzy Matching*. Medium. <https://medium.com/@m.nath/fuzzy-matching-algorithms-81914b1bc498>

26. *Matching Fuzzy Strings with Levenshtein Distance*. Medium. <https://andrew-hyde.medium.com/matching-fuzzy-strings-with-levenshtein-distance-vba-implementation-076433f3e7fe>
27. *Pearson Correlation Coefficient*. <https://www.scribbr.com/statistics/pearson-correlation-coefficient/>

Ringraziamenti

Desidero esprimere la mia sincera gratitudine al **prof. Edoardo Mollona** per la disponibilità e i preziosi consigli forniti durante tutto il percorso di realizzazione dell'elaborato.

Un ringraziamento particolare al **dott. Alessio Diana** per il supporto costante e per l'assistenza nella parte informatica e applicativa del lavoro.

Infine, un grazie di cuore alla mia **famiglia** e ai miei **amici**, per l'incoraggiamento e il sostegno dimostrato lungo tutto il percorso di studi.