

# OSSERVATORIO SULLA DISINFORMAZIONE ONLINE



Informazione e disinformazione



Categorie dell'informazione e della disinformazione



Informazione e disinformazione sulle elezioni europee



Informazione e disinformazione sui temi di rilevanza europea



Principali tematiche della disinformazione



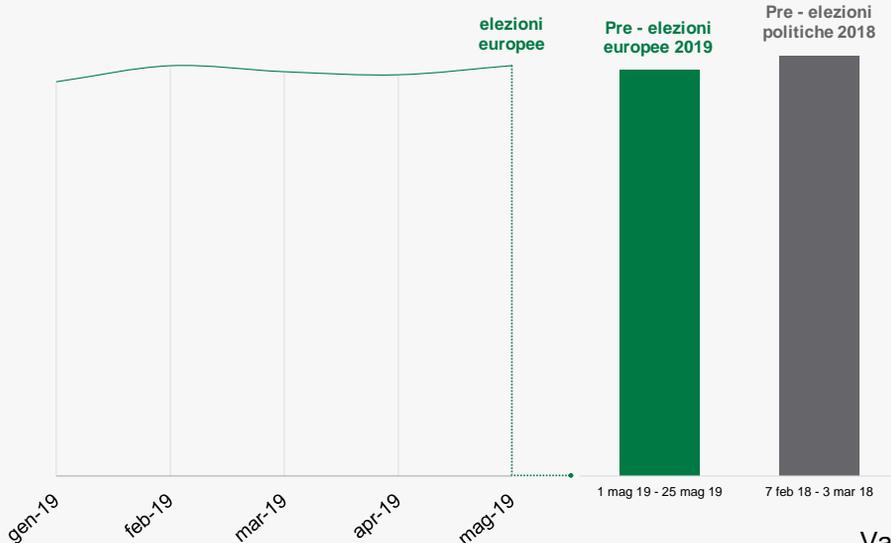
Appendice



Nota metodologica



## Totale contenuti informativi prodotti nel giorno medio



## INFORMAZIONE PRODOTTA IN ITALIA

Per i contenuti di informazione complessivamente prodotti dai media su qualsiasi genere di notizie, si riscontra:

- un incremento del 2% nel mese di maggio 2019 rispetto ad aprile
- un ammontare che, nei 25 giorni precedenti alle elezioni europee, risulta inferiore rispetto a quello prodotto nel corrispondente periodo antecedente alle elezioni politiche del 2018

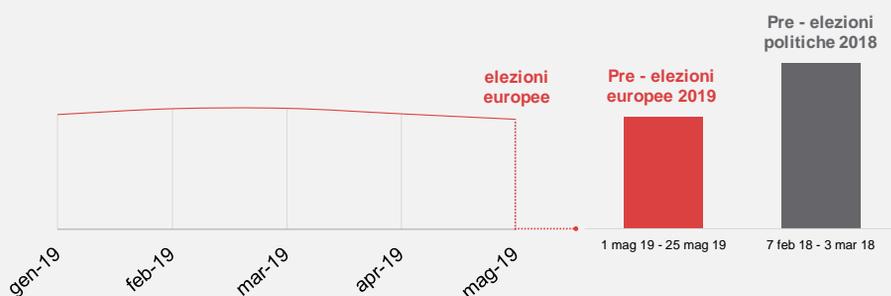
Vai all'[informazione sulle elezioni europee](#)

## DISINFORMAZIONE ONLINE PRODOTTA IN ITALIA

L'analisi del totale dei contenuti prodotti in Italia sia dai siti web che dalle pagine e account social di disinformazione mostra:

- una riduzione del 5% della disinformazione online nel giorno medio di maggio 2019 rispetto al mese precedente
- un volume di disinformazione online prodotto nei 25 giorni prima delle ultime elezioni europee nettamente inferiore a quanto registrato in concomitanza del periodo elettorale delle politiche del 2018

## Totale contenuti di disinformazione prodotti nel giorno medio



Vai alla [disinformazione sulle elezioni europee](#)

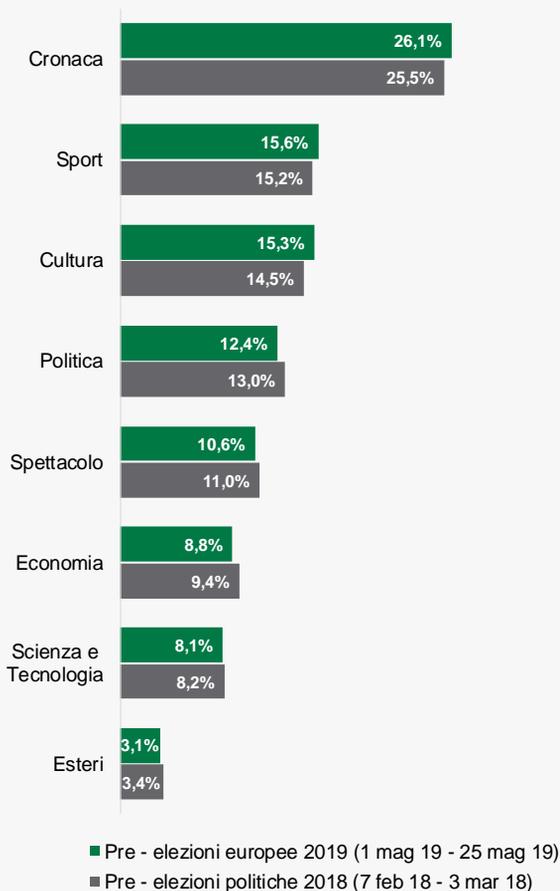
**-5%** su aprile 2019

DISINFORMAZIONE ONLINE NEL GIORNO MEDIO DI MAGGIO 2019

MENO DISINFORMAZIONE NEL MESE DELLE ELEZIONI EUROPEE RISPETTO AL MESE DELLE ELEZIONI POLITICHE 2018

# Categorie dell'informazione e della disinformazione

## Ripartizione dell'offerta di informazione per categoria



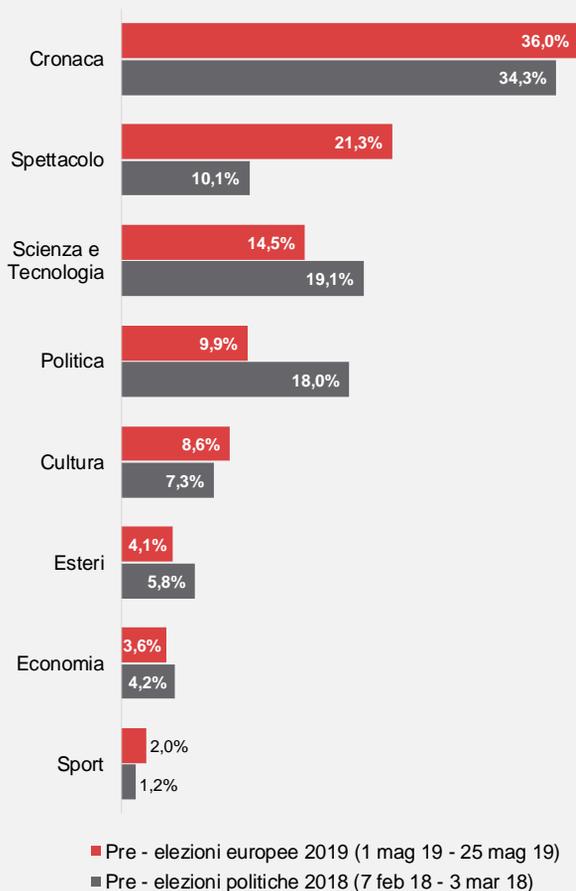
## CATEGORIE DELL'INFORMAZIONE

Esaminando i contenuti dell'offerta informativa nei 25 giorni antecedenti alle elezioni europee, si osserva che:

- la distribuzione degli stessi per categoria appare pressoché analoga a quella riscontrata nel periodo precedente alle elezioni politiche del 2018
- cronaca rimane il genere informativo prevalente (rappresentando l'oggetto di oltre un quarto delle notizie generate), seguito da sport e cultura
- permane la carenza di informazione scientifica (8% del totale)

**CRONACA SI CONFERMA LA CATEGORIA PREVALENTE DEGLI ARTICOLI DI DISINFORMAZIONE**

## Ripartizione dell'offerta di disinformazione online per categoria



## CATEGORIE DELLA DISINFORMAZIONE

L'analisi dei contenuti prodotti dai siti web di disinformazione dal 1° al 25 maggio 2019 mostra:

- una perdurante concentrazione della disinformazione online sui fatti di cronaca, trattati dal 36% degli articoli pubblicati
- una riduzione dell'incidenza sul totale della categoria "politica", che passa dal 18% del periodo pre-elettorale 2018 al 10% nel periodo precedente alle elezioni europee del 26 maggio 2019
- un deciso incremento degli articoli di disinformazione dedicati al mondo dello spettacolo

**MINORE INCIDENZA DELLA POLITICA SUI CONTENUTI DI DISINFORMAZIONE, RISPETTO AL PERIODO PRE-ELETTORALE 2018**

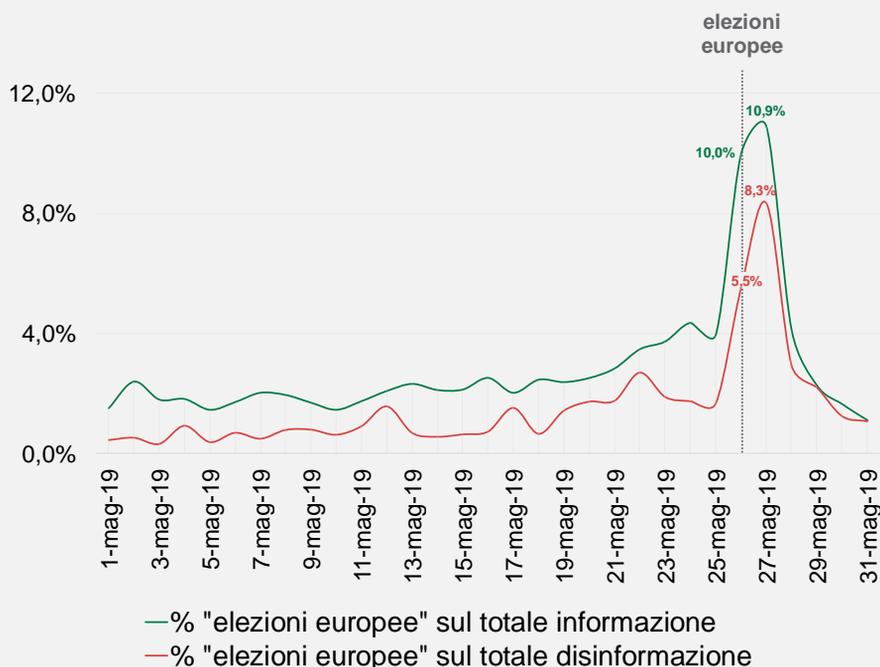
# Informazione e disinformazione sulle elezioni europee

## SPAZIO DEDICATO ALLE ELEZIONI EUROPEE

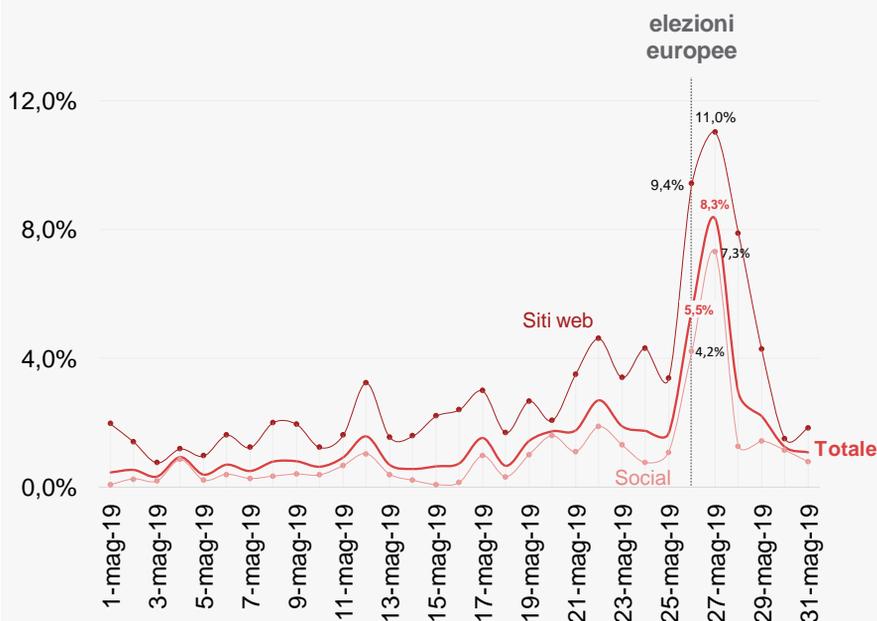
Con particolare riguardo allo spazio dedicato alle elezioni europee nel mese di maggio 2019, si riscontra:

- un andamento crescente sia per le fonti di informazione che di disinformazione in corrispondenza dell'approssimarsi e dello svolgimento della tornata elettorale
- il raggiungimento dei valori massimi nei giorni 26 (data del voto) e 27 maggio, quando l'informazione sulle elezioni europee si attesta rispettivamente sul 10% e l'11%, e la disinformazione sul 6% e l'8%
- un'incidenza media sui contenuti di disinformazione pari al 2%, a fronte del 3% nel caso dell'informazione

Incidenza dell'argomento "elezioni europee" sul totale informazione e disinformazione



Incidenza dell'argomento "elezioni europee" per tipologia di fonte di disinformazione



## SPAZIO DEDICATO ALLE ELEZIONI EUROPEE DALLE FONTI DI DISINFORMAZIONE

Concentrando l'analisi sulle due tipologie di fonti di disinformazione online, ossia siti web e pagine/account social, si rileva come:

- per i siti di disinformazione, l'incidenza dei contenuti sulle elezioni europee (pari al 3% nel giorno medio di maggio 2019) sia sempre superiore rispetto a quella riscontrata per le fonti social di disinformazione (mediamente pari all'1%)
- per entrambe, il picco di contenuti dedicati alle elezioni europee si registra il giorno successivo al voto, con valori corrispondenti all'11% (per i siti web) e al 7% (per le fonti social)

**3%** nel mese di maggio 2019

SPAZIO DEDICATO ALLE ELEZIONI EUROPEE DAI SITI DI DISINFORMAZIONE

**1%** nel mese di maggio 2019

SPAZIO DEDICATO ALLE ELEZIONI EUROPEE DALLE FONTI SOCIAL DI DISINFORMAZIONE

# Informazione e disinformazione sui temi di rilevanza europea

## Incidenza dei temi di rilevanza europea sul totale informazione (maggio 2019)



## INFORMAZIONE SUI TEMI DI RILEVANZA EUROPEA

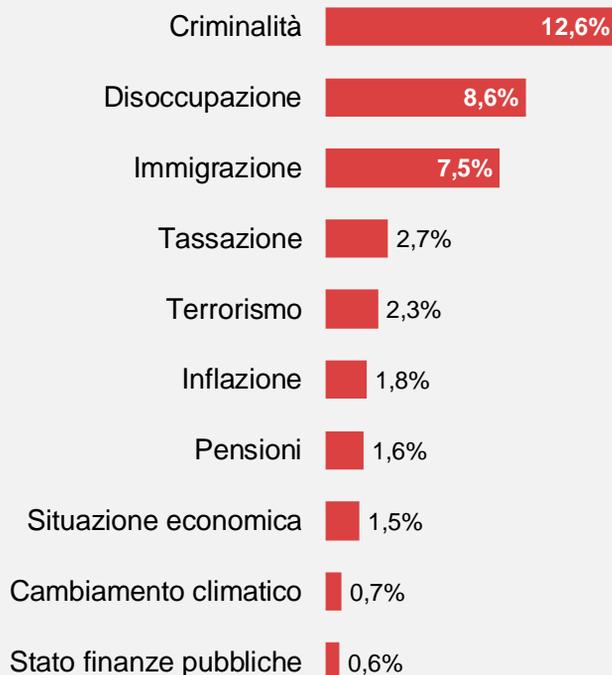
Riguardo ai temi di rilevanza europea individuati da Eurobarometro, si osserva come in Italia:

- a maggio 2019, disoccupazione e criminalità rimangono quelli prevalenti per le fonti di informazione, con un'incidenza sul totale superiore al 10% in entrambi i casi
- al terzo e quarto posto si collocano tassazione e immigrazione, con percentuali che comunque si fermano al di sotto del 5%
- il minor spazio sia attribuito al tema del cambiamento climatico (meno dell'1% del totale)

TRA I TEMI DI RILEVANZA EUROPEA:

**DISOCCUPAZIONE** SI CONFERMA IL PIÙ TRATTATO DALLE FONTI DI INFORMAZIONE, **CRIMINALITÀ** IL PIÙ TRATTATO DALLE FONTI DI DISINFORMAZIONE

## Incidenza dei temi di rilevanza europea sul totale disinformazione (maggio 2019)



## DISINFORMAZIONE SUI TEMI DI RILEVANZA EUROPEA

L'analisi della disinformazione online prodotta nel mese di maggio evidenzia un generale incremento dello spazio dedicato alle tematiche di rilevanza europea. In particolare:

- la criminalità si conferma il tema più trattato dalle fonti di disinformazione in Italia, con un'incidenza del 13% sul totale
- anche disoccupazione e immigrazione mantengono inalterate le loro posizioni, collocandosi al secondo e al terzo posto, con un'incidenza sul totale rispettivamente pari al 9% e all'8%

NEL MESE DELLE ELEZIONI EUROPEE, **CRESCE** SOPRATTUTTO LO SPAZIO DEDICATO DALLE FONTI DI DISINFORMAZIONE A **CRIMINALITÀ, DISOCCUPAZIONE E IMMIGRAZIONE**

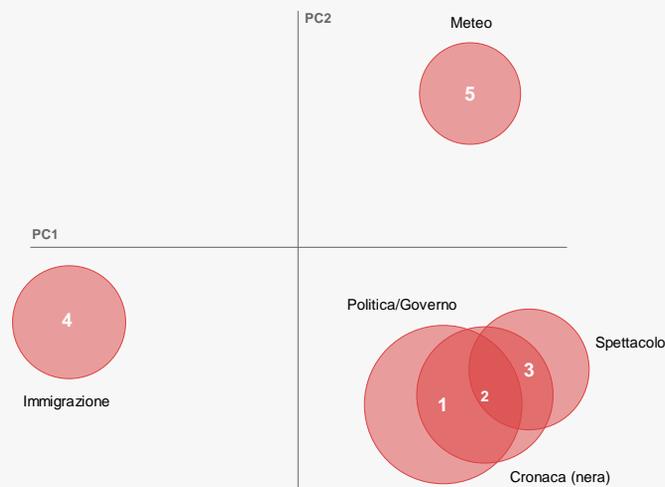
# Principali tematiche della disinformazione

## PRINCIPALI TEMATICHE DELLA DISINFORMAZIONE ONLINE

Dall'analisi testuale (*topic modeling*) degli articoli dei siti di disinformazione, che consente di individuare le macro-tematiche della disinformazione sulla base della co-occorrenza di termini all'interno dei documenti, si riscontra che:

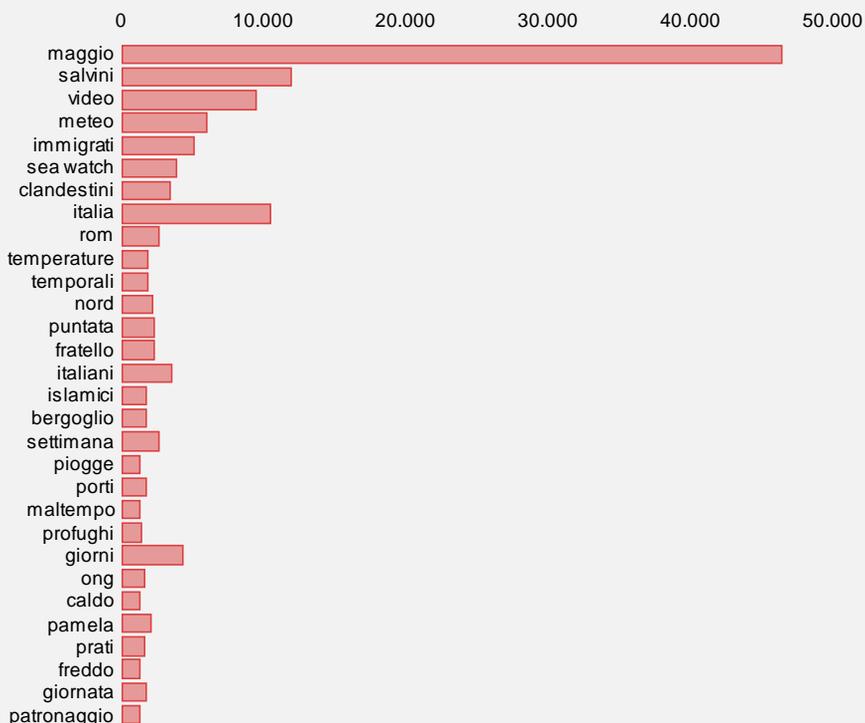
- nel mese di maggio 2019, i 5 principali argomenti oggetto di trattazione sono riconducibili alle questioni politiche e all'attività di governo (con particolare riguardo a economia, lavoro, sicurezza); ai fatti di cronaca nera; allo spettacolo e alla cronaca rosa; all'immigrazione; e alle previsioni meteorologiche

## Principali tematiche oggetto di disinformazione online (maggio 2019)



L'elenco dei 30 termini più frequenti riferiti a ciascuna tematica è riportato in [Appendice](#)

## Primi 30 termini per numero di occorrenze nei siti di disinformazione (maggio 2019)



## TERMINI SALIENTI DELLA DISINFORMAZIONE ONLINE

La distribuzione di frequenza dei termini contenuti negli articoli dei siti di disinformazione mostra come:

- circa un terzo dei 30 termini più frequenti sia relativo agli accadimenti e alle politiche legate all'immigrazione (ad esempio, "immigrati", "clandestini", "porti", "sea watch", "ong")
- tra gli altri, si distinguono quelli che richiamano le condizioni meteorologiche avverse (come "temporali", "piogge", "caldo", "maltempo"), la sfera religiosa (come "bergoglio", "islamici"), e le vicende legate a trasmissioni televisive e personaggi dello spettacolo (come "puntata", "fratello", "pamela", "prati")

**POLITICA/GOVERNO, CRONACA NERA, SPETTACOLO, IMMIGRAZIONE E METEO SONO LE TEMATICHE PRINCIPALI DELLA DISINFORMAZIONE ONLINE A MAGGIO 2019**

## TERMINI SALIENTI DELLE PRINCIPALI TEMATICHE DELLA DISINFORMAZIONE ONLINE

La composizione delle singole tematiche della disinformazione consente di ricavare, in base ai termini salienti associati alle stesse, indicazioni circa l'oggetto specifico delle notizie false diffuse.

In tal senso, per ognuna delle 5 tematiche principali della disinformazione nel mese di maggio 2019, si riporta l'elenco dei 30 termini più frequenti che le compongono:

### 1. POLITICA/GOVERNO

- euro
- anni
- italia
- governo
- ministro
- fatto
- stati
- oggi
- caso
- presidente
- paese
- anno
- interno
- legge
- persone
- europa
- politica
- milioni
- lavoro
- sicurezza
- lega
- guerra
- volta
- nazionale
- sistema
- cittadini
- tempo
- viene
- momento
- grazie

### 2. CRONACA (NERA)

- anni
- vita
- persone
- fatto
- uomo
- mondo
- napoli
- caso
- ospedale
- tempo
- salute
- viene
- giorno
- famiglia
- stati
- donna
- acqua
- morte
- polizia
- casa
- momento
- giorni
- studio
- dieta
- corpo
- durante
- capelli
- oggi
- volta
- spesso

### 3. SPETTACOLO

- casa
- fatto
- fratello
- puntata
- donne
- pamela
- prati
- uomini
- eliana
- francesca
- barbara
- momento
- foto
- anni
- michelazzo
- vita
- storia
- ragazza
- durante
- rai
- donna
- maggio
- programma
- onda
- amore
- social
- ultima
- giorgio
- giorni
- caltagirone

### 4. IMMIGRAZIONE

- maggio
- salvini
- video
- immigrati
- sea watch
- clandestini
- italia
- rom
- italiani
- islamici
- bergoglio
- porti
- milano
- sinistra
- profughi
- ong
- vuole
- foto
- migranti
- patronaggio
- africani
- lega
- sbarco
- libia
- roma
- anni
- euro
- casa
- immigrazione
- lampedusa

### 5. METEO

- meteo
- maggio
- italia
- nord
- temperature
- temporali
- giorni
- settimana
- tempo
- piogge
- mese
- giornata
- maltempo
- caldo
- freddo
- sud
- regioni
- domani
- centro
- aria
- previsioni
- condizioni
- europa
- estate
- giugno
- inizio
- oroscopo
- sotto
- neve
- valori

## COMPOSIZIONE DEL DATABASE

Le informazioni riportate nel presente numero dell'Osservatorio sono il risultato di elaborazioni svolte dall'Autorità su un database documentale costruito a partire dai dati estrapolati attraverso la piattaforma sviluppata da *Volocom Technology*.

In particolare, è stato analizzato l'intero contenuto testuale estrapolato da oltre 7 milioni di documenti generati in Italia (nei primi cinque mesi del 2019 e nel mese antecedente alle elezioni politiche del 2018) da più di 2.000 fonti informative (canali televisivi e radiofonici nazionali, quotidiani, agenzie di stampa, siti web di editori tradizionali, testate esclusivamente online, e relative pagine e account di social network), e fonti di disinformazione (siti web e pagine/account social) individuate come tali da soggetti esterni specializzati in attività di debunking. Il volume di disinformazione online prodotto in Italia è stato quindi stimato con una metodologia di tipo soggettivo, ossia considerando il numero complessivo di documenti generati mensilmente dalle predette fonti di disinformazione.

Il database utilizzato si compone di tutti i documenti generati nell'arco di una giornata da ciascuna fonte informativa e di disinformazione. Per documento si intende l'intero articolo, nel caso di quotidiani e siti web di informazione; la trascrizione di un segmento di trasmissione (servizio), nel caso di Tv e radio; l'intero tweet/post, nel caso delle piattaforme online.

## ANALISI TESTUALE ATTRAVERSO TOPIC MODELING

Per lo studio delle tematiche principali della disinformazione online, sono state condotte analisi sull'intero contenuto divulgato dai siti di disinformazione nel mese di maggio 2019, adottando un'impostazione metodologica riconducibile al c.d. *topic modeling*.

Un *topic model* è un modello statistico per l'individuazione automatica degli argomenti che compaiono in una raccolta di documenti. In particolare, la classificazione del testo dell'insieme di contenuti di disinformazione in determinati argomenti è stata ottenuta attraverso l'utilizzo di un modello prodotto via LDA (*Latent Dirichlet Allocation*) - un algoritmo di apprendimento automatico non supervisionato che tiene conto della frequenza e della co-occorrenza dei termini utilizzati nella raccolta di documenti. Lo stesso è poi alla base del sistema interattivo LDAvis (Sievert e Shirley, 2014), che fornisce una visione complessiva degli argomenti individuati (e del modo in cui differiscono l'uno dall'altro), consentendo allo stesso tempo un'analisi approfondita dei termini salienti associati a ciascuno di essi. Nello specifico, ogni tematica è rappresentata da un cerchio, la cui ampiezza segnala la frequenza di occorrenza nell'insieme complessivo dei documenti esaminati. Le tematiche, inoltre, assumono una posizione nello spazio cartesiano, collocandosi attorno agli assi, che individuano le due dimensioni principali che spiegano la variabilità dei dati.

## APPROFONDIMENTI

Per una trattazione più dettagliata sugli aspetti metodologici, si rimanda al Rapporto Agcom (2018), [News vs. fake nel sistema dell'informazione](#)

Per approfondimenti di carattere definitorio e tecnico in merito alla disinformazione online, si rinvia al Rapporto Agcom (2018), [Le strategie di disinformazione online e la filiera dei contenuti fake](#)

Per ulteriori approfondimenti e aggiornamenti sulle attività dell'Autorità in materia di disinformazione online, è possibile consultare l'area dedicata del [Tavolo pluralismo e piattaforme online](#)

Tutti i numeri dell'Osservatorio sulla disinformazione online sono consultabili nella sezione [Osservatori](#) del sito web di Agcom



[www.agcom.it](http://www.agcom.it)



[ses@agcom.it](mailto:ses@agcom.it)

**SERVIZIO ECONOMICO STATISTICO**